Molecular simulations of lipid droplet biogenesis

Vincent Nieto

To cite this version:

Vincent Nieto. Molecular simulations of lipid droplet biogenesis. Modeling and Simulation. Université Claude Bernard - Lyon I, 2023. English. NNT : 2023LYO10034 . tel-04482453

HAL Id: tel-04482453
https://theses.hal.science/tel-04482453

Submitted on 28 Feb 2024

THESE de DOCTORAT DE L'UNIVERSITE CLAUDE BERNARD LYON 1

Ecole Doctorale N° 205
Ecole Doctorale Interdisciplinaire Sciences Santé

Discipline: Biophysique Computationnelle

Soutenue publiquement le 09/03/2023, par:
Vincent Nieto

Simulations Moléculaires de la Biogénèse des Gouttelettes Lipidiques

---

Devant le jury composé de :

Gouet, Patrice	Professeur des universités – Université de Lyon	Président
Čopič, Alenka	Directrice de recherche - CNRS	Rapporteure
Li-Beisson, Yonghua	Directrice de recherche - CEA	Rapporteure
Barducci, Alessandro	Chargé de recherche - INSERM	Rapporteur
Foret, Lionel	Maître de conférence - CNRS	Examinateur
Loison, Claire	Chargée de recherche - CNRS	Examinatrice
Risselada, Jelger	Professeur - TU Dortmund University	Examinateur
Monticelli, Luca	Directeur de recherche - INSERM	Directeur de thèse

Université Claude Bernard Lyon 1

École Doctorale Interdisciplinaire Sciences Santé

Molecular Simulations of Lipid Droplet Biogenesis

by

Vincent Nieto

2023

Acknowledgements

I would like to thank the Université Claude Bernard de Lyon, l'École Doctorale Interdisciplinaire Sciences et Santé, and CNRS for making this thesis possible.

First of all, I am immensely grateful to Luca Monticelli, my supervisor, who has supported me in every possible way throughout the four years of my PhD. As a mentor or as confident you have taken the time to listen to me as your student or as your peer. You have taught me how to do science with rigor but above all with kindness. Before starting my thesis, I was advised to choose my supervisor wisely, and I can now say that I indeed picked the right one. Thank you for trusting in my abilities and pushing me to learn and grow.

Lionel Ballut and Patrick Fuchs for following my work unwaveringly over the course of the four years of my PhD, thank you for your advice, your time and your support.

A Claudia, Andrés y Jorge, gracias por hacer posible mi participación a un increíble proyecto científico y por introducirme al mundo de la investigación académica. Creyeron en mi trabajo, tuvieron la pacienzia de enseñarme el idioma y compartieron conmigo el gusto de la investivagión y la motivación de empezar una tesis doctoral.

Merci à Pascale qui m'a offert ma première opportunité dans la recherche et m'a transmis la curiosité pour la science et le goût pour l'aventure à l'international.

I am very grateful to my thesis jury, Alenka Čopič, Yonghua Li-Beisson, Alessandro Barducci, Lionel Forêt, Claire Loison, Jelger Risselada, Patrice Gouet, for their interest in my work and the time they will dedicate evaluating it.

Merci à Dorothée, Souad, Alexis et Samuel dont le support administratif ou informatique rend la vie de tous les membres du laboratoire plus facile.

Thanks to Chukiat from who I inherited the work on lipid droplets and to Jackson who will now continue after me. Good luck and may you achieve great discoveries.

Nika et Élie, merci pour l'implication dont vous avez fait preuve lors de vos stages et l'énergie que vous avez dédiée dans vos projets.

Cécile, mon adelphe, je suis content de t'avoir rencontré pendant cette thèse, et de t'avoir vu t'ouvrir et t'épanouir. Romuald, je suis ravi que tu nous aies rejoint avec ta bonne humeur et ton humour. Merci à tous les deux pour m'avoir permis d'avoir un refuge et un havre de paix dans votre bureau à tout moment.

Guillaume, Juliette, Jad, and Hafez, thank you for making the work environment so pleasant with your discussions and your kindness.

Thank you Lisbeth for sharing everything with your desk neighbour, from stupid questions, to snacks, to a postdoc offer, to finally Danish classes. Mange tak!

To everyone in the MOBI team, I truly enjoyed the years I spent working by your side, and I will miss you all.

Laurence, merci à toi qui ne manque jamais une occasion pour me soutenir et m'encourager.

À mes ami·es, Lucas, Audrey, Samuel, Aziliz, Noëlie, Théo, Rosa, Nathalie, Paul, Cédric, Doha, Jonathan, Fabien, Rémi, Perrine, Shaïnez et Cédric, merci pour votre soutien et vos encouragements.

Un grand merci à mes parents et à ma famille pour me soutenir et apprendre par cœur le titre de ma thèse, et je remercie d'avance celles et ceux qui seront présent·es et supporteront les heures de soutenance.

Et enfin, merci à Jules pour ton soutien inconditionel, pour me comprendre et m'accompagner au cours de ces derniers mois. Jeg elsker dig, tak for alt. Spis dit æble og lad os tage til Danmark, eventyret venter på os.

Résumé

Les gouttelettes lipidiques (GLs) sont des organites cellulaires omniprésents régulant le stockage et le métabolisme des lipides. Les GLs sont générées dans le réticulum endoplasmique (RE) et le mécanisme de leur biogenèse n'est pas compris car extrêmement difficile à caractériser expérimentalement, en raison de la petite taille et de la nature dynamique des GLs naissantes. Le dysfonctionnement de la biogenèse des GLs est lié à un large éventail de maladies. Comprendre la formation des GLs est l'objectif principal de cette thèse. Je me suis concentré sur trois aspects : (1) la partition des protéines dans les GLs naissantes, (2) l'analyse de la forme des GLs comme indicateur de leur tendance à bourgeonner, et (3) le mécanisme de bourgeonnement.

Les protéines jouent un rôle important dans la naissance et la fonction des organites. Les protéomes respectifs des GLs et du RE sont différents, bien que leurs membranes soient connectées, et les facteurs déterminant la distribution des protéines ne sont pas clairs. Nous avons utilisé des simulations de dynamique moléculaire pour étudier la partition des protéines du RE, en nous concentrant sur des protéines modèles pour imiter les protéines membranaires transmembranaires et monotopiques. Nous avons constaté que les propriétés chimiques des protéines affectent leur distribution autour des GLs naissantes et expliquent les différents protéosomes des deux organites.

Au cours de leur formation, les GLs subissent un changement radical de forme, passant de la forme de lentille d'une GL naissante intégrée dans une membrane à des GLs presque sphériques émergeant du RE. Des travaux théoriques ont permis une description mathématique de la forme en fonction du volume de la GL, et prédit que le changement de forme, correspondant au bourgeonnement, se produisant lorsque la GL croît au-delà d'un certain seuil, estimé à quelques dizaines de nm. Cependant, aucun changement de forme et aucun bourgeonnement n'ont été observés dans les simulations de GLs naissantes intégrées dans des membranes lipidiques plates. En comparant la forme des GLs simulées avec les formes théoriquement prédites, nous avons trouvé un très bon accord avec la théorie, indiquant que les simulations peuvent être utilisées pour obtenir un aperçu au niveau moléculaire des propriétés et de la dynamique de ces systèmes.

Le bourgeonnement des GLs se déroule sur une échelle de temps et de durée hors de portée des techniques de microscopie actuelles. De plus, il est communément admis que le bourgeonnement est également hors de portée des simulations, en raison des mêmes limitations (dans le sens opposé). Le bourgeonnement est donc la partie la moins comprise de la biogenèse des GLs. C'est aussi peut-être le plus important, car la direction du bourgeonnement détermine la couverture protéique et donc la fonction des GLs dans les cellules. Ici, nous avons développé une méthode de simulation hors équilibre, appelée POP-MD, pour émuler efficacement la synthèse des lipides et explorer les forces motrices et les conditions du bourgeonnement des GLs. Nous avons également modélisé la protéine clé de la biogenèse des GLs, seipin, dont les multiples rôles ne sont pas complètement clairs, et nous l'avons insérée des membranes tubulaires, présentant une morphologie, une taille et une composition chimique réalistes de la partie tubulaire du RE. POP-MD nous a permis de simuler, pour la première fois, le processus de bourgeonnement, et de déterminer quelles conditions sont réellement nécessaires pour l'observer. Nous avons constaté que la taille de la GL n'est pas un paramètre critique pour reproduire le bourgeonnement ; au contraire, l'asymétrie des feuillets est nécessaire et suffisante pour reproduire le bourgeonnement, indépendamment de tous les autres facteurs. Le couplage de la synthèse des triglycérides et des phospholipides est nécessaire pour produire des structures correctement bourgeonnées. Seipin échafaude le cou de la GL bourgeonnée et empêche la génération de défauts hydrophiles au niveau du cou de la GL, mais ce n'est pas suffisant pour produire le bourgeonnement. La composition lipidique complexe du RE facilite le mécanisme de bourgeonnement car elle fournit des membranes moins rigides, réduisant les défauts membranaires pendant la biogenèse. Dans l'ensemble, les simulations ont fourni une vue moléculaire sans précédent de la biogenèse GL, avec une résolution de longueur inférieure au nanomètre et une résolution temporelle de l'ordre de la picoseconde, ont permis l'interprétation de nombreuses données expérimentales et ont élucidé le rôle des différents composants nécessaires au processus de bourgeonnement.

Abstract

Lipid droplets (LDs) are ubiquitous cellular organelles regulating the lipid storage and metabolism. LDs are generated in the endoplasmic reticulum (ER), and the mechanism of their biogenesis is not understood because extremely challenging to characterize experimentally, due to the small size and dynamic nature of nascent LDs. Malfunctioning of LD biogenesis is linked to a wide range of diseases. Understanding the formation of LD is the main objective of this thesis. I focused on three aspects: (1) the partitioning of proteins in nascent LDs, (2) the analysis of LD shape as a proxy for their tendency to bud, and (3) the mechanism of budding.

Proteins play an important role in the birth and function of organelles. The respective proteome of LDs and ER are different, despite their membranes being connected, and the factors determining protein distribution are unclear. We used molecular dynamics simulations to study the partitioning of ER proteins, focusing on model proteins to mimic transmembrane and monotopic membrane proteins. We found that the chemical properties of proteins affect their distribution around nascent LDs and explain the different proteosomes of both organelle.

During their formation, LDs undergo a drastic change in shape, from the lens-like shape of nascent LD embedded in a bilayer membrane to nearly spherical LDs emerging from the ER. Theoretical work allowed a mathematical description of the shape as a function of LD volume, and predicted that the change in shape, corresponding to budding, occurs when the LD grows beyond a certain threshold, estimated at a few tens of nm. However, no change in shape and no budding was observed in simulations of nascent LDs embedded in flat, periodic lipid bilayer membranes. Comparing the shape of simulated LDs with the theoretically predicted shapes, we found very good agreement with the theory, indicating that simulations can be used to gain molecular level insight into the properties and dynamics of these systems.

LD budding takes place on a time and length scale beyond the reach for current microscopy techniques. Also, it is commonly believed that budding is out of reach for simulations too, due to the same limitations (in the opposite sense). Budding is therefore the least understood part of lipid droplet biogenesis. It is also possibly the most important, as the direction of budding determines the protein coverage and therefore the function of LDs in cells. Here we developed a non-equilibrium simulation method, coined POP-MD, to effectively emulate lipid synthesis and explore the driving forces and the conditions for LD budding. We also modeled the key protein in LD biogenesis, seipin, whose multiple roles are not completely clear, and inserted it tubular bilayer membranes, featuring realistic morphology, size, and chemical composition of the tubular part of the ER. POP-MD allowed us to simulate, for the first time, the budding process, and determine which conditions are actually necessary to observe it. We found that LD size is not a critical parameter to reproduce budding; instead, leaflet asymmetry is necessary and sufficient to reproduce budding, independently of all other factors. Coupling of triglyceride and phospholipid synthesis is necessary to produce correctly budded structures. Seipin scaffolds the neck of the budded LD and prevents the generation of hydrophilic defects at the LD neck, but it is not sufficient to produce budding. The complex ER lipid composition facilitates the budding mechanism as it provides softer membranes, reducing membrane defects during biogenesis. Overall, simulations provided an unprecedented molecular view of LD biogenesis, with sub-nanometer length resolution and picosecond time resolution, enabled the interpretation of numerous experimental data, and elucidated the role of the different components required for the budding process.

Contents

List of Figures

List of Tables

Abbreviations

ADH Alcohol Dehydrogenase.

AGPAT 1-Acylglycerol-3-Phosphate Acyltransferase.

ARF1 ADP-ribosylation factor 1.

ATGL Adipose Triglyceride Lipase.

CHOL Cholesterol.

CIDE Cell Death-Inducing DFFA-Like Effector.

CTD C-terminal domain.

DG Diacylglycerol.

DGAT Diacylglycerol Acyltransferase.

DOPA Dioleoylphosphatidic acid.

DOPC Dipalmitoylphosphatidylcholine.

DOPE Dioleoylphosphatidylethanolamine.

DOPS Dioleoylphosphatidylserine.

ER Endoplasmic Reticulum.

ERGIC ER-Golgi intermediate compartment.

FA Fatty acids.

FRET Fluorescence Resonance Energy Transfer.

GPAT Glycerol-3-Phosphate Acyltransferase.

GTP Guanosine triphosphate.

HD Hydrophobic domain.

HSL Hormone sensitive lipase.

LD Lipid Droplet.

MD Molecular Dynamics.

MGAT Monoacylglycerol Acyltransferase.

MGL Monoacylglycerol lipase.

OLE Oleic acid.

PA Phosphatidic acid.

PC Phosphatidylcholine.

PE Phosphatidylethanolamine.

PI Phosphatidylinositol.

PLIN Perilipin.

POPC Palmitoyloleoylphosphatidylcholine.

POPI Palmitoyloleoylphosphatidylinositol.

PS Phosphatidylserine.

SM Sphingomyelin.

TG Triacylglycerol.

TM Transmembrane.

TMD Transmembrane domain.

Chapter 1

Introduction

Lipid droplets are ubiquitous intracellular organelles, found in most cells of most living organisms, from bacteria to humans, responsible for lipid storage in cells (1). They are sometimes referred to as adiposomes, lipid bodies, or oil bodies. In eukaryotic cells, lipid droplets (LDs) are formed in the endoplasmic reticulum (ER) and play a major role in lipid and energy metabolism (1; 2). In LDs, different types of lipids can be converted into one another: phospholipids, that are the main constituents of cell membranes, can be converted to neutral lipids (such as triglycerides) for storage, and neutral lipids can later be converted to phospholipids or other lipids with different cellular functions. To fulfill their roles, specific mechanisms regulate their formation, growth or shrinkage, depending on the needs of the cell. The study of LDs biogenesis is relevant not only from the standpoint of fundamental biology, but also for medicine, as many metabolic diseases are linked to malfunctioning or anomalies of LD biogenesis: obesity, steatosis, lipodystrophy, diabetes, and certain types of cancer (3). LDs also act as deposits for otherwise toxic lipids, and their mechanism of formation is hijacked by certain viruses, that use LDs to replicate – hence the relevance of LDs for viral infections (1). Despite their relevance, the specific mechanisms regulating LD biogenesis are not well understood. The overall goal of my thesis is to contribute to our understanding of the molecular mechanisms of lipid droplet formation, using computer simulations.

1.1 Lipid droplet structure

From a physical-chemical standpoint, LDs are emulsions of oil drops in water (4; 2). They consist of a core of neutral lipids, mainly triacylglycerols (TGs) and cholesteryl esters; and they are bounded by a monolayer of phospholipids and proteins (5; 6). In cells, LDs are found with a wide range of sizes: from tens of nanometers to tens of millimeters, depending on the nature of the organism, the cell type, the metabolic state, and the stage of LD formation (2). For a "naked" oil-water interface, the interfacial tension is high (around $20-30$ mN/m, depending on the type of oil), which makes oil droplets unstable and prone to fusion (coalescence). In LDs, a monolayer of phospholipids and proteins acts as a surfactant, lowering the surface tension and stabilizing the emulsion. Increasing the quantity of surfactant at the interface reduces the surface area exposed to water, hence reduces the surface tension and increases the stability of the LD. However, the interfacial monolayer presents a lower phospholipid density than the bilayer (7; 8), therefore it is subject to a higher surface tension (typically around $1-2$ mN/m (7; 8)) compared to bilayer membranes, for which the surface tension is typically of the order of 10^{-2} mN/m. Such interfacial tension are critical for cell homeostasis and for proteins targeting to the surface of LDs (9).

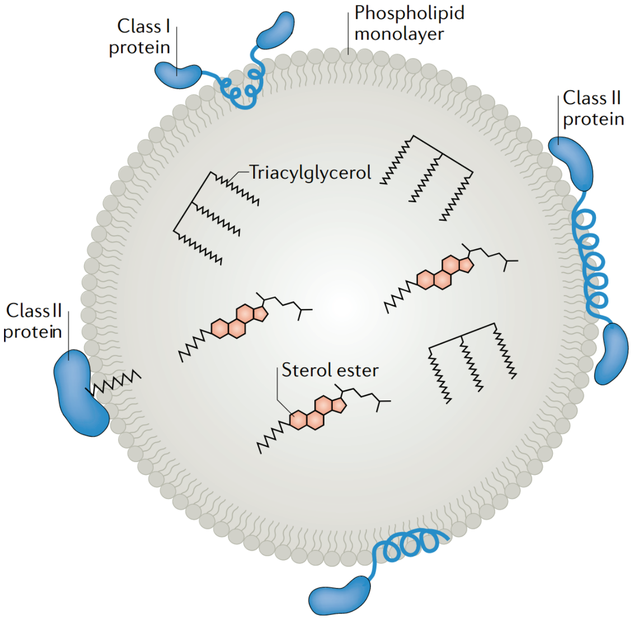


Figure 1.1: Schematic overview of lipid droplet structure. The core is composed of triacylglycerols and sterol esters. The monolayer is composed of phospholipids with their head groups oriented towards the cytosol and their acyl chains oriented toward the LD core. Proteins from various origins are also associated on the surface of LDs, described in section 1.6. Figure taken from (10).

1.2 Lipid droplet formation

LDs are substantially different from all other cellular organelles as they are bounded by a monolayer of phospholipids, instead of a bilayer (11). The unique structure of LDs stems from the process of LD formation: in eukaryotic cells, LD form in the endoplasmic reticulum (ER) bilayer. While such mechanism is not fully understood, some aspects are now rather well accepted in the biology community. At low concentration, neutral lipids are dissolved in the ER bilayer. As more neutral lipids are synthesized and their concentration rises, they phase separate, and form a lens-shaped blister in the bilayer, containing triacylglycerol and sterol esters. This event is referred to as lipid droplet nucleation (12). This symmetric lens is generally referred to as nascent lipid droplet. As the lens grows and reaches a certain size, the droplet changes in shape and becomes spherical and asymmetric, swelling toward the cytosol. This process is analogous to dewetting in systems with three fluids (12), and the emergence of the (nearly spherical) LD from the ER membrane is generally referred to as budding. The budded droplet is then bounded by the cytosolic monolayer of the ER. In theory, the budding size depends on local surface tension and line tensions as well as oil wetting properties, and surfactant type. Low tension in the surrounding bilayer membrane has been found to favor neutral lipid removal from the bilayer and to promote budding (13). The presence of positively curved phospholipids favors a lower tension and helps the budding of small LDs (14; 13). However, multiple LD proteins come into play, compensating and counterbalancing the effect of lipids during budding. Mature budded LDs can eventually detach from the ER bilayer. However, this latter step is not necessary and rarely observed, as most LDs maintain a contact point with the ER (15).

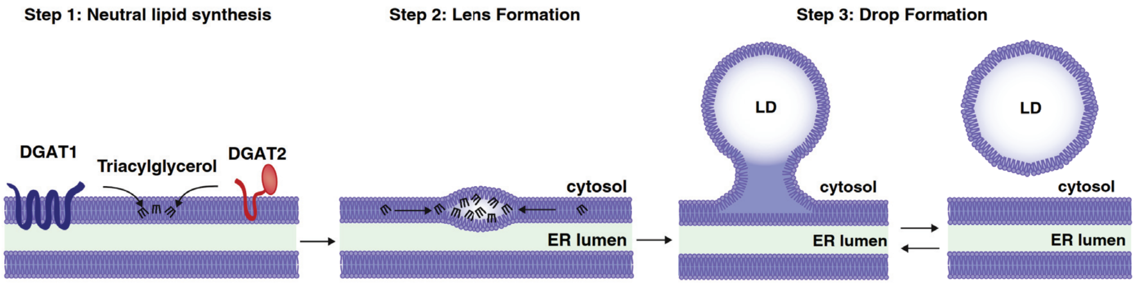


Figure 1.2: Model of lipid droplet formation. The four stages of lipid droplets formation are: nucleation, nascent LD, budded LD, and detached LD. Figure taken from (12)

1.3 Lipid droplet composition

The phospholipid monolayer on the surface of LDs consists mostly of phosphatidylcholine (PC) and phosphatidylethanolamine (PE), and also contains small amounts of phosphatidylinositol (PI), phosphatidylserine (PS), phosphatidic acid (PA), cholesterol, sphingomyelin (SM) and lysophospholipids(16). During LD biogenesis, and generally also after budding, there is a continuity between the bilayer membrane of the ER and the LD monolayer. However, somewhat surprisingly, the phospholipid composition differs in the ER bilayers and LD monolayers: the LDs surface contains higher amounts of phosphatidylcholine (PC), and lower amounts of cholesterol and sphingomyelin than ER membranes (16). The LD surface also contains cosurfactants, such as unesterified sterols, diacylglycerol (DG), or fatty acids (FAs)(17), that make the distribution of phospholipids more even, and by doing so, lower the surface tension (2).

1.4 Lipid droplet growth and shrinkage

Neutral lipids synthesis is the first event not only in LD nucleation but also LD growth. Triacylglycerol synthesis enzymes synthesize neutral lipids at LD surfaces (18). Besides oil synthesis, LD growth also relies on ripening and fusion with other LDs. Ripening occurs when multiple LDs are present in a continuous bilayer membrane, but does not require any contact between those LDs: larger LDs have lower Laplace pressure than smaller LDs, and Laplace pressure (corresponding to the difference between internal and external pressure: $\Delta P = P_{in} - P_{\theta} = 2\gamma/r$) "pushes" oil molecules from smaller LDs to larger LDs. As a consequence, ripening causes shrinkage of smaller LDs and growth of larger LDs. Ripening is a slow process and vanishing of a small LD can take several minutes (19).

Fusion of LDs, instead, requires close proximity between two LDs, and occurs when two adjacent LDs come into contact. Fusion depends on several factors, for example temperature: at high temperature, thermal fluctuations overcome the energy barrier for fusion; even though in mammalian cells, temperature is stable. Fusion also depends on the phospholipid composition, as it changes the spontaneous curvature of the monolayer curvature. An excess of negatively curved phospholipids supports pore formation and is thought to favor fusion (20). After pore opening between two LDs, fusion can occur within milliseconds (20).

LDs can be subject to destabilization even with extensive coverage in phospholipids.

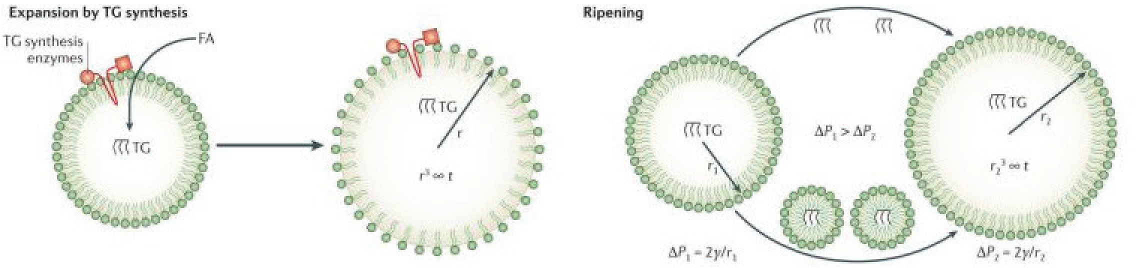


Figure 1.3: Processes governing changes in lipid droplet size. TG synthesis is mediated by proteins located in the LD monolayer or the neighboring ER membrane connected to the LD; synthezed TGs directly increase the size of the LD core. In ripening, TG travels from a smaller LD to a larger one by diffusing through the continuous ER membrane connected two neighboring LDs. Figure taken from (2)

1.5 Lipid metabolism

Lipid synthesis is a vital process since accumulations of free fatty acids (FA) is toxic in cells. FA must be converted into triacylglycerols (TGs). The glycerolipid synthesis pathway involves different enzymes: Glycerol-3-Phosphate Acyltransferase (GPAT), 1-Acylglycerol-3-Phosphate Acyltransferase (AGPAT), lipin, Monoacylglycerol Acyltransferase (MGAT), Diacylglycerol Acyltransferase (DGAT) (21).

$$\begin{tikzcd}
\text{G3P} \arrow[r, "\text{GPAT}"'] & \text{LPA} \arrow[r, "\text{AGPAT}"'] & \text{PA} \arrow[r, "\text{Lipin}"'] & \text{DG} \arrow[r, "\text{DGAT}"'] & \text{TG} \\
 & & & \text{MAG} \arrow[u, "\text{MGAT}"'] &
\end{tikzcd}$$

TGs are the main storage lipids in cells, and they are broken down in order to be used as an energy source. Lipolysis involves the hydrolysis of TGs and generally requires multiple enzymes, named lipases: adipose triglyceride lipase (ATGL or PNPLA2), hormone sensitive lipase (HLS), and monoacylglycerol lipase (MGL) (21). Lipolysis is strictly controlled in adipocytes, but the process is less understood in other tissues (21). Since some tissues can present low levels of ATGL, HSL and MGL, other lipases are likely implied in lipolysis (22).

$$\mathrm{TG} \xrightarrow[\mathrm{ATGL}]{} \mathrm{DG} + \mathrm{FA}$$
$$\mathrm{DG} \xrightarrow[\mathrm{HSL}]{} \mathrm{MAG} + \mathrm{FA}$$
$$\mathrm{MAG} \xrightarrow[\mathrm{MGL}]{} C_3H_8O_3 + \mathrm{FA}$$

These reactions show how TG is hydrolysed into diacylglycerol (DG), then into monoacylglycerol (MAG), and finally into fatty acids and glycerol ($C_3H_8O_3$).

Alternatively, PA can then be used to synthesize phospholipids, important components of cell membranes. In the process of phospholipid synthesis, PA is modified by adding additional fatty acids and polar head groups to form various types of phospholipids.

1.6 Lipid droplet proteome

The LD proteome consists of 100–150 proteins in mammalian cells, mainly enzymes involved in lipid metabolism (23). Besides the proteins responsible for the synthesis and the hydrolysis of neutral lipids, many other proteins have been identified in LDs, linked to lipid droplet functions in storage or metabolism, or to other functions, such as membrane trafficking or protein degradation. The surface of mature LDs contains proteins from two origins: Class I proteins from the ER, and class II from the cytosol. Interestingly, Class I proteins found in LDs are different from ER proteins: during the formation of LDs, only a fraction of ER proteins diffuse to the nascent lipid droplet monolayer, while others stay in the ER membrane (24). The monolayer membrane of LDs is one of the reasons for this protein non-homogeneous distribution. Bilayer membranes retain proteins with transmembrane domains and hydrophilic regions on both sides of the ER membrane. Instead, so-called monotopic membrane proteins, consisting of a hydrophobic domain bounded by hydrophilic domain(s) on the same side of the ER membrane, generally partition to the LD monolayer (25; 26; 27; 28; 29). These include hydrophobic hairpins. Class II proteins are generally partially soluble cytosolic proteins, and interact with the LD surface monolayer via amphipathic helices. These proteins also act as surfactants, reducing the surface tension of the LD monolayer, as they occupy the LD surface. And conversely, LDs presenting a high surface tension may have hydrophobic patches on their surface. The exposed oil phase may then trigger protein binding (25; 26; 27; 28; 29).

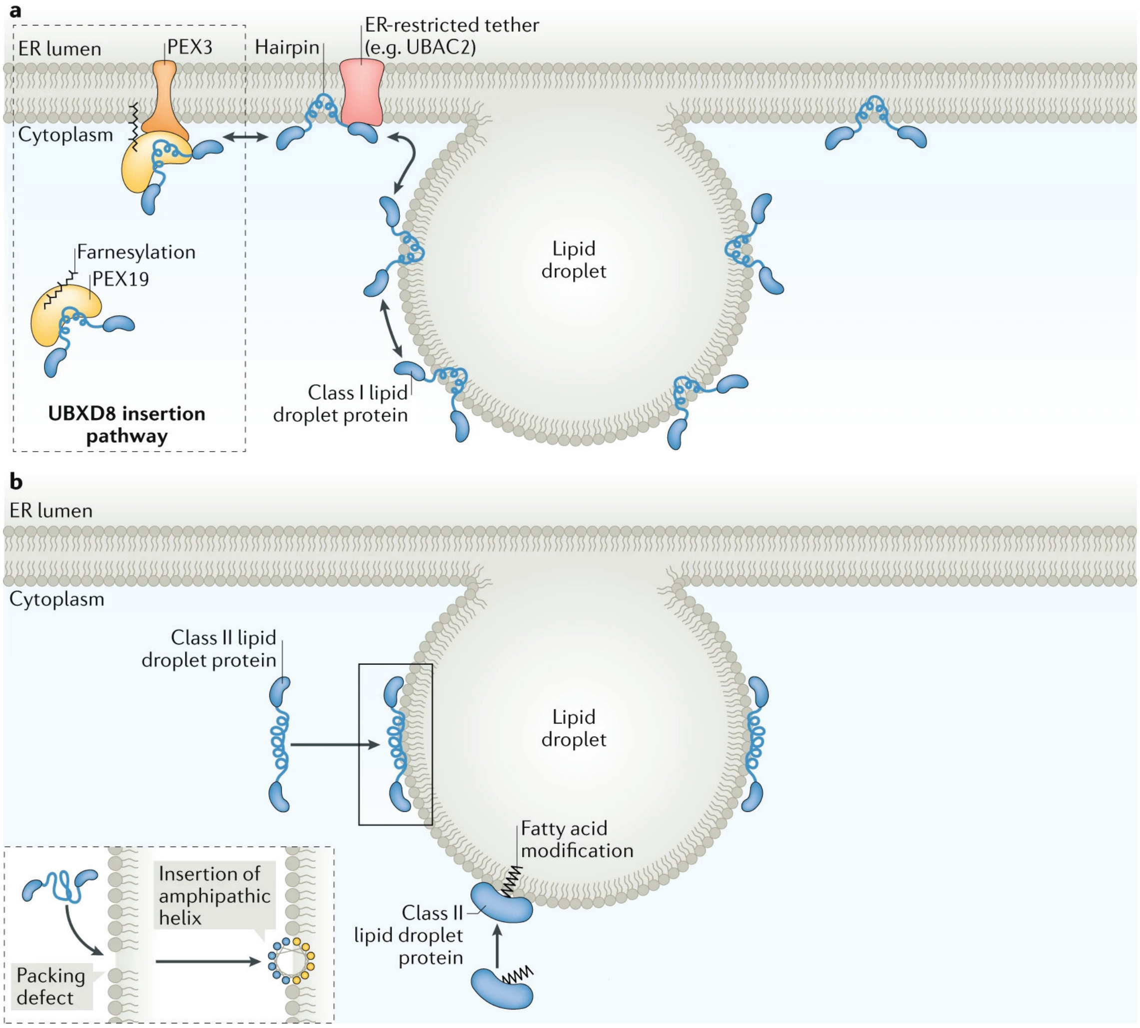


Figure 1.4: The two classes of proteins involved in the insertion process of proteins into LDs. Class I proteins are inserted into the ER and laterally diffuse into forming LDs. Their insertion can be facilitated by the association with PEX19 and its ER receptor PEX3. Class II proteins are directly inserted from the cytosol into LDs using amphipathic helices or fatty acid modifications. Figure taken from (10).

Seipin is an evolutionarily conserved transmembrane protein essential for LD birth, growth, and function. Seipin is at the center of a network of protein interactions playing important roles in LDs formations, cytoskeleton remodeling, calcium handling, lipolysis, and phospholipid homeostasis (30). Seipin is essential for normal LD formation: in the absence of Seipin, aberrant LDs form in cells, either tiny or giant LDs emerge from the ER (31; 32; 30), resulting in an abnormal LD phenotype. While many investigations have been carried out exploring the different functions of seipin, information on the molecular-level mechanisms of action and other microscopic level information on seipin is limited. Seipin structure has been determined by cryogenic electronic microscopy for four organisms: human (33), fly (34) and yeasts (35; 36). All three structures share strong similarity in the luminal domain. Human seipin forms a undecamer ring, with a diameter of 150 Å, while the Drosophila seipin forms dodecamer ring, and yeast Seipin forms a decamer ring. The largest fraction of the protein oilgomer consists of its luminal domain. Each unit of the seipin oligomer has two transmembrane helices and one cytosolic domain.

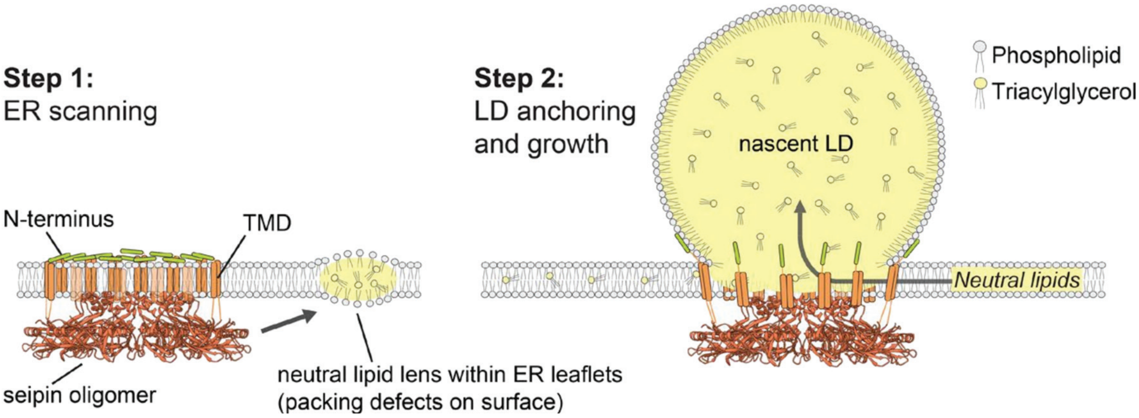


Figure 1.5: Molecular structure of seipin during LD formation. The cryo-EM structure of the luminal domain of the seipin oligomer is shown, along with the transmembrane domains (TMD), and the N-terminus segments. Step 1: seipin attracts NLs in the ER membrane with its oligomeric ring. Step 2: seipin scaffolds the neck formed at the LD-ER contact point. Figure taken from (34).

Based on the available structures, simulations have been carried out leading to the hypothesis that Seipin may promote LD nucleation by attracting and trapping NLs (37; 38). Through weak interactions with NLs, seipin TM domains reduce locally the diffusion of NLs, and increases the local concentration of NLs. At the same time, hydrophobic helices in the luminal domain have a high affinity and bind to monolayers enveloping neutral lipids (30), and also trap NLs through strong direct interactions. Subsequently, NLs cluster within the seipin ring leading to LD nucleation, lowering the NL concentration threshold necessary for LD nucleation (38; 37). Through these interactions with NLs, seipin prevents LD shrinkage by ripening (32) and promotes growth of small LDs it nucleates. Since seipin is a 15-nm-wide ring oligomer, the diameter of the LD-ER contact is predicted to be about 15 nm (32). Upon LD budding, at the LD-ER junction, the membrane forms a “neck”; seipin scaffolds the LD-ER contact point assuring stability of the neck (32). Additionally, seipin binds and regulates phosphatidic acid, a phospholipid negatively charged and cone-shaped, i.e., having a negative curvature (34). Seipin luminal domain also contains a 8-stranded β-sandwich fold that forms a C2-like domain targeting membranes binding calcium ions (34). Calcium ions may interact with two PAs to form a neutral highly negatively curved complex(34) necessary to form the neck a the LD-ER contact point.

Seipin also interacts with a number of other proteins, and forms an LD protein complex at the point of formation of the nascent LD (32; 30). Since the monolayer of LDs presents a lower phospholipid surface density compared to bilayer membranes (8), packing defects favor the binding of proteins. Unsurprisingly, seipin interacts and affects the activity of proteins responsible of lipid metabolism. Seipin may interact with GPATs, responsible of the synthesis of LPA, and antagonizes their activity (22). GPAT and DGAT, responsible of the synthesis of TG, translocates to the LD surface and supports LD expansion (22).

Proteins linked to ER shape also impact LD (39; 40; 41). For example, LDAF1 (also named Promethin), Ldo45, or TMEM159, form complexes with Seipin upon LD nucleation. Structure-wise, LDAF1 forms a double hairpin that can bind to monolayers (42). It acts to reduce the cost of forming many new LDs(30). As Seipin allows the formation of few large LDs, together, LDAF1 and Seipin allow the formation of normal sized LDs. Upon budding, LDAF1 dissociates from Seipin and coats the LD monolayer (42). In yeast, LDAF1 is required for the functioning of Seipin. Yeast Seipin is deprived of the 8-strand β-sandwich and the hydrophobic helix domains than have been found in human and drosophila Seipin. The yeast equivalent of LDAF1 forms a complex with Seipin replacing the function of these AH (35).

Fat Storage-Inducing Transmembrane protein (FITM) plays a role in fat accumulation in adipocytes and during LD biogenesis, and regulates the ER-LD crosstalk during LD budding (22). FITM is also recruited by Seipin to facilitate budding: it reduces DG levels in the ER (10) by allowing the sequestration of DG around Seipin during LD biogenesis (43), that will be converted into TGs. Seipin indeed recruits DGAT, responsible of the production of TG.

Pex is another protein affecting the ER membrane shape, and can serve as insertion pathway for Class I proteins (10). During LD biogenesis, Pex helps LD formation by deforming ER membrane subdomain (44; 45).

In yeast, Perilipins are also recruited by Seipin to LD surface (43): they play a role in scaffolding and stabilizing LD. Perilipins facilitate budding by accessing to the cytosolic side of the LD, changing balance of tension in the membrane (46). Phosphorylation of perilipin1 also activates CGI-58 that activates ATGL, the lipase responsible for the degradation of TG into DG (47).

CCTα activation is required to maintain PL homeostasis during LD growth (10). Synthesis of PC lipid is regulated by the enzyme CCTα, catalyzing PC synthesis when bound to PC-deficient LDs (48).

Above were only mentioned proteins directly involved in lipid metabolism or in LD formation. Many more proteins have an activity related to LDs. Even more considering the processes LDs take part in.

1.7 Non-canonical lipid droplet functions

LDs are involved in a wide range of functions and process, even some where they are not the primary actors, for example in the immune system. Here, LDs have been found to be involved in the innate immune system (49). LDs form a first-line intracellular defense. When presented with danger signals from the cell, LDs can respond be reprogramming the cell metabolism. LDs are also at the origin of protein response antimicrobial mechanisms. Danger-response complexes assembled on the surface of LDs. LDs also benefit to cells subjected to infections coordinating different immune systems (49). LDs attract pathogens and are able to sequestrate cytotoxic compounds. LDs are central players in metabolic adaptation to infection.

LDs play a protective role against endoplasmic reticulum (ER) stress, which occurs when there are imbalances in protein folding, calcium uptake, or lipid composition in the ER. The accumulation of unfolded proteins triggers the unfolded protein response (UPR) (50), a cellular adaptive response that aims to restore ER homeostasis. Disruption of LD biogenesis or overwhelming their storage capacity leads to UPR activation (51; 52). LDs serve to buffer excess fatty acids, prevent UPR activation, and maintain ER homeostasis (10).

During autophagy induced by nutrient deprivation, a portion of the lipids released from the breakdown of proteins and organelles is re-esterified to form triacylglycerol, which is then stored in LDs (53; 54). This process depends on the autophagy machinery and the enzyme DGAT1 (54). Without LDs, fatty acids accumulate as acylcarnitine, which disrupts mitochondrial function, impairs respiration, and leads to cell death. LDs act as a buffer, preventing the lipotoxic damage caused by high levels of acylcarnitines and maintaining mitochondrial integrity (10).

1.8 Knowledge gaps and open questions

Despite recent advances uncovering LDs role, functions and interactions, many aspects of LDs biology remain misunderstood. First and foremost, many open questions remain on the biogenesis of LDs. Nucleation cannot be observed directly via experimental techniques, but is very easy to simulate as it requires only small systems with little amount of neutral lipids; indeed LD nucleation has been simulated by several groups (38; 38; 37; 4; 55; 56). Nascent LDs budding has never been observed experimentally and experimental observations will probably remain difficult for the foreseeable future, due to limitations in the time and length scale of current microscopy techniques applied to fluid systems. Contrary to nucleation, budding is also very difficult to study by simulations, as it requires the ability to grow LDs out of equilibrium to large sizes; the threshold size for LD budding is unknown, but probably lies in the range of 30-50 nm, based on cryo-EM pictures of lens-shaped nascent LDs embedded in ER membranes (14). Change in LD shape are also very challenging for experimental methods, when referred to the early stages of biogenesis, when the nascent LDs have nanometer size. How does the change from a symmetric nascent LD to a spherical budded LD occur? Does budding occur spontaneously by addition of TG, or it does require synthesis of phospholipids, or the presence of specific proteins? Many proteins are involved in LD formation, but their molecular mechanism remains to be understood, along with the mechanism of LD budding.

Chapter 2

Methods

2.1 Molecular simulations

The most accurate and most general theory currently available to describe the motion of atoms and molecules is relativistic quantum dynamics (57). However, such general theory is very difficult to apply in the case of large molecular systems, due to the enormous computational complexity, and approximations are necessary. If a description of chemical reactions (i.e., bond formation or bond breaking) is desired, the motion of electron needs to be taken into account, and this is generally done by quantum mechanical methods (57). A vast number of approximate methods has been developed for this aim, including wavefunction methods and density functional methods. If chemical reactions are not relevant for the problem at hand, then the motion and distribution of electrons do not need to be treated explicitly, and classical mechanics methods are sufficient to predict the motion of the particles in the system (57). A large number of simulation methods has been developed based classical mechanics (58). In the following, I will provide a brief overview of common classical mechanics methods, and then I will focus on the specific methodologies used in this thesis work.

2.2 Molecular mechanics simulations methods

In molecular mechanics methods, forces among atoms are calculated via a so-called force field. Force fields generally ignore electrons, and forces among atoms are calculated based on the positions of their nuclei, and applied to the same nuclei. I will briefly describe the most common approximations made to calculated such forces in section 2.4. Once the forces are known, the system may evolve according to some equation of motion. Different possibilities exist, summarized briefly in the following section.

Energy minimization (EM) consists in finding the configuration of the particles in the system with the lowest energy possible. This is achieved by optimizing the arrangement of all particles, by finding the most favorable bonds and angles within all molecules, and reducing the electrostatics and Van der Walls forces between the particles within the system. Energy Minimization is useful to find a local energetic minimum: the gradient of the potential energy of the system in computed, then atoms are moved in the direction that reduces the potential energy the most. In such process, the energy of the system can only decrease, therefore energy barrier cannot be overcome and the system moves toward the closes local minimum of the potential energy. The evolution of the system configuration during an energy minimization process does not reflect realistically the evolution of the system, but the result leads to an energetically more favorable arrangement of the system. In practice, energy minimization is used to find configurations with close contacts among molecules, which would lead to a high energy and often numerical instability in other types of simulations.

Monte Carlo (MC) is a method that relies on statistical mechanics systems, and aims at finding stochastically the most probable configurations of a system (58; 59). From a starting conformation, a Monte Carlo simulation generates new conformations by making random changes in the positions of the atoms in the system. The moves are accepted or rejected based on an energetic criterion, usually the Monte Carlo Metropolis criterion: if the new configuration decreases the energy of the system, it will always be accepted; if the new configuration has a higher energy, it will be accepted or not depending on the probability that changes depending on the energy levels of the conformation. With this type of simulation, an effective sampling of equilibrium states can be achieved (58; 59). But because changes are stochastic, conformations are not connected in time, and no information about the dynamics of the system by be obtained for a simulation. For this reason, MC can only be used for the prediction of static, not dynamic properties.

Molecular dynamics (MD) is a popular molecular mechanics method and probably the most common types of molecular simulation in biology (58; 59). It is used to give an insight into the time evolution of molecular systems where the structure of the molecules is known. The evolution of the system is determined by solving Newton's equation of motion for all particles, which provides both positions and velocities of all atoms as a function of time. The method is applied both in the case of an explicit and an implicit representation of the solvent. Applied to systems where there are no chemical reactions, MD can give a realistic representation of structures, dynamics, and interactions. This insight allows to understand the evolution in systems interactions and changes in structures and stability of conformations. Methods based on molecular dynamics can also be used to calculate free energy changes in molecular systems (58; 59). MD is the method of choice for this thesis work, and will be described in more detail in the next section.

Langevin dynamics (LD) is a mathematical modeling method of the dynamics of molecular systems, using on simplified models based on stochastic differential equations (58; 59). The equation of motion contains a friction term, which decreases the velocities in the system, and random forces, to compensate for the (kinetic) energy lost by friction. The method is generally used when the solvent is not represented explicitly, but nothing forbids to use it also when the solvent is considered explicitly; in this case the algorithm acts as a thermostat. In both cases, the dynamics of the system is non-Newtonian and should be considered carefully.

Brownian dynamics (BD) is mathematically similar to LD, in that it includes random forces. In fact, Brownian dynamics is a particular simplification of Langevin dynamics where conservative forces (calculated via a force field) are neglected. It is used to describe the random motion of particles in a fluid.

2.3 Basic principles of molecular dynamics simulations

Equation of motion

MD simulations seek to describe the position of the atoms over time. Numerical integration of Newton's equation of motion requires discretization of time in short intervals, called time steps, indicated as δt. Starting from an initial time t_{init}, the position of the particles can be calculated at subsequent times by integrating the equation of motion in discrete time steps δt. Computing one time step after the other results in a trajectory, i.e., the position of all atoms (or particles) of the system over time. In order to deduce the motion of the particles from one moment to the next, a mathematical description needs to connect the position x of the particles between time t and time $t + \delta t$. At a moment t, the positions and velocities of all the particles are known, and the potential energy in each particle can be computed from the force field (see section 2.4). From the gradient of the potential energy, the forces F can be deduced and applied to each atom:

$$F = -\nabla V \tag{2.1}$$

where V is provided by the force field. Since $F = ma$, accelerations a of all particle can be obtained (here, m is the mass of the particle). In order to obtain positions and velocities at subsequent times, we need an algorithm that can integrate Newton's equation of motion numerically, using discrete time steps. Several algorithms have been developed to this purpose, and in general they are based on Taylor series expansions of positions and velocities:

$$x(t+\delta t) = x(t) + \dot{x}(t)\delta t + \frac{1}{2}\ddot{x}(t)\delta t^2 + \ldots \tag{2.2}$$

$$\dot{x}(t+\delta t) = \dot{x}(t) + \ddot{x}(t)\delta t + \frac{1}{2}\dddot{x}(t)\delta t^2 + \ldots \tag{2.3}$$

where $x(t)$ is the vector of the positions of all atoms at time t, $\dot{x}(t)$ is the vector of the time derivative of the positions, etc. One of the most commonly used algorithms for the numerical integration of the equation of motion is the leap-frog algorithm. Here, positions and velocities are not calculated at the same time:

$$x(t+\delta t) = x(t) + \dot{x}(t+\frac{1}{2}\delta t)\delta t \tag{2.4}$$

$$\dot{x}(t+\frac{1}{2}\delta t) = \dot{x}(t-\frac{1}{2}\delta t) + \ddot{x}(t)\delta t \tag{2.5}$$

Another widely used integration algorithm is known as Velocity-Verlet:

$$x(t+\delta t) = x(t) + \dot{x}(t)\delta t + \frac{1}{2}\ddot{x}(t)\delta t^2 \tag{2.6}$$

$$\dot{x}(t+\delta t) = \dot{x}(t) + \left(\ddot{x}(t) + \frac{1}{2}\ddot{x}(t+\delta t)\right)\delta t \tag{2.7}$$

Whichever algorithm is chosen, the length of the integration time step δt is essential for the success of the simulation. The shortest the time step, the more accurate the integration: an excessively long time step δt may results in overlapping particles, hence high forces and numerical instability. At the same time, having a too short time step makes the calculation of a trajectory more computationally expensive, as more steps imply more calculations for the same duration. The most suitable appropriate time step is a trade-off between efficiency (speed) and accuracy, and it is determined by the fastest motion in the system. The fastest motion depends on the nature of the particles in the simulation, i.e., depends on the force field. In a typical all-atom force field, the vibration of the lightest atoms is the fastest motion; for bonds involving hydrogen atoms, the period vibration is on the order of 10 fs. In order to reproduce the vibration, the time step should be on the order of 0.5 fs (59).

2.4 Force fields

Force fields are sets of equations and related parameters that allow to calculate potential energies from the positions in a system of atoms. Forces among atoms are quantum mechanical in nature, and classical force fields aim at providing accurate representations of those forces at much lower computational cost, without resorting to quantum mechanical calculations. Force fields are therefore the result of numerous approximations, and simplify the descriptions of interactions and the estimation of the forces between particles (59).

In principle, all terms in a force field can be derived from fitting forces calculated at the quantum mechanical level. In practice, for systems in condensed phases (i.e., solids and liquids), accurate quantum mechanical descriptions of forces are difficult to obtain, as they would require calculations on large number of atoms. Therefore, quite often force fields are parameterized based only in part on QM calculations. Parameters describing bonds and angles in molecules can be safely derived from QM calculations. On the other hand, forces among atoms in different molecules require fitting to experimental data, for example thermodynamic properties of simple systems (e.g., pure liquids).

Depending on the nature of the systems of interest, on the application, and on the time scale and the spatial scale, many force fields exist. All-atom force fields represent and define parameters for all the atoms in the systems; unified atoms force fields represent mostly non-hydrogen atoms: non-polar hydrogens are incorporated into the adjacent heavy atom, while only polar hydrogens (connected to nitrogen, oxygen, etc.) are represented explicitly. Coarse-grained force fields use particles or beads that represent a group of heavy atoms.

Force fields typically contain bonded terms, acting between atoms connected via 1, 2, or 3 covalent bonds, and non-bonded terms, acting among any atoms in the system, except the ones listed above (directly connected). The computation of the potential energy is decomposed into the addition of these two terms:

$$E_{\text{tot}} = E_{\text{bonded}} + E_{\text{non-bonded}} \tag{2.8}$$

Bonded interactions are the sum of intramolecular energies coming from bonds, angles and dihedrals. Non-bonded interactions are the sum of intermolecular energies coming from electrostatics and Van der Waals forces.

In bonded terms, bond stretching can be modeled using a harmonic potential or the Morse potential. The Morse potential become zero at large distances, therefore bond can break. In most cases, harmonic potentials are used instead, restraining the bond to an ideal length.

$$V_{\text{bond}} = \frac{1}{2} k_{\text{bond}} (x - x_0)^2 \tag{2.9}$$

The bending of angles is also usually modelled using harmonic functions.

$$V_{\text{angle}} = \frac{1}{2} k_{\text{angle}} (q - q_0)^2 \tag{2.10}$$

Dihedral potentials are modelled using periodic functions (e.g., combinations of trigonometric functions), which vary depending on the nature of the dihedral. Improper dihedrals can also be defined, used to conserve planarity of some molecules, such as aromatic rings, or chirality around stereocenters.

Non-bonded intermolecular terms take into account the interactions of each atom with all atoms in other molecules, and also most atoms in the same molecule (except the ones connected by 1 or 2 bonds). In the vast majority of force fields, non-bonded terms are described as a sum of pairwise interactions, i.e., interactions between two atoms. The electrostatics term can be described using Coulomb's law:

$$V_c = k_e \frac{q_1 \cdot q_2}{e_r r} \tag{2.11}$$

where k_e is Coulomb's constant, q_1 and q_2 are the charges of the particles, e_r is the dielectric constant of the medium and r the distance between the particles. Electrostatic interactions exist between ions, carrying an integer charge, but also among atoms that are part of electrically neutral molecules. This is because, in general electronic distributions in molecules are uneven, i.e., some nuclei attract more electrons (i.e., they are more electronegative) than others, which gives rise to electric dipoles (and, in some cases, higher multipoles). Since force fields do not contain an explicit representation of electrons, the uneven electronic distribution is represented via partial charges placed on the nuclei. Such partial charges are parameterized to reproduce the electrostatic potential of the molecule, as calculated via quantum mechanical methods (and possibly optimized by fitting thermodynamic properties).

In force fields, partial charges are fixed, i.e., constant in time. In real systems, electronic distributions (and therefore the electrostatic potential) can change significantly as a result of conformational changes or chenges in the chemical environment (i.e., presence of a polar vs non-polar solvent). Fixed partial charges are therefore an approximation, only valid when conformational changes or changes in chemical environment do not lead to a significantly different electronic distribution. Fortunately, this is the case for the majority of molecules of biological interest (with notable exceptions, e.g., highly charged systems, that will not be discussed here).

When a more accurate description of electrostatic interactions is needed, a polarization term can be included in the force field. This is rarely used for biological macromolecules due to its computational cost and the limited advantages, but is considered indispensable in some specific areas of chemistry and materials science (e.g., ionic liquids, polyelectrolytes, etc.).

Even in the absence of molecular dipoles and multipoles, atoms still attract each other via dispersion interactions, due to the interaction between transient dipoles (generated by the rapid motion of electrons in molecules) and induced dipoles (induced by the presence of a nearby dipole). Moreover, atoms repel each other at very short distances, since electronic distributions of different atoms cannot be overlapped (Pauli exclusion principle). Such combination of attraction (due to dispersion interactions) and repulsion is generally represented by the Lennard-Jones potential:

$$V_{VdW} = 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right) \tag{2.12}$$

where r is the distance between the 2 particles, s is related to the size of the atom, and ϵ is related to the degree of attraction between two atoms. The values of s and ϵ are specific to the nature of the two interacting atoms, and need to be defined for all possible interactions: self-interactions for the interactions of 2 atoms of the same type; and cross-interactions for the interactions for two different atoms.

2.5 Practical aspects of MD simulation methods

Besides a method to calculate forces (i.e., the force field) and a method to integrate Newton's equation of motion, a few other aspects are crucial to perform MD simulations that we can easily perform on current computer hardware and interpret. The first one is periodic boundary conditions: these are necessary to avoid that the system of interest has an interface with vacuum (59). When using periodic boundary conditions, the system is replicated in the X, Y, and Z dimension, so the particles experience forces as if they were in bulk fluid, and the system can be simulated using a relatively small number of particles.

A second crucial aspect in MD simulations is the use of cutoffs to truncate the potential energy. As a system grows, the number of interactions within the system grows as $N \cdot (N-1) \approx N^2$, where N is the number of particles in the system. Truncation of the potential energy means that only interactions within a certain cutoff are calculated. This way, as the system grows, the number of interaction scales linearly with N. This represents an enormous advantage when simulating large systems.

Truncation of the potential may lead to significant artifacts, particularly in the calculation of electrostatics interactions in systems containing many charged particles. Methods have been developed to approximate long range electrostatic interactions and avoid the most important artifacts, and the method most commonly used is named Particle Mesh Ewald (PME) summation, and allows considering all the interactions of the system as if the system was an infinite, periodic crystal (60; 61). Another method often used is named Reaction Field (62); in this case, interactions outside the cutoff are estimated assuming that the medium (solvent) has a uniform dielectric constant. The method is computationally cheaper than PME, and is often used in combination with coarse-grained models (see next section).

Another "trick" to speed up MD simulations is the use of constraints: some (or all) bond stretching terms in the force field are replaced by constraints, i.e., such bonds cannot vibrate and no energy is associated to them (i.e., they are no longer considered as degrees of freedom in the system). Eliminating fast vibrations allows the use of longer integration time steps, which speeds up the calculation.

Last but not least, to better mimic experimental conditions, MD simulations are often performed at constant temperature and pressure. This requires the use of a so-called thermostat and barostat, and a number of different algorithms have been devised to modify the integration of the equation of motion and take temperature and pressure into account. One of the most popular methods is the so-called weak coupling algorithm by Berendsen (63), whose main advantage is the great computational robustness. Unfortunately, while average temperature and pressure are correct, the Berendsen algorithm does not guarantee that fluctuations in temperature and pressure correspond to the desired thermodynamic ensemble; as a consequence, quantities depending on fluctuations in temperature (or related quantities, for instance particle velocities, etc.) and pressure are generally incorrect. If such properties are of interest, more sophisticated algorithms should be used – for example the stochastic velocity rescaling thermostat by Parrinello (64) and the Parrinello-Rahman barostat (65).

2.6 Coarse-grained force fields and the Martini force field

Despite the efforts made during the past 40 years to speed up MD simulations, sampling remained one of the most problematic aspects. Indeed, many phenomena of high biological interest occur on time scales currently out of reach for all-atom force fields. These include the folding and conformational rearrangements of large proteins and nucleic acids, protein aggregation, self-assembly of multi-protein complexes, endo- and exocytosis, the biogenesis of organelles, and many others.

Different solutions have been proposed to tackle the time and length scale problem in MD simulations. For example, a wide range of enhanced sampling techniques have been developed in order to calculate free energy changes – one of the most important and most computationally expensive types of calculations. An alternative solution is offered by coarse-grained modeling. In coarse-grained force fields, a certain number of atoms is grouped together into an effective interaction site, also referred to as particle or bead. Beads interact with a simplified effective potential, that somehow averages over all the interactions of the constituting atoms. Research in this field has been very active for at least three decades, particularly in the area of polymers and soft-matter physics, and a very large number of coarse-grained models have been developed, each one with its specific advantages and disadvantages (66).

The Martini force field is one of the very many coarse-grained force fields currently available. It was initially developed by developed by Prof. Marrink at University of Groningen (67) for modelling lipids. As other researchers joined this development, Martini was then extended to proteins (68), carbohydrates (69), nucleic acids (70), carbon nanoparticles (71; 72), polymers (73), and other materials and many other classes of molecules (74). Such developments made the force field very popular. The latest version of the force field (75) features many more particle types, distinguished by the polarity and their interactions with other building blocks.

Coarse-graining reduces the number of atoms, and hence interactions, in the system. Moreover, the functional forms used by Martini are similar to the ones used by all-atom force fields, which makes Martini usable with standard MD software, and the effect of long-range interactions is generally incorporated within the cutoff range. Altogether, this leads to a speed-up by a factor between 200 and 1000, i.e., simulations are between 2 and 3 orders of magnitude faster compared to all-atom simulations. This allows to simulate many phenomena of high biological interest, taking place on spatial and time scales not yet accessible to atomistic simulations. Martini is suitable to study systems with size up to a few hundred nanometers, on time scales up to a few hundred microseconds, or even milliseconds in the case of smaller systems. Considering the versatility and the speed of Martini CG simulations, I have chosen this force field for the vast majority of the simulations performed during my PhD thesis. Also, I have participated myself in the development of the latest version of Martini(75). For this reason, I will now provide a bit more details on the features of this force field.

In the Martini force field, each bead represents a chemical moiety and is parameterized independently of the others, to reflect the chemical and physical behavior of the moiety. For the parametrization, Martini combines top-down and bottom-up strategies: interactions among beads are parameterized considering both experimental data and atomistic simulations. Non-bonded interactions are generally based on experimental free energies of transfer from one solvent to another, that are available for a very large number of chemical moieties. Bonded interactions are generally based on computations from all-atom simulations, so that structures are reasonably reproduced.

Martini defines 4 main types of beads: polar (P), nonpolar (N), apolar (C), and charged (Q). Martini 3 introduces three new bead types corresponding to halo-compounds (X), divalent ions (D) and water (W) to model specific chemical compounds (75). Types C, N, P, Q, X are declined in 4 to 6 additional subtypes, to express more accurately the variety of interactions among common chemical building blocks. Indeed, each bead type presents a specific affinity with other beads. Martini describes these affinities with 10 different levels of interactions (20 levels in version 3). Additionally, beads are available in 3 different sizes: the “regular” bead size, with a size of 0.47 nm, corresponds to the usual 4 to 1 mapping, and is the same as in all versions of the force field; small beads, with a size of 0.41 nm, corresponding with a 3 to 1 mapping, are used in some ring and branched molecules; and tiny beads, with a size of 0.34 nm and a 2 to 1 mapping, used in most ring molecules (aromatics, nucleic acids nucleobases, etc.). Finally, Martini 3 also introduces “labels” to fine tune the behavior of beads; in addition to the “hydrogen donor” and “hydrogen acceptor” labels (already available in previous versions of the force field), electron polarizability, increased self-interaction, and partial charges are also available. Multiplying chemical types, sizes, and labels, the last version of Martini defines 843 different beads, totaling 355,746 pair interactions, providing an extensive coverage of the chemical space.

2.7 Martini Protein model

Proteins are one of the most important classes of biological macromolecules, and the richness of their chemical and physical properties makes them very challenging for MD simulations – particularly at the coarse-grained level. Indeed, the coarse-grained representation of proteins in Martini features a single bead for the amino acid backbone; as a consequence, it is difficult to represent the conformational flexibility and all properties related to the formation of hydrogen bonds. Particularly challenging is the preservation of protein structure and conformational flexibility, which depend on both bonded and non-bonded interactions. Non-bonded interaction are represented in Martini with hydrophobic forces and simplified electrostatics interactions; the over-simplified representation of hydrogen bonds makes it very difficult to preserve protein structure.

Since their introduction in 2008 (68), numerous modifications have been developed for the Martini protein model, in an attempt to address its many limitations. To reproduce protein secondary and tertiary structure, Martini uses different types of restraints. One of the first methods relied on elastic networks, i.e., a series of harmonic potentials between pairs of beads of the protein backbone (76). Typically, elastic bonds connect all backbone beads separated by a distance between 0.8 and 1 nm. The network stabilizes the structure of the protein, and the rigidity of the protein structure can be tuned by changing the force constant of the harmonic bonds. Based on specific knowledge of the protein, generally obtained via experiments, some regions of the protein can be free of elastic bond to allow more flexibility.

A different way of enforcing secondary and tertiary structure of the so-called GōMartini network. Initially developed for Martini 2 (77), a prototype version is already available in Martini 3 [manuscript in preparation]. Instead of harmonic bonds, the network consists of Lennard-Jones interactions, calculated from the atomistic structure based on the native contact map of the protein. Lennard-Jones interactions can "break", since the distance between the particles involved is to grow indefinitely. This model allows for a better structural response of the protein, allowing the protein to unfold (and, in principle, fold back again), and it is designed to allow more flexibility in the protein structure.

Chapter 3

Triacylglycerols sequester monotopic membrane proteins to lipid droplet monolayer

3.1 Abstract

LDs are generated in the endoplasmic reticulum (ER) and are coated by a monolayer of phospholipids and proteins, contiguous with the ER bilayer. The formation of LDs involves phase separation of neutral lipids, condensing in an oil droplet that generally remains connected to the ER even after budding. The LD-ER connection is important for the transfer of monotopic integral membrane proteins from the ER to the LD surface. However, how proteins partition between ER and LDs is poorly understood.

Our experimental collaborators employed model systems to study the partitioning of monotopic membrane proteins (i.e., proteins with only one hydrophobic domain, spanning a bilayer membrane only once) between the ER and the LD surface. They found that most ER proteins containing a single hydrophobic domain strongly prefer the LD monolayer. This preference appears to be due to a higher affinity of the hydrophobic domain for triglycerides over membrane phospholipids, and appears to be non-specific.

We used MD simulations to explore the partitioning of model peptides between bilayer and connected monolayer membranes. Simulations confirmed the preference of monotopic membrane proteins for the LD monolayer. Moreover, simulations allowed us to predict that

transmembrane proteins partition to the rim of LDs, monotopic membrane proteins forming antiparallel dimers mimic the behavior of transmembrane proteins. In cells, a protein editing mechanism at the ER membrane would be necessary to prevent unspecific relocation of HD-containing proteins to LDs.

3.2 Introduction

Lipid droplets (LDs) are organelles serving in lipid storage cellular energy metabolism (78). The endoplasmic reticulum (ER) membrane is the birthplace of LD, during high energy or stress conditions (10). LD biogenesis is a multiple steps process, starting with the synthesis of neutral lipids, the main component of LDs, such as triacylglycerols (TG) or sterol esters. At low concentration, neutral lipids are dissolved in the ER membrane. As their concentration raises, neutral lipids phase-separate from the membrane phospholipids, forming a lens-shaped nascent droplet embedded in the bilayer (79) (Fig. 3.4a). Along with the synthesis of the neutral lipids, the droplet grows and emerges from the ER membrane into the cytosol as a mature LDs. The surface of mature LDs is covered with embedded proteins, as throughout the stages of LD formation many proteins dynamically bind to the LD surface or the surrounding ER membrane (24; 80; 81). Proteins on the surface of LDs play a role in LD formation and functioning (10; 82; 32; 83; 84). It is not known how proteins specifically target the LD surface, however its unique structure, consisting of a phospholipid monolayer, plays a key role in protein binding (25). Understanding the driving forces of protein targeting to LDs would provide valuable knowledge on lipid metabolism and protein cellular activity (85; 86).

The phospholipid composition and the physical properties of the membrane are key factors of protein spatial distribution and activity (87; 88; 89). The LD surface has unique properties: it presents a lower lipid packing compared to lipid bilayer membranes (25; 2; 26; 27); and the oil core represents a large hydrophobic region in comparison to the hydrophobic thickness of a bilayer (90; 91). These differences between the LD monolayer and bilayer membranes determine the binding of proteins to one or the other surface.

Proteins found on the surface of LDs have two origins: either the ER membrane (class I proteins) or the cytosol (class II proteins) (24). Class 1 proteins are monotopic integral membrane domains, containing hydrophobic helical domains (HDs) (24), helical hairpins (i.e., two hydrophobic helices connected by a loop), or transmembrane domains that do not fully cross bilayers (24; 92; 93; 94). Class II proteins are cytosolic soluble proteins and generally bind membranes through amphipathic helices (AHs)(24). Since AHs play the role of surfactants, decreasing the interfacial energy from the oil/water interface of LDs (2; 95), their binding has been studied both in vitro and in vivo (25; 26; 27; 28; 29; 96). Proteins containing AHs can often detect membrane curvature, surface charges, packing defects, and the presence of neutral lipids (25; 26; 27; 28; 29). In comparison, the role of HDs targeting LDs from the ER membrane is less understood (27; 97; 98). Our interest is to better understand the factors determining the partitioning of HD-containing proteins to the surface of LDs.

Insertion of HD domain into membranes can generate local perturbations, which translates into an energy cost (99; 100; 101; 102; 103). Concerning their insertion into LD surfaces, little is known about the energetic cost, or the perturbation imposed to the LD surface.

Along with the work of our experimental collaborators, we studied the partioning of LD proteins between a bilayer and a LD monolayer, focusing on monotopic HD-containing proteins, and transmembrane proteins. Our collaborators used the droplet interface bilayer (DIB) system (104) (Fig. 3.4a), while we used molecular dynamics simulations to interpret experimental findings. We found that proteins containing hydrophobic domains partition preferentially to the LD monolayer while proteins containing TM domain remains at the edge of nascent LDs.

3.2.1 Characterization of the droplet interface bilayer (DIB) system

In order to reproduce the continuous monolayer-bilayer environment specific to nascent LDs, our collaborators used a droplet interface bilayer (DIB system) (104; 105). To build this system, buffer droplets were introduced in a trioctanoate phase, whose interfacial energy is similar to triolein, and were coated by a monolayer of phospholipids. DIB systems were prepared by bringing into contact two buffer droplets (coated by lipid monolayers) in oil: a bilayer is formed at the interface between the droplets, continuously connected to the two monolayers bounding the buffer droplets (Figure 3.10a). Since the size of the DIB was on the order of few μm to tens of μm, each membrane is flat at the scale of protein. DIBs are a particularly convenient model system, as they can be built with different phospholipid compositions, including pure DOPE, that is not a bilayer-forming lipid. Here, PC and PE lipids were used, since they are the mayor ER phospholipids (5) (Figure 3.10b).

PE usually does not form bilayers on its own due to its structure (106; 107). Unlike other common phospholipids, PE has a smaller headgroup and lacks a charged group, such as a phosphate or amino group. These result in a reduced electrostatic repulsion between PE molecules, making it difficult for them to form stable bilayers through self-assembly. However, a pure PE bilayer can be formed in a DIB by exploiting the interfacial tension between two immiscible aqueous phases(104; 105). When the droplets come into contact, the interfacial tension at the water-oil interface promotes the formation of a stable bilayer between them. It is possible to create a pure PE bilayer in the DIB system, enabling the study of PE-specific properties and interactions.

We used all-atom MD simulations to show that the TG molecules dissolved in a PE bilayer do not affect significantly the physical properties of the bilayer: the thickness (Figure 3.10c), the PC/PE ratio (Figure 3.10d), and the bulk composition of the bilayer.

To provide a molecular insight to the findings of the experiments of our collaborators, we ran MD simulations of phospholipids bilayers with embedded nascent LDs, reproducing the specific monolayer-bilayer environment of the experiments. We analyzed three systems, containing a different phospholipid compositions: pure DOPC, a 80:20 mixture of DOPC:DOPE, or a 60:40 mixture of DOPC:DOPE. Analysis the contact fraction DOPC and DOPE showed that DOPC and DOPE mix ideally (Figure 3.10d), and that their distribution between monolayer and bilayer region is nearly homogeneous, with only a slightly higher fraction of DOPC in the monolayer compared to the bilayer. The difference the 80:20 mixture and the 60:40 mixture was respectively of 2.4% and 5.4%. Conversely, the fraction of DOPE was slightly higher in the monolayer. Our results shows that PE/PC mixtures behave closely to ideal mixtures, and the distribution of the phospholipids in the monolayer and in the bilayer are almost identical.

The results from our collaborator's experiments and from our simulations suggest that DIB systems mimic the environment of nascent lipid droplet embedded in a bilayer (13), make it an appropriate model system to study the partitioning of protein in nascent LDs.

3.2.2 Monotopic integral membrane proteins partition more favorably to TG-covering monolayers over bilayers

Using the DIB system, our collaborators evaluated the partitioning of 2 types of proteins: monotopic integral membrane proteins representing class I proteins from the ER, and soluble protein representing class II proteins from the cytosol. Proteins are incorporated into the DIB system during its preparation, when forming the buffer in oil droplets, that are later assembled in DIBs (Figure 3.5a). The partitioning of the protein is measured later after equilibration of the system (Figure 3.5b).

Seven monotopic integral membrane proteins and six soluble LD proteins (108; 109; 110; 111)(Fig. 3.5d, 3.11a, 3.11c) were tested experimentally by our collaborators. Monotopic integral membrane proteins bind to membranes thanks to helical hairpin or hydrophobic helix domains.

Both the monotopic membrane proteins and soluble protein showed a preference for the monolayer rather than the bilayer. This preferential partitioning was higher for the monotopic proteins and independent of the membrane phospholipid composition.

Partitioning to the monolayer was lowest in the case of soluble proteins in PE composed membrane, even if the monolayer was still preferred; while monotopic membrane proteins partitioned strongly to the monolayer.

3.2.3 KWALP peptides recapitulate the global behavior of HD proteins

All the tested proteins preferentially partitioned to the LD, but they all contained both HDs and AHs (109; 111; 112), therefore experiments on those proteins could not reveal the behavior of isolated HD. In order to focus exclusively on HDs, we study several model peptides from the KWALP family. KWALP peptides consist of alternating alanine (A) and leucine (L) residues with two tryptophan (W) residues at the C-terminus and three lysine (K) residues at the N-terminus. Additionally, in experiments, a dye is attached to the N-terminus. Such peptides form helical hydrophobic domains, partition to lipid membranes and can be used as a model for transmembrane protein domains (90; 99; 113). In experiments, KWALP20 was used, containing 16 hydrophobic residues out of 20. Its length is close to the ER bilayer thickness (90). In order to focus on pure AHs, the model protein PL108 was used. It is composed from the repetition of the AH motif from Perilipin1 (96) (Figure 3.6a).

As previously, the peptides were added to DIB systems under the same experimental conditions, with varying phospholipid composition. While in all conditions the peptides partitioned preferentially to the monolayer, partitioning of PL108 to the bilayer increased in PE membranes, whereas partitioning of KWALP20 to the bilayer increased in PE/PC membranes (Figure 3.6b).

KWALP reproduced the behavior of HD proteins, and can consequently be used for further investigate the partitioning of HD proteins on LDs (Fig. 3.5e, 3.5f, 3.11e, 3.11f).

3.2.4 PC/PE ratio regulates the partitioning of KWALP

KWALP20 was completely absent from the bilayer in PE DIBs. In order to understand the role of hydrophobic mismatch, a longer peptide was used, KWALP28, containing 24 hydrophobic residues. The experiment repeated with KWALP28 gave the same result, ruling out any effect hydrophobic mismatch between the bilayer and the peptide in partitioning (Fig 3.12a).

Our experimental collaborators showed that the distribution of the peptide was dynamic and can be varied by changing the PC/PE ratio. For example, in pure PE DIBs, KWALP are absent from the bilayer, but adding PC induced partitioning of the peptide to the bilayer (Figure 3.6d 3.6e, 3.6f).

3.2.5 Phospholipid shape defines KWALP partitioning

As different phospholipid compositions altered the partitioning of HDs, the affinity between HDs and lipids could be one of the driving forces for partitioning.

Our collaborators tested the effect of phospholipids by adding KWALP20 peptides to the surface of TG droplets without phospholipids or with PC and PE. The peptide distribution was uniform on the LDs surface without phospholipids (Figure 3.7a S4b), it was mostly uniform with PE, and clustered with PC (Figure 3.7a, 3.7b). This result suggests than KWALP has a higher affinity for TG than for acyl chains.

In LDs embedded in bilayer, the monolayer has a looser phospholipid packing11. The peptide-phospholipids acyl chains contact in also significantly reduced in monolayer. In bilayer, the HDs are in contact with acyl chains from the two leaflets along the entire length of the peptide. In comparison, in monolayers, approximately half of the peptide length is in contact with TGs. HDs higher affinity for monolayers over bilayers can be due to a preference for TGs over phospholipids.

Additionally, in a monolayer, the phospholipid type covering the TGs can regulate HD-TG contact. PC cylindrical shape leads to tighter lipid packing(13) and increased HD-acyl chain contact. PE conical shape(89; 114) leads to looser lipid packing and increased HD-TG contact. Hence HD proteins have a preference for PE monolayer over PC monolayer (Figure 3.7c).

Experiments using KWALP were repeated using different phospholipids. PA with an inverted conical shape reproduced the behavior of PE. N-methyl-PE and N,N-dimethyl-PE phospholipids which are structurally intermediates between PE and PC, reproduced the behavior of PC/PE mixtures (Figure 3.7d, 3.7e, 3.14c, 3.14d).

These results confirm that HDs prefer mixing with TGs over phospholipids, explaining the differential partitioning of HDs proteins in LD monolayers.

3.2.6 Preferential interactions with TG retains HDs to the monolayer interface

In order to test if KWALP preference for monolayer is due to HDs-TG interactions, a different oil was tested, namely silicone oil, very different chemically and structurally from TG, but still forming droplets in bilayers.

While KWALP20 partitioned to PE DIBs monolayer with TG, with silicone oil KWALP20 stays in the bilayer. We concluded that the peptide has a higher affinity for the phospholipids than the silicone oil (Figure 3.8).

This experiment with silicone oil was repeated with one protein, VAMP2. In TG, VAMP2, had a very clear preference for the monolayer. In silicon oil, VAMP2 partitions to the bilayer, similar to KWALP20 (Figure 3.14e, 3.14f, 3.14g, 3.14h).

These results show that molecular affinity drives protein partitioning. Since HD-proteins have the highest affinity for TG, they preferentially partition to monolayers of TG LDs.

3.2.7 KWALP exits the bilayer to accumulate in model lipid droplets

The DIB system serves to mimic the nascent LD environment, i.e., the continuity between a bilayer and the monolayer of a nascent LD. Hence the experiments realized in the DIB suggest that HD proteins bind to nascent LDs, as observed previously (83; 98; 108; 111). Since multiple HD proteins show this behavior, our collaborators could reproduce this observation in experiments.

Nascent LDs were formed by incorporating TG droplets in the bilayer of a giant unilamellar vesicle (GUV) (Figure 3.9a)(13; 115). The KWALP peptide was then incorporated into the droplet-embedded vesicle. The monolayer of the droplet was highly enriched with the peptide, thus confirming the results found with the DIBs systems (Figure 3.9b).

3.2.8 Molecular dynamics to study the partitioning of proteins in lipid droplets

MD simulations can give a molecular insight into protein partitioning in LDs, and even predict the distribution of proteins in nascent LDs. To achieve that, we focused on the same simple a model proteins, the KWALP family. KWALP peptides have a stable helical secondary structure, hence it is simple and efficient to study their properties and peptide-lipid interactions using a coarse-grained force field such as Martini. Moreover, KWALP properties can be tuned, for instance by adding charges on the termini or changing the length. In simulations, we can also easily tune their ability to aggregate. Here, aggregation refers to their ability to form small oligomers.

The possibility of creating several models of peptide allow us to model different types of proteins. Indeed, the charges added on the termini mimic hydrophilic domains. Indeed, charged termini are highly polar and their insertion itno lipid membranes presents a high energy cost. Consequently, we can mimic a transmembrane protein simply by including two charged termini. With one charged terminus, instead, we mimic a monotopic membrane protein, as the uncharged, non-polar terminus can freely enter the bilayer (Figure 3.1a).

We designed 8 experiments, with 8 types of proteins in order to mimic the different types of proteins present in the ER, varying 4 parameters: peptide length (20 or 28 residues), ability to oligomerize, charged termini, and relative peptides orientation (Figure 3.1b). We created bilayers with an embedded LD and placed 16 or 32 copies of the peptide in the bilayer. As a control, simulations were repeated in the absence of LDs.

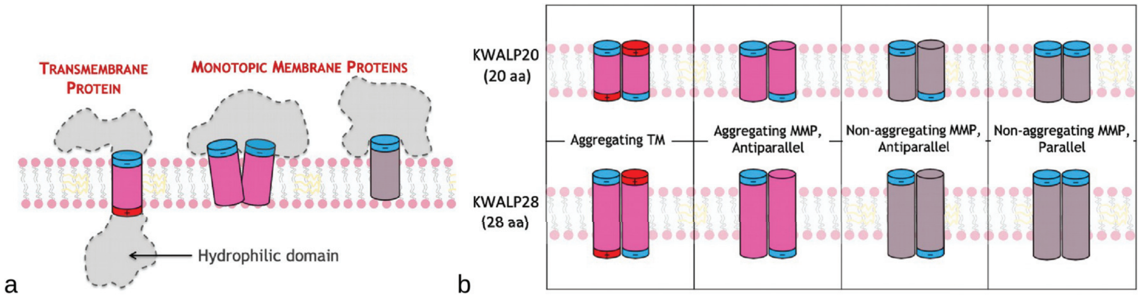


Figure 3.1: KWALP can mimic ER membrane proteins (a) Illustration of the membrane proteins modelled using the KWALP peptide. Positively charged termini shown in red; negatively charged termini in blue. Peptide with high or low ability to form clusters respectively in pink or grey. (b) Proteins used in simulations: on top KWALP20, bottom, KWALP28. From left to right, first aggregating TM proteins; aggregating Monotopic Membrane proteins, in antiparallel conformation; Non-aggregating monotopic membrane proteins, in antiparallel conformation; Non-aggregating monotopic membrane proteins, in parallel conformation.

3.3 Results

3.3.1 Transmembrane proteins partition to the edge of lipid droplets

In the absence of LD, simulations showed that the peptides distributes randomly, in agreement with our expectations. In the following simulations including LDs embedded in bilayer, any different distribution would be the consequence of the presence of the nascent LD.

Simulations including aggregating transmembrane (TM) proteins have no experimental equivalent that can be used for validation. Indeed, the DIB systems used experimentally do not allow the insertion of transmembrane proteins. Hence, only HD proteins and AH proteins could be tested experimentally.

For both KWALP20 and KWALP28, the analysis of our simulations shows that transmembrane proteins partition to the edge of LDs (Figure 3.2a). This a new result, because there is no example of TM proteins distribution in LDs in the literature. This can be easily explained by the fact that proteins still favor interactions with TG over phospholipids, but the presence of hydrophilic charges on both sides of the TM domain prevent the peptide from entering the LD, since the energy cost of embedding a charge in the oil phase is very high (Figure 3.2b). When the peptide partitions to the edge to LD, contact with TG molecules is allowed while charges are solvated, therefore the monolayer-bilayer junction is the most

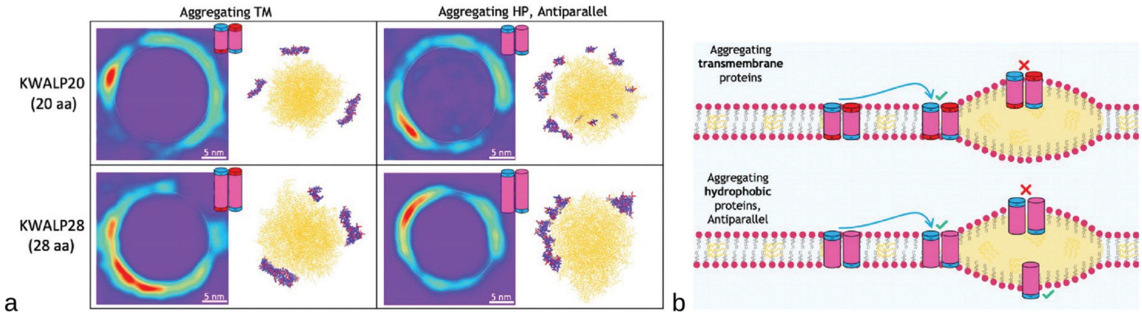


Figure 3.2: Transmembrane proteins partition to the edge of lipid droplets(a) Four panels corresponding to the four simulations using aggregating TM proteins and aggregating HD proteins. For each simulation, the left panel is the density landscape of the peptide, representing the spatial distribution of the peptide around LD averaged over the entire simulation time. The center of the LD is used as a reference. A high density value shows where the protein is more likely to be found. The right panels show a representative snapshot of each simulation, view from the top, with the LD in yellow and the peptides in purple. (b) Basic model showing for aggregating TM proteins and aggregating HD proteins , how the peptides charges or hydrophilic domains allow the partitioning the peptide around the LD.

favorable environment for TM proteins.

A similar behavior was observed for the aggregating HD proteins, which also partitioned to the edge of LDs (Figure 3.2a). With the stronger aggregation, the proteins form oligomers, and because of the antiparallel conformation, these oligomers mimic the same hydrophilic domains as TM proteins. Consequently, oligomers cannot enter the LDs because of the polar charges on both sides of the bilayer. So, they distribute at the edge of the LD, with the same driving forces as for TM proteins (Figure 3.2b).

However, in the event of a single peptide escaping the aggregation with other peptides, nothing would prevent such a peptide from partitioning to the LD surface. Although unlikely, this event was observed in the simulation containing HD-KWALP20, and single peptides could be observed in the monolayer regions for a short time before aggregating with other peptides. This event show is interest of simulating proteins that do not aggregate, in order to observed their behaviors as monomers.

3.3.2 Monotopic membrane proteins partition to the monolayer of lipid droplets

The previous simulations were repeated with non-aggregating HD proteins, both for KWALP20 and KWALP28. Since monotopic HD peptides only have one charged terminus, we inserted them in two different orientations: antiparallel, with the half of charged termini on one side of the bilayer and the other haft on the other side; or parallel, with all charged termini on the same side of the bilayer.

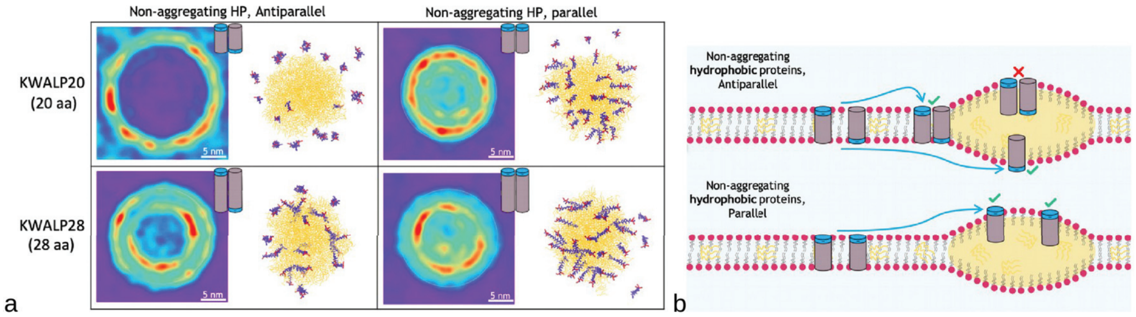


Figure 3.3: Monotopic membrane proteins partition to the monolayer of lipid droplets. (a) Four panels corresponding to the four simulations using non-aggregating HD proteins in antiparallel or parallel conformations. For each simulation, the left panel is the density landscape of the peptide, representing the spatial distribution of the peptide around LD averaged over the entire simulation time. The center of the LD is used as a reference. A high density value shows where the protein is more likely to be found. The right panels show a representative snapshot of each simulation, view from the top, with the LD in yellow and the peptides in purple. (b) Cartoon showing, for non-aggregating HD proteins in antiparallel or parallel conformations, how the peptide charges (or hydrophilic domains) allow the partitioning the peptide around the LD.

In the case of non-aggregating KWALP20 in antiparallel conformation, the majority of the peptide form antiparallel dimers, with charges on both sides of the membrane. Antiparallel dimers partition at the edge of the LD (Figure 3.3a). These dimers also mimic the same hydrophilic domains as TM proteins. Consequently, the dimers cannot enter the LDs and distribute at the edge of the LD, with the same driving forces as for TM proteins (Figure 3.3b).

In the three simulations with non-aggregating KWALP28 in antiparallel conformation, and non-aggregating KWALP20 or KWALP28 in parallel conformation, the peptides partitioned to the monolayer of the LD (Figure 3.3a). Because they mostly remained monomeric, the peptides were allowed to enter the LD monolayer, due to their favorable interactions with the oil (Figure 3.3b). This result is in agreement with the experiments, showing that monotopic membrane proteins partition largely to the monolayer region of the LD.

3.4 Discussion

We were able to demonstrate that TM proteins partition at the edge of LDs, and antiparallel peptide dimers also show the same behavior. Using molecular dynamics simulations, we predicted a specific peptide behavior that could not be observed experimentally, since the relevant systems could not be prepared. Indeed, because of the protocol, the experiments can only be done with monotopic membrane proteins. The way proteins are inserted in the experimental system does not allow to use transmembrane proteins: protein-containing liposomes simmply would not fuse with the DIB system. It is possible that such experiments will be achieved in the future.

Simulations also reproduced the partitioning of monomeric monotopic membrane proteins to LD monolayers. This result confirms the experimental findings.

One of the key driving forces responsible for the partitioning of protein in LDs is protein-lipid interactions. These interactions control the location of the protein in their environment.

3.5 Supplementary Information

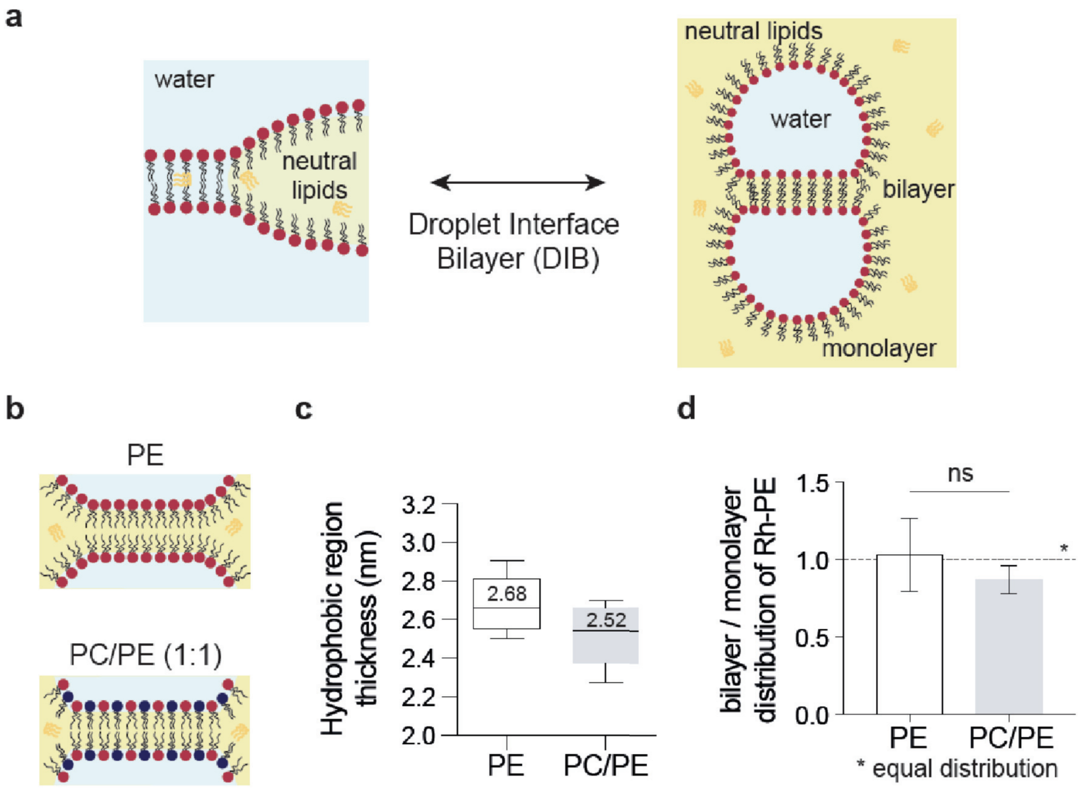


Figure 3.4: Droplet interface bilayers (DIBs) mimic the ER-LD continuity. (a) Schematic representation of ER phospholipids bilayer and contiguous LD monolayer, as observed in a triglyceride lens in the ER (left side), and the corresponding DIB model system, containing also contiguous bilayer and monolayers (right side). The water phase is represented in light blue and the oil phase (neutral lipids, i.e. TG) in yellow. (b) Drawings of a bilayer of DOPE (top) and DOPC/DOPE (1:1) (bottom) in the DIB system, possibly containing TG in between the leaflets. (c) Hydrophobic thickness of DOPE (white) and DOPC/DOPE (1:1) (grey) bilayers determined by electrical measurements. Results are shown as box-plots (box limits, upper and lower quartiles; middle line, median; whiskers, minimum and maximum value; the mean is indicated) from n=5 independent experiments. (d) Distribution of Rh-PE between bilayer and monolayers in DOPE (white) and DOPC/DOPE (1:1) (grey) DIBs. The results are the mean ± sd of respectively n=10 and n=5 independent measurements. Significance was determined by Welch's t test (unpaired parametric test, two-tailed P value) and is indicated by ns (not significant): $p > 0.05$

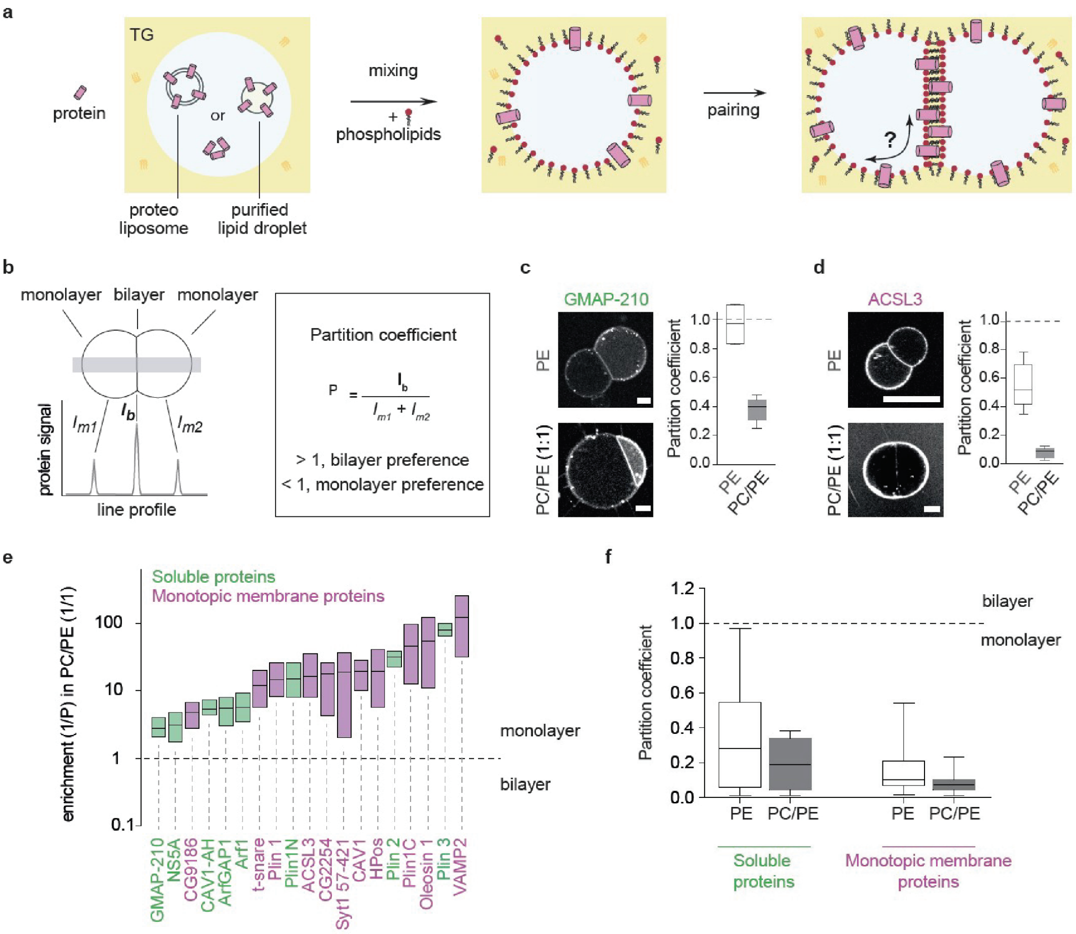


Figure 3.5: Hydrophobic domains (HDs) and amphipathic helices (AHs) partitioning in droplet interface bilayers (DIBs). (a) Formation of protein-containing DIBs: soluble protein or proteoliposome or purified lipid droplet are added in the buffer (left). Droplets are formed by mixing the protein with TG-containing phospholipids, allowing the relocalization of proteins at the lipid interface (middle). When two drops come close together, and they zip to form a bilayer at the junction of the two drops. The proteins can then stay in the monolayer or relocalize to the bilayer (right). The water phase is represented in light blue and the oil phase (neutral lipids, i.e. TG) in yellow. (b) Protein distribution between bilayer and monolayer is characterized by the partition coefficient P, which corresponds to the protein signal in the bilayer divided by the protein signal in the monolayers. When a protein is found preferentially in the bilayer, P>1; when it distributes preferentially in the monolayer, P<1. (c, d) Distribution of GMAP-210-AH (soluble protein) and ACSL3 (monotopic membrane protein) respectively, in DOPE or DOPC/DOPE (1:1) DIBs. Scale bar: 20 µm. The partition coefficient is represented for each condition as box-plots (box limits, upper and lower quartiles; middle line, median; whiskers, minimum and maximum value), from n=5 independent measurements (excepted for GMAP-210-AH in PE, n=4). (e) Enrichment parameter in DOPC/DOPE (1:1) membranes for AH- (green) and HD- (pink) containing proteins, shown as floating bars (bar limits, min to max values; central line, mean), $2 \leq n$ independent measurements ≤ 8 for each protein. HD-containing proteins coming from lipid droplets (Plin 1, Plin 1C, ACSL3, CG2254, CG9186, Oleosin 1, Caveolin 1, HPos) or proteoliposomes (Syt1 57-421, t-snare, VAMP2); AH-containing proteins coming from LDs are Plin3, Plin2, Plin1N; the other AHs are added soluble. (f) Average partition coefficient of soluble and monotopic membrane proteins groups in DOPE or DOPC/DOPE (1:1) DIBs. Results on proteins presented in (Fig. S2e) were used to generate this box-plot (box limits, upper and lower quartiles; middle line, median; whiskers, minimum and maximum value).

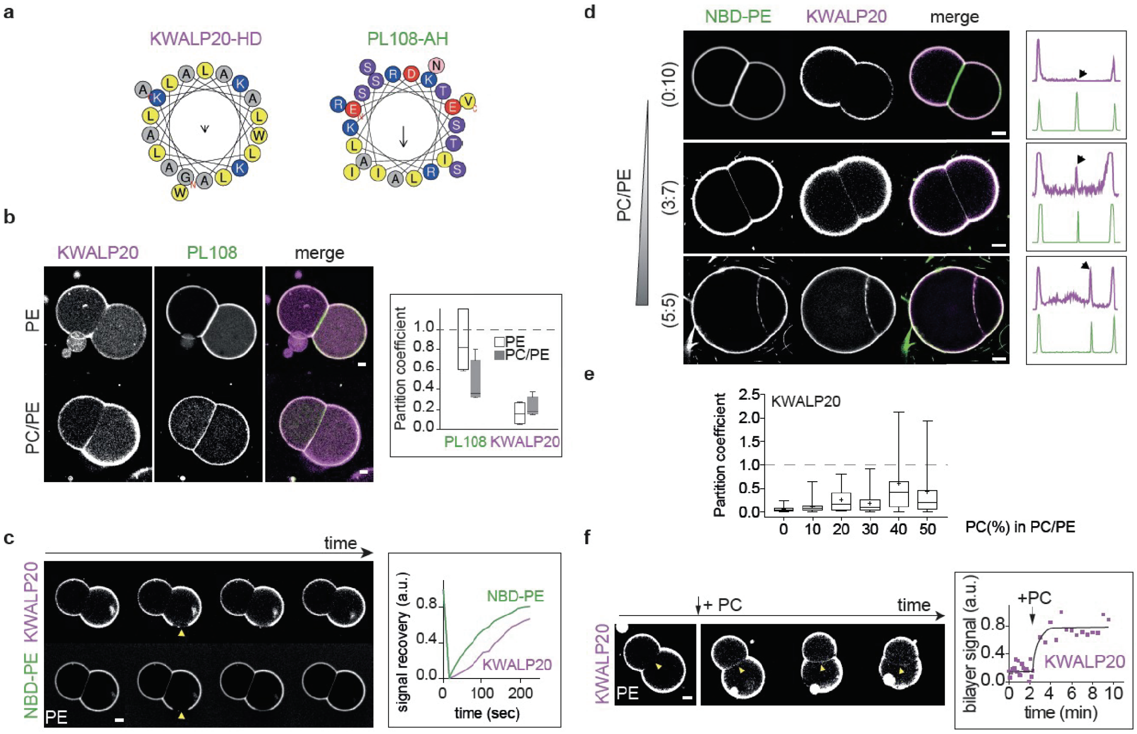


Figure 3.6: PC/PE ratio modulates the partitioning of AH and HD model peptides. (a) Helical wheel representation of KWALP20 and PL108-AH, generated with HeliQuest51. (b) Distribution of KWALP20 and PL108 in DOPE or DOPC/DOPE (1:1) DIBs. KWALP20 is labelled with Rh-B and PL108 with NBD. Scale bar: 10 µm. The partition coefficient is plotted for both peptides as box-plots (box limits, upper and lower quartiles; middle line, median; whiskers, minimum and maximum value), from n=4 independent measurements in each condition. (c) FRAP experiment showing the recovery of KWALP20 (purple) and lipids (green) signals after fluorescence bleaching. Lipids are labelled with 0.2% NBD-PE. Yellow arrows indicate the area bleached. Scale bar: 10 µm. (d) KWALP20 in DIBs of different compositions. The amount of DOPC is gradually increased. Line profile is represented for each image. Black arrows indicate the bilayer signal. Scale bar: 20 µm. (e) Partition coefficient of KWALP20 in DIBs of different DOPC/DOPE ratios is represented as box-plots (box limits, upper and lower quartiles; middle line, median; +, mean; whiskers, minimum and maximum value). Sample size was n=31 for 0% PC, n=14 for 10% PC, n=18 for 20% PC, n=41 for 30% PC, n=16 for 40% PC and n=28 for 50% PC. (f) Relocalization of KWALP20 after addition of DOPC in DOPE DIBs. The bilayer signal is plotted against time. Scale bar: 10 µm.

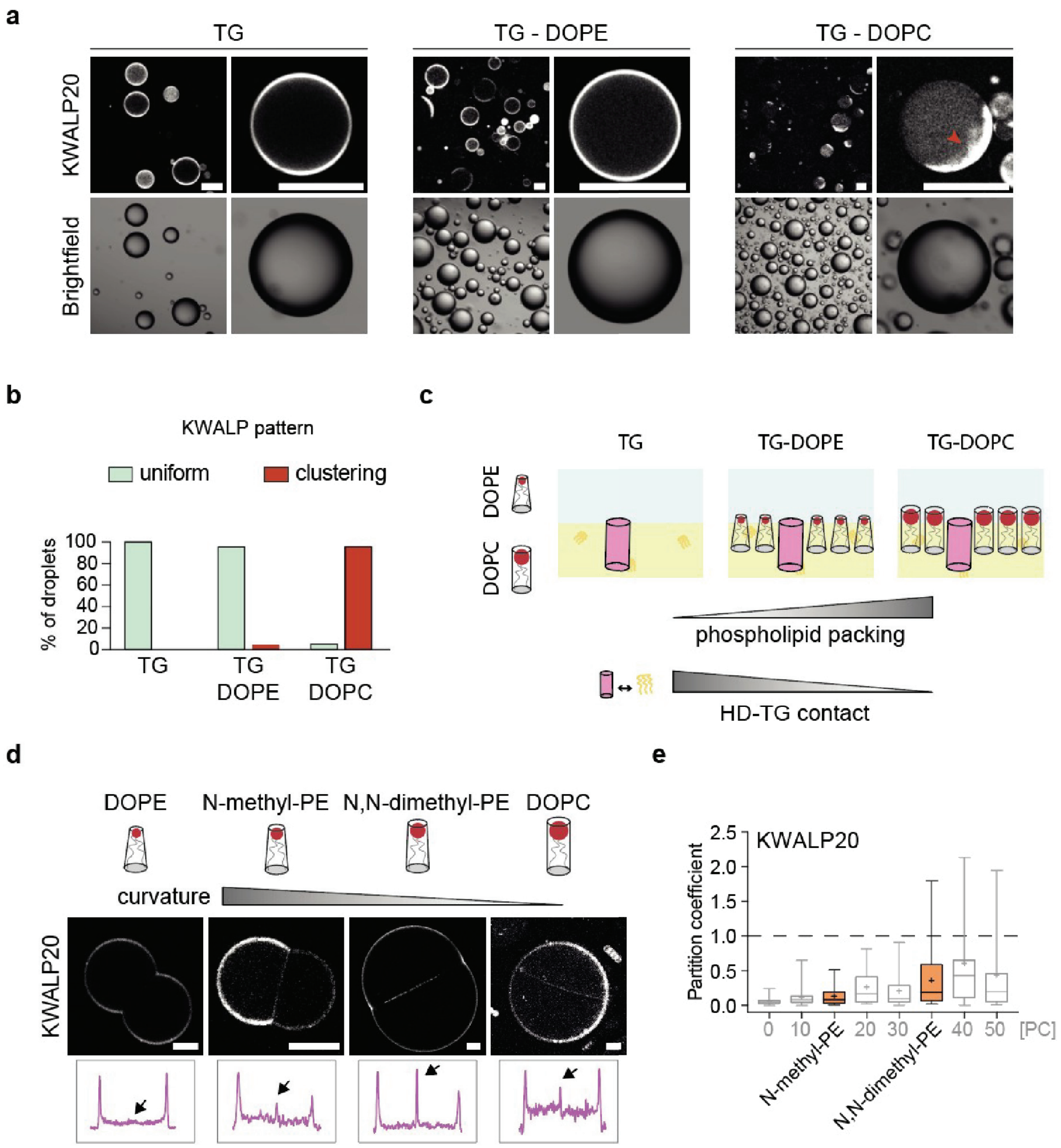


Figure 3.7: Monolayer-bilayer distribution of KWALP is affected by phospholipid shape. (a) KWALP20 distribution in TG droplets without or with DOPE or DOPC lipids. Red arrow highlights peptide accumulation as a cluster. Scale bar: 100µm. (b) Quantification of KWALP pattern, i.e. uniform (light green) or clustering (red) signal, in TG (n=8), TG + DOPE (n=47) and TG + DOPC (n=19), from n independent measurements. (c) Schematic representation of the difference in phospholipid packing, and thus in HD-TG interaction, when DOPE (cone shape) or DOPC (cylinder shape) are present. Increasing DOPC concentration in a DOPC/DOPE monolayer increases the lipid packing and decreases the interactions between HDs and TG. (d) Distribution of KWALP20 in DOPE, N-methyl-PE, N,N-dimethyl-PE, and DOPC DIBs. The average shape differs in these phospholipids, and consequently their curvature too. KWALP20 is labeled with Rh-B. The line profile of peptide is shown for each condition. Arrows indicate the bilayer signal. Scale bar: 20 µm. € Partition coefficient of KWALP20 in DIBs of different compositions is represented as box-plots (box limits, upper and lower quartiles; middle line, median; +, mean; whiskers, minimum and maximum value). Sample size is n=29 and 20 for KWALP20 in N-methyl-PE and N,N-dimethyl-PE respectively. Previous results of varying PC/PE ratios (Fig. 3e) are reported in light grey.

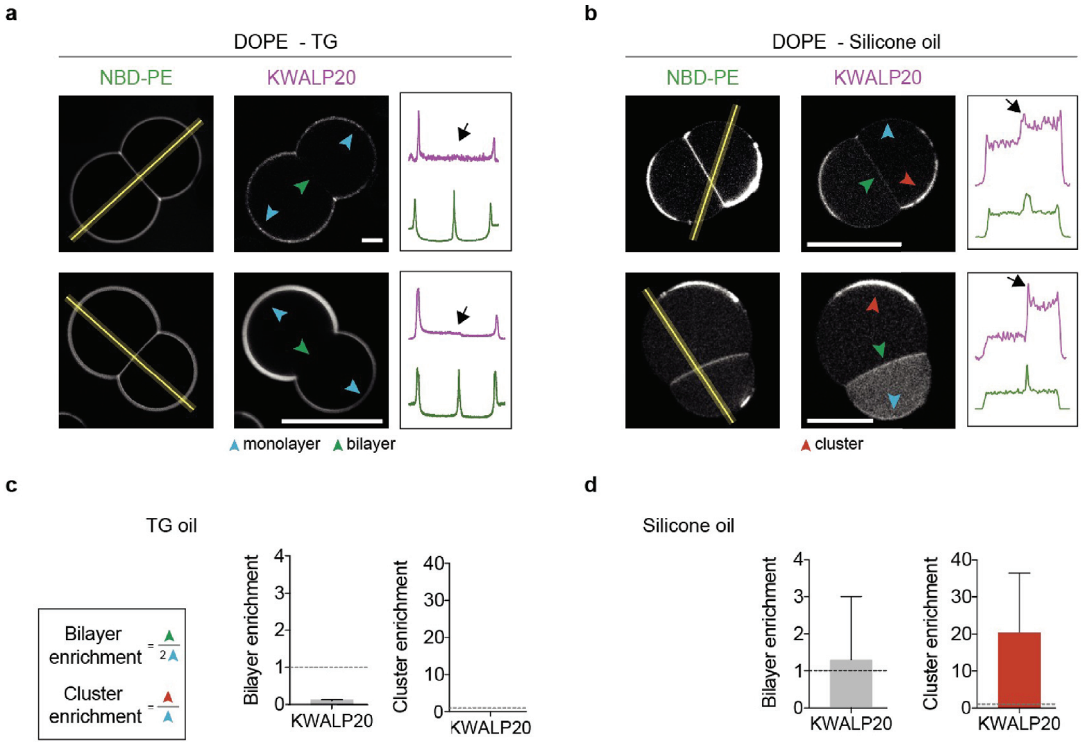


Figure 3.8: Monolayer-bilayer distribution of KWALP is affected by chemical nature of the oil. (a,b) Distribution of KWALP20 in DOPE DIBs. KWALP20 is labelled with Rh-B. Oil phase was TG (a) or silicone oil (b). Blue arrows indicate monolayers, green ones the bilayers, and red ones the clusters areas. Plot profiles are determined using the yellow lines. The bilayer signal is indicated by a black arrow. Scale bar: 20 µm. (c, d) Partition coefficient is reported in grey in TG (c) and silicone oil (d), as mean ± sd (n=2 and 8 independent measurements respectively). A cluster enrichment coefficient (red) is determined for the experiment in silicone oil to take into account aggregation, and is shown as mean ± sd (n=8 independent measurements).

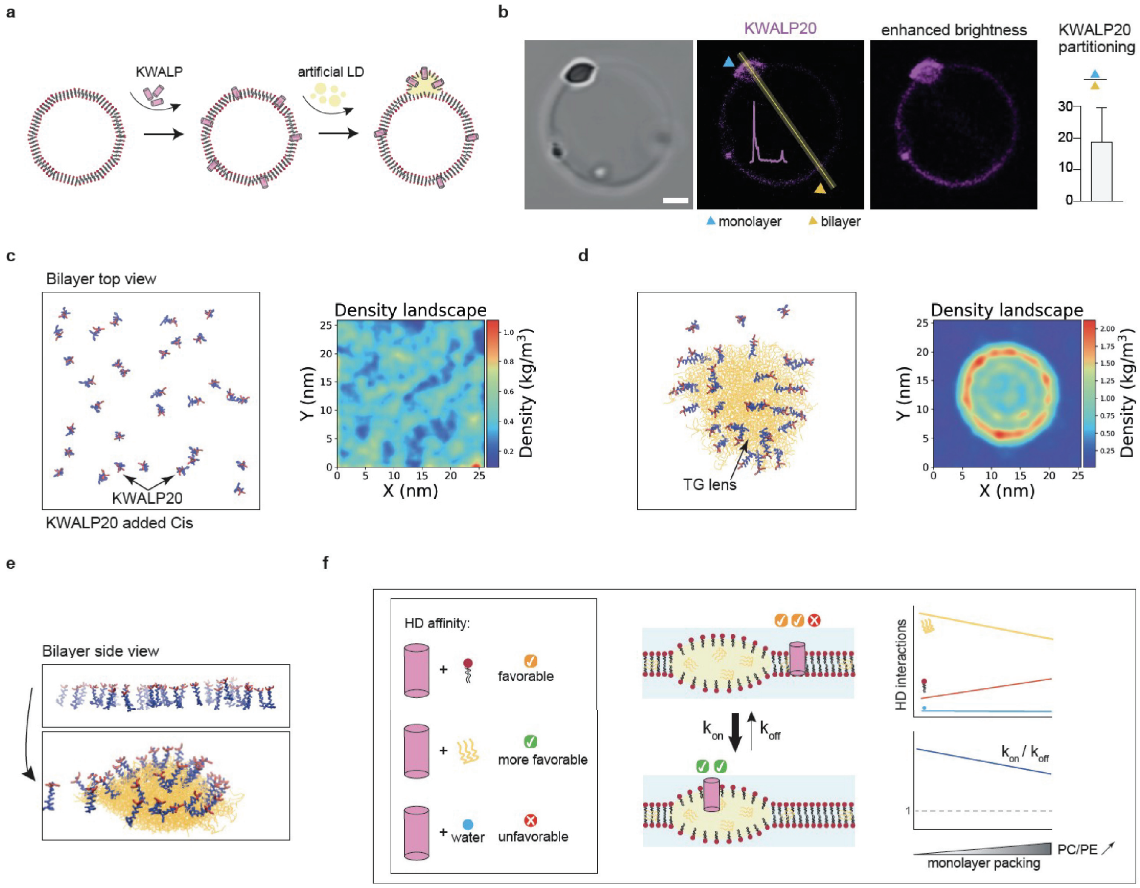
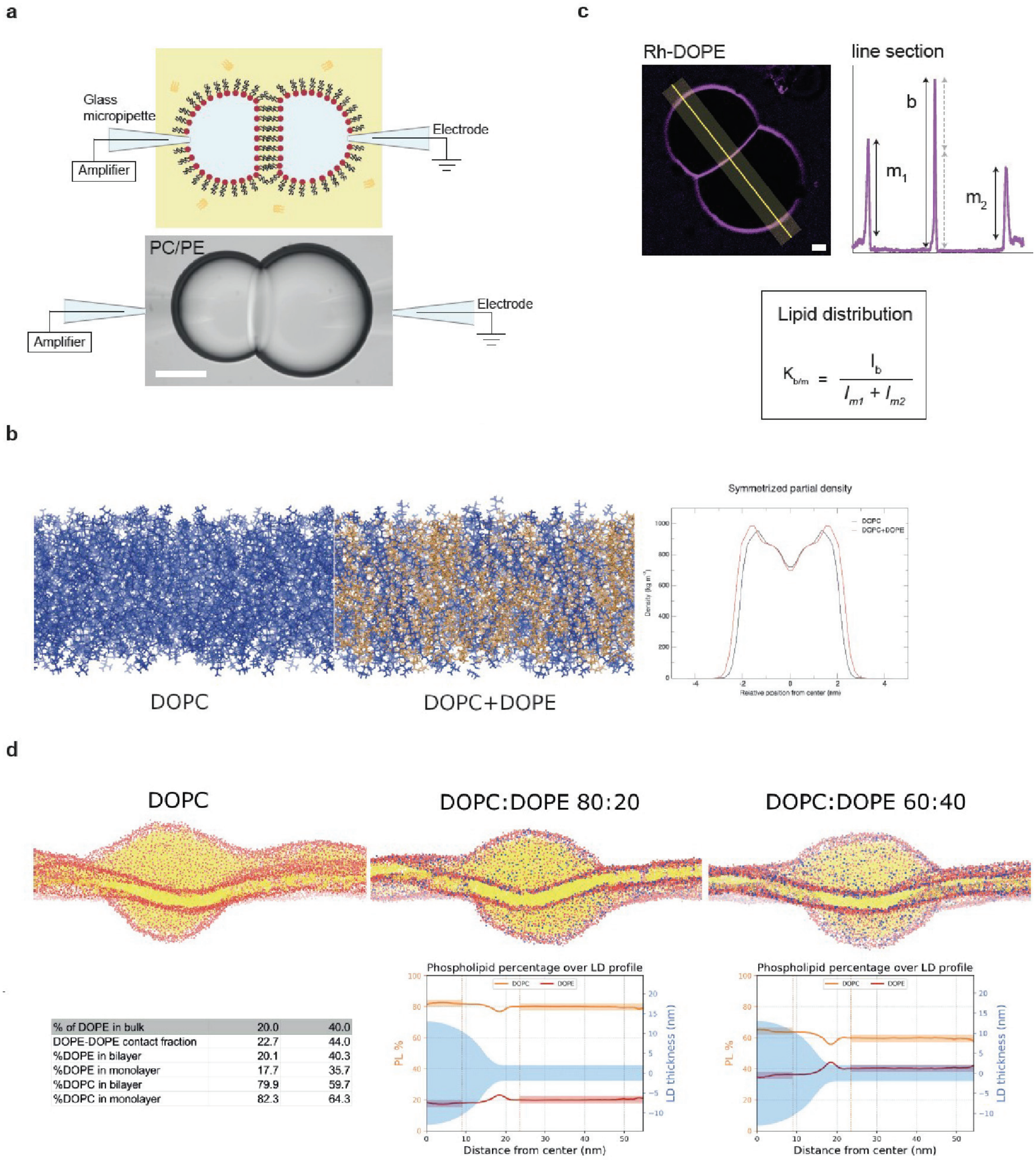


Figure 3.9: HDs accumulate in LD. (a) DEVs-containing KWALP20 formation: PC/PE (7/3) GUVs are incubated with KWALP20, so the peptide is incorporated in the bilayer. Then, KWALP20-containing GUVs are mixed with a TG emulsion, leading to incorporation of TG droplets in the bilayer. (b) KWALP20 distribution in DEV system. The peptide is labeled with Rh-B. An image with enhanced brightness, in which GUV bilayer is more visible, is presented. Scale bar: 2 μm. The plot profile is presented, indicating the monolayer (blue arrows) and bilayer (green arrows) signals. The ratio between monolayer and bilayer signals is plotted as mean ± sd (n=5 independent measurements). (c, d) Left: snapshot from molecular dynamics simulations of a bilayer with 32 KWALP20 peptides in parallel orientation, in the absence (c) and the presence (d) of a TG lens. Hydrophobic amino-acids are represented in blue and charged ones (lysines) in red. Right: average protein density profiles in the bilayer plane, averaged over the entire MD simulation (20 ms), in the absence (c) and in the presence (d) of a TG lens. (f) Basic model of how the different interactions of an HD might favor its LD monolayer localization. HD contacts with TG, phospholipids or water according to the monolayer packing are represented, as well as the kon/koff ratio.



% of DOPE in bulk	20.0	40.0
DOPE-DOPE contact fraction	22.7	44.0
%DOPE in bilayer	20.1	40.3
%DOPE in monolayer	17.7	35.7
%DOPC in bilayer	79.9	59.7
%DOPC in monolayer	82.3	64.3

Figure 3.10 *(previous page)*: Droplet interface bilayers (DIBs) characterization. (a) Illustration of electrical measurements. Aqueous droplets were made at the tip of micropipettes containing electrodes, in oil containing phospholipids. Lipids adsorb at the oil/water interface and assemble as a monolayer. When two drops are brought together, they assemble to form a bilayer. The water phase is represented in light blue and the oil phase (neutral lipids, i.e. TG) in yellow. Electrical measurements are made between the two electrodes. DIB image microscopy is shown at the bottom. Scale bar: 100 μm. (b) In silico study of the impact of PE on bilayer thickness. Left, snapshots from all-atom simulation of DOPC and DOPC/DOPE (1/1) bilayers; DOPC is shown in blue, DOPE in orange. Right, density profiles for both bilayers against position of the phospholipids, showing an increase in overall bilayer thickness when DOPE is present. (c) Rh-DOPE distribution between bilayer and monolayer in DOPE DIBs. Plot profile is determined using the yellow line and is shown on the right. The bilayer signal intensity is equal to the sum of the monolayers signal intensity. Lipid distribution is characterized by the partition coefficient Kb/m. Scale bar: 20μm. (d) Phospholipid distribution between a monolayer and a bilayer in continuity, calculated from simulations of systems containing DOPC, DOPC/DOPE 80:20 and DOPC/DOPE 60:40 systems. The lateral size of the simulated boxes is ca. 80 nm for all three systems. DOPC head groups are represented in red, DOPE head groups in blue, TO in yellow. The average fraction of DOPC and DOPE in DOPC/DOPE systems is represented as a function of the distance from the center of the nascent LD, and reported in the table. In the diagram, the light blue shaded area shows the average shape of the nascent LD. DOPE-DOPE contact fraction indicates the probability that DOPE molecules make contact with other DOPE molecules; if mixing is ideal, this corresponds exactly to the fraction of DOPE in the mixture (i.e., 20% or 40%).

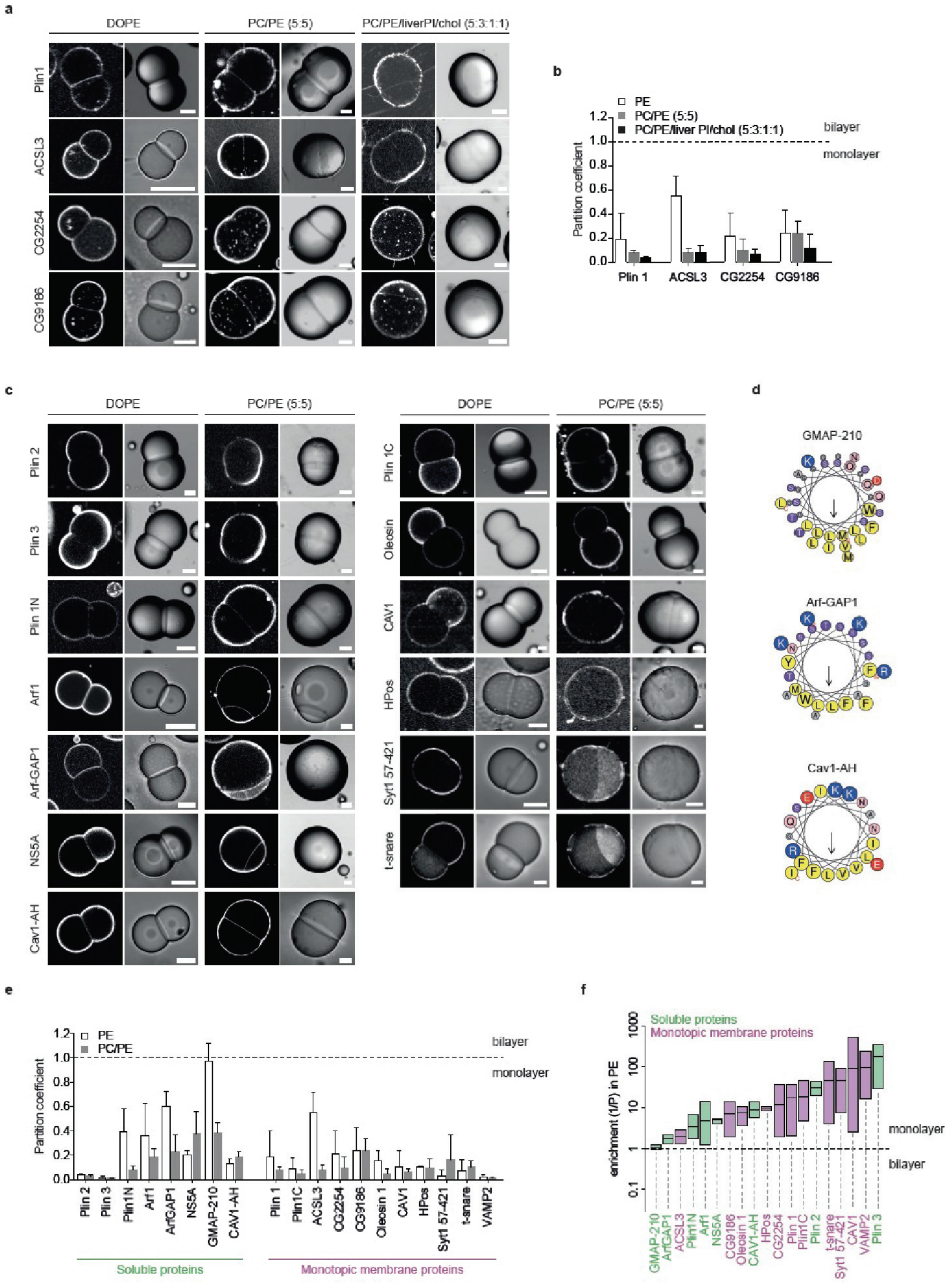
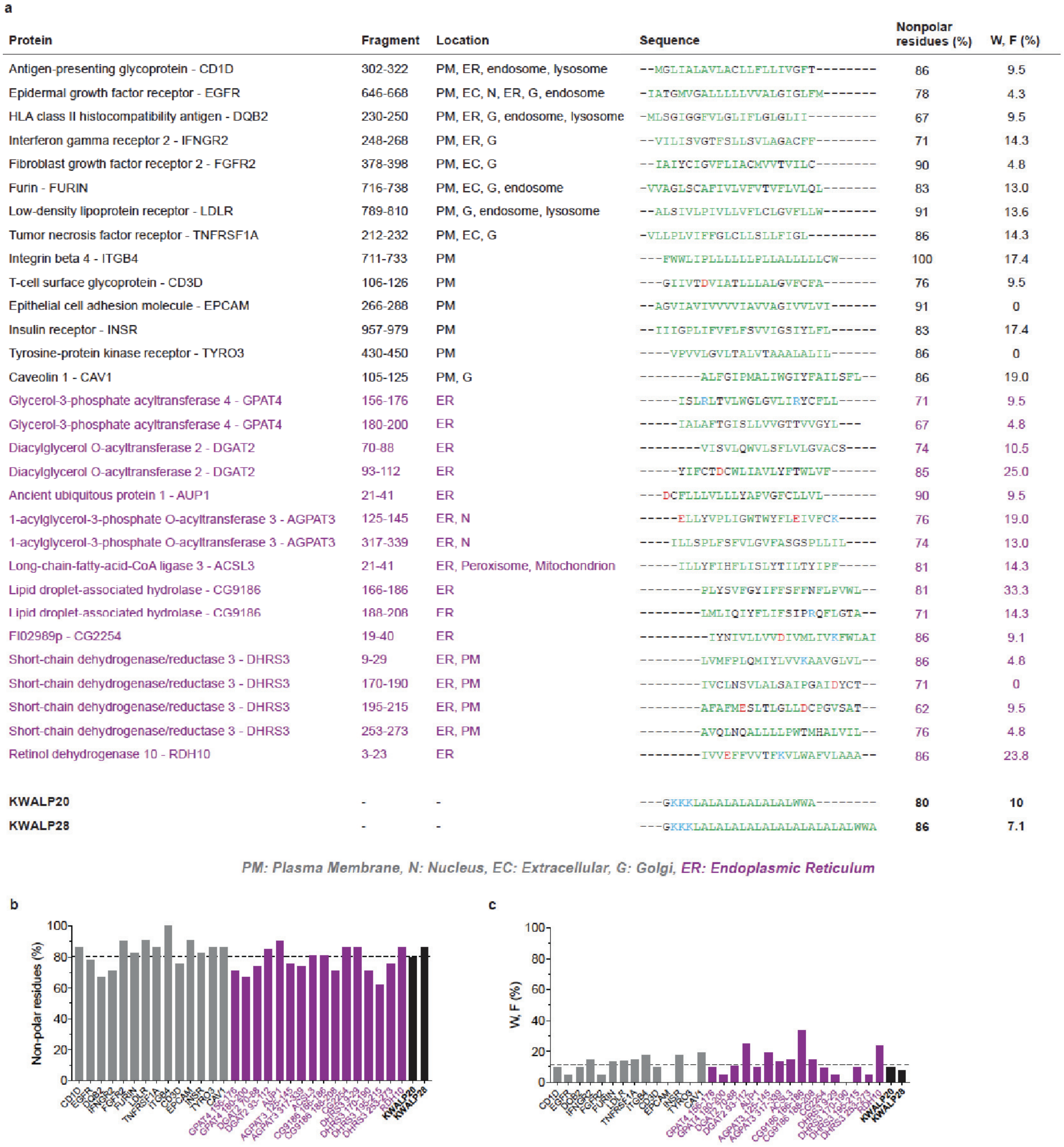

a
DOPE
PC/PE (5:5)
PC/PE/liverPI/chol (5:3:1:1)
Plin1
ACSL3
CG2254
CG9186
b
PE
PC/PE (5:5)
PC/PE/liver PI/chol (5:3:1:1)
bilayer
monolayer
Partition coefficient
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Plin 1
ACSL3
CG2254
CG9186
c
DOPE
PC/PE (5:5)
Plin 2
Plin 3
Plin 1N
Arf1
Arf-GAP1
NS5A
Cav1-AH
Plin 1C
Oleosin
CAV1
HPos
Syt1 57-421
t-snare
d
GMAP-210
Arf-GAP1
Cav1-AH
e
PE
PC/PE
bilayer
monolayer
Partition coefficient
Plin 2
Plin 3
Plin1N
Arf1
ArfGAP1
NS5A
GMAP-210
CAV1-AH
Plin 1
Plin1C
ACSL3
CG2254
CG9186
Oleosin 1
CAV1
HPos
Syt1 57-421
t-snare
VAMP2
Soluble proteins
Monotopic membrane proteins
f
Soluble proteins
Monotopic membrane proteins
enrichment (1/P) in PE
1000
100
10
1
0.1
monolayer
bilayer
GMAP-210
ArfGAP1
ACSL3
Plin1N
Arf1
NS5A
CG9186
Oleosin 1
CAV1-AH
HPos
CG2254
Plin 1
Plin1C
Plin 2
t-snare
Syt1 57-421
CAV1
VAMP2
Plin 3

Figure 3.11 *(previous page)*: AH and HD proteins/peptides distribution in DIBs. (a) Distribution of some LD proteins (Plin1, ACSL3, CG2254, CG9186) in DOPE, DOPC/DOPE (1:1) and DOPC/DOPE/liverPI/cholesterol (5:3:1:1) DIBs. Scale bar: 20µm. (b) Partition coefficient of proteins from experiments shown in (a), shown as mean ± sd from at least 5 independent measurements. (c) Distribution of AHs- (left) and HDs- (right) containing proteins and peptides in DOPE and DOPC/DOPE (1:1) DIBs. Scale bar: 20 µm. HD-containing proteins were coming from lipids droplets (Plin 1C, Oleosin 1, CAV1, HPos) or proteoliposomes (Syt1 57-421, t-snare). (d) Helical wheel representation of the peptides GMAP-210, ArfGAP1 and CAV1-AH, generated with HeliQuest51. (e) Partition coefficient of soluble (green) and monotopic membrane proteins (pink) in DOPE and DOPC/DOPE (1:1) DIBs, represented as mean ± sd. For each condition, $2 \leq$ n independent measurements ≤ 8 were done. (f) Enrichment parameter in PE membranes for AH (green) and HD (pink) proteins shown as floating bars (bar limits, min to max values; central line, mean), $2 \leq$ n independent measurements ≤ 8 for each protein.



Protein	Fragment	Location	Sequence	Nonpolar residues (%)	W, F (%)
Antigen-presenting glycoprotein - CD1D	302-322	PM, ER, endosome, lysosome	--MGLIALAVLACLLFLLIVGFT--------	86	9.5
Epidermal growth factor receptor - EGFR	646-668	PM, EC, N, ER, G, endosome	-IATGMVGALLLLLVVALGIGLFM-------	78	4.3
HLA class II histocompatibility antigen - DQB2	230-250	PM, ER, G, endosome, lysosome	-MLSGIGGFVLGLIFLGLGLII---------	67	9.5
Interferon gamma receptor 2 - IFNGR2	248-268	PM, ER, G	--VILISVGTFSLLSVLAGACFF--------	71	14.3
Fibroblast growth factor receptor 2 - FGFR2	378-398	PM, EC, G	--IAIYCIGVFLIACMVVTVILC--------	90	4.8
Furin - FURIN	716-738	PM, EC, G, endosome	-VVAGLSCAFIVLVFVTVFLVLQL-------	83	13.0
Low-density lipoprotein receptor - LDLR	789-810	PM, G, endosome, lysosome	--ALSIVLPIVLLVFLCLGVFLLW-------	91	13.6
Tumor necrosis factor receptor - TNFRSF1A	212-232	PM, EC, G	-VLLPLVIFFGLCLLSLLFIGL---------	86	14.3
Integrin beta 4 - ITGB4	711-733	PM	---FWWLIPLLLLLLPLLALLLLLCW-----	100	17.4
T-cell surface glycoprotein - CD3D	106-126	PM	---GIIVTDVIATLLLALGVFCFA-------	76	9.5
Epithelial cell adhesion molecule - EPCAM	266-288	PM	--AGVIAVIVVVVIAVVAGIVVLVI------	91	0
Insulin receptor - INSR	957-979	PM	--IIIGPLIFVFLFSVVIGSIYLFL------	83	17.4
Tyrosine-protein kinase receptor - TYRO3	430-450	PM	----VPVVLGVLTALVTAAALALIL------	86	0
Caveolin 1 - CAV1	105-125	PM, G	--------ALFGIPMALIWGIYFAILSFL--	86	19.0
Glycerol-3-phosphate acyltransferase 4 - GPAT4	156-176	ER	-----ISLRLTVLWGLGVLIRYCFLL-----	71	9.5
Glycerol-3-phosphate acyltransferase 4 - GPAT4	180-200	ER	-----IALAFTGISLLVVGTTVVGYL---	67	4.8
Diacylglycerol O-acyltransferase 2 - DGAT2	70-88	ER	--------VISVLQWVLSFLVLGVACS----	74	10.5
Diacylglycerol O-acyltransferase 2 - DGAT2	93-112	ER	-----YIFCTDCWLIAVLYFTWLVF------	85	25.0
Ancient ubiquitous protein 1 - AUP1	21-41	ER	---DCFLLLVLLLYAPVGFCLLVL-------	90	9.5
1-acylglycerol-3-phosphate O-acyltransferase 3 - AGPAT3	125-145	ER, N	-----ELLYVPLIGWTWYFLEIVFCK-----	76	19.0
1-acylglycerol-3-phosphate O-acyltransferase 3 - AGPAT3	317-339	ER, N	----ILLSPLFSFVLGVFASGSPLLIL----	74	13.0
Long-chain-fatty-acid-CoA ligase 3 - ACSL3	21-41	ER, Peroxisome, Mitochondrion	-----ILLYFIHFLISLYTILTYIPF-----	81	14.3
Lipid droplet-associated hydrolase - CG9186	166-186	ER	--------PLYSVFGYIFFSFFNFLPVWL--	81	33.3
Lipid droplet-associated hydrolase - CG9186	188-208	ER	--------LMLIQIYFLIFSIPRQFLGTA--	71	14.3
FI02989p - CG2254	19-40	ER	---------IYNIVLLVVDIVMLIVKFWLAI	86	9.1
Short-chain dehydrogenase/reductase 3 - DHRS3	9-29	ER, PM	--------LVMFPLQMIYLVVKAAVGLVL--	86	4.8
Short-chain dehydrogenase/reductase 3 - DHRS3	170-190	ER, PM	--------IVCLNSVLALSAIPGAIDYCT--	71	0
Short-chain dehydrogenase/reductase 3 - DHRS3	195-215	ER, PM	--------AFAFMESLTLGLLDCPGVSAT--	62	9.5
Short-chain dehydrogenase/reductase 3 - DHRS3	253-273	ER, PM	--------AVQLNQALLLLPWTMHALVIL--	76	4.8
Retinol dehydrogenase 10 - RDH10	3-23	ER	--------IVVEFFVVTFKVLWAFVLAAA--	86	23.8
KWALP20	-	-	---GKKKLALALALALALALWWA--------	**80**	**10**
KWALP28	-	-	---GKKKLALALALALALALALALALALWWA	**86**	**7.1**

PM: Plasma Membrane, N: Nucleus, EC: Extracellular, G: Golgi, ER: Endoplasmic Reticulum

Figure 3.12: Comparison of KWALP peptide with transmembrane domains of several proteins. (a) Sequence alignment of KWALP peptide with transmembrane domains of plasma membrane (black) and LD (purple) proteins. Hydrophobic amino acids are indicated in green, polar ones in black, and the positively and negatively charged in blue and red respectively. The transmembrane fragments, protein location and percentage of non-polar residues are indicated. The proteins CD1D, EGFR, DQB2, IFNGR2, FGFR2, FURIN, LDLR, TNFRSF1A, ITGB4, CD3D, EPCAM, INSR, TYRO3 were used in the analysis which allowed the design of transmembrane model peptides70,71; (b, c) Plot of the percentage of non-polar residues (b) or very hydrophobic residues W and F (c) for each fragment. The dashed line represents the mean of all these fragments.

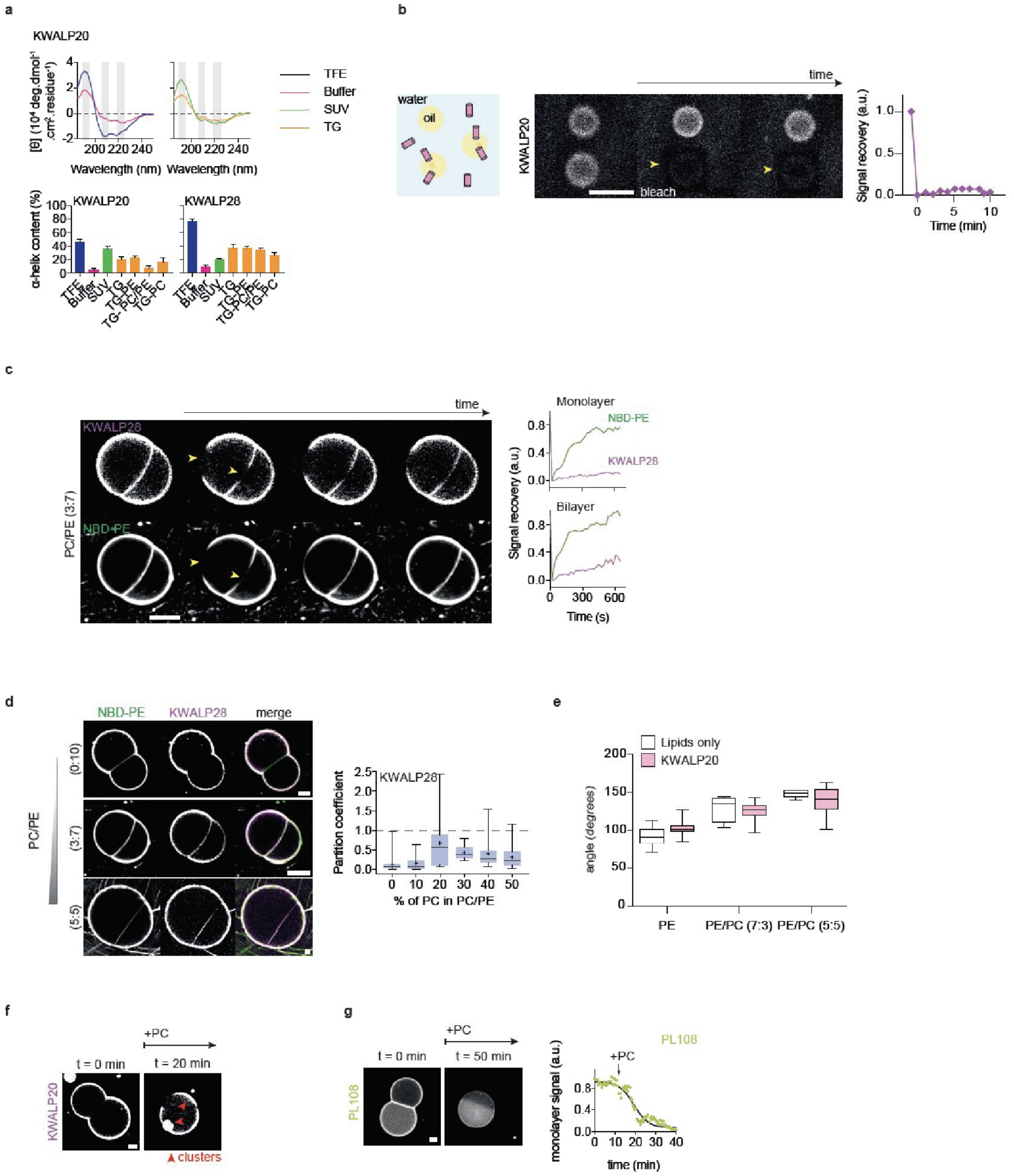

a
KWALP20
[θ] (10⁴ deg.dmol⁻¹.cm².residue⁻¹)
Wavelength (nm)
TFE
Buffer
SUV
TG
α-helix content (%)
KWALP20
KWALP28
TFE Buffer SUV TG TG-PE TG-PC/PE TG-PC
b
water
oil
time
KWALP20
bleach
Signal recovery (a.u.)
Time (min)
c
time
KWALP28
NBD-PE
PC/PE (3:7)
Monolayer
Bilayer
NBD-PE
KWALP28
Signal recovery (a.u.)
Time (s)
d
NBD-PE
KWALP28
merge
PC/PE
(0:10)
(3:7)
(5:5)
KWALP28
Partition coefficient
% of PC in PC/PE
e
Lipids only
KWALP20
angle (degrees)
PE
PE/PC (7:3)
PE/PC (5:5)
f
+PC
t = 0 min
t = 20 min
KWALP20
clusters
g
+PC
t = 0 min
t = 50 min
PL108
PL108
+PC
monolayer signal (a.u.)
time (min)

Figure 3.13 *(previous page)*: KWALP folds in a-helix, PC/PE ratio modulates AH and HD localization. (a) Circular dichroism spectra (top) of KWALP20 in various environments. Trifluoroethanol (TFE) is known to favor the a-helix structuration, the resulting graph (black) shows indeed the characteristics of an a-helix signal: a maximum around 192 nm, and two minima around 208 and 222 nm. These features are found also in buffer (purple), liposomes (green), TG droplets (blue), with a signal less intense. The spectra were deconvoluted to extract a-helix content (bottom). The result is the mean ± sd from 21 deconvolution programs. (b) TG in water emulsion. Rh-KWALP20 in solution relocalizes to oil/water interface. The fluorescence of the bottom droplet is bleached (yellow arrow). After 10 minutes, no fluorescence recovery was observed. Scale bar: 20 μm. (c) FRAP experiment showing the recovery of KWALP28 and lipids signal after fluorescence bleaching. Yellow arrows indicate the bleached areas. The signal recovery in the monolayer and in the bilayer is plotted against time (right). Scale bar: 20 μm. (d) KWALP28 in DIBs of varying DOPC/DOPE ratios. KWALP28 is labelled with Rh-B and lipids with 0.2% NBD-PE. Scale bar: 20 μm. The partition coefficient for each lipid composition is plotted (right) as box-plots (box limits, upper and lower quartiles; middle line, median; +, mean; whiskers, minimum and maximum value). Sample size was n=19 for 0% PC, n=10 for 10% PC, n=19 for 20% PC, n=14 for 30% PC, n=16 for 40% PC and n=24 for 50% PC. (e) DIBs contact angle is represented as a function of DOPC concentration with box-plots (box limits, upper and lower quartiles; middle line, median; whiskers, minimum and maximum value). Sample size is n=8, 5, 7 respectively for PE, PE/PC (7:3) and PE/PC (5:5) in the absence of KWALP20, and n=23, 24, 20 in the presence of KWALP20. (f, g) Relocalization of KWALP20 (f) and PL108 (g) after addition of DOPC in DOPE DIBs. After 20 minutes KWALP20 shows some aggregates (red arrows). Scale bar: 20 μm. The monolayer signal of PL108 is plotted against time (g)

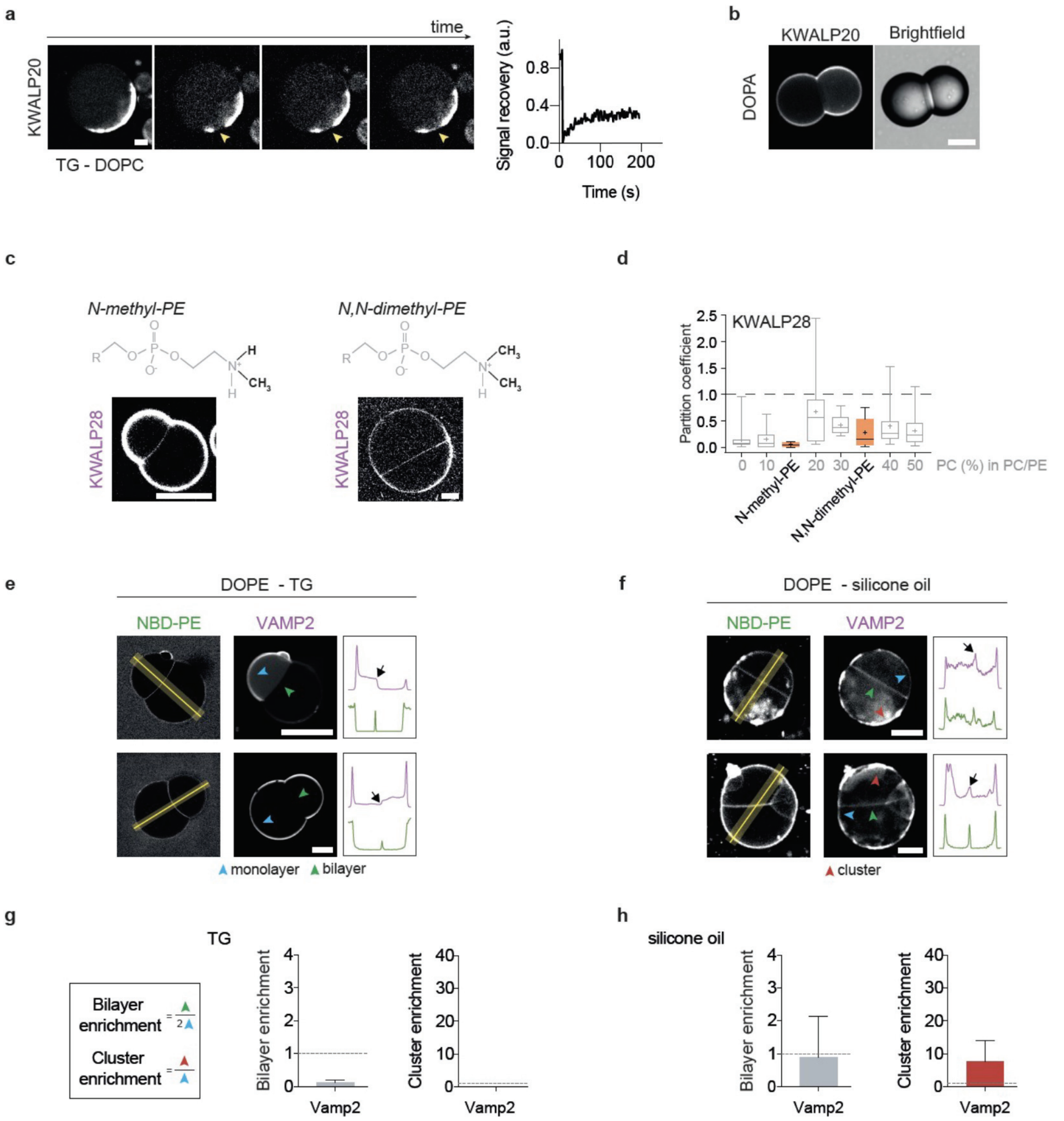


Figure 3.14: Influence of phospholipid shape and oil nature on KWALP distribution. (a) FRAP experiment on a cluster of KWALP20 in a DOPC droplet. Yellow arrow indicates the bleached area. Signal intensity was plotted against time. Scale bar: 20 µm. (b) KWALP20 in DOPA DIBs. Scale bar: 20 µm. (c) KWALP28 distribution in N-methyl-PE and N,N-dimethyl-PE DIBs. The headgroup methylation differs in these phospholipids. KWALP28 is labelled with Rh-B. Scale bar: 20 µm. (d) Partition coefficient of KWALP28 in DIBs of different compositions is represented as box-plots (box limits, upper and lower quartiles; middle line, median; +, mean; whiskers, minimum and maximum value). Sample size is n=8 and 9 for KWALP28 in N-methyl-PE and N,N-dimethyl-PE respectively. Previous results of varying PC/PE ratios (Fig. S4d) are reported in light grey. (e, f) Distribution of VAMP2 in DOPE DIBs. VAMP2 is labelled with Atto565. Oil phase was TG (e) or silicone oil (f). Blue arrows indicate monolayers, green ones the bilayers, and red ones the clusters areas. Plot profiles are determined using the yellow lines. The bilayer signal is indicated by a black arrow. Scale bar: 20 µm. (g, h) Partition coefficient is reported in grey in TG (g) and silicone oil (h) as mean ± sd (n=2 and 8 independent measurements respectively). A cluster enrichment coefficient (red) was also determined for the experiment in silicone oil to take account of the aggregation and is shown as mean ± sd (n=8 independent measurements).

Chapter 4

Lipid droplet shape and tendency towards budding: insight from theory and molecular simulations

Vincent Nieto[1], Jackson Crowley[1], Francois Deslandes[2], Abdou Rachid Thiam[3], Lionel Foret[3,*], Luca Monticelli[1,*]

[1] University of Lyon, CNRS, Molecular Microbiology and Structural Biochemistry (MMSB, UMR 5086), F-69007, Lyon, France

[2] Université Paris-Saclay, INRAE, MaIAGE, 78350, Jouy-en-Josas, France

[3] Laboratoire de Physique de l'École Normale Supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris Cité, F-75005 Paris, France

4.1 Abstract

Lipid droplets (LDs) are cellular organelles responsible for lipid storage and metabolism. They resemble oil-in-water emulsions, coated with a monolayer of phospholipids and proteins. The mechanism of biogenesis of LDs is poorly understood, but it is well established that it involves phase separation of oil (consisting of neutral lipids, such as triglycerides) from the surrounding phospholipids, which generates oil lenses embedded in lipid bilayers, also known as nascent LDs. As nascent LDs grow, at some point they bud out of the bilayer membrane, forming nearly spherical droplets. Nascent LDs have different propensity to bud, and it has been

proposed that their shape provides information on such propensity to bud. Here we studied the shape of lipid droplets as obtained from MD simulations at the coarse-grained level, and compared it to the predictions by an established theory. We find that the shape of simulated LDs resembles a spherical cap (i.e., it has constant curvature) already for very small droplet sizes, in excellent agreement with the theory. The aspect ratio (height/radius) of nascent LDs increases with increasing LD volume, increasing membrane softness, and increasing surface tension between oil and water, in agreement with theoretical predictions. Fitting the simulated LD shapes with a theoretical shape equation suggests that either a negative line tension exists at the monolayer-bilayer junction or a non-zero surface tension is present in the bilayer region. We find that the line tension at the monolayer-bilayer junction is positive, not negative, and increases with the size of the LD, suggesting that it originates from the deformation (curvature) of the bilayer as it unzips into two monolayers. A positive line tension can act as a driving force for LD budding, but cannot explain the shape of simulated LDs. LD shape is instead well explained by the high surface tension in the bilayer region, as confirmed by local stress calculations. Such high surface tension in the bilayer is due to the mechanical equilibrium condition imposed by the pressure coupling algorithm. Our results show that simulations of nascent LD systems provide a microscopic view into the properties of droplet embedded vesicles.

4.2 Introduction

Lipid droplets (LDs) are generated from the endoplasmic reticulum (ER) membrane, and perform their functions in lipid storage and metabolism when they are budded out of the ER membrane, towards the cytosolic compartment (2; 12; 78). Therefore, it is very important to understand what determines the direction of budding and the propensity of nascent LDs to bud. Theoretical models can provide useful insight into the driving forces for LD budding (14; 116; 117). Available theories are based on a continuous media description (Helfrich membrane theory (118)) and take into account mechanical properties of membranes, such as their elastic energy, the surface tension of the interfaces, and the spontaneous curvature of each membrane leaflet. Thiam and Foret developed a theory describing the mechanics of LD budding starting from a nascent LD, i.e., an oil lens sandwiched between the ER leaflets

(116). According to this theory, the oil lens is flat at small oil volume, when the typical droplet size is smaller than the elastic length $\lambda = \sqrt{\frac{\kappa_m}{\gamma_m}}$ where κ_m is the bending rigidity of the monolayer at the droplet interface and γ_m is surface tension. At larger size, the surface tension dominates and imposes a nearly spherical shape. In this situation, the lens can become unstable leading to the spontaneous formation of a spherical protrusion because of the monolayer deformation near the droplet edge that gives rise to an effective line tension. The analysis of LD shape is particularly interesting in the context of LD biogenesis, as shape is a proxy for the tendency of a nascent LD to bud out of the bilayer membrane: nascent LDs with a shape close to a sphere have a high tendency to bud, while flatter ones tend to remain embedded in the bilayer membrane (79). Such correspondence between shape and tendency to bud has been observed experimentally, both in synthetic model systems (82) and in biological systems (119).

As stated above, theories are based on a continuum description of the system and contains several approximations (14; 116; 117). On the other hand, the early stages of LD budding are difficult to study experimentally, and obtaining high-resolution images of budding LDs is very challenging for current microscopy techniques. To provide a microscopic interpretation of the theory and a molecular view into the mechanism of budding, here we perform molecular dynamics simulations at the coarse-grained level for nascent LDs of different size. We compare LD shapes predicted by the theory and calculated directly from large scale MD simulations. We find that the shape of simulated LDs has constant curvature, i.e., resembles a spherical cap, already for droplets of very small size (10 nm), indicating that the shape is largely dominated by surface tension. On the other hand, the aspect ratio (height/radius) of small nascent LDs is small, i.e., small LDs are rather "flat", indicating that they are not prone to budding. Line tension at the junction between the nascent LD and the bilayer region is positive and increases with LD size, suggesting that the origin of such line tension is the deformation of the junction, where the bilayer unzips into two monolayers. We also find that simulations feature a high surface tension in the bilayer region, not only in the monolayer region, and that is due to the mechanical equilibrium condition imposed by the pressure coupling algorithm. Our results indicate that simulations of nascent LD systems may be used to provide a microscopic view into the properties of droplet embedded vesicles.

4.3 Methods

4.3.1 System setup

To generate lipid droplets in bilayer membranes we used molecular dynamics simulations at the coarse-grained level (Martini force field (67), see next section for details). In the majority of the simulations, we used dioleoylphosphatidylcholine (DOPC) as phospholipid and triolein (TG) as neutral lipid. We initially built phospholipid bilayers containing 2016 or 18144 lipids using the Insane software (120), then separated the two leaflets and added a layer of TG (325 to 7500 TG molecules) in between them. This resulted in trilayer systems, with a uniform layer of TG sandwiched between two layers of DOPC. The systems were solvated with water particles. The exact composition of each system is reported in Table 4.1. Hereafter, systems with 2016 phospholipids will be referred to as "small", the other as "large" systems. After equilibration, these systems rapidly form lens-shaped nascent LDs (Figure 4.1a). Additional small systems were generated containing 625 TG molecules and different phospholipid types: DLPC, DPPC, POPC, and a mixture of 50% POPC and 50% DPPC. Finally, additional small systems were generated using two different neutral lipids, namely trilauryl glycerol (12 carbon atoms) and trioctanoyl (8 carbon atoms) glycerol. Table 4.1 contains a full list of the simulated systems.

To calculate line tension at the bilayer-monolayer junction, we built tube-shaped lipid droplets (hereafter referred to as lipid tubules or oil tubules), i.e., DOPC bilayer systems with embedded TG droplets connected in one direction (the y axis) with their periodic images, therefore forming infinite TG tubules running along the Y direction (see Figure 4.1c). Oil tubules were built using the Insane software (120) along with the Gromacs software (121). Six oil tubules were built, with three different amounts of TG and two different lengths, 10 or 20 nm (along the Y axis). All six systems contained DOPC phospholipids and TG as neutral lipid. In all cases we used semi-isotropic pressure coupling; compressibility was $3 \cdot 10^{-4}\text{bar}^{-1}$) in the Z direction, and either the same or zero in the XY plane; the second amounts to no pressure coupling in the XY plane, yielding constant XY size in the system. In this case, the only forces in the direction of the tube are resulting from the line tension, which can be calculated easily, based on the mechanical equilibrium condition, from the values of the pressure tensor:

$$\tau_y = L_x \cdot L_y \cdot \frac{\frac{1}{2}(P_{xx} + P_{zz}) - P_{yy}}{2}$$

To check for simulation artifacts in the calculations of surface tension, we built (periodic) bilayer stripes from a 200 DOPC bilayer using the Insane code (120). The bilayer was initially periodic along the X and Y axes, and we extended the X dimension and added solvent, so that two sides of the bilayer were not in contact with their periodic images (see Figure 4.1d). The bilayer stripes, running along the Y dimension, were simulated with fixed length in Y by setting the compressibility along Y to zero; pressure coupling was applied along X and Z dimensions. These bilayer systems have zero surface tension as they can freely adjust their area (they are not subject to any restraint or pressure coupling along the x dimension) and possess no curvature. The same system was used for the calculation of line tension along the bilayer edge.

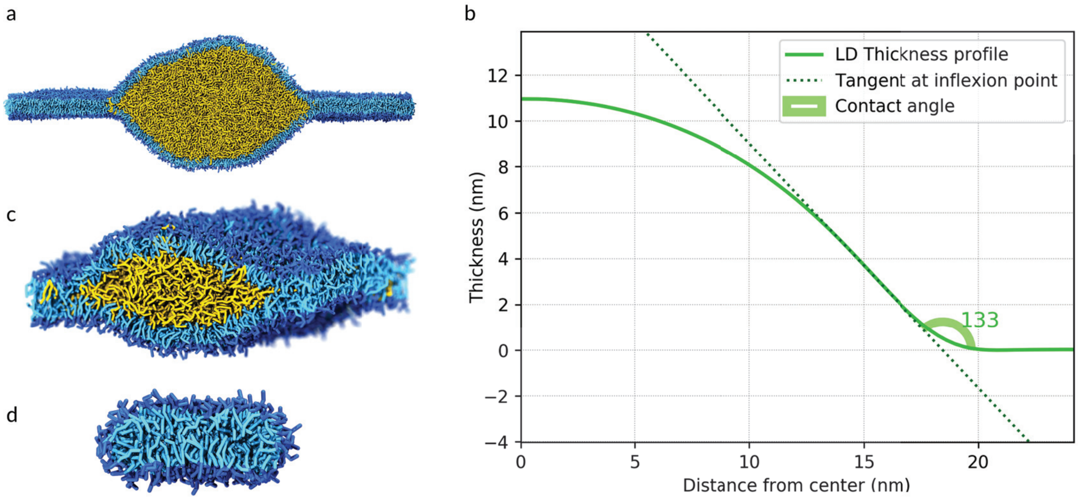


Figure 4.1: (a) Nascent LD as simulated in this work, and used for the calculation of LD shape and contact angle. (b) Illustration of the methodology to calculate the LD contact angle. (c) Oil tubules, used to calculate the line tension at the LD-bilayer connection. (d) Bilayer stripes, used to calculate the line tension at the bilayer edge.

4.3.2 Simulation details

All simulations were performed with the Gromacs software (v.2016) (121), using the coarse-grained description provided by the Martini force field (version 2)(67). The Martini model

Name	Description	Temperature (K)	Box-X	Box-Y	Box-Z	Production run (µs)	Growth (µs)	Equilibium (µs)
2016DOPC+325TO	Flat bilayer with circular lipid droplet	300	26,6	26,6	18,2	10	2	8
2016DOPC+625TO	Flat bilayer with circular lipid droplet	300	26,1	26,1	19,6	12	2	10
4050DOPC+1250TO	Flat bilayer with circular lipid droplet	300	37,1	37,1	26,0	20	4	16
18144DOPC+2500TO	Flat bilayer with circular lipid droplet	300	78,9	78,9	23,8	40	0	40
18144DOPC+5000TO	Flat bilayer with circular lipid droplet	300	77,0	77,0	28,7	60	56	4
18144DOPC+7500TO	Flat bilayer with circular lipid droplet	300	78,1	78,1	36,0	60	38	22
2016DLPC+625TO	Flat bilayer with circular lipid droplet	300	25,7	25,7	19,5	40	4	36
2016DLPC+625TO	Flat bilayer with circular lipid droplet	323	25,7	25,7	19,5	40	2	38
2016DOPC+625TO	Flat bilayer with circular lipid droplet	323	26,8	26,8	18,5	32	2	30
2016DPPC+625TO	Flat bilayer with circular lipid droplet	323	26,8	26,8	18,5	40	1	39
2016POPC+625TO	Flat bilayer with circular lipid droplet	300	26,8	26,8	18,5	40	4	36
2016POPC+625TO	Flat bilayer with circular lipid droplet	323	26,8	26,8	18,5	39	4	35
2016POPC-DPPC+625TO	Flat bilayer with circular lipid droplet	300	26,2	26,2	19,4	40	2	38
2016POPC-DPPC+625TO	Flat bilayer with circular lipid droplet	323	26,2	26,2	19,4	40	2	38
200DOPC stripe	Bilayer stripe: Semiisotropic PC, CompXY=0	300	16,0	8,0	16,0	10		10
200DOPC stripe	Bilayer stripe: Semiisotropic PC; CompXY=3e^{-4}	300	16,0	8,0	16,0	10		10
200DOPC stripe	Bilayer stripe: SurfaceTension; CompXY=0	300	16,0	8,0	16,0	10		10
200DOPC stripe	Bilayer stripe: SurfaceTension; CompXY=3e^{-4}	300	16,0	8,0	16,0	10		10
200DOPC+CHOL stripe	Bilayer stripe: Semiisotropic; CompXY=0	300	16,0	8,0	16,0	10		10
200DOPC+CHOL stripe	Bilayer stripe: Semiisotropic; CompXY=3e^{-4}	300	16,0	8,0	16,0	10		10
200DOPC+CHOL stripe	Bilayer stripe: SurfaceTension; CompXY=0	300	16,0	8,0	16,0	10		10
200DOPC+CHOL stripe	Bilayer stripe: SurfaceTension; CompXY=3e^{-4}	300	16,0	8,0	16,0	1		1
728DOPC+419TO TO	Tubular droplet along Y=10nm, CompY=0	300	24,0	10,0	17,0	15		15
1456DOPC+838TO TO	Tubular droplet along Y=20nm, CompY=0	300	24,0	20,0	17,0	15		15
1977DOPC+1250TO TO	Tubular droplet along Y=10nm, CompY=0	300	65,0	10,0	25,0	15		15
3954DOPC+2500TO TO	Tubular droplet along Y=20nm, CompY=0	300	65,0	20,0	25,0	15		15
2498DOPC+3434TO TO	Tubular droplet along Y=10nm, CompY=0	300	80,0	10,0	40,0	15		15
4996DOPC+6868TO TO	Tubular droplet along Y=20nm, CompY=0	300	80,0	20,0	40,0	15		15

Table 4.1: Composition and size of the simulated systems. The growth phase corresponds to 5τ (where τ is the autocorrelation time of LD size, see Methods), the equilibrium is used for the analysis of shape and equilibrium properties and corresponds to the difference between the total duration of the run and growth phase.

has been widely used in simulations of various biomolecular systems containing membranes, proteins (68), nucleic acids (70), as well as synthetic polymers (73) and nanoparticles (71), and has been shown to provide a realistic description of lipid aggregates and their properties (74). A cutoff at 1.1 nm was applied for the calculation of non-bonded interactions with the Verlet cutoff scheme, together with the reaction field for electrostatics interactions. The potential was shirted to zero between 0.9 and 1.1 nm with the Potential-shift-Verlet function (121). The same parameters were used in the calculation of local stresses with the Gromacs-LS software (122). MD simulations were carried out with the leap-frog integrator and a time step of 20 fs. Temperature was kept constant with the Bussi-Donadio-Parrinello (v-rescale) thermostat (64) (time constant of 1 ps). Pressure was kept constant in equilibration with the Berendsen barostat (63) (time constant of 20 ps, compressibility $4 \cdot 10^{-5}$bar^{-1}) and in production runs with the Parrinello-Rahman barostat (65) (time constant of 20 ps, compressibility $4 \cdot 10^{-5}\text{bar}^{-1}$). Semi-isotropic pressure coupling was applied in all cases. Systems were energy minimized using the steepest descent algorithm for 500 steps. After energy minimization, the systems were equilibrated for 10 ns in the NPT ensemble. During equilibration neutral lipids started to phase-separate and formed lens-like nascent lipid droplets embedded in the phospholipid bilayer. Then, production runs were carried out for 10-60 μs for each system (duration reported for each system in Table 4.1).

4.3.3 Simulation analysis

The number of TG molecules in the LD was monitored over time using our g_aggregate code (123). A modification of this code also allowed rewriting the trajectory with the center of mass of the LD translated at the center of the simulation box, resulting in trajectories with a virtually immobile LD. Shape analysis was carried out with our g_thickness code (90) the thickness of the membrane, defined as the distance between the phospholipid heads (PO4 beads) of each leaflet, was computed over an XY-grid, and averaged over time, providing a thickness landscape. The shape of the LD or thickness profile was obtained by averaging the landscape around the axis of symmetry (vertical axis) of the LD. A python code was developed to extract the geometric properties of the LD shape from the thickness profile. Briefly, the LD shape was fit to a third-degree polynomial function. The radius at which the second derivative of the polynomial approximation cancels out gives the coordinates of the inflection point of the shape curve. The tangent line to the curve at this inflection point was computed, and the intersection of the tangent with the line defining the bilayer was used to estimate the LD radius. The angle between these two lines was defined as "contact angle", in analogy with the contact angle defined in the theory of wetting/dewetting (see Figure 4.1b). We used the software Gromacs-LS (122) to compute the local stress tensor in simulations. For the local stress tensor grid, we used a cubic voxel of 0.5 nm corresponding to the size of Martini beads. Since our nascent LD systems are large (almost 80 nm lateral size), undulations of the bilayer and monolayer are very substantial; as a consequence, averaging the local stress tensor over the XY plane and over time requires particular care to avoid that undulations cancel out contributions from subsequent time frames and from regions in the XY plane that have different position with respect to the membrane (e.g., different distance from a reference point, such as the center of the bilayer, or the head group region). To this end, we developed a specific code to take into account bilayer and monolayer fluctuations, and extract the pressure profile from the local stress tensor. Averaging of the local stress tensor was carried out using in-house software.

4.4 Results

4.4.1 Assessment of equilibrium thermodynamics and kinetics

We first generated trilayer systems, with a layer of TG sandwiched between two phospholipid layers (see Methods). The trilayer systems are unstable and evolve rapidly by phase separation of TG into a compact droplet, that is the nascent lipid droplet, embedded into the lipid bilayer; part of the TG remains dispersed in the bilayer, indicating that phase separation is dynamic, as expected. Similar observations have been reported before, based on MD simulations (4; 56; 124). Starting from a uniform distribution of TG molecules, we generally observed nucleation of one or more aggregates, and then LD growth during several microseconds. If we want to analyze LD shape, we need to consider only the fraction of simulations representing an equilibrium ensemble, hence we need a quantitative way to assess the equilibration of the system. To this aim, we calculated the size of the largest TG cluster (that is the lens-shaped nascent LD) over time. We observed that, after an initial period of growth, the number of TG molecules in the nascent LDs becomes approximately constant, fluctuating around an average value that depends on system composition, size, and temperature. We define the initial period as the growth phase, and the following as the equilibrium phase. Examples illustrating growth and equilibrium phases are shown in Figure 4.2.

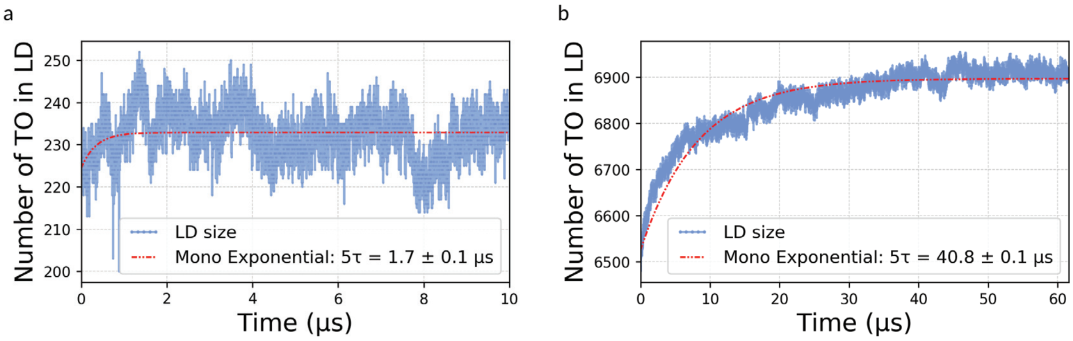


Figure 4.2: Assessment of system equilibration. (a) Number of TG molecules in the nascent LD in simulation with 2016 DOPC and 325 TG. (b) Number of TG molecules in the nascent LD in simulation with 18144 DOPC and 7500 TG.

The duration of the growth phase was different in each simulation, and appeared to depend mainly on the specific initial conditions of the system and the size of the bilayer. As expected, larger systems required much longer equilibration, due to the longer time required for diffusion of TG across the system. For instance, systems of ∼80 nm in lateral size required equilibration times around 50 μs. Our estimate cannot be precise as we simulated only 1 replica of the larger system, due to the large computational cost; therefore, we refrain from providing any quantitative interpretation of growth kinetics. However, available data suggests that the duration of the growth phase does not depend strongly on the number of TG molecules, nor on the nature of the phospholipids (which all form fluid bilayers in the simulation conditions), but rather on the surface area of the system, suggesting that equilibration times are mostly limited by diffusion. Having estimated the equilibration time for each simulation of interest, we extended the duration of the simulation beyond the growth phase (see Table 4.1). We then analyzed the shape of the nascent LD only considering the equilibrium phase of the simulations, when the number of TG molecules in the nascent LD is approximately constant.

4.4.2 Simulations are in qualitative agreement with the theory

One of the effects predicted by theory is that LDs with an increasing volume should become more spherical (116). One parameter that indicates whether the LD shape is close to the spherical shape is the curvature of the LD profile: constant curvature is characteristic of spherical shapes. Another parameter indicating similarity to a spherical shape is the LD aspect ratio, i.e., LD height over LD radius (116). This is also related to the contact angle, in analogy with dewetting theory (79). Here we define the contact angle as the angle formed by the membrane plane with the tangent to the droplet profile at its inflection point (Figure 4.1b). The contact angle decreases (to approach the value of 90 degrees, in the case of a symmetric bilayer) as the LD aspect ratio increases.

We calculated LD curvature, aspect ratio, and contact angle for six nascent LDs with increasing volume and the same phospholipid composition, namely pure DOPC. Shape analysis shows that the curvature is constant at the center of the spherical caps for all simulated LDs; for large LDs, the curvature is constant over a large fraction of the LD surface (Figure 4.3c-e). Constant curvature indicates that the LD shape is controlled by surface tension of the monolayer region (116), that become prevalent already at relatively small sizes (20 nm in diameter). At the same time, the aspect ratio of all simulated nascent LDs is small and the contact angle is large, indicating that the droplets are not close to budding. The contact angle decreases as LDs get larger, confirming that larger LDs are closer to the budding transition, as expected (Figure 4.3b). Such change in shape is due to the increase in surface energy relative to elastic energy, as the radius of the LD increases. At large enough sizes, the contribution of bending energy to the shape of the LD becomes negligible compared to the contribution from surface tension, and we expect that the main contribution to LD shape comes from the increased surface energy. These findings agree well with theoretical predictions. We also notice that, beyond the spherical cap region, LDs are "stretched", suggesting the existence of forces pulling on the edges of the nascent LD. We will explore such forces later in the paper.

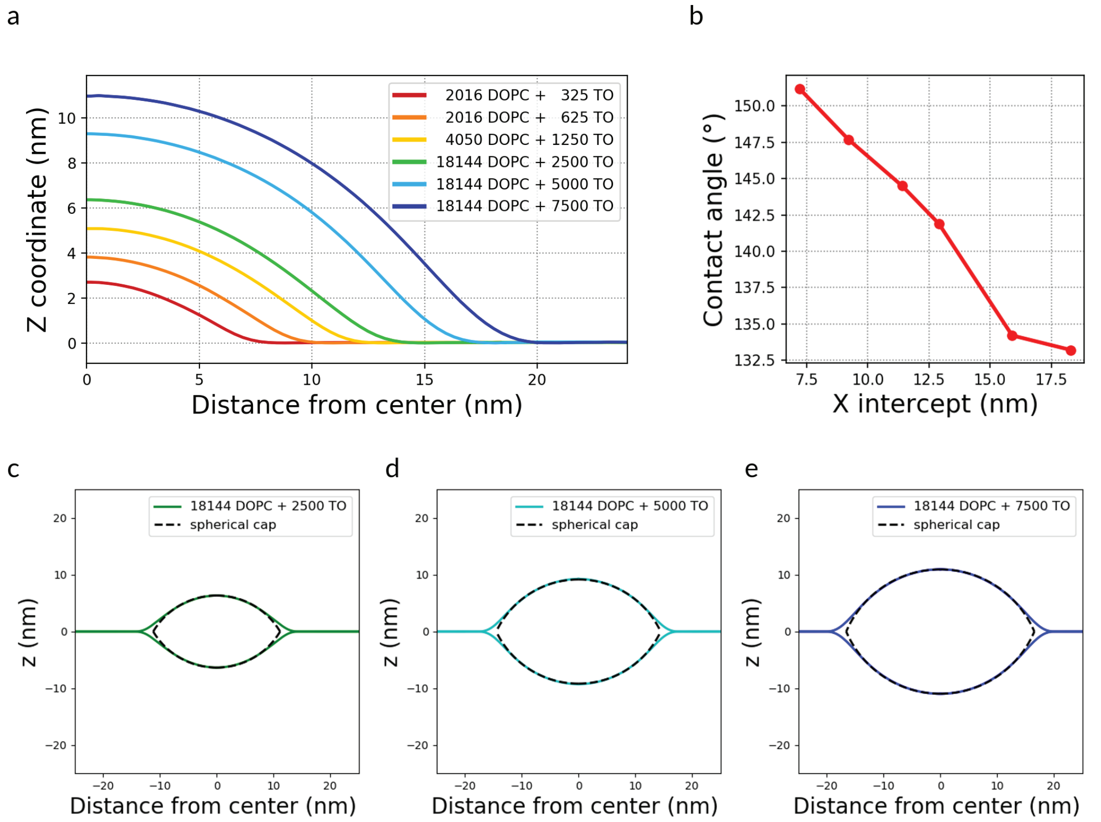


Figure 4.3: Shape analysis for nascent LDs. (a) Shape profile, after averaging around the axis of symmetry of the LD and over time, for LDs of different size. (b) Contact angle for the nascent LDs in panel B, as a function of the size (expressed as the x intercept). (c) Fitting the simulated LD shape with a spherical cap, for simulated nascent LDs containing 2500 TO lipids, (d) 5000 TO lipids, (e) 7500 TO lipids.

LD shape theoretically depends on the interplay between the monolayer bending modulus and surface tension (116): higher monolayer surface tension favor spherical shapes, while higher rigidity disfavors them. We therefore aimed at quantitively evaluating the contribution of each of these factors to LD shape. Measuring the bending moduli of phospholipid monolayers at the TG/water interface is problematic, as the values should depend on the level of coverage of the monolayers (related to their surface tension) and on the amount of TG dissolved in the monolayer. Calculations of the same quantities from MD simulations are also problematic: computational methods to calculate the bending modulus are well tested for lipid bilayers (125), not on monolayers. However, it is reasonable to assume that, to a first approximation, the trends in bending moduli vs lipid composition should be the same for bilayers and monolayers (i.e., if a bilayer consisting of lipid X has bending modulus larger than a bilayer made from lipid Y, then the same will be true also for monolayers). It is also reasonable to assume that the effect of TG on the bending modulus is approximately the same on monolayers with different phospholipid composition. In other words, it is reasonable to assume that the bending modulus of lipid monolayers at the oil-water interface are approximately half of the bending modulus of bilayers in contiguity, with the same chemical composition.

To understand the effect of bending rigidity on LD shape, we simulated nascent LDs with the same size (i.e., same number of phospholipids and TG molecules) and different phospholipid composition, namely DLPC, DOPC, DPPC, POPC, POPC+DPPC. Shape analysis (Figure 4.4a) shows that LDs made from lipid bilayers with a high bending modulus have a higher contact angle (Figure 4.4c), i.e., LDs embedded in more rigid bilayers have a lower aspect ratio, as predicted by continuum theory (116). Our results suggest that nascent LD are closer to a spherical shape when formed in softer membranes.

To understand the influence of surface tension on LD budding, we simulated nascent LDs containing different triglycerides, having different interfacial tensions with water. We chose trimyristin and tricaprin, which have a lower tension than triolein. The simulated nascent LDs contained the same number of neutral lipids and phospholipids and the same phospholipid composition. Shape analysis shows a correlation between surface tension and contact angle: nascent LDs with lower surface tension have lowed aspect ratio and lower contact angle (Figure 4.4b). TG, the most abundant neutral lipids present in LDs in cells, has a higher surface tension and therefore LDs composed of TG are more prone to budding.

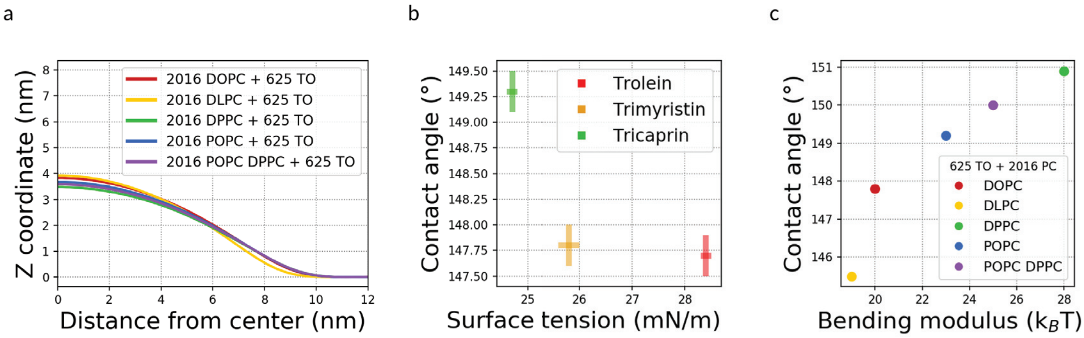


Figure 4.4: LD shape and bilayer rigidity. (a) LD shape profiles for LDs with the same number of neutral lipids and phospholipids, but different nature of the phospholipids. (b) Contact angle as a function of surface tension of the oil-water interface. (c) Contact angle as a function of pure lipid bilayer bending modulus of the bilayer, as calculated in ref. (125).

4.4.3 Quantitative comparisons between simulations and continuum theory

The shape profiles of nascent LDs simulated with our coarse-grained model comply qualitatively with the predictions of continuum theory, but do they reproduce quantitatively those predictions? We addressed this question by fitting our simulation-based shape profiles with theoretical shape from ref. (116). To this end, we used the same assumptions as in the theoretical analysis; in particular, long-range molecular interactions contribution to the line tension and the spontaneous curvature of the monolayers were neglected, simplifying the theoretical shape equations. Under those assumptions, the LD shape equations contain only 3 parameters: the elastic length λ defined as $\lambda^3 = (\kappa_m/\gamma_m)^{1/2}\ V/\lambda^3$ (where V is the LD volume), and γ_b/γ_m (surface tension of the bilayer and the monolayer). The volume of the droplet being known, we used λ and γ_b/γ_m as fitting parameters. Values of surface tension for the monolayer γ_m can be estimated from λ (that is obtained from the fitting procedure), with some assumptions on the bending rigidity of the monolayer κ_m. Bending rigidities have been measured at $\sim 20 k_B T$ for pure DOPC bilayers in numerous studies, using different experimental and simulation methods (125; 126; 127). Furthermore, Santinho et al. showed that the addition of a small amount of TG decreased the bending modulus by nearly a factor of 2, from $21 \pm 4 k_B T$ to $12 \pm 4 k_B T$ (7). Assuming that the bending rigidity of monolayers is close to half of the value for the bilayer (128), we estimate an approximate value of $\kappa_m = 6\ k_B T$. This is in good agreement with values estimated for DOPC monolayers (128).

Simulations were carried out using the semi-isotropic pressure coupling method, with pressure of 1 bar in the XY plane and in the Z dimension. The same setup is generally used

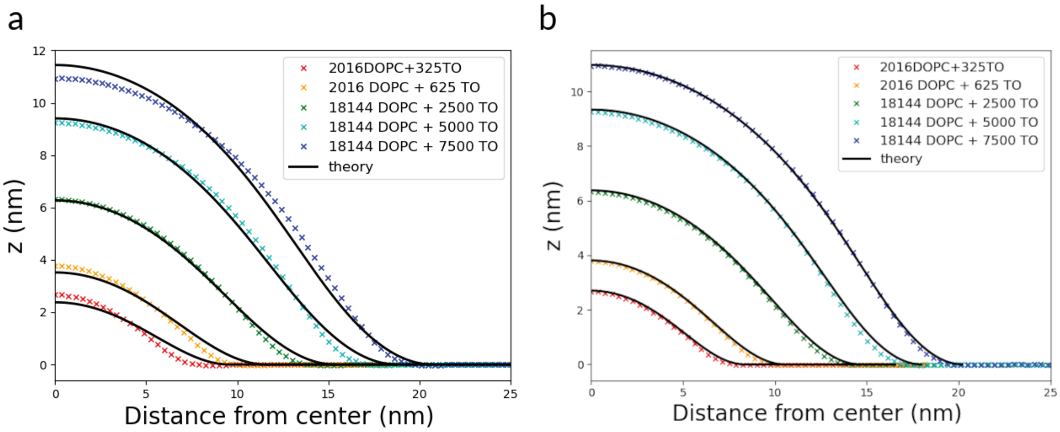


Figure 4.5: Fitting the simulated shape profiles with the theoretical shape equation. Continuous lines represent the shape predicted by the theory, symbols the simulated shape. (a) Fit obtained by imposing zero surface tension in the bilayer. The fitting procedure yields λ = 6.877 nm, from which we calculate $\gamma_m = 0.51$ mN/m. (b) Fit obtained with λ = 3.308 nm, and $\gamma_b/\gamma_m = 0.834$ (7500 TG), 0.844 (5000 TG), 0.944 (2500 TG), 0.956 (625 TG), 0.824 (325 TG).

in simulations of bilayer membranes, where it provides zero surface tension for the membrane system. Based on the expectation of a tensionless bilayer, we first fitted the simulation-based shapes imposing zero surface tension in the bilayer, i.e., using a single fitting parameter, λİn this case, the results were unsatisfactory: the simulated shapes had approximately constant curvature in the central region, while theoretical shapes did not (Figure 4.5a). Indeed, in the absence of bilayer tension, shapes with constant curvature over a large fraction of the LD surface (obtained from the theory) would have an aspect ratio close to one. Several possibilities exist to explain the mismatch between simulated shapes and theoretical shapes: (a) the surface tension is actually non-zero in the simulations, both in the monolayer and in the bilayer regions; (b) a non-negligible, negative line tension exists in the system; (c) the theory is inadequate to describe our nano-sized system. We do not comment here on the possibility that the theory breaks down for our specific systems, and in the following we explore the other two possibilities.

Given than the system is not semi-isotropic in the presence of the oil droplet, we first considered the possibility that the bilayer tension is not vanishing in our system, and fitted the simulated shapes using both λ and γ_b/γ_m as variables (i.e., two fitting parameters). In this case, the fit was much more accurate (Figure 4.5b). Assuming the same value of bending rigidity as above ($\kappa_m = 6k_BT$), we obtained $\gamma_m = 2.19$ mN/m and then γ_b

close to 2 mN/m for all LD volumes. The value of the bilayer tension obtained from the fit is surprisingly large, knowing that the tension of the ER is very small, nearly vanishing (129). However, this is not surprising if we consider that the monolayer tension is also higher than experimentally measured in micrometer-sized droplet embedded vesicles: in the case of tensionless droplet embedded vesicles made of DOPC and TG, the monolayer tension γ_m is about 1.4 mN/m (8). In experiments on droplet embedded vesicles, bilayer and monolayer surface tension are linearly related (8) ($\gamma_m = 0.45\gamma_b + 1.32$ mN/m), and a monolayer tension of 2.19 mN/m would correspond to a bilayer tension of 1.93 mN/m, very close to the one predicted by fitting the simulated LD shapes. Hence, the present findings are fully compatible with available experimental data on droplet embedded vesicles with non-vanishing surface tension (8) and with the theoretical model (116).

4.4.4 Line tension in nascent LDs

In principle, a high surface tension in the monolayer region of nascent LDs is also compatible with a negative line tension at the junction between the monolayer and bilayer region (henceforth, the LD junction). Line tension indicates a free energy cost for the LD junction: a positive line tension would act as an effective force to reduce the length of the LD junction, hence favor LD budding; a negative line tension, instead, would act as a force extending the length of the LD junction, and hence increase the surface tension in the LD monolayer region. As a consequence, a negative line tension would also explain the simulated LD shapes, "stretched" in the proximity of the LD junction. An estimate of the line tension at the LD junction would therefore be important to improve our understanding of the physics of the system. However, to the best of our knowledge, such line tension has bever been measured experimentally, nor it has been calculated from simulations.

To estimate the line tension in our nascent LD systems, we set up simulations of tubular oil droplets, similar to oil tubules embedded in bilayer membranes (Figure 4.1c). Such systems are easily produced in simulations by exploiting periodic boundary conditions, but may be difficult to produce experimentally. We simulated tubular oil droplets with different sizes (see methods), and for each system we calculated the contact angle, aspect ratio (height/X intercept), and line tension (Figure 4.6).

First of all, we notice that the length of the tubular oil droplets (along the Y axis) has only a minor effect on the value of its physical properties: the values of the contact angle, aspect ratio, and line tension are within the statistical uncertainties. We conclude that finite size effects in our simulation setup are negligible. Second, for any given length of tubular droplet, as the tubules increases in size (i.e., the TG/DOPC ratio increases) the aspect ratio (height/radius) increases and the contact angle decreases, i.e., the shape of the tubular droplet becomes more similar to a cylinder. This is consistent with findings on isolated, lens-shaped droplets, showing higher aspect ratio as their size increases, both in simulations (Figure 4.3a,b) and in the theory (116). Third, we observe that the values of line tension are always positive, and they increase with the size of the oil tubule, from ∼17 pN (for the smallest size) to over 40 pN (for the largest size). Only a negative, not a positive line tension can explain the difference between simulated and theoretical LD shapes obtained with tensionless bilayers (Figure 4.5a). Since our calculations predict a positive line tension around nascent LDs, we conclude that non-vanishing surface tension in the bilayer is the most satisfactory way to explain LD shapes from simulations. Besides elucidating the physical determinants of LD shape in simulations, our predictions also provide microscopic insight into the driving forces for LD budding:

- The increase in line tension with LD size has never been reported before, to the best of our knowledge, and suggests that line tension is due to the deformation of the unzipping bilayer, that is highly curved at the LD junction, and becomes more highly curved as the size of the tubular oil droplet increases. The larger the LD, the higher the (negative) curvature of the phospholipid leaflets at the junction. Line tension would therefore be due to a local deformation, not to long-range interactions (116).

- While the tubular oil droplets used in our calculations were periodic, we infer that most likely also isolated, lens-shaped nascent LDs possess a similar, positive line tension. In isolated LDs, a positive line tension will act as a driving force for LD budding, as predicted by theoretical studies (116): the higher line tension of larger LDs would favor their budding.

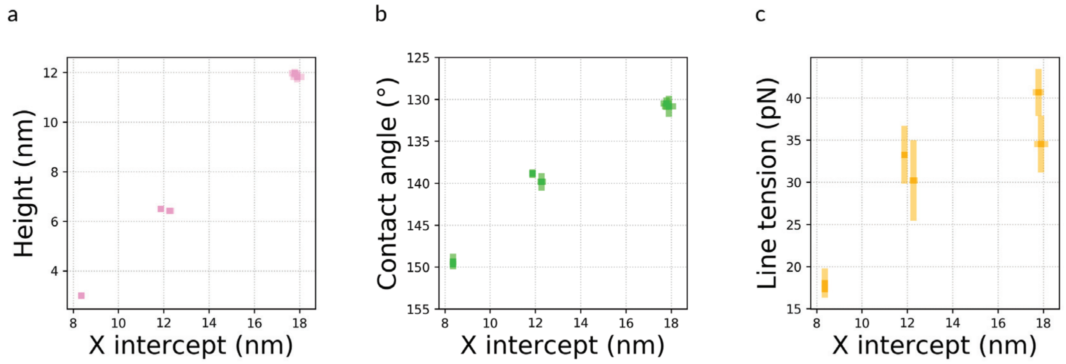


Figure 4.6: Properties of periodic oil tubules embedded in flat DOPC bilayer membranes. Height (a), contact angle (b), and line tension (c) in the oil tubule as a function of the LD size (expressed as the X intercept).

4.4.5 Surface tension in nascent LDs

Since line tension does not provide an interpretation for the shape of nascent LDs, and since the fitting results indicate that significant surface tension should be present in the simulated systems, we explore in more detail the origin of such surface tension. It is possible, in molecular dynamics simulations, to calculate directly the mechanical tension in molecular systems based on the analysis of the local stress tensor. In materials undergoing strain, neighboring particles exert forces on each other to counteract deformations. The stress tensor σ describes these internal resistive forces. In simulations, we can calculate the local stress tensor directly from kinetic energy and the virial, i.e., from velocities, positions, and forces in the system (130). From the local stress tensor, in simulations of flat, periodic lipid bilayer systems, we can then calculate the lateral stress profile σ_L, describing the internal forces in the bilayer as a function of the position along the membrane normal; this is defined as

$$\sigma_L = \sigma_T - \sigma_N$$

where σ_L is the lateral stress, σ_T is the tangential stress (i.e., the stress in the plane of the membrane), and σ_N is the normal stress (i.e., in the direction normal to the membrane). In the case of a flat membrane oriented with the normal in the direction of the Z axis, this can be written as:

$$\sigma_L = \frac{\sigma_{XX} + \sigma_{YY}}{2} - \sigma_{ZZ}$$

where σ_{XX}, σ_{YY}, σ_{ZZ} are the diagonal components of the stress tensor. The stress tensor can also be calculated separately in sub-cells of the simulation cell, expressed as a function of the Z coordinate, and averaged over the simulation time and over the XY plane (Figure 4.7a). In this case, the lateral pressure profile $\pi(z)$ can be calculated as:

$$\pi(z) = -\sigma_L(z)$$

The lateral pressure profile has been reported in the literature for a number of different bilayer systems (131; 132; 122), and it typically presents a characteristic large negative peak corresponding to the interface between the hydrophobic interior of the membrane and the hydrated head group region (close to the glycerol backbone), and smaller positive peaks in the head group region and the membrane interior (131; 132; 122). In the case of a flat membrane, the surface tension matches exactly the mechanical tension obtained from the simulations; if the membrane normal is oriented as the Z axis, the surface tension can be computed easily by integrating the pressure profile:

$$\gamma = \int \pi(z) dz$$

Similarly, the surface tension can also be calculated easily from the local stress tensor in systems with spherical symmetry, such as (spherical) liposomes, or even cylindrical symmetry (131). However, the nascent LD systems simulated here do not have planar, nor spherical, nor cylindrical symmetry, and calculations of local stress or lateral pressure have not been reported so far on similar systems, to the best of our knowledge. Calculation of the local stress tensor is straightforward (Figure 4.7a), but interpretation in terms of surface tension requires separate procedures for averaging over the region corresponding to the bilayer plane (that is parallel to the XY plane, and has planar symmetry) and over the monolayer region corresponding to the spherical cap (that has spherical symmetry). Based on such calculations (Figure 4.7b), we find that the surface tension in the bilayer region is indeed positive and large (Table S2), of the order of a several mN/m, well beyond the statistical uncertainty. This is in agreement with the predictions obtained by our fitting procedure (Figure 4.5), and raises questions on the origin of such large surface tension in the bilayer region, with a value close to the surface tension in the monolayer region. The simplest explanation is the mechanical equilibrium of the simulation box: the pressure coupling algorithm imposes that tangential and normal component of the stress tensor, averaged over the entire simulation box (including the droplet embedded bilayer and the water solvent), be equal to 1 bar, so that the difference, that is mathematically equivalent to the lateral stress over the simulation box, be zero; this is achieved by having the same mechanical tension in the monolayer and bilayer region. The non-zero mechanical tension in the monolayer region is an intrinsic property of the oil-phospholipid-water interface, as shown experimentally; the non-zero mechanical tension in the bilayer region, instead, is simply a requirement imposed by the pressure coupling algorithm.

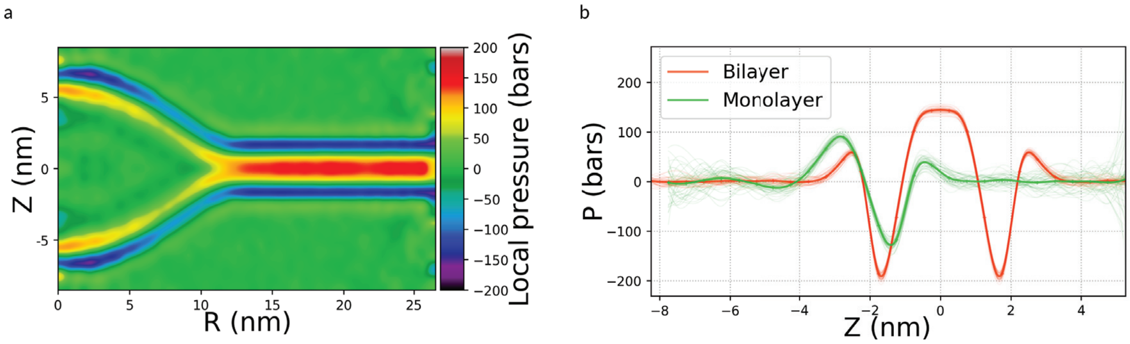


Figure 4.7: (a) Average lateral pressure in a nascent LD system (4050 DOPC+1250 TG). The characteristic negative peak corresponds to the head groups and positive peak to the acyl chains. (b) Pressure profile in the monolayer (green) and the bilayer (red) region as a function of the distance from the center of the bilayer or the monolayer-oil interface.

4.5 Discussion

The shape of nascent lipid droplets can be considered as a proxy for their tendency to bud out of the membrane. According to an established theory (116), the shape of nascent LDs should depend on the LD size and the chemical nature of the oil and the phospholipids. Here we simulated nascent LDs, i.e., isolated TG droplets embedded in lipid bilayers, with different size and different chemical composition, calculated parameters related to their shape, and compared the shape of the simulated LDs with the shapes predicted by the theory. We found good agreement between simulations and theory, with LDs becoming more spherical (a) as they grow in size, (b) as their oil-water surface increases, and (c) as the phospholipids form softer membranes. Also, we noticed that the LD shape resembles a spherical cap, with constant curvature over a large fraction of the LD surface, already for relatively small nascent LD sizes (10 nm in radius). This strongly indicates that the shape is mostly determined by surface tension, more than membrane bending modulus. We then fitted the simulated LD shapes in order to interpret simulation results in terms of elastic properties of the systems. Excellent quality fits were obtained when using monolayer and bilayer surface tensions as free fitting parameters. In this case, the fitting procedure predicted monolayer and bilayer surface tensions of the same order of magnitude, both around 2 mN/m, with the monolayer tension only marginally higher than the bilayer tension. Alternatively, large negative values of the line tension at the LD junction could compensate for the tension in the monolayer, and explain the shape of the simulated LDs. A negative line tension would introduce a force acting to increase the linear interface at the junction between bilayer and monolayers. In the specific case of circular LDs, increasing the length of the contact line would result in spreading the LD and increasing the surface tension in the LD monolayer, thus making the nascent LD more "flat" or less spherical. On the other hand, a positive line tension would tend to reduce the line interface. To discriminate between the two possibilities, we carried out simulations to predict the line tension at the monolayer-bilayer junction. Simulation results are very clear: line tension at the LD junction is positive, and increases with the size of the LD, as the deformation of the unzipping bilayer becomes more important. As a result, the shape of simulated LDs must result from non-vanishing surface tension in the bilayer region. Such result was confirmed by calculations of the local stress tensor, which are independent of the fitting procedure.

Large values of the surface tension (∼2 mN/m) in bilayer membranes may appear surprising, as very small values (∼0.01 mN/m) are generally measured experimentally in the bilayer region for macroscopic (micrometer-sized) droplet embedded vesicles (13; 115). However, the large mechanical tension predicted by the simulations is the result of the specific simulation setup, based on the NPT ensemble with semi-isotropic pressure coupling. Such thermodynamic ensemble is generally applied in simulations of flat, periodic bilayer systems, as it generates configurations for tensionless membranes. In the case of nascent LD systems, the ensemble imposes zero overall mechanical tension in the simulation box. A non-vanishing surface tension is an intrinsic property of the monolayer region, that consists of an oil-phospholipid-water interface (8). Surface tension for such interface can be lowered via compression of the interfacial monolayer, which achieves high surface density of phospholipids, as recently shown experimentally; however, even strong compression cannot achieve a vanishing surface tension for such interface (8). In our NPT simulations, with semi-isotropic pressure coupling, mechanical equilibrium is achieved when the tangential component and the normal component of the stress tensor over the entire simulation box are equal; this is achieved when the mechanical tension in the monolayer and bilayer region are equal. Based on recent measurements in droplet embedded vesicles (8), tension in the bilayer region should be lower than the one predicted by the simulations, even while tension in the monolayer region remains significant ($\sim$1 mN/m). It is unclear if this is achievable in MD simulations, considering that simulations of laterally compressed nascent LD systems would still need to fulfill the condition of mechanical equilibrium in a periodic system.

One limitation of our approach is that we could only determine the line tension in periodic oil tubules, not in isolated nascent LDs. Tubular droplets are not identical to isolated, nascent LDs, in that the local curvature at the monolayer-bilayer junction is different. However, for a tubular vs isolated droplets with the same contact angle, curvature at the monolayer-bilayer junction is in fact higher in isolated LDs; therefore, we expect the estimates obtained in tubular droplets to be a lower bound for line tension in isolated droplets.

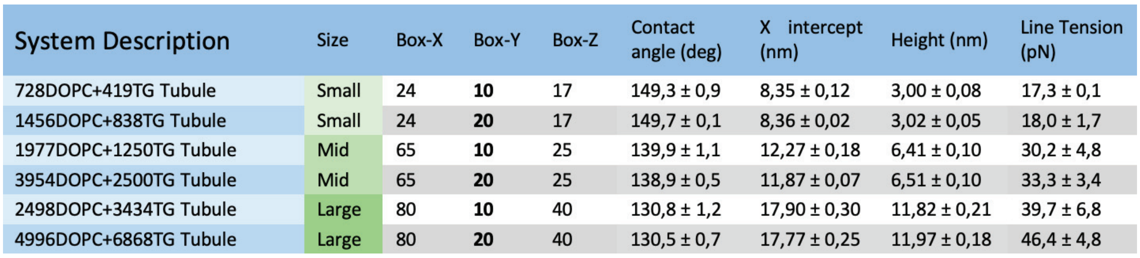
4.6 Conclusions

We studied the shape of lipid droplets as obtained from MD simulations at the coarse-grained level, and compared it to the predictions by an established theory (116). We identified the parameters affecting LD shape: LD volume, as larger LDs are more spherical in shape; membrane softness, as LDs embedded in less rigid bilayers are more spherical; surface tension between oil and water, as a higher surface tension makes LDs more spherical. These trends are consistent with theoretical predictions. Fitting the simulated LD shapes with a theoretical shape equation suggests that either a negative line tension exists at the monolayer-bilayer junction or a non-zero surface tension is present in the bilayer region. We find that the line tension at the monolayer-bilayer junction is positive, not negative, and increases with the size of the LD, suggesting that it originates from the deformation (curvature) of the bilayer as it unzips into two monolayers. A positive line tension can act as a driving force for LD budding, as predicted by theoretical studies. Local stress calculations confirm that surface tension is present in the simulated bilayers. This can be explained by the mechanical equilibrium condition: zero mechanical tension over the simulation box implies that equal and opposing mechanical tensions exist in the monolayer and bilayer region. Our results indicate that simulations of nascent LD systems may be used to provide a microscopic view into the properties of droplet embedded vesicles.

4.7 Acknowledgement

LM acknowledges funding from the Institut national de la santé et de la recherche médicale (INSERM). Molecular dynamics calculations were carried out at CINES, GENCI grant no. A0080710138, A0100710138, and A0120710138 to LM. This work was supported the ANR-NANODROP (ANR-17-CE11-0003) and ANR-LIPRODYN (ANR-21-CE11-0032) to ART and LM.

4.8 Supporting information



System Description	Size	Box-X	Box-Y	Box-Z	Contact angle (deg)	X intercept (nm)	Height (nm)	Line Tension (pN)
728DOPC+419TG Tubule	Small	24	**10**	17	149,3 ± 0,9	8,35 ± 0,12	3,00 ± 0,08	17,3 ± 0,1
1456DOPC+838TG Tubule	Small	24	**20**	17	149,7 ± 0,1	8,36 ± 0,02	3,02 ± 0,05	18,0 ± 1,7
1977DOPC+1250TG Tubule	Mid	65	**10**	25	139,9 ± 1,1	12,27 ± 0,18	6,41 ± 0,10	30,2 ± 4,8
3954DOPC+2500TG Tubule	Mid	65	**20**	25	138,9 ± 0,5	11,87 ± 0,07	6,51 ± 0,10	33,3 ± 3,4
2498DOPC+3434TG Tubule	Large	80	**10**	40	130,8 ± 1,2	17,90 ± 0,30	11,82 ± 0,21	39,7 ± 6,8
4996DOPC+6868TG Tubule	Large	80	**20**	40	130,5 ± 0,7	17,77 ± 0,25	11,97 ± 0,18	46,4 ± 4,8

Table 4.2: Box size, contact angles, LD size (X intercept and height), and line tension estimated from simulations of periodic oil droplets.

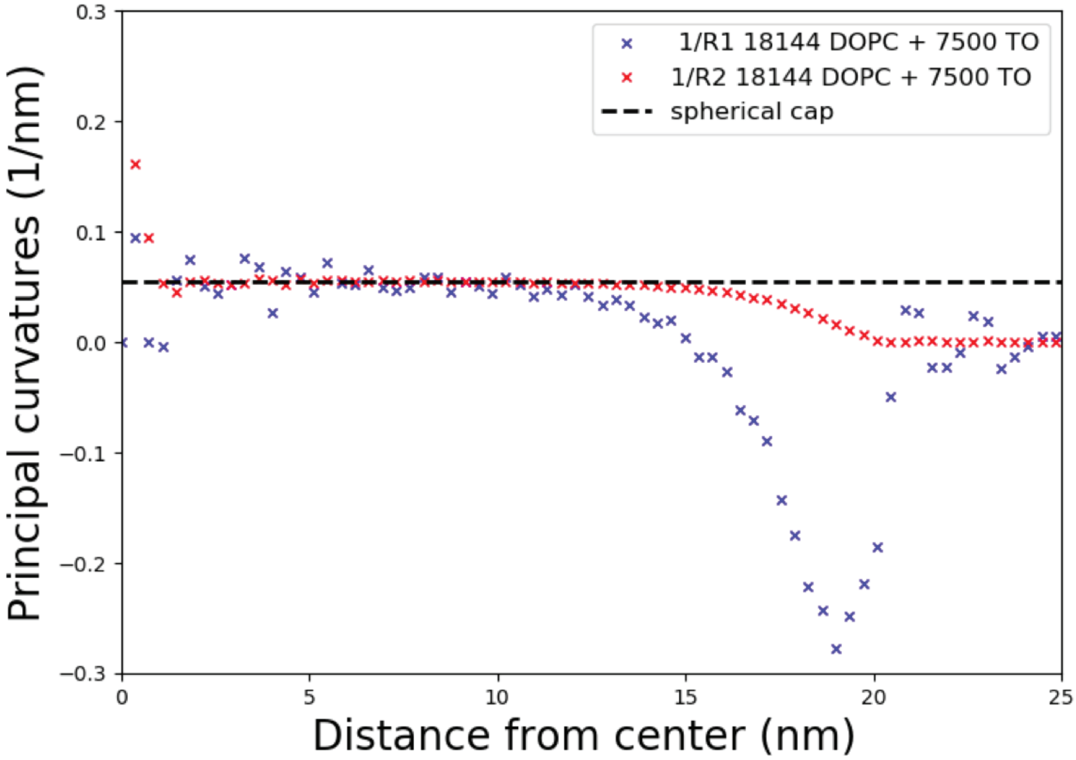


Figure 4.8: Principal curvatures of simulated nascent LDs containing 7500 TG lipids.

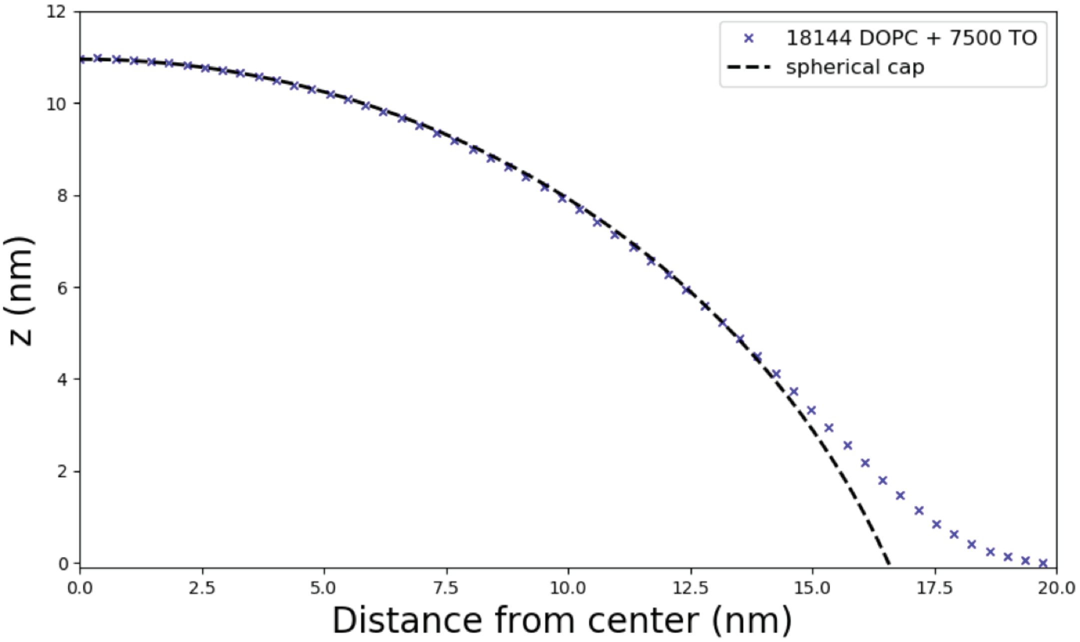
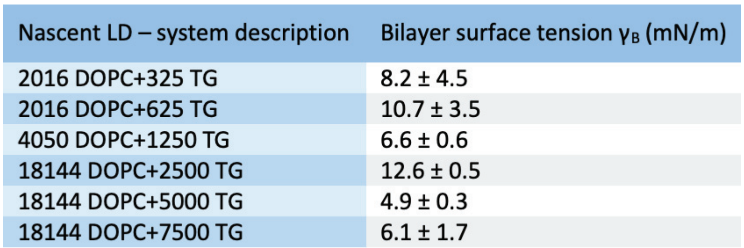


Figure 4.9: Fitting the simulated LD shape with a spherical cap, for simulated nascent LDs containing 7500 TG lipids.



Nascent LD – system description	Bilayer surface tension γ_B (mN/m)
2016 DOPC+325 TG	8.2 ± 4.5
2016 DOPC+625 TG	10.7 ± 3.5
4050 DOPC+1250 TG	6.6 ± 0.6
18144 DOPC+2500 TG	12.6 ± 0.5
18144 DOPC+5000 TG	4.9 ± 0.3
18144 DOPC+7500 TG	6.1 ± 1.7

Table 4.3: Bilayer surface tension computed used the local stress in the simulated LD systems.

Chapter 5

Birth of an organelle: molecular mechanism of lipid droplet biogenesis

Vincent Nieto[1], Jackson Crowley[1], Denys Santos[2], Luca Monticelli[1,*]

[1] University of Lyon, CNRS, Molecular Microbiology and Structural Biochemistry (MMSB, UMR 5086), F-69007, Lyon, France

[2] Federal University of Pernambuco, Department of Fundamental Chemistry, Recife, Brazil.

5.1 Abstract

Lipid droplets (LDs) are cellular organelles regulating energy and lipid metabolism. They are generated mostly in the tubular regions of the endoplasmic reticulum (ER) by the accumulation of neutral lipids, such as triglycerides, into lens-shaped blisters. When the lens-shaped nascent lipid droplet grows beyond a certain threshold, it emerges from the ER bilayer membrane, generally towards the cytosol – a process known as budding. Images of mature, micrometer-sized LDs are available, as well as a few images of lens-shaped nascent LDs trapped in the ER bilayer, but the mechanism of biogenesis has never been observed experimentally. Here we developed a novel, robust computational methodology to simulate the initial steps in lipid droplet biogenesis, from the nucleation of the nascent LD to the budding towards the cytosol. Our simulations show that LDs do not bud in the absence of sufficient asymmetry between the two leaflets of a membrane, independently of membrane

morphology; leaflet asymmetry is necessary and sufficient to promote the budding transition. Seipin, a membrane-anchored protein essential for the correct functioning of LDs, promotes an asymmetric shape of nascent LDs but it is not sufficient per se to promote the budding transition. However, we predict that seipin increases the mechanical stability to the LD-tubule connection. The simulations also allow to understand the role of the complex ER composition in allowing for a stable, defect-less, high curvature LD-tubule connection, and suggest that regulation of the oil/phospholipid synthesis ratio is crucial to preserve the mechanical stability of the ER network. Our new methodology paves the way to simulations of complex transformations of membrane systems, including the biogenesis of other organelles, the formation of viral envelopes, and bacterial division.

5.2 Introduction

Adenosine triphosphate (ATP) is the molecule providing energy to drive most chemical reactions in living organisms(133). Humans produce daily approximately their own body weight in ATP, and consume about the same amount, as ATP cannot easily be stored(134). Lipids, instead, are a convenient way to store energy, and they are stored in lipid droplets (LDs). LDs are crucial hubs of lipid metabolism, central in membrane biogenesis and several non-metabolic processes (e.g., gene expression, or protein quality control, viral infections, etc. (10)). Deficiencies in LD formation or functioning result in impaired cell metabolism, cell stresses, and numerous diseases, e.g., lipodystrophies, liver steatosis, type II diabetes, neurodegeneration (135; 12). LDs are generated in the endoplasmic reticulum (ER), the prime site for lipid synthesis, and indeed LD biogenesis starts with the synthesis of neutral lipids (oils, such as triglycerides and sterol esters). When the concentration of neutral lipids in the ER membrane reaches a certain threshold, they phase-separate from the phospholipids constituting the ER bilayer. Phase separation is thought to start with the nucleation of a lens-shaped nascent LD, initially sandwiched between the two leaflets of the ER bilayer; the nascent LD grows as more neutral lipids are synthesized, and eventually buds out of the ER towards the cytosol, coated solely by the cytosolic leaflet of the ER bilayer. Mature LDs resemble oil-in-water emulsions, i.e., oil droplets coated by a mono-molecular layer of phospholipids and proteins. Interestingly, they generally remain attached to the ER via LD-ER contact sites marked by seipin, an oligomeric protein with an important role in preventing pathological phenotypes (84; 136). However, the mechanism of LD biogenesis has never been observed, due to the fluid nature of membrane and LD systems and the insufficient spatial and time resolution of current microscopy techniques: observable LDs have a size of hundreds of nanometers or more, but nascent LDs are much smaller, and isolating the different steps of the process is proving a formidable challenge for structural biology. In the absence of a high-resolution view of the mechanism of biogenesis, several questions remain open on the driving forces and the molecular factors determining LD budding. For instance, continuum theory predicts that LDs should spontaneously bud off from a symmetric bilayer when their size becomes large enough (116), but validation of the theory is problematic as it would require fine control over the synthesis of neutral and polar lipids. Contrasting results have been reported on the role and the nature of ER membrane asymmetry: Choudhary et al. proposed that directional budding is determined by intrinsic curvature of ER phospholipids (14), while Chorley et al. proposed that it depends on the asymmetry between ER leaflets (82). The function of seipin is also not completely clear: simulations and experiments suggested that it may trap triglycerides (37; 38), therefore affecting LD nucleation and growth by ripening, but its exquisite localization at the LD-ER contact site raises questions on a possible role also in the budding process. Here we address the questions above by simulating the entire mechanism of LD biogenesis at the molecular level. To this end, we develop a protocol that allows molecular simulations with growing number of particles, and we use it to emulate the synthesis of different lipids in membranes mimicking ER tubules. The outcome is, for the first time, a molecular view of the birth of a cellular organelle, with nanosecond time resolution and sub-nanometer spatial resolution. Simulations show that triglyceride synthesis results in LD nucleation, growth, and budding of stable and defect-free LDs only under specific conditions: active generation of asymmetry between the ER leaflets, presence of seipin, a specific composition of the ER membrane, and a precise regulation of both triglyceride and phospholipid synthesis. The picture resulting from our molecular movies provides an unprecedented view of the actual process leading to the birth of lipid droplets as organelles, an interpretation for published experimental data, and specific predictions amenable to experimental validation.

5.3 Methods

All simulations used the Martini coarse-grained force field, either version 2 (67; 68) or version 3 (75), as specified in Table 5.1, and were carried out with the Gromacs (121) software (version 2018 or 2019). In the following, we first describe the simulation setup for the different systems studied here, and then we provide the simulation parameters.

5.3.1 Setup of nascent LDs in flat bilayer systems

We generated flat periodic lipid bilayer systems using the Insane software [ref]. The dimensions of the bilayer were 27×27 nm (small bilayer) or 75×75 nm (large bilayer). To generate nascent LDs, we initially built tri-layer systems, with a layer of oil sandwiched between two layers of phospholipids. Pure DOPC was used for the bilayer, while pure TG was used for the oil phase. The TG layer contained 325, 650 (small bilayer), 1250, 2500, 5000, or 7500 molecules (large bilayer). The tri-layer systems spontaneously evolved to lens-shaped nascent LDs during unbiased MD simulations, on time scales of a few microseconds.

5.3.2 Generation of membrane pores

Pores were generated in lipid membranes using flat-bottomed potentials, restraining lipid particles out of a part of the simulation volume. The potentials acted only on lipid acyl chain particles and only if they entered a user-defined volume, in this case a cylindrical region in the membrane away from the location of the nascent lipid droplet. Acting only on acyl chains, the potentials induce the formation of hydrophilic pores, with lipid head groups in contact with water. Pores had a diameter of 5 nm. To prevent LDs from approaching the pores, cylindrical flat-bottomed potentials were applied to keep TG molecules 7.5 nm away from the pore center (2.5 nm away from the pore edge). In all cases, we used a force constant of 5000 KJ mol^{-1} nm^{-2}.

5.3.3 Setup of vesicular and tubular systems

A vesicle with an embedded LD (18144 DOPC and 7500 TG) was generated from one of the flat systems described above by removing membrane periodicity; in this case, line tension rapidly transforms a square membrane in a circular one, that bends out of plane to reduce the perimeter of the open edge and eventually closes to form a vesicle. Tubular membranes were then generated from the vesicles by generating large water pores on opposite sides of the vesicle (along the x axis) and then creating a connection with the periodic image. The external diameter of the tubules was 35 nm, the length was 75 nm, and tubules were periodic in the direction of their main axis. All tubules contained 4 circular pores (5 nm in diameter, generated as detailed above), allowing rapid flip-flops between inner and outer leaflet, preventing the build-up of asymmetric stresses and differences in pressure in/out of the tube. In the case of vesicular systems, we only used a simple composition consisting of pure DOPC as phospholipid and pure TG as the oil component. In the case of tubular systems, we also used richer compositions (see Table 5.1). Besides the DOPC/TG composition, we built systems with DOPC/POPC 1:1 mixtures, and systems with a complex composition mimicking the ER composition (137; 138). The complex systems contained 25% DOPC, 25% POPC, 23% DOPE, 3% DOPS, 2% DOPA, 10% POPI, 10% cholesterol, and 2% of DG. The resulting tubular membrane system was composed of 22482 phosopholipids: 5620 DOPC, 5620 POPC, 5170 DOPE, 676 DOPS, 450 DOPA, 2248 POPI cholesterol, and 450 DG.

5.3.4 Building models of seipin

We built a model of seipin dodecamer based on the Drosophila monomer structure (pdb code 6MLU (34)), adding the transmembrane (TM) and the N-terminal helices based on secondary structure prediction algorithms. The putative N-terminal helix has been predicted but not resolved in any of the cryo-EM structures. The two transmembrane domains, instead, are resolved in the cryo-EM structure of yeast seipin (35). Comparison with the structure of yeast seipin shows high similarity of both the luminal domain and the TM segments (see Supporting Information). In the model, secondary and tertiary structure of the protein were maintained using elastic networks (76) in the luminal domain, while no elastic network was required to fix the distance among 24 TM helices, because the loops connecting them to the luminal domain are short (11 amino acid residues).

5.3.5 Iterative Slow Growth (POP) simulations

Current MD software does not allow, to the best of our knowledge, changes in system composition "on-the-fly". However, a simple way to achieve this is via a "stop-and-go" procedure: a simulation is performed in the desired thermodynamic ensemble (NPT in our case), then stopped after a defined time; particles are added to the system, and the simulation is resumed. We devised an Iterative Slow Growth (POP) procedure to simulate LD systems with increasing amounts of lipids over time, by iteratively adding lipids to the system. Our POP protocol works as follows. First, we choose the approximate positions of insertion of the molecules, taking into account the exact instantaneous geometry of the membrane. Second, lipids are inserted using gmx insert-molecules (part of the Gromacs package), that allows the insertion of molecules in a user-defined position even when particles partially overlap with the particles in input structure (by setting the radius of the inserted particles to zero). We generally imposed insertion of each new molecule in positions close to an existing lipid, allowing a maximum displacement of 0.4 nm in x, y, z, and random rotation in the case of TG (but not when inserting phospholipids). Third, the newly added molecules are fixed using spherical flat-bottomed position restraints, and their interaction with the rest of the system is switched on progressively using the slow growth algorithm in Gromacs. Here, the new molecules are initially decoupled from the system and get progressively coupled during 100 ps. Fourth, the system is energy minimized to reduce the probability of particles overlaps. Fifth, the system is relaxed using standard MD, allowing adaptation to the new conditions (e.g., volume expansion). After relaxation, the cycle is repeated. The number of lipids added at each cycle and the duration of the relaxation phase are user-defined. In principle, addition of lipids could be achieved at the same rate as determined by lipid synthesis in cells. However, considering a turnover rate of 104 s^{-1} for a typical enzyme, this is much too slow for present-day supercomputers. On the other hand, very short run times between subsequent POP steps pose problems in terms of computational stability, as systems need to adjust their volume after each addition of lipids. Moreover, lipids and water need to redistribute by diffusion to dissipate internal stresses in the membranes and possible differences in pressure in/out of the tubule. Therefore, a compromise needs to be found between realistic growth rates, computational efficiency, numerical stability, and relaxation of stresses. We found that, for the tubular membranes described above, addition of 50 to 150 lipids (considering the sum of both TG and PL) per POP cycle with a run time of 50-500 ns per cycle results in stable simulations. We carried out POP simulations both starting from the flat bilayer setup and starting from tubular membrane setup. In the case of the flat bilayer, we only increased the number of oil (TG) molecules in the system, and added them in random positions near other TG molecules. In the POP of tubular systems, we added either only TG, or only phospholipids, or both (see Table 5.1). We tested different ways of adding lipids. When adding only TG, the oil molecules can be added either in random positions within the tubular bilayer, or in the proximity of the nascent LD. The first mode requires very long equilibration times, therefore was rapidly discarded in favor of the second. Addition of TG directly in the LD is realistic, because TG synthesis takes place in the proximity of the LD; indeed, TG-synthase is recruited at LD formation sites by seipin in yeasts (43). When adding PL, on the contrary, faster equilibration was achieved by adding PL in random positions in the outer leaflet of the tubular membrane. We used random placement of PL to improve the stability of our POP procedure, even if PL synthesis is probably also occurring close to the LD formation sites (DG-synthase is activated at LD formation sites; in yeasts, DG production by Pah1 is activated by Nem1, and Nem1 is activated by seipin (43). The same POP procedure was also adapted to the presence of seipin. In this case, TG was added only within the nascent LD and PL in random positions in the outer leaflet. A list of the simulations carried out and the specific modes of addition of lipids is found in Table 5.1.

5.3.6 LD pulling simulations

Two additional simulations were performed with LDs embedded in a vesicle, pulling the LDs away from the vesicle. Center of mass (COM) pulling (as implemented in Gromacs) was used, allowing to apply a force to the center of mass a group of particles. In the first simulation, a harmonic potential with fixed position was applied to the COM of DOPC (force constant of 2200 KJ mol^{-1} nm^{-2}) to keep the vesicle at an approximately fixed position in the simulation box. A second harmonic potential was applied to the COM of TG, while increasing the distance between the COM of TG and the COM of DOPC by 0.5 nm/μs (force constant of 1000 KJ mol^{-1} nm^{-2}), to pull the LD away from the vesicle. The simulation was run for 20 μs. The second simulation started from the final frame of the first pulling simulation.

system #	**System**	Seipin	FF	Pores	ISG type	Initial #TO	Inital #PL	**TO/ cycle**	location of TO insertion	**PL/ cycle**	location of PL insertion	# cycles	relaxation phase (ns) [in each cycle]	Total simulation time (µs)
1	Flat bilayer (DOPC)	No	M2	--	--	325	2016	--	--	--	--	--	--	20
2	Flat bilayer (DOPC)	No	M2	--	--	625	2016	--	--	--	--	--	--	20
3	Flat bilayer (DOPC)	No	M2	--	--	1250	4050	--	--	--	--	--	--	20
4	Flat bilayer (DOPC)	No	M2	--	--	2500	18144	--	--	--	--	--	--	40
5	Flat bilayer (DOPC)	No	M2	--	--	5000	18144	--	--	--	--	--	--	40
6	Flat bilayer (DOPC)	No	M2	--	--	7500	18144	--	--	--	--	--	--	40
7	Flat bilayer (DOPC)	No	M2	1	--	7500	18144	--	--	--	--	--	--	25
8	Flat bilayer (DOPC)	No	M2	-	DOPC	7500	18144	--	-	100	Edge box	50	100	5
9	Flat bilayer (DOPC/POPC)	Yes	M2	1	TO	350	18144	50	Random	-	-	51	250	12,75
10	DET (DOPC)	No	M2	4	TO	7625	22482	50	LD	-	-	446	50	22,3
11	DET (DOPC)	No	M2	4	--	29925	22482	--	-	-	-	-	-	20
12	DET (DOPC)	No	M2	-	TO+DOPC (1:2)	7625	22482	50	LD	100	Bottom	64	300	19,2
13	DET (DOPC)	No	M2	-	TO+DOPC (1:1)	7625	22482	50	LD	50	Bottom	76	300	22,8
14	DET (DOPC/POPC)	No	M2	-	TO+DOPC/POPC (1:1)	7625	22482	50	LD	50	Bottom	83	300	24,9
15	DET (DOPC/POPC)	No	M2	4	TO+DOPC/POPC (1:1)	7625	22482	50	LD	50	Bottom	0	200	
16	equilibration	No	M2	4	--	12000	22482	--	-	--	-	-	-	
17	equilbration	No	M3	4	--	12000	22482	--	-	--	-	-	-	
15	DET-PULL (DOPC/POPC)	No	M3	4	--	12000		--	-	--	-	-	-	
16	DET (DOPC/POPC)	Yes	M3	4	TO	350	22482	50	Random	-	-	86	250	21,5
17	DET (DOPC/POPC)	Yes	M3	4	--	4650	22482	--	-	-	-	-	-	10
18	DET (DOPC/POPC)	Yes	M3	4	TO	350	22482	50	LD	-	-	50	250	12,5
19	DET (DOPC/POPC)	Yes	M3	4	--	2850	22482	--	-	-	-	-	-	10
20	DET (DOPC/POPC)	Yes	M3	4	TO+DOPC/POPC	350	22482	50	LD	50	Bottom	215	200	43
21	DET-PULL (DOPC/POPC)	Yes	M3	4	--	12000		--	-	--	-	-	-	
22	DET (ER composition)	No	M3	--	TO	300	23182	50	LD	-	-	272	50	13.6
23	DET (ER composition)	No	M3	4	TO	300	23182	50	LD	-	-	240	50	12
24	DET (ER composition)	No	M3	4	--	12300	23182	--	-	-	-	-	-	10
25	DET (ER composition)	Yes	M3	4	8 PL + TO	300	23182	50	LD	50	Random	38	500	19
26	DET (ER composition)	Yes	M3	4	8 PL + TO	300	23182	50	LD	50	Bottom	53	250	13,25
27	DET (ER composition)	Yes	M3	4	8 PL + TO	300	23182	50	LD	50	LD	256	250	64
28	DET (ER composition)	Yes	M3	4	8 PL + TO	300	23182	100	LD	50	LD	206	250	26,5
29	DET-PULL (ER composition)	Yes	M3	4	--	12000		--	--	--	-	-	-	
30	DET (ER composition)	Yes	M3	4	8 PL + TO	300	23182	25	LD	50	LD	163	200	32,6
31	DET (ER composition)	Yes	M3	4	8 PL + TO	300	23182	50	LD	25	LD	52	200	10,4

Table 5.1: Summary of the main simulations described in this work.

In this case, the distance between the COM of DOPC and the COM of TG was kept fixed and both potentials used a force constant of 2200 KJ mol^{-1} nm^{-2}. In both simulations, hydrophilic pores were kept open to avoid asymmetric stresses and allow equilibrium between the two leaflets of the vesicle.

5.3.7 Simulations parameters

Standard Martini simulation parameters were used in all simulations; the Verlet neighborlist scheme (139) was applied, with a cutoff of 1.1 nm for non-bonded interactions and the Verlet cutoff scheme for potential shift (140). Reaction field (141) was used in all simulations, with $\epsilon_\rho = 15$. Relaxation runs were carried out with the leap-frog integrator (142) and a time step of 20 fs. Temperature was kept constant with the Bussi-Donadio-Parrinello (v-rescale) thermostat (64) (time constant of 1 ps). Pressure was kept constant in equilibration with the Berendsen barostat (63), with semi-isotropic pressure coupling (time constant of 20 ps, compressibility of 0 in the XY plane and $4 \cdot 10^{-5} \text{bar}^{-1}$) in the Z direction).

Slow growth simulations were carried out by coupling the interactions (both Van der Waals and Coulomb) of the newly inserted molecules with all other molecules in the system. Coupling was achieved by growing linearly the coupling parameter, λ from an initial state

$\lambda = 0$ incremented by 0.0002 per time step (20 fs), so full coupling was achieved in 100 ps. The soft-core alpha and sigma parameters were set to 1.3 and 0.47 nm, respectively, while the soft-core power was set to 1.

5.3.8 Simulation analysis

Analysis of the number of oil molecules in the nascent LDs was carried out with the g_aggregate software (123), while the number of lipids in each leaflet was calculated with the SuAVE software (143; 144). In SuAVE, the fitting process established in radial basis functions was used with the same parameters developed for closed surfaces (143). A roughness parameter and a point resolution were used to define the grid points that make up the fitting surfaces. Then, a classification process was performed based on the distance between the lipids and the bidimensional fitted surfaces. A 2.5 nm cutoff was applied to distinguish the lipids from the different leaflets.

5.4 Results

5.4.1 Lipid droplets do not spontaneously bud off from symmetric bilayers

According to a theory of lipid droplet budding, lipid droplets can spontaneously bud off from a symmetric bilayer simply upon increasing their volume beyond a certain threshold (116). Based on estimates of the physical properties of the membrane, spontaneous budding is predicted to occur at sizes of a few tens of nm (116). In the order to test this hypothesis, we performed coarse-grained molecular dynamics simulations of dioleoylphosphatidylcholine (DOPC) lipid bilayers with increasing LD volume. Starting from bilayer systems with high concentration of triglycerides (TG), well above the critical concentration (Table 5.1, systems 1 to 7), initially well-dispersed oil molecules rapidly formed one or more lens-shaped nascent LDs; nucleation of LDs occurred on time scales of hundreds of ns. Then the nascent LDs grew by attracting free TG molecules or by fusing with other LDs, until about 1.5% of TG was left free in the bilayer. The shape of the LD evolved during the LD growth phase and eventually reached an equilibrium. Adding up to 7500 TG molecules in a bilayer (lateral size of 78×78 nm) did not result in LD budding (Figure 5.1). This is expected, because budding requires a larger surface area of one leaflet over the other; this can be reached, in principle, via area fluctuations and/or flip-flop. Flip-flops occur on time scales of hours for phosphatidylcholine (PC) lipid bilayers (145), too slow to support budding even in biological systems. Significant fluctuations in the area of each bilayer leaflet require large lipid reservoirs, not available in our simulation setup (but available in the ER). We then devised a simulation setup allowing large fluctuations in the content of each membrane leaflet. Starting from the largest system equilibrated above, we generated a large, stable pore in the bilayer membrane, away from the nascent LD (see Methods). Lipid flip-flop along the pore edge presents no energy barrier, hence phospholipids can freely flow from one leaflet to the other and, in principle, allow the build-up of significant asymmetry between the leaflets, necessary for LD budding (82). Flip-flops were observed indeed, but the maximum imbalance observed in our simulations was less than 2% (Figure 5.1d), while an imbalance of at least 5% is necessary to generate an asymmetric LD shape in systems of the same size (82). Larger fluctuations would appear on longer time scales, but the probability distribution of fluctuations suggests that much larger membrane reservoirs are required to allow sufficient asymmetry and spontaneous budding. Indeed, theoretical works predict spontaneous budding with relatively small LDs and infinite membrane reservoirs(14; 116; 117). Systems with two nascent LDs could, in principle, allow much larger leaflet imbalance by budding each LD in opposite directions: budding of one nascent LD would trigger budding of the other one in the opposite direction, to conserve the surface area of each leaflet. However, such double LD setup provided similar results as the single LD, and no budding was observed on the simulation time scales. Limitations in system size are a possible culprit for the failure to reproduce LD budding, but they are difficult to overcome in simulations, particularly with respect to the size of the membrane reservoir. At this stage, lack of budding may also be the result of other limitations: the flat morphology of the bilayer membrane, the absence of seipin at the LD-bilayer junction, or the insufficient asymmetry between bilayer leaflets. In the following, we explore these possibilities, to understand which driving forces are essential for LD to bud.

5.4.2 Lipid droplets in tubular membranes spontaneously emerge towards the lumen

In cells, LDs generally bud off from tubular regions of the ER towards the cytosol(36), and ER curvature has been proposed to catalyze LD assembly (146). So, we wondered whether

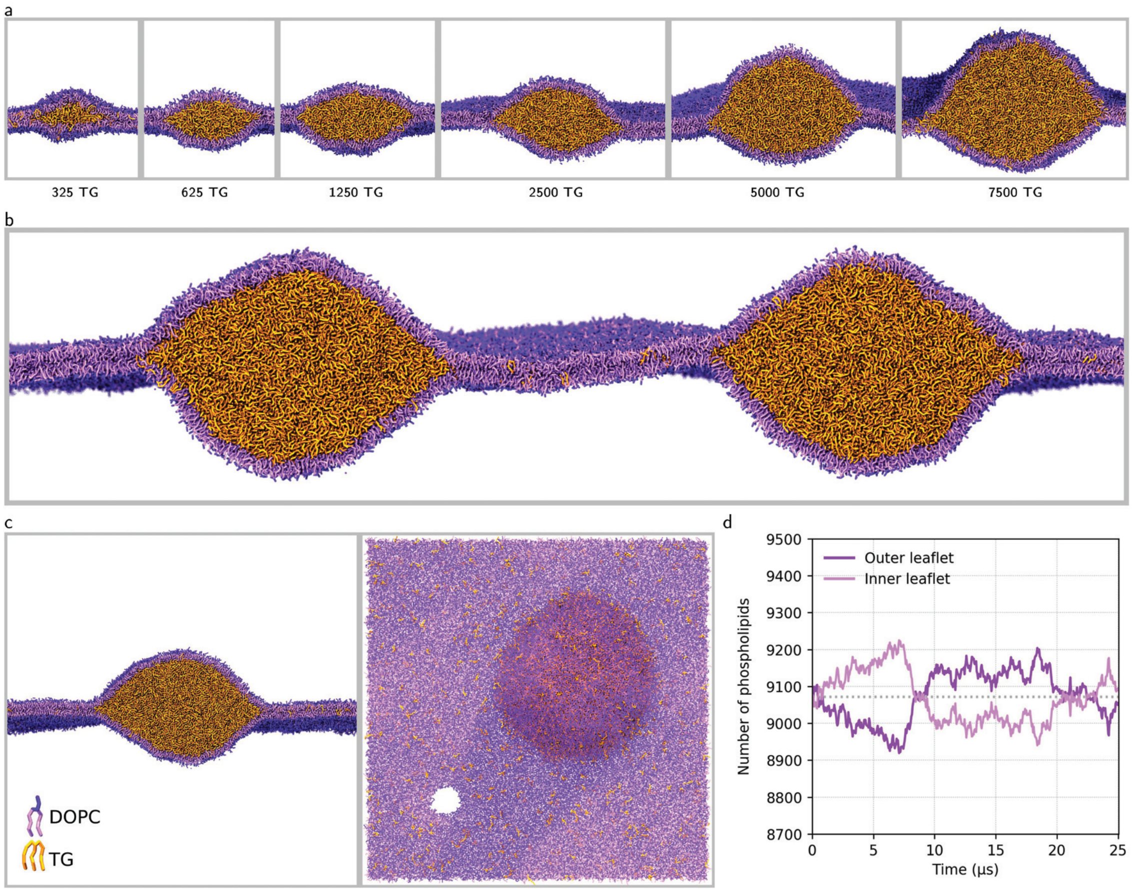


Figure 5.1: Coarse-grained molecular models of nascent lipid droplets in flat DOPC bilayer membranes. (a) Snapshots of nascent LDs containing different number of TG lipids, side view cut across the LD. (b) Snapshot from simulation of a system with two nascent LDs, side view. (c) Side view and top view of nascent LD with 7500 TG molecules and with a hydrophilic pore in the bilayer region. (d) Number of phospholipids in each leaflet during a simulation of the system shown in (c), with the hydrophilic pore.

membrane curvature could be sufficient to induce budding in nascent LDs. To address this question, we built tubular membranes with an external diameter of 35 nm, comparable to the diameter of ER tubules (147; 148), and a fixed length of 75 nm (see Methods), with a nascent LD and TG molecules dispersed in the bilayer, and 4 hydrophilic pores, to allow for build-up of an imbalance between the leaflets by (barrierless) lipid flip-flop. At the same time, the pores allow dissipation of asymmetric stresses and differences in pressure in/out of the tubule (water can freely flow in/out), as well as changes in tubule diameter and surface area (Figure 5.7). This setup allows to achieve chemical equilibrium in the system, provided sufficiently long simulation times, and effectively mimics the biological system. Simulations of membrane tubules with embedded LDs with diameter up to 35 nm (over 7000 TG molecules) did not show any tendency towards budding. The simplest possible explanation is the limited size of the LD itself. We therefore devised a procedure to allow simulations of LDs with growing size by having TG molecules pop into the system at user-defined locations. The new procedure, coined POP-MD, consists of a growth step, in which lipids (or any other molecules) are added to the system in user-defined positions, and a relaxation step, that is a standard molecular dynamics simulation; the first step takes the system out of equilibrium, while during the second step the system drifts towards a new equilibrium (without reaching it, due to limited sampling). The biological systems we aim to understand are also pushed out of equilibrium by lipid synthesis, without which no LD can be formed. Starting from a membrane tubule with a large embedded LD, we used POP-MD to add overall 22,300 TG molecules within the LD during 22.30 μs (system #17 in Table 5.1). We observed growth of the LD volume but no budding transition. In fact, opposite to most observations in biological systems, the LD protruded towards the lumen of the tubule, instead of the external (cytosolic) compartment (Figure 5.2). Similar results were obtained The presence of pores did allow for the spontaneous buildup of leaflet imbalance: the number of phospholipids in the outer leaflet grew (Figure 5.2d), but the growth was not sufficient to achieve a symmetrical shape of the LD. The symmetrical shape of the LD was not rescued even by long relaxation runs: the leaflet imbalance plateaued within 5 μs and the LD shape remained strongly asymmetric, protruding towards the lumen (Figure 5.2c).

Another possible reason for the failure to observe LD budding might lie in differences in membrane composition. ER membranes have a composition much more complex than the one used in simulations to this point, and that might have an effect on the propensity of LDs to bud out of the ER – for example, due to the presence of lipids with positive intrinsic curvature (14). To address this point, we repeated the POP-MD simulations using a membrane tubule with more realistic composition mimicking the ER, consisting of 8 different lipids (137; 138) (see Methods). We used the same composition for both inner and outer leaflet, and such composition did not change significantly during LD growth nor the subsequent equilibrium run. Insertion of TG promoted leaflet asymmetry, with the outer leaflet continuously acquiring lipids by flip-flop (Figure 5.2f), but no budding was observed (Figure 5.2e). As in the case of pure PC membranes, the LD shape remained largely asymmetric, protruding towards the lumen of the tubule instead of the cytosolic compartment.

These results are opposite to observations in biological systems and might therefore appear surprising, but they can be understood simply based on the geometrical constraints imposed by the tubule morphology. When an LD forms between the leaflets of a tubular membrane, the bilayer unzips; protrusion of the LD outwards requires an increase in the surface area of the outer leaflet, while protrusion inwards does not require any change in surface area of any leaflet. Simulations show that emergence of cellular LDs towards the cytosol is not due to the ER tubule morphology, and suggest that asymmetry-generating structures or processes determine the direction of budding.

5.4.3 Seipin scaffolds nascent LDs but is not sufficient to induce directional budding

A possible source of asymmetry is the presence of proteins. Seipin is an oligomeric ER membrane protein with a major role in LD biogenesis. The structure of seipin has been solved at high resolution by cryo-EM for three species: human (33), drosophila (34), and yeast(36; 35) and consists of a highly conserved luminal domain, two transmembrane domains forming a ring of ∼15 nm in diameter, and one N-terminal cytosolic domain. Seipin is critical for normal LD formation, as it defines LD nucleation sites and maintains LD-ER contact after budding (43). It has been suggested that seipin favors LD nucleation by catalyzing TG aggregation and trapping TGs within the perimeter defined by its TM helices (55; 32). It has also been proposed that seipin plays a major role in preventing abnormal phenotypes, with few super-sized LDs, by decreasing the probability of LD growth by ripening (37; 32). The relevance of this protein in many aspects of LD biogenesis raises questions on its role in the budding process. To understand if seipin has any role in the budding process, we built a complete model of the dodecameric protein (Figure 5.3a) and inserted the transmembrane portion in the bilayer of our tubular system with the ER-mimicking lipid composition. Then we simulated LD nucleation and growth with the POP-MD protocol; since TG synthesis in cells is localized near seipin (43; 149), we emulated the process by adding TG molecules only in the proximity of seipin. In the initial stages of growth, the nascent LD was scaffolded by seipin and acquired an asymmetric shape protruding towards the cytosolic compartment, not observed in the absence of the protein (Figure 5.3c). However, once the size of the nascent LD exceeded the diameter of the dodecamer, TG molecules leaked out of the seipin scaffold

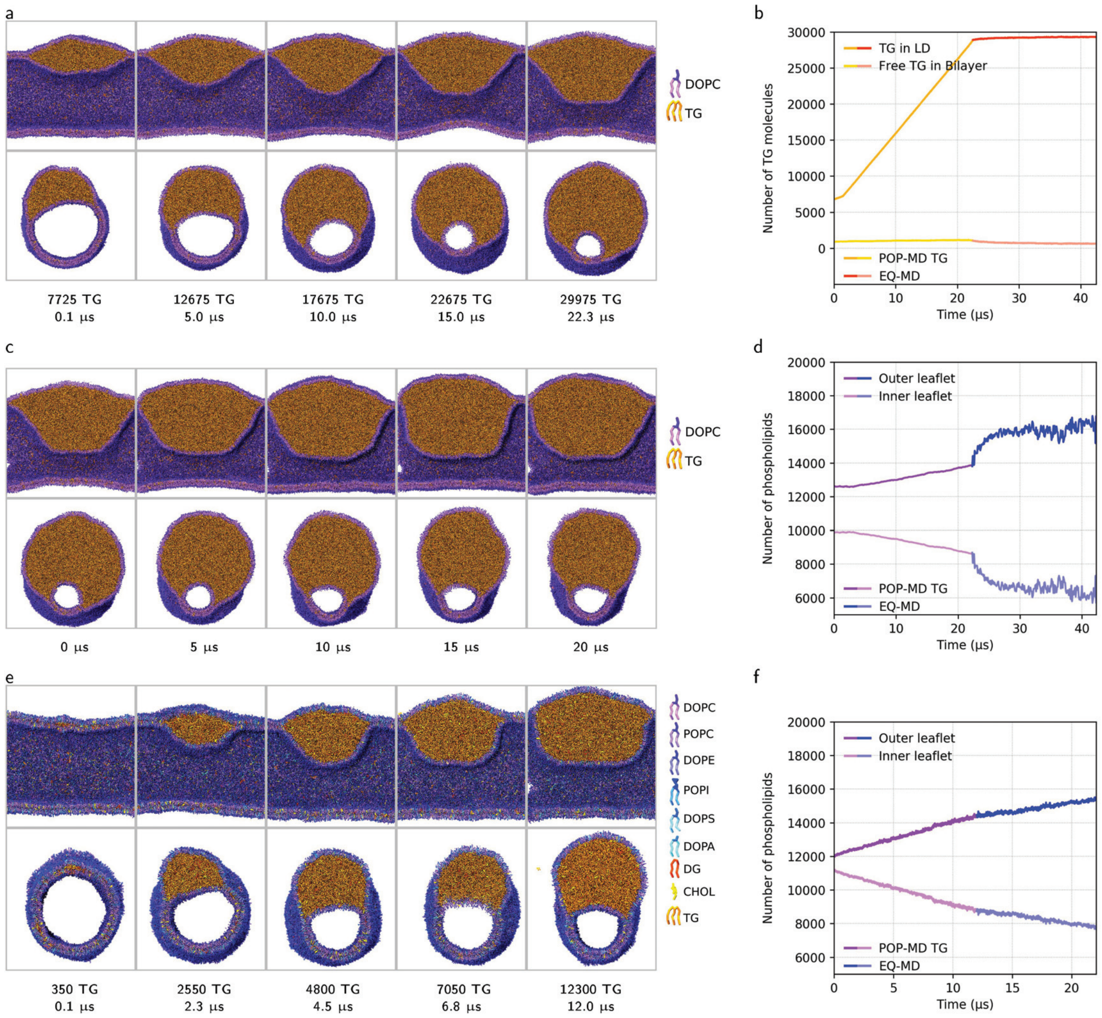


Figure 5.2: Growing LDs in tubular membranes by TG synthesis. (a) Snapshots from POP-MD simulations with a pure PC membrane tubule: side view cut along the tubule and perpendicular to the tubule, and (b) size of the nascent LD during POP-MD (23.3 µs in total) and during the subsequent equilibration phase (EQ-MD, 20 µs in total). (c) Snapshots from the equilibration phase (EQ-MD, no lipids added), side views cut along the tubule and perpendicular to the tubule, and (d) analysis of leaflet imbalance. (e) Snapshots from POP-MD simulations performed on a tubule with ER-mimicking bilayer composition, side views cut along the tubule and perpendicular to the tubule, and (f) analysis of leaflet imbalance during POP-MD and the subsequent equilibration phase (EQ-MD).

(Figure 5.3c, 9.8 µs) and the LD leaned towards the interior of the ER tubule. Leakage of TG was reproduced in different conditions and was enabled by the high flexibility of seipin in the region connecting the luminal and transmembrane domains, consistent with experimental data(36). Extension of the simulation without addition of TG promoted further buildup of leaflet asymmetry but did not rescue directional budding towards the cytosolic compartment (Figure 5.3d). Large membrane reservoirs are available in the ER and, in principle, they could allow large enough leaflet asymmetry simply via fluctuations in the area of individual leaflets. However, active generation of leaflet asymmetry appears, at this point, as a simpler explanation.

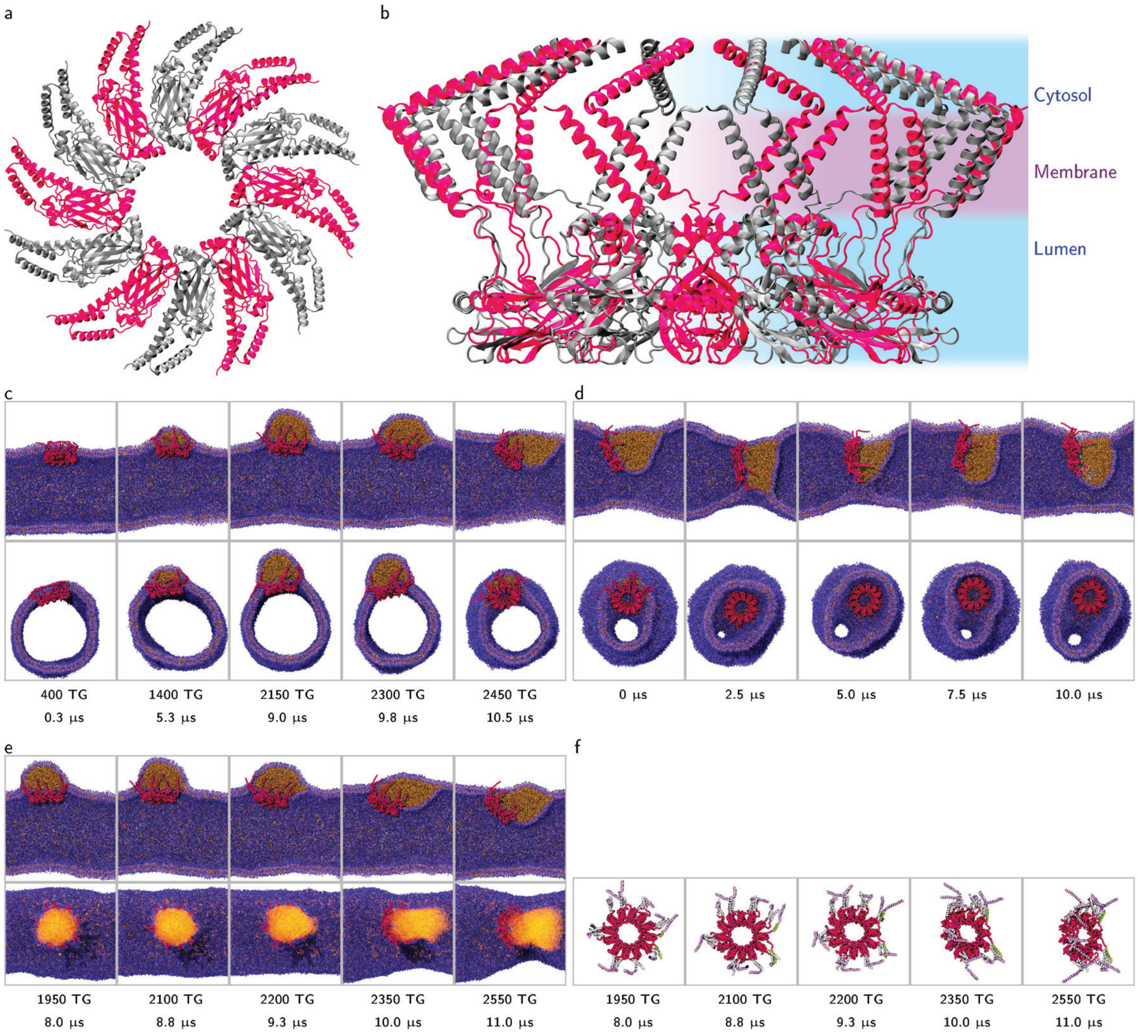


Figure 5.3: Growing a TG droplet in a tubular membrane in the presence of seipin. (a) Complete model of seipin, based on cryo-EM structure from Drosophila(34), top view and (b) side view. (c) Snapshots of the LD growth with seipin (number of TG molecules added by POP-MD is indicated for each snapshot), side views cut along the tubule and perpendicular to the tubule. (d) Snapshots from the equilibration phase. (e) Snapshots of the LD growth with seipin at the moment when seipin opens and let the LD out, side views cut along the tubule and top view. (f) Snapshots of seipin only at the moment when it opens and let the LD out, top view. Luminal domain in pink, N-ter AH in purple, TMD in white. The two TM helices the open up are highlighted in green.

5.4.4 Generation of leaflets asymmetry is necessary and sufficient for LD budding

According to experiments, theory, and simulations, leaflets asymmetry is necessary for LD directional budding(82; 14). This can be achieved by adding lipids with positive intrinsic curvature to the outer leaflet of the ER membrane(14), by adding proteins that bind the outer LD monolayer(82), or by adding any lipids to the outer ER leaflet to reduce its surface tension(82). In the ER, asymmetry can be generated by the asymmetric synthesis of PC lipids, that is regulated by the enzyme CCTα, catalyzing PC synthesis when bound to (PC-deficient) LDs(48). To emulate the biological process, we adapted our POP-MD algorithm to allow simultaneous insertion of TG (in the proximity of the seipin ring, to reduce equilibration times) and PC lipids (to the outer leaflet of the ER-mimicking tubule, in random positions), and therefore actively generate leaflet asymmetry. Adding a 1:2 ratio of TG and PC, we finally did observe, for the first time, a true budding transition. The budding transition generated structures qualitatively different from the ones observed so far, with the LD volume entirely outside the idealized cylindrical shape of the tubule (Figure 5.4a). Such structures were never observed without active generation of leaflet asymmetry, indicating that active buildup of leaflet asymmetry is necessary and sufficient for LD budding. One drawback of the budding process observed here is that the size of the LD-tubule connection is very large, only limited by the size of the tubule ($\sim$35 nm); the budded structure cannot accommodate the presence of seipin. Another drawback is that the excess surface of the outer layer causes large deformations of the LD-tubule connection, with water-filled defects protruding into the core of the budded LD (Figure 5.4a). We repeated the simulation adding a 1:1 TG:PC ratio, so leaflet asymmetry increases more slowly (Figure 5.4d). After the budding transition, deformations appeared again, but they were smaller and appeared later in the budding process (Figure 5.4c). These membrane deformations were not specific to tubular membranes systems, as they also appeared in analogous LD growth simulations of flat bilayer systems of different size, always in proximity of the LD-bilayer connection and always protruding into the liquid core of the LD (Figure 5.9). The consequences of defects in the LD neck region are unclear, but we can speculate that water-filled defects make the LD-tubule connection mechanically weak. Also, they may interfere with the flow of proteins and lipids from the ER to the LD, because both proteins and lipids diffuse from the ER to the LD (and vice versa) via the LD neck region. Finally, alterations in the morphology of the LD neck may interfere with protein binding.

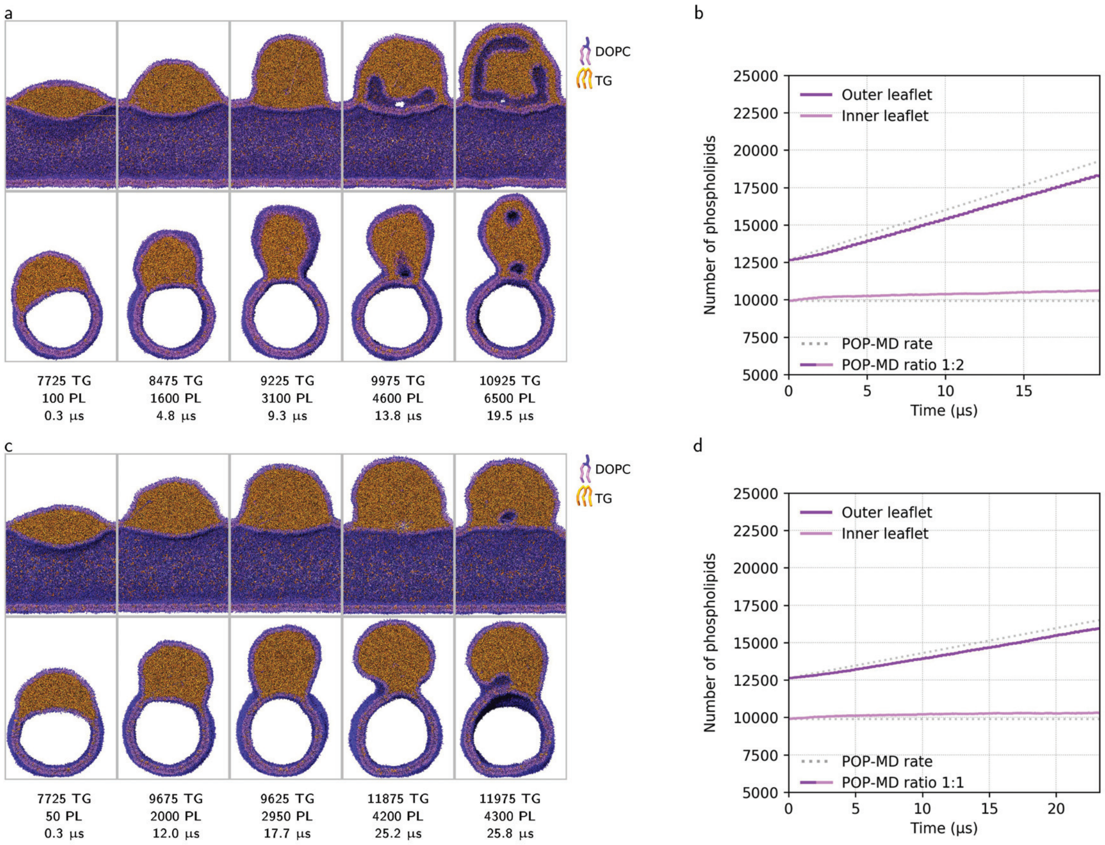


Figure 5.4: (a) Growing a TG droplet in a tubular membrane by adding both TG and PC lipids, in the absence of seipin, using POP-MD; side views cut along the tubule and perpendicular to the tubule. Time and the number of TG and PC molecules inserted are indicated, the ratio is 1:2. (b) Build-up of leaflet asymmetry during the POP-MD simulation. (c) Same as (a), but with 1:1 ratio. (d) Build-up of leaflet asymmetry during the POP-MD simulation.

5.4.5 Combining generation of leaflet asymmetry with the presence of seipin yields a robust budding mechanism

While simulations show that leaflet asymmetry is necessary and sufficient for LD budding, the simulated budding mechanism appears to be fragile, as it requires fine tuning of TG:phospholipid (PL) ratio to avoid water-filled membranous defects protruding into the LD neck, potentially impairing its mechanical stability, as well as lipid and protein diffusion and protein binding. Also, the LD-tubule connection is too large, incompatible with the presence of seipin. Seipin provides a scaffold for the LD neck, but per se it is not sufficient to induce directional budding. We hypothesize that combining leaflet asymmetry with the seipin scaffold might be necessary to achieve directional budding with a more robust budding mechanism, less dependent on the precise TG:PL ratio, yielding a stable, defect-free LD neck. To test this hypothesis, we carried out simulations of LD growth with active generation of leaflet asymmetry and in the presence of seipin, using the same TG/PL ratio as in the absence of seipin. Contrary to simulations without PL synthesis, a budding transition was observed when the nascent LD became too large to be contained entirely within the ring formed by seipin transmembrane domain. Seipin remained at the LD-ER contact site, fixing its size (15 nm) and shape, even after the budding transition was completed and the LD diameter became much larger than the seipin TM domain, effectively exerting a constraint on the LD and limiting its growth in the plane of the ER membrane. The constraint originates from the protein amino acid sequence: hydrophobic transmembrane helices are linked to luminal domains by short loops, so the transmembrane ring cannot expand; also, amino acid sequences of the transmembrane helices are bounded by charged residues or hydrophilic loops at each end, which prevents both ends from crossing the bilayer and ultimately prevents bilayer unzipping into two monolayers (observed in the absence of seipin), in agreement with recent simulations(55). Remarkably, no water-filled membrane defects were formed, neither in the proximity of the LD neck nor in other regions of the tubule. Instead, the tubule underwent a different type of deformation, with the cylindrical shape squeezed to avoid the formation of high-curvature regions in the proximity of LD neck (Figure 5.5a). As leaflet imbalance increased, when the LD was larger than the LD neck, tether-like deformations appeared in the tubular bilayer membrane, away from the LD neck and the LD monolayer, due to the excessive surface area of the outer leaflet. Overall, seipin had a clear effect on the mechanism of budding, yielding an approximately spherical budded LDs with a narrow, stable, and defect-free neck, different from all observations in the absence of seipin.

5.4.6 The complex ER composition and precise regulation of lipid synthesis stabilize LD-ER contact sites

We simulated LD budding by active generation of leaflet asymmetry in the presence of seipin, using a very simple lipid composition: only one type of oil and one or two types of phos-

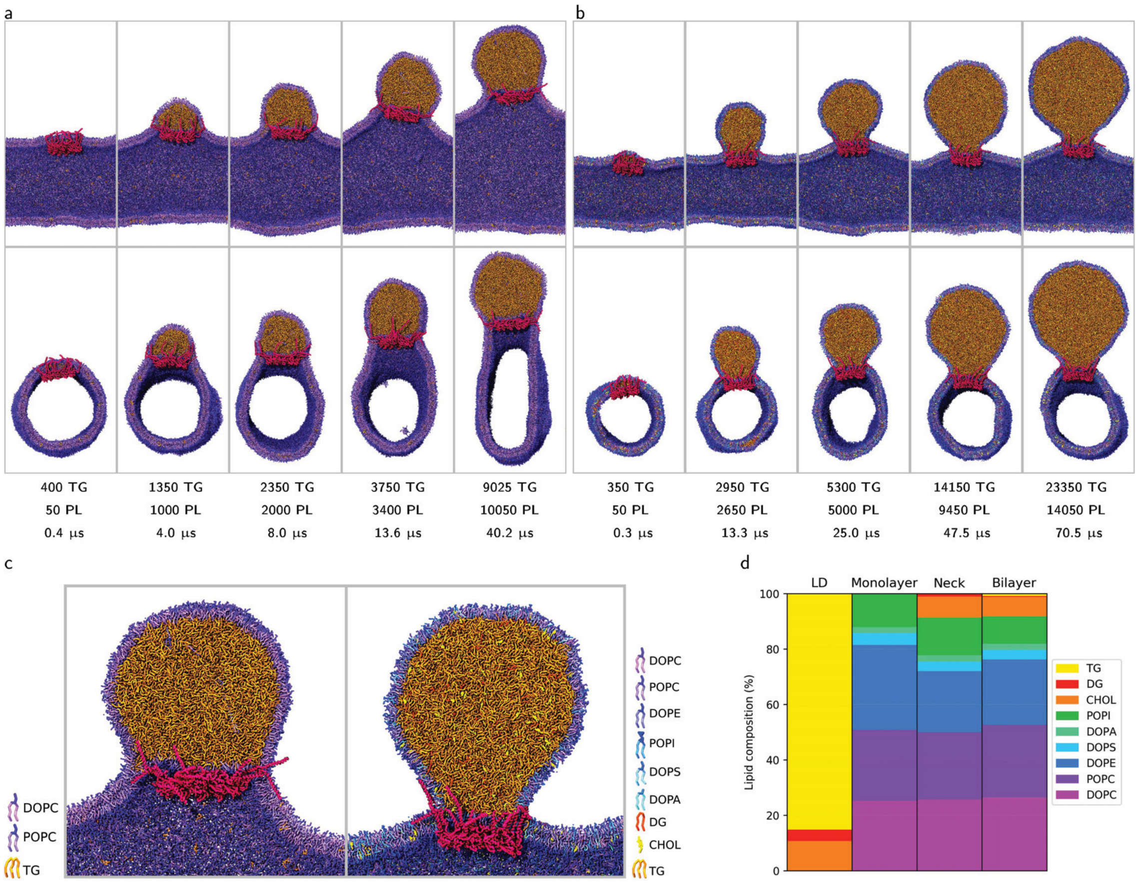


Figure 5.5: (a) Growing a TG droplet in a tubular membrane by adding both TG and PL (1:1 ratio), in the presence of seipin. Snapshots from POP-MD simulation with simple composition for the membrane tubule (DOPC:POPC 1:1), the number of TG and PL molecules added and the corresponding simulation time are indicated below. (b) Same as (a), but from simulation with the complex ER lipid mixture; all 8 lipid types are added during POP-MD. (c) Close-up of the LD-tubule connection, i.e., the LD neck. (d) Average lipid composition of the different regions in the system, showing local enrichment in Cholesterol, DO, and PI in the LD neck compared to the bulk bilayer region.

pholipids. So, what is the role of the complex ER composition in the budding process? To address this question, we repeated the simulations of LD growth on tubular membranes mimicking the complex ER phospholipid composition(137; 138) (Table 5.1). Nucleation and growth showed no qualitative difference compared to simulations with simpler lipid compositions. However, a substantial difference appeared in the budding transition: the complex membrane yielded high-curvature LD necks (Figure 5.5b,c), and the tubule remained almost perfectly cylindrical throughout the budding process, in contrast with the squeezed tubular shape observed in pure PC tubules (Figure 5.5a). As with the pure PC membrane, also with the complex ER composition tether-like defects formed at very high leaflet imbalance; however, the ER tubule conserved a nearly ideal cylindrical shape and high-curvature, defect-free neck. The same morphology could be reproduced independently of the localization of PL synthesis on the outer leaflet of the tubule: addition of PL close to the seipin ring, opposite to the seipin ring, and in random positions yielded the same remarkable result, indicating that the regular shape of the tubule depends on ER membrane composition (Supporting Information, Figure 5.13). We found two factors concurring to this effect: first, the membrane with the complex ER composition is softer than both pure DOPC and the DOPC:POPC 1:1 mixture (bending modulus of $11 \pm 1 k_B T$, compared to $18 \pm 2 k_B T$ for DOPC and $16 \pm 2 k_B T$ for DOPC/POPC), hence it is easier to bend. Second, the distribution of PL around the protein is not homogeneous: a significant enrichment in cholesterol, dioleoylglycerol (DG), and phosphatidylinositol (PI) is noticeable around the protein, as well as a depletion of POPC (Figure 5d). Interaction of seipin with PI and PA lipids was observed experimentally(33), and is probably related to the positive electrostatic charge on residues adjacent to the protein transmembrane region. Cholesterol and DO, having a smaller polar head, fit better than other lipids in the highly negatively curved neck region, stabilizing it.

In the simulations above, we added phospholipids at a constant rate, using a 1:1 TG:PL ratio. In cells, lipid synthesis is regulated through a feedback loop by the CCTα protein, and is tuned based on the PL coverage of the ER membrane(48). In other words, PC production in cells depends on the needs of the growing LD. While it would be difficult to reproduce the feedback loop in simulations, our POP-MD algorithm allows setting different PL synthesis rates, and even variable PL synthesis rates. The simulations above showed that, in the absence of PL synthesis, LD budding does not take place, due to insufficient leaflet asymmetry (Figure 5.1, 5.2, 5.3, 5.7, 5.9). On the other hand, the 1:1 TG:PL ratio causes deformations of the LD neck (in the absence of seipin, Figure 5.4) or the ER tubule (in the presence of seipin, Figure 5.6). But what happens if the TG:PL synthesis ratio is lower or higher? We repeated the simulations of LD growth using a 2:1 TG:PL synthesis ratio, so leaflet asymmetry grows more slowly compared to simulations above (Figure 5.6). In this case, an lens-shaped LD formed initially within the seipin ring; the nascent LD was asymmetric and protruding towards the cytosol, as imposed by the presence of the protein. Then a second one formed outside the seipin ring, also asymmetric but protruding towards the lumen, indicating that leaflet asymmetry generated by PL synthesis was insufficient for directional budding of both nascent LDs. Extending the simulation, the two LDs fused, but the fused LD (still in contact with seipin) could not bud towards the cytosolic compartment, due to insufficient leaflet asymmetry. We then repeated the simulations using a 1:2 TG:PL synthesis ratio (i.e., with leaflet asymmetry growing twice as fast). In this case, large tether-like deformations appeared in the ER tubule, away from the LD neck (Figure 5.6c). Interestingly, in the presence of Seipin, overgrowth of the outer leaflet never affects the shape of the budded LD, confirming the high stability of the LD neck in the presence of seipin. Deformations in the membrane tubule appear to depend solely on excessive leaflet imbalance. If this is the case, then it should be possible to rescue the "normal phenotype" (i.e., absence of deformations in the tubule) by tuning the TG:PL ratio during the simulations; this is analogous to the way cells operate, by regulating PC synthesis based on the needs of the growing LD(48). To confirm this, we performed a simulation with variable PL synthesis rate, starting from a 1:1 TG:PL rate and then increasing to 2:1 TG:PL rate after the budding transition. In this case, tubule deformations were completely avoided (Figure 5.6). While the actual rates of addition of TG and PL used in the simulations are perhaps inconsequential to biological systems, the high sensitivity of the simulated budding mechanism to lipid addition rates provides further support to our prediction, that precise regulation of lipid synthesis in cells is crucial for LD biogenesis, as well as for the stability of LD-ER contact sites.

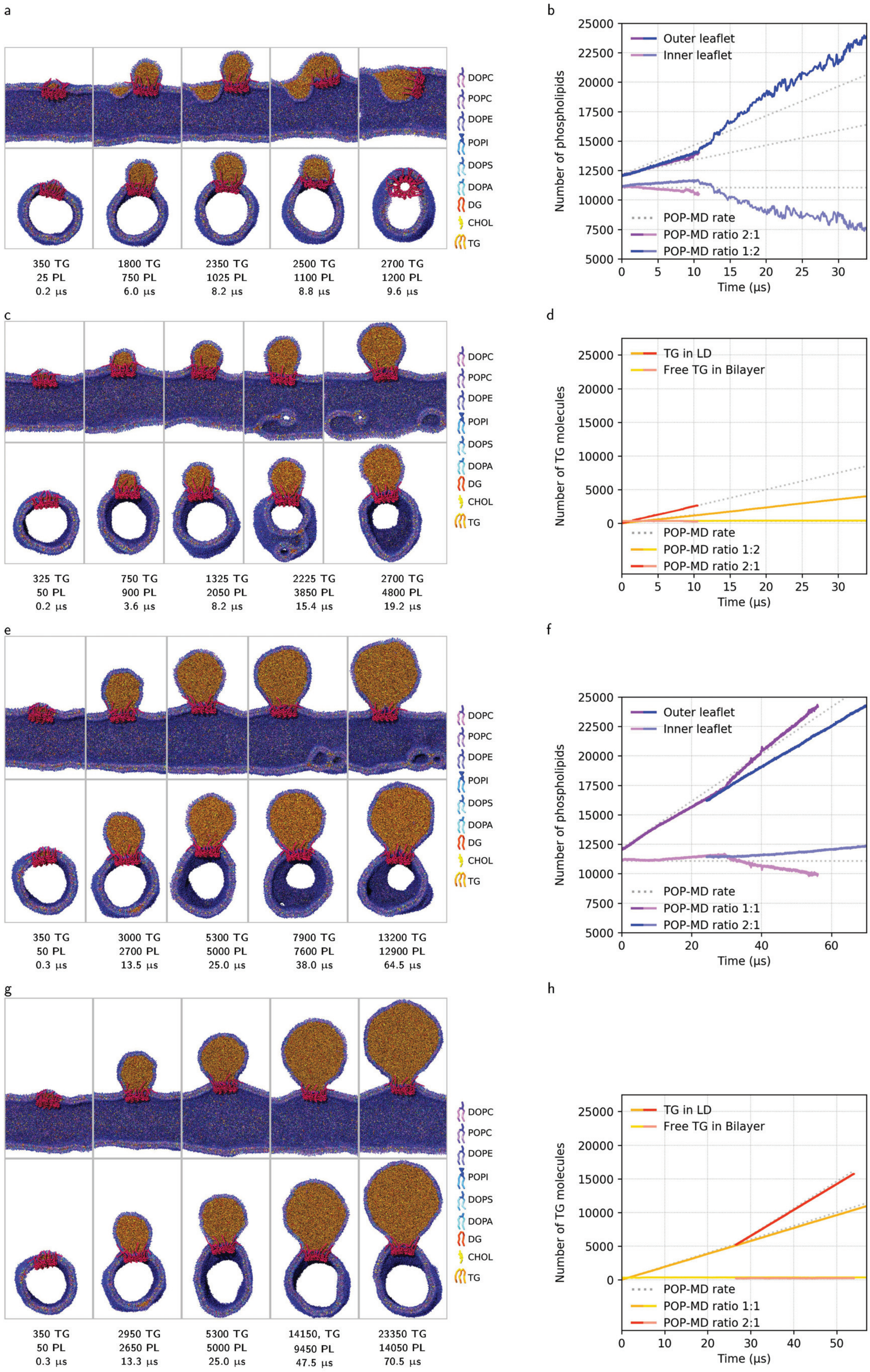

a
DOPC
POPC
DOPE
POPI
DOPS
DOPA
DG
CHOL
TG
350 TG 25 PL 0.2 µs
1800 TG 750 PL 6.0 µs
2350 TG 1025 PL 8.2 µs
2500 TG 1100 PL 8.8 µs
2700 TG 1200 PL 9.6 µs
b
Number of phospholipids
Time (µs)
Outer leaflet
Inner leaflet
POP-MD rate
POP-MD ratio 2:1
POP-MD ratio 1:2
25000
22500
20000
17500
15000
12500
10000
7500
5000
0
5
10
15
20
25
30
c
325 TG 50 PL 0.2 µs
750 TG 900 PL 3.6 µs
1325 TG 2050 PL 8.2 µs
2225 TG 3850 PL 15.4 µs
2700 TG 4800 PL 19.2 µs
d
Number of TG molecules
Time (µs)
TG in LD
Free TG in Bilayer
POP-MD rate
POP-MD ratio 1:2
POP-MD ratio 2:1
25000
20000
15000
10000
5000
0
e
350 TG 50 PL 0.3 µs
3000 TG 2700 PL 13.5 µs
5300 TG 5000 PL 25.0 µs
7900 TG 7600 PL 38.0 µs
13200 TG 12900 PL 64.5 µs
f
Number of phospholipids
Time (µs)
Outer leaflet
Inner leaflet
POP-MD rate
POP-MD ratio 1:1
POP-MD ratio 2:1
20
40
60
g
350 TG 50 PL 0.3 µs
2950 TG 2650 PL 13.3 µs
5300 TG 5000 PL 25.0 µs
14150, TG 9450 PL 47.5 µs
23350 TG 14050 PL 70.5 µs
h
Number of TG molecules
Time (µs)
TG in LD
Free TG in Bilayer
POP-MD rate
POP-MD ratio 1:1
POP-MD ratio 2:1
10
20
30
40
50

Figure 5.6 *(previous page)*: Growing a TG droplet in a tubular membrane with a complex ER mixture by adding both TG and 8 different PL types, in the presence of seipin. (a) Snapshots from simulation with TG:PL 2:1 insertion ratio, side views cut along the tubule and perpendicularly to the tubule, and (b) build-up of leaflet asymmetry. (c) Snapshots from simulation with TG:PL 1:2 insertion ratio, and (d) size of the nascent LD during POP-MD. (e) Snapshots from simulation with TG:PL 1:1 insertion ratio (f) build-up of leaflet asymmetry (g) Snapshots from simulation with variable TG:PL insertion ratio, starting with 1:1 and continuing with 2:1, and (h) size of the nascent LD during POP-MD.

5.5 Discussion

The mechanism of LD biogenesis is difficult to explore via current structural biology techniques, because of the fluid nature of biological membranes and LDs, as well as limitations in spatial and temporal resolution. Theory and simulations provided important contributions to our understanding of LD biogenesis (14; 116; 37; 124; 38), but the current picture is far from complete. A theoretical study proposed that the synthesis of neutral lipids, driving the initial stages of biogenesis (nucleation and growth), is sufficient also for driving the budding process (116), but experimental validation is lacking. In the theory, the bilayer is assumed to be very large (in principle, infinite) compared to the size of the LD, hence it can act as a reservoir of phospholipids and allows the buildup of leaflet asymmetry. In simulations, the reservoir is small, which makes it necessary to actively generate leaflets asymmetry. Asymmetric synthesis and degradation of phospholipids provide or destroy the material that coats outwards-budding LDs; is the assumption of a very large reservoir realistic for the ER tubular network, or rather asymmetric synthesis explains directional budding? We contend that simulations offer a more realistic picture, for two reasons. First, in the absence of leaflet asymmetry, LDs spontaneously should remain trapped in the ER membrane and protrude towards the lumen of tubular membranes; this event is rare in cells, and is associated with conditions that alter phospholipid synthesis or degradation (119). Second, the composition of LD monolayers has been shown to be different from the composition of ER bilayers (5), which can be easily explained by an important role of asymmetric PC synthesis during budding, consistent with our findings.

One of the main results of our simulations is the demonstration that generation of leaflet asymmetry (e.g., by PC synthesis) is necessary and sufficient for budding. This result is consistent with previous experimental data, showing that membrane asymmetry determines the direction of LD budding from the ER (82). But simulations show a much more detailed picture of the budding mechanism, and predict the existence of a budding transition as the key step in the process – not yet observed in cryo-EM images of the early stages of LD budding. The prediction is amenable to experimental validation, although trapping the different stages of budding is challenging for current experimental techniques, due to limitations in time resolution. A corollary to this result is the absence of a well-defined size for the budding transition: it is not the LD size that drives budding, but leaflet asymmetry. In our simulations, large LDs (up to 30 nm in diameter) would not undergo a budding transition in the absence of leaflet asymmetry, while smaller ones ($\sim$15 nm) showed a clear transition when leaflet asymmetry was sufficient.

In the simulations, the budding transition appeared within less than 1 μs after the system reached sufficient leaflet imbalance. However, predictions on the budding kinetics have limited value for two main reasons: first, the coarse-grained methodology used here yields faster dynamics compared to real processes; second, most important, the synthesis rates used in the simulations are extremely fast compared to reaction rates in cells, of the order of 1 lipid per ns; even considering that multiple enzymes may concur to asymmetry generation, rates used in the simulations exceed the turnover of lipid-synthesizing enzymes by several orders of magnitude. We believe our predictions on the driving forces for budding are instead more meaningful: the fast response to the buildup of leaflet asymmetry, observed in simulations, suggests that, in vivo, LD budding rate will be essentially limited by phospholipid synthesis.

The presence of seipin substantially alters the mechanism of LD budding in two ways: by determining the size of the LR-tubule connection, and by avoiding defects in the proximity of such connection. The prediction of a marked difference in the size of the LD-tubule connection is also amenable to experimental verification in vivo; differences in size by a few tens of nm are perfectly visible in cryo-EM (119). Defects at the LD neck were observed in all simulations in the absence of seipin; they formed faster (at an earlier stage in the simulation) with higher leaflet imbalance, and grew larger as LD size increased. Their presence was due to an excess of surface area of the outer leaflet of the tubule. It is conceivable that very tight regulation of the TG-PL ratio could avoid the formation of such defects, but the need for such tight regulation makes the budding mechanism fragile. In the presence of seipin, deformations appeared in the tubular membrane, instead, far from the LD-tubule connection.

We hypothesize that the different localization of defects has an effect on the stability of the LD neck, that would be relevant for the stability of the LD-ER contact site in vivo. Further work needed to test our hypothesis is in progress.

Finally, our simulations also allow some predictions on the role of lipid composition. Addition of lipid mixture supports a more robust mechanism of budding, without deformations in the tubule in the proximity of the LD-tubule contact site. It is possible that such differences in the morphology of the LD necks be linked with differences in the stability of the LD-tubule connections, but this hypothesis has not been verified. It should be noticed that our simulation results do not imply that synthesis of all phospholipids is important for the budding mechanism. In fact, in biological systems, synthesis of a single type of phospholipid (e.g., PC) would generate leaflet asymmetry without significantly altering the overall ER composition, because the extent of the ER membrane is far superior to the surface area of the LD (which is not the case in the simulations). The effect of asymmetry buildup by synthesis of a single phospholipid would be the alteration of the composition of the LD monolayer compared to the ER tubule, fully compatible with the differences between the composition of LD monolayers and ER bilayers recorded by mass spectrometry (5). On the other hand, regulation of the TG/PL ratio is predicted to be important, as it determines the appearance of deformations and defects (depending on phospholipid composition and on the presence of absence of seipin). In the presence of seipin, defects were only observed in the tubular membrane, but the morphology of the LD neck and the tubule were significantly different compared to the pure PC case; we hypothesize that such differences in morphology may reflect differences in the stability of the LD-tubule connection.

Last but not least, a brief comment on the capabilities and limitations of the POP algorithm devised here, allowing non-equilibrium simulations of LD growth. The algorithm makes it simple to add any molecular species in user-defined regions of the system, independently of the morphology of the system (flat, vesicle, tubule) and the presence of proteins, with a user-defined, variable rate, and hence allows to emulate the growth of any membrane system. Depending on the synthesis rate, the simulated systems may or may not reach equilibrium during the subsequent relaxation. The lack of equilibrium is problematic for the calculation of any thermodynamic properties of the simulated systems, which is an important limitation. On the other hand, the morphological features observed for our nascent LDs proved to be robust and reproducible, and we could directly link different features to differences in protein and lipid composition among the systems. We anticipate that the same methodology will enable simulations of even more complex transformations involving membrane systems, including the biogenesis of other organelles, the formation of viral envelopes, and bacterial division.

5.6 Supporting Information

5.6.1 Membrane tubules with pores allow changes in tube diameter and length

The presence of stable hydrophilic pores in membrane tubules allows barrierless flip-flop, dissipation of asymmetric stresses and differences in pressure in/out of the tubule (water can freely flow in/out), and changes in tubule diameter and length. We generated tubules of different diameter by pulling on them with external forces.

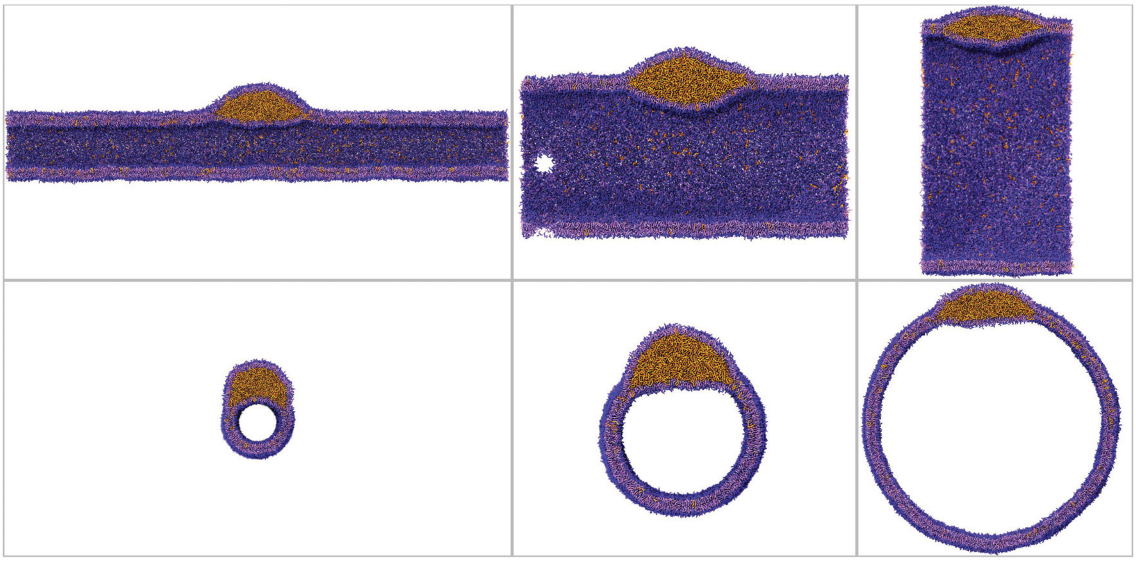


Figure 5.7: Membrane tubules with a diameter of 10 nm, 35 nm, and 50 nm, generated by extending the simulation box.

5.6.2 Membrane tubules with pores allow phospholipid asymmetry

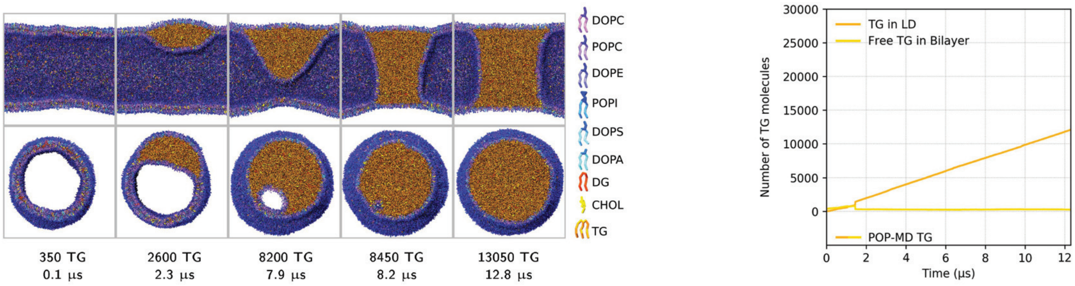


Figure 5.8: Growing LDs in tubular membranes by TG synthesis without pores. (a) Snapshots from POP-MD simulations performed on a tubule with ER-mimicking bilayer composition, side views cut along the tubule and perpendicular to the tubule (b) size of the nascent LD during POP-MD.

5.6.3 Uncontrolled phospholipid synthesis in the outer leaflet generates defects in the LD neck

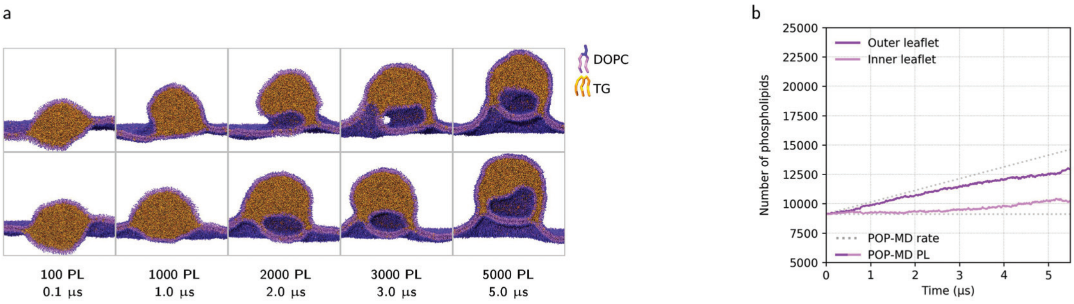


Figure 5.9: Snapshots from POP-MD in flat bilayer systems, with only phospholipids added to the system, only in the outer leaflet. When the surface area of the outer leaflet is beyond a certain threshold, the excess surface folds and generates defects. Defects are always in the proximity of the LD neck.

5.6.4 Leaflet imbalance is necessary and sufficient for LD budding independently of the membrane morphology, presence of seipin, and size of the LD

Budding simulations were achieved using tubular membranes and very large nascent LDs, so the role of tubular geometry and LD size remained unclear. To address this question, we repeated the POP-MD simulations on much smaller LDs embedded in planar membranes. Surprisingly, LD budding was observed even on tiny nascent LDs, as small as 10 nm in diameter (corresponding to 625 TG molecules), indicating that leaflet imbalance is necessary and sufficient for LD budding independently of the membrane morphology, the presence of seipin, and even the size of the LD.

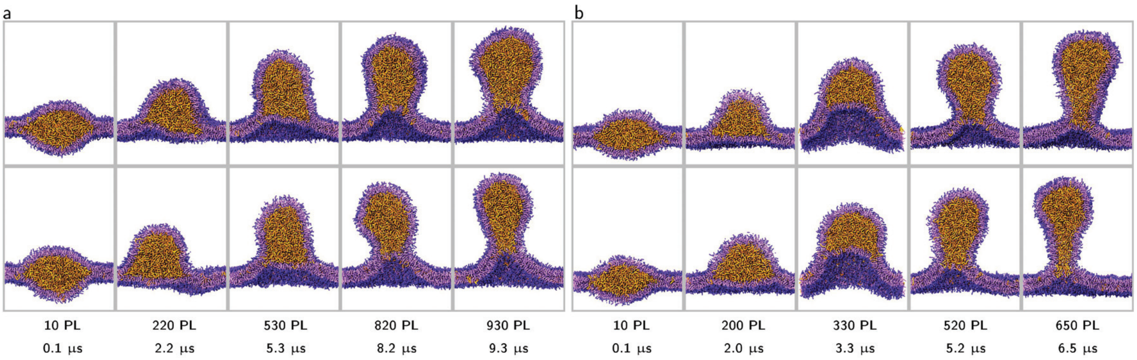


Figure 5.10: Snapshots from POP-MD in flat bilayer systems, with only phospholipids added to the system, only in the outer leaflet, (a) imposing a constant area to the bilayer, (b) with varying bilayer area.

5.6.5 Growing a TG droplet in a tubular membrane in the presence of seipin

We added the TG molecules in the tubular membrane in random positions. Although seipin has been suggested to favor LD nucleation, we observed 4 nucleation events happening outside the seipin ring. This can be explained by the long diffusion times required for TG molecules to travel to the seipin ring, due to the large size of the system. To overcome this barrier, we repeated the simulations by adding TG molecules close to the seipin ring.

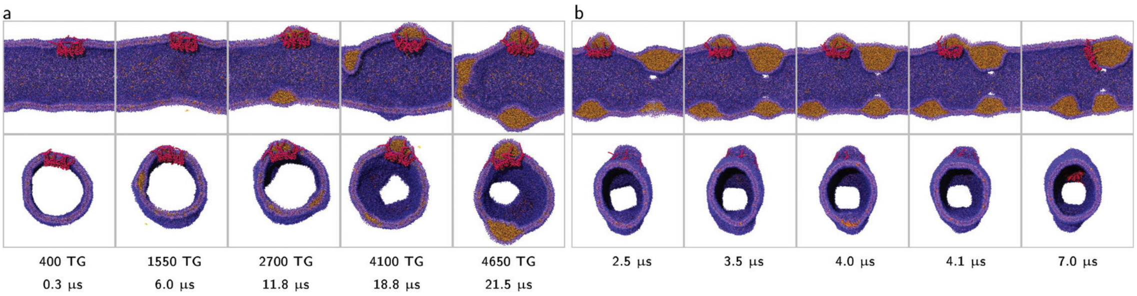


Figure 5.11: Randomly adding TG in a tubular membrane in the presence of seipin. (a) Snapshots of the LDs nucleations with Seipin after addition of 50, 1200, 2350, 3750, 4300 TG molecules, side views cut along the tubule and (b) perpendicularly to the tubule. (c) Snapshots from the equilibration phase after 2.5 µs, 3.5 µs, 4 µs, 4.1 µs and 7 µs, side views cut along the tubule and (d) perpendicularly to the tubule.

5.6.6 Seipin alone does not trap nascent lipid droplets subject to fusion

In membranes containing TG molecules, seipin prevents TG from diffusing out of is ring. Seipin TM-helices have been reported to have affinity for TG, and the $\alpha2$-$\alpha3$ helix in luminal domain shows an even higher affinity. The $\alpha2$-$\alpha3$ helix is necessary for seipin functioning in forming TG clusters [Prasanna 2021] [Zoni 2021]. We inserted seipin in the bilayer of flat membrane system, and repeated the ISG procedure to add TG molecules (see Methods), starting from a membrane containing only dispersed TG molecules. We added the TG molecules randomly in the membrane. In the early stages of growth (Figure 5.12), we can observe up to 6 LD nucleation events happening outside the Seipin ring, due to the large size of the system and the high rate of addition of TG.

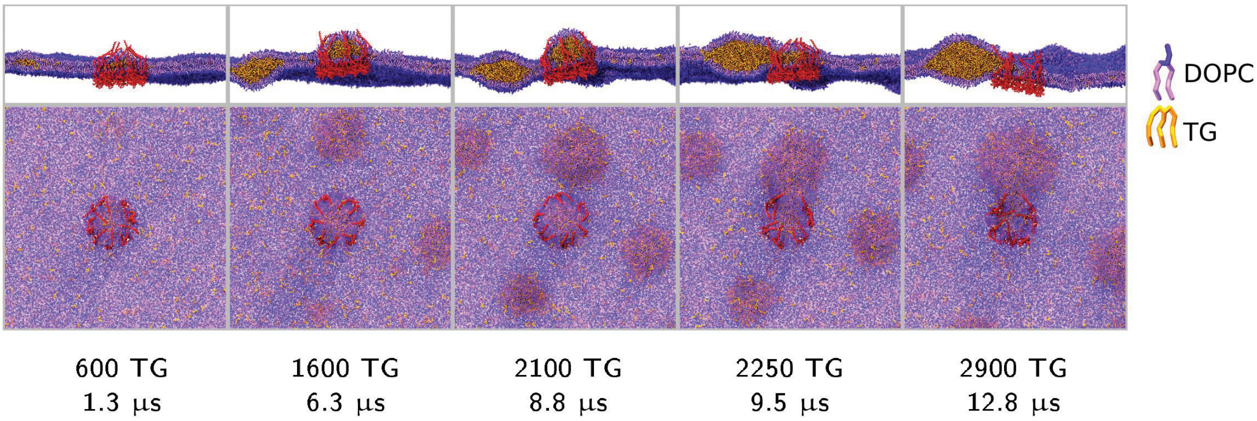


Figure 5.12: Growing a TG droplet in a flat membrane in the presence of seipin. (A) Snapshots of the LD growth with Seipin after addition of 250, 1250, 1750, 1900, 2550 TG molecules, side views cut along the LD and (B) view for the top.

The asymmetric droplet scaffolded by Seipin needs more phospholipids on the upper leaflet than the lower leaflet. The symmetric lens-shaped droplet uses approximately the same number of phospholipids on each leaflet. When the asymmetric droplet fuses with the symmetric droplet, there is not enough phospholipids available on the upper leaflet to support the asymmetry than would be need to bud a larger LD contained in Seipin. As a consequence, without additional asymmetry generation, the result is comparable to result in figure 5. These findings on dissolved TG molecules [Prasanna 2021] [Zoni 2021] do not transfer to whole nascent LDs. Indeed, the volume of a small accumulation of TG is small enough not to require asymmetry. The volume would start to be significant after the addition of hundreds of TG molecule.

It is unlikely that LDs escaping the radius of Seipin would happened in cells. As other proteins are recruited by Seipin at the LD-ER contact point, the structure formed by the complex coats the neck and could prevent fusion of the budded LD with other LDs.

5.6.7 Excessive leaflet asymmetry leads to major deformation in membrane tubules

Addition of PC lipids to the outer leaflet of the membrane tubule resulted in severe deformations of the tubule, due to the excess outer membrane surface (Figure 5.6c). We added TG and PC lipids at different ratios, while at the same time maintaining the membrane pores open. Although the use of hydrophilic pores may appear to defeat the purpose of asymmetric synthesis, the pores do not allow for rapid dissipation of leaflet asymmetry, just like they do not allow for rapid enough generation of asymmetry (see above). The net result is a rapid buildup of leaflet asymmetry with somewhat limited, self-regulated, buildup of asymmetric stresses.

First, we started from the tubular system containing 7625 TG and 22482 PL, and added only PL. The system did not contain hydrophilic pores. In this case, we observed a rapid deformation of the LD shape, that becomes asymmetric and buds toward the cytosolic compartment. The width of the neck connecting the LD to the tubule decreases as the LD grows, until it deforms significantly (Figure 5.6): a water-filled invagination is formed within the LD, to accommodate the excess area of the outer leaflet. A similar result was also obtained in the case of a flat system. Instability of the budded LD depends on the excessive outer membrane surface, that has grown too much with respect to the volume of the LD.

These membranes deformations are not specific to tubular membranes systems. We repeat an ISG procedure in a flat membrane system with an embedded nascent lipid droplet of 7500 TG molecules. Only phospholipids were added, to the upper leaflet of the membrane to create asymmetry. In this simulation as well, shown in figure 8, the LD buds, but does not form a stable nor narrow neck at the LD-ER contact point. The leaflet asymmetry ends up creating membranes folds bellow the LD. We repeated this ISG procedure with a smaller lipid droplet of 2500 TG molecules which gave the same result.

5.6.8 The localization of phospholipids synthesis does not affect LD budding mechanism

We repeated the simulations of LD growth in the tubular membranes mimicking the complex ER composition, in the presence of seipin, by adding oil and phospholipids in a 1:1 ratio in 3 different ways: phospholipids were added in the proximity of the seipin ring (Figure 5.13g,h), on the opposite side of the seipin ring (Figure 5.13a,b), or in random locations of the outer leaflet (Figure 5.13d,e). In all three cases, LD budding followed the same mechanism, resulting in a regular cylindrical shape of the ER tubule, with no deformations in the LD monolayer nor close to the LD neck.

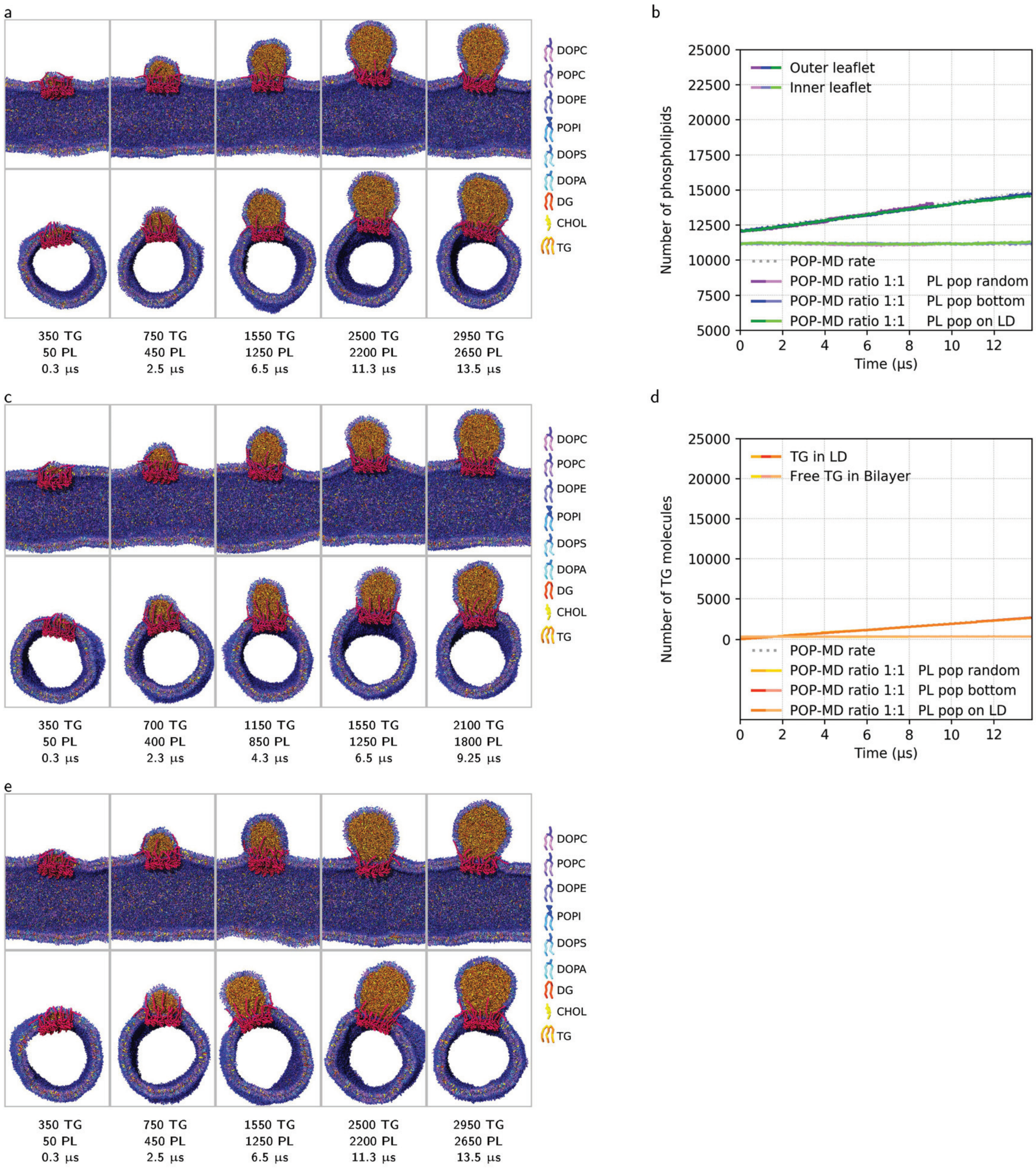


Figure 5.13: Growing a TG droplet in a tubular membrane with a complex ER mixture by adding both TG and 8 different PL types, in the presence of seipin. (a) Snapshots of the LD growth after addition of 50, 450, 1250, 2200, 2650 TG and 50, 450, 1250, 2200, 2650 PL molecules at the bottom of the outer leaflet of the tubular membrane, side views cut along the tubule and perpendicularly to the tubule. (b) Build-up of leaflet asymmetry. (c) Snapshots of the LD growth after addition of 50, 400, 850, 1250, 1800 TG and 50, 400, 850, 1250, 1800 PL molecules at any location in the outer leaflet of the tubular membrane, side views cut along the tubule and perpendicularly to the tubule. (d) Size of the nascent LD during POP-MD. (e) Snapshots of the LD growth after addition of 50, 450, 1250, 2200, 2650 TG and 50, 450, 1250, 2200, 2650 PL molecules at the top of the outer leaflet of the tubular membrane, close to the nascent LD, side views cut along the tubule and perpendicularly to the tubule.

5.6.9 LD-tubule connection is extremely stable

Having simulated LD budding under different conditions, we never observed instability of the LD neck nor any tendency of the budded LD to detach from the membrane tubule, not even when the LD neck was highly distorted. This is rather expected, as it has been shown that mature, fully budded LDs generally remain in contact with the ER via an LD-ER contact site (usually punctuated by seipin). However, it is conceivable that the great stability of the LD neck is the result of insufficient simulation time. While it is difficult to directly tackle the problem of time scales with current computer hardware, it is possible to probe the stability of the LD-tubule connection by stressing the system with mechanical forces. To this end, we simulated a vesicle with an embedded LD and added a strong external force pulling the LD away from the vesicle at constant speed (0.5 nm/μs), to induce fission. After pulling for 10 μs, the center of mass of the TG LD was displaced by 10 nm away from the center of mass of the vesicle. During pulling, the neck of the LD became longer and thinner, reaching an outer diameter of 9 nm (Figure 5.14), very close to the membrane thickness. Despite the extremely small diameter, the tubular neck did not rupture. In a second simulation, starting from the final configuration of the previous, the same forces were applied on TG and DOPC, but the position of the applied forces was fixed for 12 s (Figure 5.14). The neck became slightly wider, and did not rupture. Finally, all external forces were set to zero, allowing relaxation for 20 μs. This resulted in a fast recovery of the original shape: the neck becomes shorter and wider, reaching a final diameter of 30 nm (same as the initial diameter), the vesicle becomes spherical again, and the LD adopted a lens shape. These results confirm the very high mechanical stability of the LD-ER connection, and suggest that the lightbulb morphology of budded LDs, with a neck narrower than the LD diameter, is essentially due to the presence of seipin.

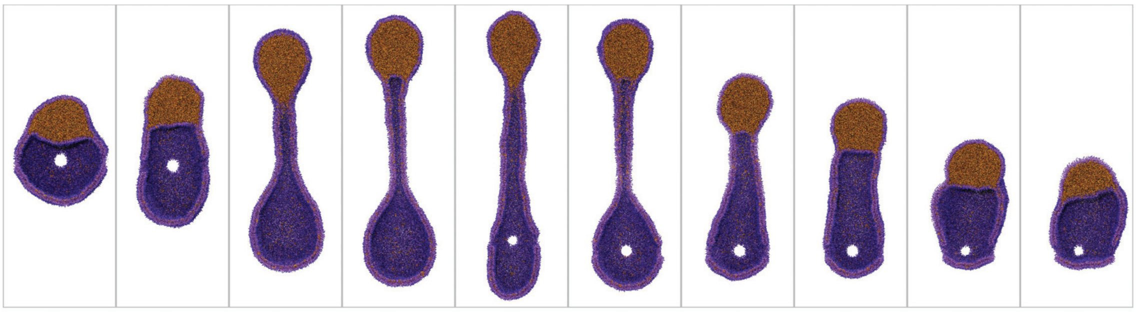


Figure 5.14: Pulling and releasing of the lipid droplet in a vesicle. Pulling simulation of the LD after 0 µs, 4 µs, 8 µs, 12 µs, pulling at rate of 5 nm/µs. Holding the pulling forces in place for 22 µs. Unbiased simulation of the vesicle after 0 µs, 1 µs, 2.5 µs, 11 µs, 20 µs.

Chapter 6

Conclusion

In conclusion, this PhD thesis aimed to investigate the complex process of LD biogenesis, focusing on understanding the partitioning of proteins in nascent LDs, analyzing LD shape as a proxy for budding, and investigating the mechanism of LD budding. Through molecular dynamics simulations, this work has provided valuable insights into these fundamental aspects of LD formation.

The first aspect addressed in this thesis was the partitioning of proteins in nascent LDs. It was demonstrated experimentally that monotopic membrane proteins tend to accumulate at the surface of LDs, indicating a preference for the LD monolayer. Molecular dynamics simulations successfully reproduced the partitioning of monotopic membrane proteins to LD monolayers, confirming the experimental observations. Simulations further supported these experimental findings and revealed that protein-lipid interactions play a crucial role in determining protein distribution in LDs. Additionally, simulations showed for the first time that transmembrane proteins tend to accumulate at the edge of LDs; an result which had never been tested experimentally. These results contribute to a better understanding of the factors influencing protein distribution in LDs and highlight the importance of protein-lipid interactions in organelle biogenesis.

The second aspect investigated the relationship between LD shape and budding. By simulating nascent LDs embedded in lipid bilayers of varying sizes and compositions, it was found that LDs become more spherical as they grow in size, as their oil-water surface increases, and as the bilayer membranes become softer. Theoretical predictions regarding LD shape were validated through simulations, revealing that surface tension plays a dominant role in determining LD shape, rather than membrane bending modulus. Furthermore, the fitting of simulated LD shapes to theoretical equations provided insights into the elastic properties of these systems, suggesting the presence of surface tension in the bilayer region. These findings shed light on the factors influencing LD shape and provide a microscopic view of droplet embedded vesicles.

The third and most challenging aspect addressed in this thesis was the mechanism of LD budding. LD budding is a crucial step in determining the protein coverage and function of LDs, yet it remains poorly understood due to the limitations of current experimental techniques. In this thesis, a novel non-equilibrium simulation method called POP-MD was developed to emulate lipid synthesis and explore the driving forces and conditions for LD budding. By incorporating seipin, a key protein involved in LD biogenesis, into realistic tubular bilayer membranes, the simulations successfully replicated the budding process. The results revealed that LD size is not the critical parameter for budding, but rather leaflet asymmetry is necessary and sufficient to induce budding. The presence of seipin was found to scaffold the neck of the budded LD, preventing the generation of hydrophilic defects. Additionally, the complex lipid composition of the endoplasmic reticulum (ER) facilitated the budding mechanism by providing softer membranes and reducing membrane defects during biogenesis. These insights into the molecular mechanisms of LD budding offer a significant advancement in our understanding of this crucial process.

Overall, this PhD thesis has made significant contributions to the field of LD biogenesis. By combining molecular dynamics simulations, theoretical modeling, and experimental insights, it has provided unprecedented molecular-level views of LD formation and dynamics. The findings regarding protein partitioning, LD shape, and the mechanism of budding have shed light on the intricate processes involved in LD biogenesis. This research not only expands our knowledge of cellular organelles but also has important implications for understanding the role of LDs in various diseases. Further experimental validations of the theoretical and simulation predictions presented in this thesis will continue to enhance our understanding of LD biogenesis and open up new avenues for therapeutic interventions targeting LD-associated diseases.

The comprehensive understanding of LD biogenesis, protein partitioning, and the mechanism of LD budding opens up exciting avenues for further investigation. Future studies could explore the specific protein-lipid interactions governing protein localization, investigate the dynamics of LD shape in different cellular contexts, and explore the regulatory factors and signaling pathways involved in LD budding. Together, experimental approaches, computational modeling and theoretical predictions can unravel the intricacies of LD formation and dynamics, leading to the development of novel therapeutic strategies for LD-associated diseases and a deeper understanding of LD's broader implications in cellular physiology and pathology.

References

[1] Robert V Farese Jr and Tobias C Walther. Lipid droplets finally get a little respect. *Cell*, 139(5):855–860, 2009.

[2] Abdou Rachid Thiam, Robert V Farese Jr, and Tobias C Walther. The biophysics and cell biology of lipid droplets. *Nature reviews Molecular cell biology*, 14(12):775–786, 2013.

[3] E Currie and A Schulze. zechner r, walther tc and farese rv jr: Cellular fatty acid metabolism and cancer. *Cell Metab*, 18:153–161, 2013.

[4] Himanshu Khandelia, Lars Duelund, Kirsi I Pakkanen, and John H Ipsen. Triglyceride blisters in lipid bilayers: implications for lipid droplet biogenesis and the mobile lipid signal in cancer cell membranes. *PloS one*, 5(9):e12811, 2010.

[5] Kumi Tauchi-Sato, Shintaro Ozeki, Toshiaki Houjou, Ryo Taguchi, and Toyoshi Fujimoto. The surface of lipid droplets is a phospholipid monolayer with a unique fatty acid composition. *Journal of Biological Chemistry*, 277(46):44507–44512, 2002.

[6] Darcy L Johannsen, Yourka Tchoukalova, Charmaine S Tam, Jeffrey D Covington, Wenting Xie, Jean-Marc Schwarz, Sudip Bajpeyi, and Eric Ravussin. Effect of 8 weeks of overfeeding on ectopic fat deposition and insulin sensitivity: testing the "adipose tissue expandability" hypothesis. *Diabetes care*, 37(10):2789–2797, 2014.

[7] Alexandre Santinho, Aymeric Chorlay, Lionel Foret, and Abdou Rachid Thiam. Fat inclusions strongly alter membrane mechanics. *Biophysical journal*, 120(4):607–617, 2021.

[8] Aymeric Chorlay, Lionel Forêt, and Abdou Rachid Thiam. Origin of gradients in lipid density and surface tension between connected lipid droplet and bilayer. *Biophysical journal*, 120(24):5491–5503, 2021.

[9] Christoph Thiele and Johanna Spandl. Cell biology of lipid droplets. *Current opinion in cell biology*, 20(4):378–385, 2008.

[10] James A Olzmann and Pedro Carvalho. Dynamics and functions of lipid droplets. *Nature reviews Molecular cell biology*, 20(3):137–155, 2019.

[11] Juergen Zanghellini, F Wodlei, and HH Von Grünberg. Phospholipid demixing and the birth of a lipid droplet. *Journal of theoretical biology*, 264(3):952–961, 2010.

[12] Florian Wilfling, Joel T Haas, Tobias C Walther, and Robert V Farese Jr. Lipid droplet biogenesis. *Current opinion in cell biology*, 29:39–45, 2014.

[13] Kalthoum Ben M'barek, Dalila Ajjaji, Aymeric Chorlay, Stefano Vanni, Lionel Forêt, and Abdou Rachid Thiam. Er membrane phospholipids and surface tension control cellular lipid droplet formation. *Developmental cell*, 41(6):591–604, 2017.

[14] Vineet Choudhary, Gonen Golani, Amit S Joshi, Stéphanie Cottier, Roger Schneiter, William A Prinz, and Michael M Kozlov. Architecture of lipid droplets in endoplasmic reticulum is determined by phospholipid intrinsic curvature. *Current Biology*, 28(6):915–926, 2018.

[15] Alex M Valm, Sarah Cohen, Wesley R Legant, Justin Melunis, Uri Hershberg, Eric Wait, Andrew R Cohen, Michael W Davidson, Eric Betzig, and Jennifer Lippincott-Schwartz. Applying systems-level spectral imaging and analysis to reveal the organelle interactome. *Nature*, 546(7656):162–167, 2017.

[16] Gerrit Van Meer, Dennis R Voelker, and Gerald W Feigenson. Membrane lipids: where they are and how they behave. *Nature reviews Molecular cell biology*, 9(2):112–124, 2008.

[17] Aldert R van Buuren, D Peter Tieleman, Jacob de Vlieg, and Herman JC Berendsen. Cosurfactants lower surface tension of the diglyceride/water interface: A molecular dynamics study. *Langmuir*, 12(10):2570–2579, 1996.

[18] Margret Paar, Christian Jüngst, Noemi A Steiner, Christoph Magnes, Frank Sinner, Dagmar Kolb, Achim Lass, Robert Zimmermann, Andreas Zumbusch, Sepp D Kohlwein, et al. Remodeling of lipid droplets during lipolysis and growth in adipocytes. *Journal of Biological Chemistry*, 287(14):11164–11173, 2012.

[19] Suwimon Ariyaprakai and Stephanie R Dungan. Influence of surfactant structure on the contribution of micelles to ostwald ripening in oil-in-water emulsions. *Journal of colloid and interface science*, 343(1):102–108, 2010.

[20] Nicolas Bremond, Abdou R Thiam, and Jérôme Bibette. Decompressing emulsion droplets favors coalescence. *Physical review letters*, 100(2):024501, 2008.

[21] Florian Wilfling, Huajin Wang, Joel T Haas, Natalie Krahmer, Travis J Gould, Aki Uchida, Ji-Xin Cheng, Morven Graham, Romain Christiano, Florian Fröhlich, et al. Triacylglycerol synthesis enzymes mediate lipid droplet growth by relocalizing from the er to lipid droplets. *Developmental cell*, 24(4):384–399, 2013.

[22] Zhe Cao and Ho Yi Mak. Married at birth: regulation of cellular fat metabolism by er–lipid droplet crosstalk. *Contact*, 3:2515256420934671, 2020.

[23] Kirill Bersuker, Clark WH Peterson, Milton To, Steffen J Sahl, Victoria Savikhin, Elizabeth A Grossman, Daniel K Nomura, and James A Olzmann. A proximity labeling strategy provides insights into the composition and dynamics of lipid droplet proteomes. *Developmental cell*, 44(1):97–112, 2018.

[24] Nora Kory, Robert V Farese Jr, and Tobias C Walther. Targeting fat: mechanisms of protein localization to lipid droplets. *Trends in cell biology*, 26(7):535–546, 2016.

[25] Aymeric Chorlay and Abdou Rachid Thiam. Neutral lipids regulate amphipathic helix affinity for model lipid droplets. *Journal of Cell Biology*, 219(4), 2020.

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

[27] Abdou Rachid Thiam, Bruno Antonny, Jing Wang, Jérôme Delacotte, Florian Wilfling, Tobias C Walther, Rainer Beck, James E Rothman, and Frédéric Pincet. Copi buds 60-nm lipid droplets from reconstituted water–phospholipid–triacylglyceride interfaces, suggesting a tension clamp function. *Proceedings of the National Academy of Sciences*, 110(33):13244–13249, 2013.

[28] Alenka Čopič, Sandra Antoine-Bally, Manuel Giménez-Andrés, César La Torre Garay, Bruno Antonny, Marco M Manni, Sophie Pagnotta, Jeanne Guihot, and Catherine L Jackson. A giant amphipathic helix from a perilipin that is adapted for coating lipid droplets. *Nature communications*, 9(1):1–16, 2018.

[29] Dalila Ajjaji, Kalthoum Ben M'barek, Michael L Mimmack, Cheryl England, Haya Herscovitz, Liang Dong, Richard G Kay, Satish Patel, Vladimir Saudek, Donald M Small, et al. Dual binding motifs underpin the hierarchical association of perilipins1–3 with lipid droplets. *Molecular biology of the cell*, 30(5):703–716, 2019.

[30] Veijo T Salo, Maarit Hölttä-Vuori, and Elina Ikonen. Seipin-mediated contacts as gatekeepers of lipid flux at the endoplasmic reticulum–lipid droplet nexus. *Contact*, 3:2515256420945820, 2020.

[31] Bethany R Cartwright and Joel M Goodman. Seipin: from human disease to molecular mechanism. *Journal of lipid research*, 53(6):1042–1055, 2012.

[32] Veijo T Salo, Shiqian Li, Helena Vihinen, Maarit Hölttä-Vuori, Abel Szkalisity, Peter Horvath, Ilya Belevich, Johan Peränen, Christoph Thiele, Pentti Somerharju, et al. Seipin facilitates triglyceride flow to lipid droplet and counteracts droplet ripening via endoplasmic reticulum contact. *Developmental cell*, 50(4):478–493, 2019.

[33] Renhong Yan, Hongwu Qian, Ivan Lukmantara, Mingming Gao, Ximing Du, Nieng Yan, and Hongyuan Yang. Human seipin binds anionic phospholipids. *Developmental cell*, 47(2):248–256, 2018.

[34] Xuewu Sui, Henning Arlt, Kelly P Brock, Zon Weng Lai, Frank DiMaio, Debora S Marks, Maofu Liao, Robert V Farese Jr, and Tobias C Walther. Cryo–electron microscopy structure of the lipid droplet–formation protein seipin. *Journal of Cell Biology*, 217(12):4080–4091, 2018.

[35] Yoel A Klug, Justin C Deme, Robin A Corey, Mike F Renne, Phillip J Stansfeld, Susan M Lea, and Pedro Carvalho. Mechanism of lipid droplet formation by the yeast sei1/ldb16 seipin complex. *Nature communications*, 12(1):1–13, 2021.

[36] Henning Arlt, Xuewu Sui, Brayden Folger, Carson Adams, Xiao Chen, Roman Remme, Fred A. Hamprecht, Frank DiMaio, Maofu Liao, Joel M. Goodman, Robert V. Farese, and Tobias C. Walther. Seipin forms a flexible cage at lipid droplet formation sites. *Nature Structural Molecular Biology*, 29(3):194–202, 2022.

[37] Xavier Prasanna, Veijo T Salo, Shiqian Li, Katharina Ven, Helena Vihinen, Eija Jokitalo, Ilpo Vattulainen, and Elina Ikonen. Seipin traps triacylglycerols to facilitate their nanoscale clustering in the endoplasmic reticulum membrane. *PLoS biology*, 19(1):e3000998, 2021.

[38] Valeria Zoni, Wataru Shinoda, and Stefano Vanni. Seipin accumulates and traps diacylglycerols and tryglycerides in its ring-like structure. *Biophysical Journal*, 120(3):231a, 2021.

[39] Robin W Klemm, Justin P Norton, Ronald A Cole, Chen S Li, Seong H Park, Matthew M Crane, Liying Li, Diana Jin, Alexandra Boye-Doe, Tina Y Liu, et al. A conserved role for atlastin gtpases in regulating lipid droplet size. *Cell reports*, 3(5):1465–1475, 2013.

[40] Julia Falk, Magdalena Rohde, Mohamed M Bekhite, Sophie Neugebauer, Peter Hemmerich, Michael Kiehntopf, Thomas Deufel, Christian A Hübner, and Christian Beetz. Functional mutation analysis provides evidence for a role of reep 1 in lipid droplet biology. *Human mutation*, 35(4):497–504, 2014.

[41] Chrisovalantis Papadopoulos, Genny Orso, Giuseppe Mancuso, Marija Herholz, Sentiljana Gumeni, Nimesha Tadepalle, Christian Jüngst, Anne Tzschichholz, Astrid Schauss, Stefan Höning, et al. Spastin binds to lipid droplets and affects lipid metabolism. *PLoS genetics*, 11(4):e1005149, 2015.

[42] Jeeyun Chung, Xudong Wu, Talley J Lambert, Zon Weng Lai, Tobias C Walther, and Robert V Farese Jr. Ldaf1 and seipin form a lipid droplet assembly complex. *Developmental cell*, 51(5):551–563, 2019.

[43] Vineet Choudhary, Ola El Atab, Giulia Mizzon, William A Prinz, and Roger Schneiter. Seipin and nem1 establish discrete er subdomains to initiate yeast lipid droplet biogenesis. *Journal of Cell Biology*, 219(7), 2020.

[44] Amit S Joshi, Benjamin Nebenfuehr, Vineet Choudhary, Prasanna Satpute-Krishnan, Tim P Levine, Andy Golden, and William A Prinz. Lipid droplet and peroxisome biogenesis occur at the same er subdomains. *Nature communications*, 9(1):2940, 2018.

[45] Sihui Wang, Fatima-Zahra Idrissi, Martin Hermansson, Alexandra Grippa, Christer S Ejsing, and Pedro Carvalho. Seipin and the membrane-shaping protein pex30 cooperate in organelle budding from the endoplasmic reticulum. *Nature communications*, 9(1):2939, 2018.

[46] Qiang Gao, Derk D Binns, Lisa N Kinch, Nick V Grishin, Natalie Ortiz, Xiao Chen, and Joel M Goodman. Pet10p is a yeast perilipin that stabilizes lipid droplets and promotes their assembly. *Journal of Cell Biology*, 216(10):3199–3217, 2017.

[47] Caleb C Lord, Daniel Ferguson, Gwynneth Thomas, Amanda L Brown, Rebecca C Schugar, Amy Burrows, Anthony D Gromovsky, Jenna Betters, Chase Neumann, Jessica Sacks, et al. Regulation of hepatic triacylglycerol metabolism by cgi-58 does not require atgl co-activation. *Cell reports*, 16(4):939–949, 2016.

[48] Natalie Krahmer, Yi Guo, Florian Wilfling, Maximiliane Hilger, Susanne Lingrell, Klaus Heger, Heather W Newman, Marc Schmidt-Supprian, Dennis E Vance, Matthias Mann, et al. Phosphatidylcholine synthesis for lipid droplet expansion is mediated by localized activation of ctp: phosphocholine cytidylyltransferase. *Cell metabolism*, 14(4):504–515, 2011.

[49] Marta Bosch, Miguel Sánchez-Álvarez, Alba Fajardo, Ronan Kapetanovic, Bernhard Steiner, Filipe Dutra, Luciana Moreira, Juan Antonio López, Rocío Campo, Montserrat Marí, et al. Mammalian lipid droplets are innate immune hubs integrating cell metabolism and host defense. *Science*, 370(6514):eaay8085, 2020.

[50] Peter Walter and David Ron. The unfolded protein response: from stress pathway to homeostatic regulation. *science*, 334(6059):1081–1086, 2011.

[51] Romain Volmer, Kattria van der Ploeg, and David Ron. Membrane lipid saturation activates endoplasmic reticulum unfolded protein response transducers through their transmembrane domains. *Proceedings of the National Academy of Sciences*, 110(12):4628–4633, 2013.

[52] Thanyarat Promlek, Yuki Ishiwata-Kimata, Masahiro Shido, Mitsuru Sakuramoto, Kenji Kohno, and Yukio Kimata. Membrane aberrancy and unfolded proteins activate the endoplasmic reticulum stress sensor ire1 in different ways. *Molecular biology of the cell*, 22(18):3520–3532, 2011.

[53] Angelika S Rambold, Sarah Cohen, and Jennifer Lippincott-Schwartz. Fatty acid trafficking in starved cells: regulation by lipid droplet lipolysis, autophagy, and mitochondrial fusion dynamics. *Developmental cell*, 32(6):678–692, 2015.

[54] Truc B Nguyen, Sharon M Louie, Joseph R Daniele, Quan Tran, Andrew Dillin, Roberto Zoncu, Daniel K Nomura, and James A Olzmann. Dgat1-dependent lipid droplet biogenesis protects mitochondrial function during starvation-induced autophagy. *Developmental cell*, 42(1):9–21, 2017.

[55] Siyoung Kim, Jeeyun Chung, Henning Arlt, Alexander J. Pak, Robert V. Jnr Farese, Tobias C. Walther, and Gregory A. Voth. Seipin transmembrane segments critically function in triglyceride nucleation and lipid droplet budding from the membrane. *eLife*, 11:e75808, 2022.

[56] Siyoung Kim, Jessica M. J. Swanson, and Gregory A. Voth. Computational studies of lipid droplets. *The Journal of Physical Chemistry B*, 126(11):2145–2154, 2022.

[57] Herman Berendsen. *Simulating the physical world*. Cambridge University Press Cambridge, 2007.

[58] Michael P Allen and Dominic J Tildesley. *Computer simulation of liquids*. Oxford university press, 2017.

[59] Andrew R Leach and Andrew R Leach. *Molecular modelling: principles and applications*. Pearson education, 2001.

[60] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald: An n log (n) method for ewald sums in large systems. *The Journal of chemical physics*, 98(12):10089–10092, 1993.

[61] Ulrich Essmann, Lalith Perera, Max L Berkowitz, Tom Darden, Hsing Lee, and Lee G Pedersen. A smooth particle mesh ewald method. *The Journal of chemical physics*, 103(19):8577–8593, 1995.

[62] Ilario G Tironi, René Sperb, Paul E Smith, and Wilfred F van Gunsteren. A generalized reaction field method for molecular dynamics simulations. *The Journal of chemical physics*, 102(13):5451–5459, 1995.

[63] Herman JC Berendsen, JPM van Postma, Wilfred F Van Gunsteren, ARHJ DiNola, and Jan R Haak. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81(8):3684–3690, 1984.

[64] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *The Journal of chemical physics*, 126(1):014101, 2007.

[65] Michele Parrinello and Aneesur Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, 52(12):7182–7190, 1981.

[66] Gregory A Voth. *Coarse-graining of condensed phase and biomolecular systems*. CRC press, 2008.

[67] Siewert J Marrink, H Jelger Risselada, Serge Yefimov, D Peter Tieleman, and Alex H De Vries. The martini force field: coarse grained model for biomolecular simulations. *The journal of physical chemistry B*, 111(27):7812–7824, 2007.

[68] Luca Monticelli, Senthil K Kandasamy, Xavier Periole, Ronald G Larson, D Peter Tieleman, and Siewert-Jan Marrink. The martini coarse-grained force field: extension to proteins. *Journal of chemical theory and computation*, 4(5):819–834, 2008.

[69] Cesar A López, Andrzej J Rzepiela, Alex H de Vries, Lubbert Dijkhuizen, Philippe H Hunenberger, and Siewert J Marrink. Martini coarse-grained force field: extension to carbohydrates. *Journal of Chemical Theory and Computation*, 5(12):3195–3210, 2009.

[70] Jaakko J Uusitalo, Helgi I Ingolfsson, Parisa Akhshi, D Peter Tieleman, and Siewert J Marrink. Martini coarse-grained force field: extension to dna. *Journal of chemical theory and computation*, 11(8):3932–3945, 2015.

[71] Jirasak Wong-Ekkabut, Svetlana Baoukina, Wannapong Triampo, I-Ming Tang, D Peter Tieleman, Luca Monticelli, et al. Computer simulation study of fullerene translocation through lipid membranes. *Nature nanotechnology*, 3(6):363–368, 2008.

[72] Luca Monticelli. On atomistic and coarse-grained models for c60 fullerene. *Journal of Chemical Theory and Computation*, 8(4):1370–1378, 2012.

[73] Giulia Rossi, Luca Monticelli, Sakari R Puisto, Ilpo Vattulainen, and Tapio Ala-Nissila. Coarse-graining polymers with the martini force-field: polystyrene as a benchmark case. *Soft Matter*, 7(2):698–708, 2011.

[74] Siewert J Marrink, Luca Monticelli, Manuel N Melo, Riccardo Alessandri, D Peter Tieleman, and Paulo CT Souza. Two decades of martini: Better beads, broader scope. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, page e1620, 2022.

[75] P. C. T. Souza, R. Alessandri, J. Barnoud, S. Thallmair, I. Faustino, F. Grunewald, I. Patmanidis, H. Abdizadeh, B. M. H. Bruininks, T. A. Wassenaar, P. C. Kroon, J. Melcr, V. Nieto, V. Corradi, H. M. Khan, J. Domanski, M. Javanainen, H. Martinez-Seara, N. Reuter, R. B. Best, I. Vattulainen, L. Monticelli, X. Periole, D. P. Tieleman, A. H. de Vries, and S. J. Marrink. Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nature methods*, 18(4):382–388, 2021.

[76] Xavier Periole, Marco Cavalli, Siewert-Jan Marrink, and Marco A Ceruso. Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *Journal of chemical theory and computation*, 5(9):2531–2543, 2009.

[77] Adolfo B Poma, Marek Cieplak, and Panagiotis E Theodorakis. Combining the martini and structure-based coarse-grained approaches for the molecular dynamics studies of conformational transitions in proteins. *Journal of Chemical Theory and Computation*, 13(3):1366–1374, 2017.

[78] Tobias C Walther and Robert V Farese Jr. Lipid droplets and cellular lipid metabolism. *Annual review of biochemistry*, 81:687, 2012.

[79] Abdou Rachid Thiam and Lionel Forêt. The physics of lipid droplet nucleation, growth and budding. *Biochimica Et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*, 1861(8):715–722, 2016.

[80] Abdou Rachid Thiam and Mathias Beller. The why, when and how of lipid droplet diversity. *Journal of cell science*, 130(2):315–324, 2017.

[81] W Mike Henne, Michael L Reese, and Joel M Goodman. The assembly of lipid droplets and their roles in challenged cells. *The EMBO journal*, 37(12):e98947, 2018.

[82] Aymeric Chorlay, Luca Monticelli, Joana Veríssimo Ferreira, Kalthoum Ben M'barek, Dalila Ajjaji, Sihui Wang, Errin Johnson, Rainer Beck, Mohyeddine Omrane, Mathias Beller, et al. Membrane asymmetry imposes directionality on lipid droplet emergence from the er. *Developmental cell*, 50(1):25–42, 2019.

[83] Huajin Wang, Michel Becuwe, Benjamin E Housden, Chandramohan Chitraju, Ashley J Porras, Morven M Graham, Xinran N Liu, Abdou Rachid Thiam, David B Savage, Anil K Agarwal, et al. Seipin is required for converting nascent to mature lipid droplets. *elife*, 5:e16582, 2016.

[84] Veijo T Salo, Ilya Belevich, Shiqian Li, Leena Karhinen, Helena Vihinen, Corinne Vigouroux, Jocelyne Magré, Christoph Thiele, Maarit Hölttä-Vuori, Eija Jokitalo, et al. Seipin regulates er–lipid droplet contacts and cargo delivery. *The EMBO journal*, 35(24):2699–2716, 2016.

[85] Nurulain Ho, Chengchao Xu, and Guillaume Thibault. From the unfolded protein response to metabolic diseases–lipids under the spotlight. *Journal of cell science*, 131(3):jcs199307, 2018.

[86] Ravi Dhiman, Stefanie Caesar, Abdou Rachid Thiam, and Bianca Schrul. Mechanisms of protein targeting to lipid droplets: A unified cell biological and biophysical perspective. In *Seminars in Cell & Developmental Biology*, volume 108, pages 4–13. Elsevier, 2020.

[87] Robert Ernst, Stephanie Ballweg, and Ilya Levental. Cellular mechanisms of physicochemical membrane homeostasis. *Current opinion in cell biology*, 53:44–51, 2018.

[88] Erdinc Sezgin, Ilya Levental, Satyajit Mayor, and Christian Eggeling. The mystery of membrane organization: composition, regulation and roles of lipid rafts. *Nature reviews Molecular cell biology*, 18(6):361–374, 2017.

[89] Derek Marsh. Lateral pressure profile, spontaneous curvature frustration, and the incorporation and conformation of proteins in membranes. *Biophysical journal*, 93(11):3884–3899, 2007.

[90] Norberto Castillo, Luca Monticelli, Jonathan Barnoud, and D Peter Tieleman. Free energy of walp23 dimer association in dmpc, dppc, and dopc bilayers. *Chemistry and physics of lipids*, 169:95–105, 2013.

[91] Barbara A Lewis and Donald M Engelman. Lipid bilayer thickness varies linearly with acyl chain length in fluid phosphatidylcholine vesicles. *Journal of molecular biology*, 166(2):211–217, 1983.

[92] Abdou Rachid Thiam and Isabelle Dugail. Lipid droplet–membrane contact sites–from protein binding to function. *Journal of Cell Science*, 132(12):jcs230169, 2019.

[93] Camille I Pataki, João Rodrigues, Lichao Zhang, Junyang Qian, Bradley Efron, Trevor Hastie, Joshua E Elias, Michael Levitt, and Ron R Kopito. Proteomic analysis of monolayer-integrated proteins on lipid droplets identifies amphipathic interfacial α-helical membrane anchors. *Proceedings of the National Academy of Sciences*, 115(35):E8172–E8180, 2018.

[94] Tarik Exner, Inés Romero-Brey, Eden Yifrach, Jhon Rivera-Monroy, Bianca Schrul, Christos C Zouboulis, Wolfgang Stremmel, Masanori Honsho, Ralf Bartenschlager, Einat Zalckvar, et al. An alternative membrane topology permits lipid droplet localization of peroxisomal fatty acyl-coa reductase 1. *Journal of cell science*, 132(6):jcs223016, 2019.

[95] Donald M Small, Libo Wang, and Matthew A Mitsche. The adsorption of biological peptides and proteins at the oil/water interface. a potentially important but largely unexplored field. *Journal of lipid research*, 50:S329–S334, 2009.

[96] Emily R Rowe, Michael L Mimmack, Antonio D Barbosa, Afreen Haider, Iona Isaac, Myriam M Ouberai, Abdou Rachid Thiam, Satish Patel, Vladimir Saudek, Symeon Siniossoglou, et al. Conserved amphipathic helices mediate lipid droplet targeting of perilipins 1–3. *Journal of Biological Chemistry*, 291(13):6664–6678, 2016.

[97] Nicolas Jacquier, Vineet Choudhary, Muriel Mari, Alexandre Toulmay, Fulvio Reggiori, and Roger Schneiter. Lipid droplets are functionally connected to the endoplasmic reticulum in saccharomyces cerevisiae. *Journal of cell science*, 124(14):2424–2437, 2011.

[98] Florian Wilfling, Abdou Rachid Thiam, Maria-Jesus Olarte, Jing Wang, Rainer Beck, Travis J Gould, Edward S Allgeyer, Frederic Pincet, Jörg Bewersdorf, Robert V Farese Jr, et al. Arf1/copi machinery acts directly on lipid droplets and enables their connection to the er for protein targeting. *Elife*, 3:e01607, 2014.

[99] Olaf S Andersen and Roger E Koeppe. Bilayer thickness and membrane protein function: an energetic perspective. *Annu. Rev. Biophys. Biomol. Struct.*, 36:107–130, 2007.

[100] James U Bowie. Solving the membrane protein folding problem. *Nature*, 438(7068):581–589, 2005.

[101] Rob Phillips, Tristan Ursell, Paul Wiggins, and Pierre Sens. Emerging roles for lipids in shaping membrane-protein function. *Nature*, 459(7245):379–385, 2009.

[102] Anthony G Lee. How lipids affect the activities of integral membrane proteins. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1666(1-2):62–87, 2004.

[103] J Antoinette Killian. Hydrophobic mismatch between proteins and lipids in membranes. *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes*, 1376(3):401–416, 1998.

[104] Michael J Booth, Vanessa Restrepo Schild, Florence G Downs, and Hagan Bayley. Functional aqueous droplet networks. *Molecular BioSystems*, 13(9):1658–1691, 2017.

[105] Abdou R Thiam, Nicolas Bremond, and Jerome Bibette. From stability to permeability of adhesive emulsion bilayers. *Langmuir*, 28(15):6291–6298, 2012.

[106] Elizabeth Calzada, Ouma Onguka, and Steven M Claypool. Phosphatidylethanolamine metabolism in health and disease. *International review of cell and molecular biology*, 321:29–88, 2016.

[107] Jelske N van der Veen, John P Kennelly, Sereana Wan, Jean E Vance, Dennis E Vance, and René L Jacobs. The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1859(9):1558–1572, 2017.

[108] Chien-Yu Huang and Anthony HC Huang. Unique motifs and length of hairpin in oleosin target the cytosolic side of endoplasmic reticulum and budding lipid droplet. *Plant physiology*, 174(4):2248–2260, 2017.

[109] Anne G Ostermeyer, James M Paci, Youchun Zeng, Douglas M Lublin, Sean Munro, and Deborah A Brown. Accumulation of caveolin in the endoplasmic reticulum redirects the protein to lipid storage droplets. *The Journal of cell biology*, 152(5):1071–1078, 2001.

[110] Peter J Thul, Kirsten Tschapalda, Petra Kolkhof, Abdou Rachid Thiam, Monika Oberer, and Mathias Beller. Targeting of the drosophila protein cg2254/ldsdh1 to a subset of lipid droplets. *Journal of Cell Science*, 130(18):3141–3157, 2017.

[111] Adam Kassan, Albert Herms, Andrea Fernández-Vidal, Marta Bosch, Nicole L Schieber, Babu JN Reddy, Alba Fajardo, Mariona Gelabert-Baldrich, Francesc Tebar, Carlos Enrich, et al. Acyl-coa synthetase 3 promotes lipid droplet biogenesis in er microdomains. *Journal of Cell Biology*, 203(6):985–1001, 2013.

[112] Dawn L Brasaemle. Thematic review series: adipocyte biology. the perilipin family of structural lipid droplet proteins: stabilization of lipid droplets and control of lipolysis. *Journal of lipid research*, 48(12):2547–2559, 2007.

[113] J Antoinette Killian and Thomas KM Nyholm. Peptides in lipid bilayers: the power of simple models. *Current opinion in structural biology*, 16(4):473–479, 2006.

[114] Leonid V Chernomordik and Michael M Kozlov. Protein-lipid interplay in fusion and fission of biological membranes. *Annual review of biochemistry*, 72(1):175–207, 2003.

[115] Aymeric Chorlay and Abdou Rachid Thiam. An asymmetry in monolayer tension regulates lipid droplet budding direction. *Biophysical journal*, 114(3):631–640, 2018.

[116] F. Deslandes, A. R. Thiam, and L. Foret. Lipid droplets can spontaneously bud off from a symmetric bilayer. *Biophysical Journal*, 113(1):15–18, 2017.

[117] Meng Wang and Xin Yi. Bulging and budding of lipid droplets from symmetric and asymmetric membranes: competition between membrane elastic energy and interfacial energy. *Soft Matter*, 17(21):5319–5328, 2021.

[118] W. Helfrich. Elastic properties of lipid bilayers - theory and possible experiments. *Zeitschrift Fur Naturforschung C - a Journal of Biosciences*, C 28(11-1):693–703, 1973.

[119] Vineet Choudhary, Namrata Ojha, Andy Golden, and William A. Prinz. A conserved family of proteins facilitates nascent lipid droplet budding from the er. *Journal of Cell Biology*, 211(2):261–271, 2015.

[120] Tsjerk A. Wassenaar, Helgi I. Ingolfsson, Rainer A. Boeckmann, D. Peter Tieleman, and Siewert J. Marrink. Computational lipidomics with insane: A versatile tool for generating custom membranes for molecular simulations. *Journal of Chemical Theory and Computation*, 11(5):2144–2155, 2015.

[121] Mark James Abraham, Teemu Murtola, Roland Schulz, Szilárd Páll, Jeremy C Smith, Berk Hess, and Erik Lindahl. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1:19–25, 2015.

[122] Juan M. Vanegas, Alejandro Torres-Sánchez, and Marino Arroyo. Importance of force decomposition for local stress calculations in biomembrane molecular simulations. *Journal of Chemical Theory and Computation*, 10(2):691–702, 2014.

[123] J. Barnoud, G. Rossi, and L. Monticelli. Lipid membranes as solvents for carbon nanoparticles. *Physical Review Letters*, 112:068102, 2014.

[124] V. Zoni, R. Khaddaj, P. Campomanes, A. R. Thiam, R. Schneiter, and S. Vanni. Pre-existing bilayer stresses modulate triglyceride accumulation in the er versus lipid droplets. *Elife*, 10, 2021.

[125] Jad Eid, Hafez Razmazma, Alia Jraij, Ali Ebrahimi, and Luca Monticelli. On calculating the bending modulus of lipid bilayer membranes from buckling simulations. *The Journal of Physical Chemistry B*, 124(29):6299–6311, 2020.

[126] J. F. Nagle. Experimentally determined tilt and bending moduli of single-component lipid bilayers. *Chem Phys Lipids*, 205:18–24, 2017.

[127] W. Rawicz, K. C. Olbrich, T. McIntosh, D. Needham, and E. Evans. Effect of chain length and unsaturation on elasticity of lipid bilayers. *Biophys. J.*, 79(1):328–339, 2000.

[128] Derek Marsh. Elastic curvature constants of lipid monolayers and bilayers. *Chemistry and Physics of Lipids*, 144(2):146–159, 2006.

[129] Arpita Upadhyaya and Michael P. Sheetz. Tension in tubulovesicular networks of golgi and endoplasmic reticulum membranes. *Biophysical Journal*, 86(5):2923–2928, 2004.

[130] 2017.

[131] O. H. S. Ollila, H. J. Risselada, M. Louhivuori, E. Lindahl, I. Vattulainen, and S. J. Marrink. 3d pressure field in lipid membranes and membrane-protein complexes. *Physical Review Letters*, 102(7):078101, 2009.

[132] Alejandro Torres-Sánchez, Juan M. Vanegas, and Marino Arroyo. Examining the mechanical equilibrium of microscopic stresses in molecular simulations. *Physical Review Letters*, 114(25):258102, 2015.

[133] J. R. Knowles. Enzyme-catalyzed phosphoryl transfer reactions. *Annual Review of Biochemistry*, 49(1):877–919, 1980.

[134] Jerry J. Zimmerman, Amélie von Saint André-von Arnim, and Jerry McLaughlin. *Chapter 74 - Cellular Respiration*, pages 1058–1072. Mosby, Saint Louis, 2011.

[135] N. Krahmer, R. V. Farese, and T. C. Walther. Balancing the fat: lipid droplets and human disease. *Embo Molecular Medicine*, 5(7):973–983, 2013.

[136] V. T. Salo and E. Ikonen. Moving out but keeping in touch: contacts between endoplasmic reticulum and lipid droplets. *Current Opinion in Cell Biology*, 57:64–70, 2019.

[137] Julie Jacquemyn, Ana Cascalho, and Rose E. Goodchild. The ins and outs of endoplasmic reticulum-controlled lipid biosynthesis. *EMBO reports*, 18(11):1905–1921, 2017.

[138] Gerrit van Meer and Anton I. P. M. de Kroon. Lipid map of the mammalian cell. *Journal of Cell Science*, 124(1):5–8, 2011.

[139] Szilard Pall and Berk Hess. A flexible algorithm for calculating pair interactions on simd architectures. *Computer Physics Communications*, 184(12):2641–2650, 2013.

[140] Loup Verlet. Computer "experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Physical Review*, 159(1):98–103, 1967.

[141] Ilario G. Tironi, René Sperb, Paul E. Smith, and Wilfred F. van Gunsteren. A generalized reaction field method for molecular dynamics simulations. *The Journal of Chemical Physics*, 102(13):5451–5459, 1995.

[142] R. W. Hockney, S. P. Goel, and J. W. Eastwood. Quiet high-resolution computer models of a plasma. *Journal of Computational Physics*, 14(2):148–158, 1974.

[143] Denys E. S. Santos, Kaline Coutinho, and Thereza A. Soares. Surface assessment via grid evaluation (suave) for every surface curvature and cavity shape. *Journal of Chemical Information and Modeling*, 62(19):4690–4701, 2022.

[144] Denys E. S. Santos, Frederico J. S. Pontes, Roberto D. Lins, Kaline Coutinho, and Thereza A. Soares. Suave: A tool for analyzing curvature-dependent properties in chemical interfaces. *Journal of Chemical Information and Modeling*, 60(2):473–484, 2020.

[145] D. P. Tieleman and S. J. Marrink. Lipids out of equilibrium: Energetics of desorption and pore mediated flip-flop. *Journal of the American Chemical Society*, 128(38):12462–12467, 2006.

[146] A. Santinho, V. T. Salo, A. Chorlay, S. Q. Li, X. Zhou, M. Omrane, E. Ikonen, and A. R. Thiam. Membrane curvature catalyzes lipid droplet assembly. *Current Biology*, 30(13):2481–2494, 2020.

[147] Yoko Shibata, Tom Shemesh, William A. Prinz, Alexander F. Palazzo, Michael M. Kozlov, and Tom A. Rapoport. Mechanisms determining the morphology of the peripheral er. *Cell*, 143(5):774–788, 2010.

[148] Matt West, Nesia Zurek, Andreas Hoenger, and Gia K. Voeltz. A 3d analysis of yeast er structure reveals how er domains are organized by membrane curvature. *Journal of Cell Biology*, 193(2):333–346, 2011.

[149] Martin Pagac, Daniel E. Cooper, Yanfei Qi, Ivan E. Lukmantara, Hoi Yin Mak, Zengying Wu, Yuan Tian, Zhonghua Liu, Mona Lei, Ximing Du, Charles Ferguson, Damian Kotevski, Pawel Sadowski, Weiqin Chen, Salome Boroda, Thurl E. Harris, George Liu, Robert G. Parton, Xun Huang, Rosalind A. Coleman, and Hongyuan Yang. Seipin regulates lipid droplet expansion and adipocyte development by modulating the activity of glycerol-3-phosphate acyltransferase. *Cell Reports*, 17(6):1546–1559, 2016.

Appendix A

To Bud or Not to Bud: A Perspective on Molecular Simulations of Lipid Droplet Budding

frontiers
in Molecular Biosciences

OPINION
published: 13 November 2019
doi: 10.3389/fmolb.2019.00124

To Bud or Not to Bud: A Perspective on Molecular Simulations of Lipid Droplet Budding

Valeria Zoni[1], *Vincent Nieto*[2], *Laura J. Endter*[3], *Herre J. Risselada*[3], *Luca Monticelli*[2]* and *Stefano Vanni*[1,4]*

[1] Department of Biology, University of Fribourg, Fribourg, Switzerland, [2] Molecular Microbiology and Structural Biochemistry, UMR 5086 CNRS, Universitè de Lyon, Lyon, France, [3] Department of Theoretical Physics, Georg-August University Göttingen, Göttingen, Germany, [4] CNRS, Institut de Pharmacologie Moléculaire et Cellulaire, Université Côte d'Azur, Valbonne, France

Keywords: lipid droplets, molecular dynamics simulations, budding, membrane asymmetry, lipid droplet proteome

Edited by:
Valentina Tozzini,
National Research Council, Italy

Reviewed by:
Giacomo Fiorin,
Temple University, United States

***Correspondence:**
Luca Monticelli
luca.monticelli@inserm.fr
Stefano Vanni
stefano.vanni@unifr.ch

Specialty section:
This article was submitted to Biological Modeling and Simulation, a section of the journal Frontiers in Molecular Biosciences

Received: *22 August 2019*
Accepted: *24 October 2019*
Published: *13 November 2019*

Citation:
Zoni V, Nieto V, Endter LJ, Risselada HJ, Monticelli L and Vanni S (2019) To Bud or Not to Bud: A Perspective on Molecular Simulations of Lipid Droplet Budding. Front. Mol. Biosci. 6:124. doi: 10.3389/fmolb.2019.00124

INTRODUCTION

Fat storage is an essential mechanism whereby cells store energy that can be later used to perform basal functions when food intake is reduced or insufficient. In cells, fat is deposited in organelles called lipid droplets (LDs). LDs are not mere inert storage pools, but they are active sites of lipid metabolism and remodeling. Furthermore, they are involved in numerous diseases, such as obesity, diabetes, cancer, and viral infection (Welte and Gould, 2017).

Despite this central role in important physiological and pathological processes, the general biology of LDs is poorly understood. This is due to the unique structure of LDs, featuring a core of neutral lipids (NLs), surrounded by a monolayer of phospholipids (PLs). As a consequence of this peculiar composition and organization, the mechanism of LD formation remains largely unclear.

The general consensus is that NLs are produced and stored between the two leaflets of the endoplasmic reticulum (ER) bilayer (**Figure 1A**); as the concentration of NLs exceeds a certain threshold, they aggregate in lenses (**Figure 1B**), that grow into nascent LDs (**Figure 1C**). Subsequently LDs bud from the ER bilayer toward the cytosol (**Figure 1D**) and, depending on the organism, they can either stay connected to the ER (**Figure 1E**) or detach in the cytosol (**Figures 1E,F**) (Wilfling et al., 2014b).

LD BUDDING: EVIDENCES AND CHALLENGES

The budding step (**Figures 1D,E**) is crucial for proper LD maturation, and it has important physiological consequences. For example, a budded LD has a higher cytosolic surface that can thus be more efficiently exposed to enzymes, such as lipases, the proteins involved in the catabolism of NLs. Also, enrichment of NLs in the ER is toxic for the cell and formation and budding of LDs might provide an effective mechanism to remove NLs from the ER bilayer (Wilfling et al., 2014b). However, the main forces and molecular actors responsible for the regulation of LD budding are still unknown. Of note, the classical machineries for vesicle budding, such as COPI and COPII, have been ruled out, since, even if COPI can bind to LDs and detach nanodroplets *in vivo* (Thiam et al., 2013), its activity affects protein targeting rather than LD budding (Wilfling et al., 2014a).

On the other hand, regulation of both ER and LD surface tension has been shown to play a crucial role in modulating LD budding (Ben M'barek et al., 2017; Chorlay and Thiam, 2018; Chorlay et al., 2019). To this end, two main mechanisms have been demonstrated to modulate LD budding *in vivo* and *in vitro* by acting on surface tension: (i) protein binding to LDs (Chorlay et al., 2019) and (ii) PL composition (Ben M'barek et al., 2017; Choudhary et al., 2018)

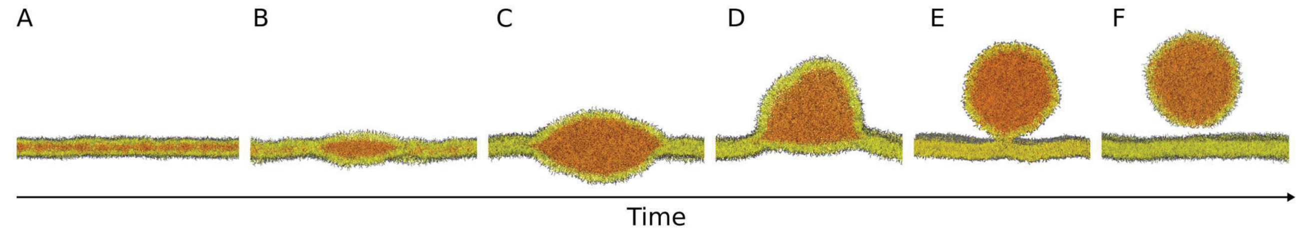


FIGURE 1 | Molecular models of the main steps of lipid droplet biogenesis. Neutral lipids are produced and stored between the two leaflets of the endoplasmic reticulum bilayer **(A)** and they aggregate in lenses **(B)** that grow into nascent lipid droplets **(C)**. Subsequently, lipid droplets bud from the bilayer **(D)** and they can either stay connected to the endoplasmic reticulum **(E)** or detach in the cytosol **(E,F)**. Images are snapshots from molecular dynamics simulations. NL: orange; PL polar heads: gray; PL acyl chains: yellow.

and abundance (Chorlay et al., 2019). For example, asymmetry in the PL coverage of the NL core has been shown to favor emergence of LDs, promoting budding toward the side with the higher number of PLs (Chorlay et al., 2019). However, potential mechanisms leading to PL asymmetry between the two ER leaflets, and specifically at sites where nascent LDs are present, are currently not well-understood. Alternatively, asymmetry can also be promoted by protein binding, whereby proteins inserting in the PL monolayer, increase NL coverage and favor budding toward the side where binding occurs (Chorlay et al., 2019). At the same time, PL composition of the ER bilayer can modulate the emergence of LDs from the ER via two distinct mechanisms: PL shape and PL-induced membrane tension. In fact, PLs with intrinsic positive curvature have been shown to favor budded states (Choudhary et al., 2018), as do PLs that are able to reduce ER tension (Ben M'barek et al., 2017).

In parallel, several proteins localized at LDs have also been shown to regulate LD budding. Two such proteins are seipin and Pex30. Pex30 is a membrane shaping protein that can tubulate the ER (Joshi et al., 2016) and that is present only transiently at LDs (Wang et al., 2018). Simultaneous deletion of Pex30 and seipin leads to an impairment in LD budding (Wang et al., 2018). Seipin is a transmembrane ER protein that forms ring-shaped homo-oligomers (Sui et al., 2018; Yan et al., 2018) that can been found stably at ER-LD contact sites (Salo et al., 2016). Cryo-EM structures (Sui et al., 2018; Yan et al., 2018) suggest that the luminal portion of seipin, by covering most of the inner LD monolayer, hinders binding of peripheral proteins toward that side. Therefore, the outer monolayer can be covered by a larger number of proteins, including possibly seipin cytosolic loops, and budding would be favored toward the cytosolic side (Chorlay et al., 2019). Furthermore, electron microscopy images reveal that LD-ER contact sites have a well-defined neck-like topology, and the size of the observed membrane neck is compatible with one ring-shaped seipin oligomer (Salo et al., 2019), suggesting that seipin is crucial to maintain this structure. At the same time, the tertiary structure of the ER domain of the protein is very similar to that of some lipid binding proteins, and it has been shown that the luminal portion can bind phosphatidic acid (PA), suggesting that it could sequester it from the bilayer and possibly present it to metabolic enzymes to form either PLs or diacylglycerols (DAGs) (Yan et al., 2018).

Another family of proteins that is necessary for LD budding is the FIT family (Choudhary et al., 2015). FITs are phosphatases that convert PA into DAG (Hayes et al., 2017; Becuwe et al., 2018), a lipid that not only presents a very low energy barrier for bilayer flip-flop, but that can be also partially stored in the middle of the bilayer, like NLs (Campomanes et al., 2019). Since FITs act only on lipids in the luminal leaflet of the ER, production of DAGs could occur asymmetrically and consequently promote LD asymmetry and budding. At the same time, the high intrinsic curvature of DAG lipids, together with the presence of several transmembrane helices in FIT proteins (Gross et al., 2010), might lead to deformation in the ER cytosolic monolayer generating positive curvature (Thiam and Forêt, 2016). The relevance of deformations in the ER bilayer for LD budding has been proposed also for other proteins that target LDs through a hairpin domain and that, consequently, can impose high positive curvature to the bilayer. An example of this class of proteins is caveolins, also found at LDs (Ostermeyer et al., 2001) and known to deform the membrane at sites of vesicle formation (Parton and Collins, 2016).

HOW CAN MOLECULAR DYNAMICS HELP UNDERSTANDING THIS PROCESS?

From the evidence in the literature, it appears that a combination of protein activity together with changes in membrane properties (such as surface tension, lipid composition, and surface coverage) is key in controlling LD budding. However, several aspects of this process are difficult to address with state-of-the-art experimental methods. Most notably, a detailed characterization of the molecular structures along the budding pathway remains unaddressed and difficult to achieve using current structural biology methods, due to the liquid nature of lipid aggregates, the small size of early-stage nascent LDs (well below optical resolution), and the transient nature of budding intermediates.

Molecular dynamics (MD) simulations are optimally suited to investigate the structural and dynamic properties of liquids, and they are particularly promising for the study of molecular mechanisms underlying LD budding (Soares et al., 2017). Notably, MD simulations have already been successfully applied to interpret and corroborate several experimental findings. For instance, MD simulations clarified how changes in bilayer surface tension alter the concentration of NLs stored in a LD lens (Ben M'barek et al., 2017). Also, MD simulations showed that asymmetry in monolayer coverage (hence asymmetry in surface tension) is able to control budding directionality independently of the lipid spontaneous curvature (Chorlay et al., 2019).

However, several questions remain on the mechanism and the energetics of budding, as well as on the role of different proteins in the process; we foresee that MD simulations will be instrumental in addressing such questions. First of all, MD simulations can be used to explore the structural role of PLs and how the distribution of different lipids influences budding. In particular it will be interesting to understand the role of PLs, such as phosphatidic acid, lysolipids, and DAG, during all the stages of LD growth and budding, since they seem to largely influence budding and protein recruitment (Ben M'barek et al., 2017; Choudhary et al., 2018). Second, MD simulations may help elucidating the energetic requirements associated with various steps of the budding process (depicted in **Figure 1**). Theoretical studies of LD budding suggest that, in order to achieve LD fission, the NL phase should completely dewed from either the inner or the outer leaflet of the ER, a mechanism that requires external energy, possibly controlled by surface tension (Thiam and Forêt, 2016). We envisage that MD simulations may allow detailed predictions on the energetics of LD budding under different and controlled conditions, therefore clarifying which of the proposed budding stages are spontaneous and which ones require external energy. Third, for those steps requiring external energy input, MD simulations will enable predictions of the molecular mechanisms by which proteins regulate LD budding. For example, how Pex30 and seipin promote concertedly budding is not understood. While it has been shown that seipin imposes a distinct topology to LD-ER contact sites (Salo et al., 2019) it remains unclear if, in order to achieve a fully budded state with a well-defined neck (**Figure 1E**), the LD lens needs to reach a certain size or if this topology is already stable in the early stages of LD formation (Deslandes et al., 2017).

More generally, open questions remain on the relevance of protein-induced membrane deformations in LD budding as well as on the influence of LD-binding proteins, and MD simulations can greatly contribute to address such questions, particularly as high-resolution structures of the proteins involved become available. Overall, MD simulations can help unveiling which morphologies are more energetically favorable for lipid aggregates with different compositions (e.g., different concentrations of NLs), and which transformations are more likely.

Finally, even though the mechanism of LD formation and budding showed in **Figure 1** is generally accepted, whether the final step of the process actually happens *in vivo* remains controversial. Of particular concern, no fission machinery leading to LD detachment from the ER has been identified so far, and it is unclear whether LD detachment could be promoted by membrane physical properties alone. MD simulations should be able to provide an estimate of the energetic cost of breaking the LD-ER neck and to clarify whether the process is driven only by surface tension or if protein activity is necessary to detach LDs from the ER.

CONCLUSIONS

In this Opinion, we illustrate the main unanswered questions regarding LD budding that can be investigated using MD simulations. One of the challenges of simulating such systems is their computational cost, since LDs have diameters of hundreds of nanometers and their growth takes place on time scales of seconds (Salo et al., 2019). The employment of chemical-specific coarse-grained models, such as MARTINI (Marrink et al., 2007; Monticelli et al., 2008) and SDK (Shinoda et al., 2007), has recently allowed simulating some aspects of LD budding using realistic sizes and timescales. However, simulations representing the complexity of LD formation (that involves multiple lipid species and proteins throughout the process) might be beyond the current capabilities and accuracy of available CG models. Equilibrium CG simulations might not be sufficient to explore the key aspects of LD budding, and enhanced-sampling strategies might be required. Thus, even though pioneering simulations have started highlighting important aspects of LD biology (Khandelia et al., 2010; Bacle et al., 2017; Ben M'barek et al., 2017; Vanni, 2017; Pezeshkian et al., 2018; Chorlay et al., 2019; Zoni et al., 2019), we foresee that further developments in molecular modeling techniques will be required to advance our understanding of the mechanisms of LD biogenesis.

AUTHOR CONTRIBUTIONS

VZ and SV wrote the article with help from VN, LE, HR, and LM.

ACKNOWLEDGMENTS

VZ and SV acknowledge support from the Swiss National Science Foundation grant #163966 and from the Swiss National Supercomputing Centre (CSCS) under project ID s726 and s842. VZ and SV acknowledge PRACE for awarding us access to Piz Daint, ETH Zurich/CSCS, Switzerland. VN and LM acknowledge funding from the Agence Nationale de la Recherche (grant ANR-17-CE11-0003-01) and Grand Èquipement National de Calcul Intensif (GENCI, grant A0060710138). LM acknowledges funding from the Institut National de la Sante et de la Recherche Medicale (INSERM).

REFERENCES

Bacle, A., Gautier, R., Jackson, C. L., Fuchs, P. F. J., and Vanni, S. (2017). Interdigitation between triglycerides and lipids modulates surface properties of lipid droplets. *Biophys. J.* 112, 1417–1430. doi: 10.1016/j.bpj.2017.02.032

Becuwe, M., Bond, L. M., Mejhert, N., Boland, S., Elliott, S. D., Cicconet, M., et al. (2018). FIT2 is a lipid phosphate phosphatase crucial for endoplasmic reticulum homeostasis. *bioRxiv* 291765. doi: 10.1101/291765

Ben M'barek, K., Ajjaji, D., Chorlay, A., Vanni, S., Forêt, L., and Thiam, A. R. (2017). ER membrane phospholipids and surface tension control cellular lipid droplet formation. *Dev. Cell* 41, 591–604.e7. doi: 10.1016/j.devcel.2017.05.012

Campomanes, P., Zoni, V., and Vanni, S. (2019). Local accumulation of diacylglycerol alters membrane properties nonlinearly due to its transbilayer activity. *Commun. Chem.* 2:72. doi: 10.1038/s42004-019-0175-7

Chorlay, A., Monticelli, L., Veríssimo Ferreira, J., Ben M'barek, K., Ajjaji, D., Wang, S., et al. (2019). Membrane asymmetry imposes directionality on lipid droplet emergence from the ER. *Dev. Cell* 50, 25–42.e7. doi: 10.1016/j.devcel.2019.05.003

Chorlay, A., and Thiam, A. R. (2018). An asymmetry in monolayer tension regulates lipid droplet budding direction. *Biophys. J.* 114, 631–640. doi: 10.1016/j.bpj.2017.12.014

Choudhary, V., Golani, G., Joshi, A. S., Cottier, S., Schneiter, R., Prinz, W. A., et al. (2018). Architecture of lipid droplets in endoplasmic reticulum is determined by phospholipid intrinsic curvature. *Curr. Biol.* 28, 915–926.e9. doi: 10.1016/j.cub.2018.02.020

Choudhary, V., Ojha, N., Golden, A., and Prinz, W. A. (2015). A conserved family of proteins facilitates nascent lipid droplet budding from the ER. *J. Cell Biol.* 211, 261–71. doi: 10.1083/jcb.201505067

Deslandes, F., Thiam, A. R., and Forêt, L. (2017). Lipid droplets can spontaneously bud off from a symmetric bilayer. *Biophys. J.* 113, 15–18. doi: 10.1016/j.bpj.2017.05.045

Gross, D. A., Snapp, E. L., and Silver, D. L. (2010). Structural insights into triglyceride storage mediated by fat storage-inducing transmembrane (FIT) protein 2. *PLoS ONE* 5:e10796. doi: 10.1371/journal.pone.0010796

Hayes, M., Choudhary, V., Ojha, N., Shin, J. J., Han, G. S., Carman, G. M., et al. (2017). Fat storage-inducing transmembrane (FIT or FITM) proteins are related to lipid phosphatase/phosphotransferase enzymes. *Microb. Cell.* 5, 88–103. doi: 10.15698/mic2018.02.614

Joshi, A. S., Huang, X., Choudhary, V., Levine, T. P., Hu, J., and Prinz, W. A. (2016). A family of membrane-shaping proteins at ER subdomains regulates pre-peroxisomal vesicle biogenesis. *J. Cell Biol.* 215, 515–529. doi: 10.1083/jcb.201602064

Khandelia, H., Duelund, L., Pakkanen, K. I., and Ipsen, J. H. (2010). Triglyceride blisters in lipid bilayers: implications for lipid droplet biogenesis and the mobile lipid signal in cancer cell membranes. *PLoS ONE* 5:e12811. doi: 10.1371/journal.pone.0012811

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

Monticelli, L., Kandasamy, S. K., Periole, X., Larson, R. G., Tieleman, D. P., and Marrink, S. J. (2008). The MARTINI coarse-grained force field: extension to proteins. *J. Chem. Theory Comput.* 4, 819–834. doi: 10.1021/ct700324x

Ostermeyer, A. G., Paci, J. M., Zeng, Y., Lublin, D. M., Munro, S., and Brown, D. A. (2001). Accumulation of caveolin in the endoplasmic reticulum redirects the protein to lipid storage droplets. *J. Cell Biol.* 152, 1071–1078. doi: 10.1083/jcb.152.5.1071

Parton, R. G., and Collins, B. M. (2016). Unraveling the architecture of caveolae. *Proc. Natl. Acad. Sci. U.S.A.* 113, 14170–14172. doi: 10.1073/pnas.1617954113

Pezeshkian, W., Chevrot, G., and Khandelia, H. (2018). The role of caveolin-1 in lipid droplets and their biogenesis. *Chem. Phys. Lipids* 211, 93–99. doi: 10.1016/j.chemphyslip.2017.11.010

Salo, V. T., Belevich, I., Li, S., Karhinen, L., Vihinen, H., Vigouroux, C., et al. (2016). Seipin regulates ER –lipid droplet contacts and cargo delivery. *EMBO J.* 35, 2699–2716. doi: 10.15252/embj.201695170

Salo, V. T., Li, S., Vihinen, H., Hölttä-Vuori, M., Szkalisity, A., Horvath, P., et al. (2019). Seipin facilitates triglyceride flow to lipid droplet and counteracts droplet ripening via endoplasmic reticulum contact. *Dev. Cell.* 50, 1–16. doi: 10.1016/j.devcel.2019.05.016

Shinoda, W., DeVane, R., and Klein, M. L. (2007). Multi-property fitting and parameterization of a coarse grained model for aqueous surfactants. *Mol. Simul.* 33, 27–36. doi: 10.1080/08927020601054050

Soares, T. A., Vanni, S., Milano, G., and Cascella, M. (2017). Toward chemically resolved computer simulations of dynamics and remodeling of biological membranes. *J. Phys. Chem. Lett.* 8, 3586–3594. doi: 10.1021/acs.jpclett.7b00493

Sui, X., Arlt, H., Brock, K. P., Lai, Z. W., DiMaio, F., Marks, D. S., et al. (2018). Cryo–electron microscopy structure of the lipid droplet–formation protein seipin. *J. Cell Biol.* 217, 4080–4091. doi: 10.1083/jcb.201809067

Thiam, A. R., Antonny, B., Wang, J., Delacotte, J., Wilfling, F., Walther, T. C., et al. (2013). COPI buds 60-nm lipid droplets from reconstituted water-phospholipid- triacylglyceride interfaces, suggesting a tension clamp function. *Proc. Natl. Acad. Sci. U.S.A.* 110, 13244–13249. doi: 10.1073/pnas.1307685110

Thiam, A. R., and Forêt, L. (2016). The physics of lipid droplet nucleation, growth and budding. *Biochim. Biophys. Acta* 1861, 715–722. doi: 10.1016/j.bbalip.2016.04.018

Vanni, S. (2017). Intracellular lipid droplets: from structure to function. *Lipid Insights* 10, 1–3. doi: 10.1177/1178635317745518

Wang, S., Idrissi, F. Z., Hermansson, M., Grippa, A., Ejsing, C. S., and Carvalho, P. (2018). Seipin and the membrane-shaping protein Pex30 cooperate in organelle budding from the endoplasmic reticulum. *Nat. Commun.* 9:2939. doi: 10.1038/s41467-018-05278-2

Welte, M. A., and Gould, A. P. (2017). Lipid droplet functions beyond energy storage. *Biochim. Biophys. Acta* 1862, 1260–1272. doi: 10.1016/j.bbalip.2017.07.006

Wilfling, F., Haas, J. T., Walther, T. C., and Farese, R. V. (2014b). Lipid droplet biogenesis. *Curr. Opin. Cell Biol.* 29, 39–45. doi: 10.1016/j.ceb.2014.03.008

Wilfling, F., Thiam, A. R., Olarte, M. J., Wang, J., Beck, R., Gould, T. J., et al. (2014a). Arf1/COPI machinery acts directly on lipid droplets and enables their connection to the ER for protein targeting. *Elife* 2014:e01607. doi: 10.7554/eLife.01607

Yan, R., Qian, H., Lukmantara, I., Gao, M., Du, X., Yan, N., et al. (2018). Human SEIPIN binds anionic phospholipids. *Dev. Cell* 47, 248–256.e4. doi: 10.1016/j.devcel.2018.09.010

Zoni, V., Campomanes, P., Khaddaj, R., Thiam, A. R., Schneiter, R., and Vanni, S. (2019). Lipid droplet biogenesis is a liquid phase separation spatially regulated by seipin and membrane curvature. *bioRxiv* 777466. doi: 10.1101/777466

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix B

Triacylglycerols sequester monotopic membrane proteins to lipid droplets

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-020-17585-8 OPEN

Triacylglycerols sequester monotopic membrane proteins to lipid droplets

Lucie Caillon[1], Vincent Nieto[2], Pauline Gehan[3,4], Mohyeddine Omrane[1], Nicolas Rodriguez[3,4], Luca Monticelli[2] & Abdou Rachid Thiam[1]✉

Triacylglycerols (TG) are synthesized at the endoplasmic reticulum (ER) bilayer and packaged into organelles called lipid droplets (LDs). LDs are covered by a single phospholipid monolayer contiguous with the ER bilayer. This connection is used by several monotopic integral membrane proteins, with hydrophobic membrane association domains (HDs), to diffuse between the organelles. However, how proteins partition between ER and LDs is not understood. Here, we employed synthetic model systems and found that HD-containing proteins strongly prefer monolayers and returning to the bilayer is unfavorable. This preference for monolayers is due to a higher affinity of HDs for TG over membrane phospholipids. Protein distribution is regulated by PC/PE ratio via alterations in monolayer packing and HD-TG interaction. Thus, HD-containing proteins appear to non-specifically accumulate to the LD surface. In cells, protein editing mechanisms at the ER membrane would be necessary to prevent unspecific relocation of HD-containing proteins to LDs.

---

[1] Laboratoire de Physique de l'École Normale Supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, F-75005 Paris, France. [2] University of Lyon, CNRS, Molecular Microbiology and Structural Biochemistry (MMSB, UMR 5086), F-69007 Lyon, France. [3] Sorbonne Universités, UPMC Univ Paris 06, Ecole Normale Supérieure, CNRS, Laboratoire des Biomolécules (LBM), 4 place Jussieu, 75005 Paris, France. [4] Département de Chimie, Ecole Normale Supérieure, PSL Research University, UPMC Univ Paris 06, CNRS, Laboratoire des Biomolécules (LBM), Paris, France. ✉email: thiam@ens.fr

Lipid droplets (LDs) are lipid storage organelles primarily functioning in cellular energy metabolism[1]. LD biogenesis occurs at the endoplasmic reticulum (ER) membrane during energy rich or stress conditions. LD biogenesis starts with the synthesis of neutral lipids, such as triacylglycerols (TG) or sterol esters, which, at low concentration, are dissolved in the ER bilayer[2]. Upon increase in concentration, neutral lipids demix from membrane phospholipids to form an oil lens or a nascent droplet within the bilayer[3] (Fig. 1a). The lens grows and emerges in the cytosol as a mature LD: an oil-in-water droplet covered by a phospholipid monolayer with proteins embedded. Indeed, throughout the steps of LD emergence, many proteins target to the surface and around the LD[4–7]. Proteins targeting the LD surface essentially come from the ER membrane or from the cytosol[7], and ensure proper LD budding[8]. How proteins bind and accumulate to LDs is not well understood but the neutral lipid chemistry is determinant to these processes[9]. Specificity of protein targeting to LDs is at the heart of LD biology, and understanding its principles will provide fundamental knowledge on lipid metabolism and cellular proteostasis[5–7,10,11].

Membrane physicochemical properties regulate the protein distribution at bilayer-encircled organelles[12–14]. The LD-water interface is distinguishable from a bilayer-water interface by several features: it can sustain a loose lipid packing[9,15,16]; the thickness of the underlying hydrophobic region, up to hundreds nm, is much larger than the hydrophobic thickness of a bilayer (~3 nm)[17,18]; the hydrophobic core consists of neutral lipids, instead of phospholipid acyl chains. Considering these discrepancies in physical chemistry, it may not be surprising that proteins show preference for one interface over the other.

Most proteins physically associating with LD surfaces are either peripheral or monotopic[6,7,11] and do not fully cross bilayer membranes. Proteins moving from the ER to LD surface, contain helical hydrophobic domains (HDs), which are monotopic integral membrane domains embedded only in one face of the membrane. These HDs include helical hairpins, hydrophobic helices, and possibly transmembrane domains not fully crossing a bilayer[11,19–21]. In contrast, soluble proteins often use amphipathic helices (AHs) for binding to LDs.

The binding of AHs to LDs is more documented both in vitro and in vivo[9,16,19,22–24]: AHs act as surfactants, favorably adsorbed to the oil/water interface of LDs to decrease the interfacial energy. AHs recognizes a variety of membrane features, such as surface charges, curvature, phospholipid packing defects, and neutral lipids[9,16,19,22,23]. In contrast, much less is known about HDs which target to LDs mostly from the ER membrane through ER-LD connecting bridges[25–27]. Neither the energetics involved in their binding to LDs nor the parameters controlling their ER-to-LD partitioning are known.

The inclusion of HD-containing proteins into lipid bilayers can cause local perturbation to the bilayer properties, which translates into an energy penalty[28–31]. For instance, proteins can locally perturb the organization of the phospholipids and enhance exposure of the hydrophobic core of the bilayer to water[28–32]. The extent of membrane perturbation depends on the amino acid sequence, and is for instance important when the mismatch between the bilayer thickness and the HD length is significant[31,32]. As for protein insertion into LD surfaces, no information is available regarding the energy cost of the process, nor the type and the extent of the perturbation generated in the surrounding lipids.

Here, we study how LD proteins, and particularly monotopic HD-containing proteins, partition between a bilayer and an LD in contiguity. We employ the droplet interface bilayer (DIB) system[33] (Fig. 1a) to study the partitioning of proteins and peptides bearing HDs, as compared with AH-containing proteins. We find that all proteins investigated partition preferentially to the LD monolayer surface, but HD-containing proteins display a higher

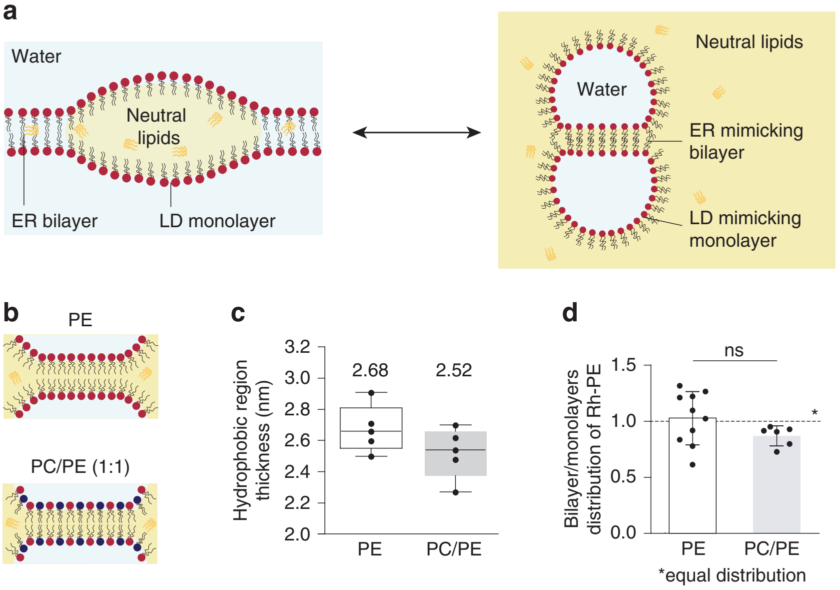


Fig. 1 Characterization of droplet interface bilayers. a Schematic representation of the ER phospholipid bilayer contiguous with the monolayer of a nascent LD (left side); the corresponding DIB system reproducing contiguous bilayer and monolayers is shown on the right side. The water phase is represented in light blue and the oil phase in yellow (neutral lipid, e.g. triglycerides (TG)). **b** Drawings of a DIB bilayer of DOPE (top) and DOPC/DOPE (1:1) (bottom). **c** The thickness of the hydrophobic region of the DIB bilayer in DOPE (white) and DOPC/DOPE (1:1) (gray) is determined by capacitance measurement. Results are shown as box-plots (box limits, upper and lower quartiles; middle line, median; whiskers, minimum and maximum value; the mean is indicated) from $n = 5$ independent experiments. Each point is represented as a black dot. **d** Distribution of Rh-PE between the bilayer and the monolayers in DOPE (white) and DOPC/DOPE (1:1) (gray) DIBs. The results are the mean ± SD of respectively $n = 10$ and $n = 5$ independent measurements. Each point is represented as a black dot. Significance was determined by Welch's *t*-test (unpaired parametric test, two-tailed *p*-value) and is indicated by ns (not significant): $p > 0.05$. Source data are provided as a Source Data file.

enrichment in the monolayer than AH-containing ones. Relocation of HD proteins to the bilayer is unfavorable, while moving from the bilayer to the monolayer is spontaneous. We also found that protein distribution is altered by the ratio between PC and PE phospholipids by regulating the extent of HD-TG contact at the LD surface.

Results

Characterization of the droplet interface bilayer system. To determine the partition coefficient of proteins capable of binding a monolayer and a bilayer in contiguity, we decided to employ the droplet interface bilayer (DIB) system[33,34]. DIBs consist of two micrometric buffer-in-oil droplets covered by a phospholipid monolayer (Fig. 1a). The oil phase used here was trioctanoate, a triglyceride with similar interfacial energy as triolein[8], the major cellular neutral lipid. Contact of the droplets induces the formation of a bilayer in contiguity with the two monolayers (Fig. 1a). Thus DIBs mimic ER-LD contiguity (Fig. 1a) without curvature considerations; the different interfaces are flat at the protein scale and the concavity of the monolayer surfaces is irrelevant with respect to curvature. For phospholipids, we used dioleoyl phosphatidylethanolamine (termed PE) and dioleoyl phosphatidylcholine (termed PC) (Fig. 1b). Phospholipids were added to the oil phase and were recruited to the surface of the aqueous droplets whose contact generates within 5 min an equilibrated DIB[34,35].

DIBs can be generated with almost any phospholipids[35]. In the case of non-bilayer phospholipids, such as DOPE, a PE-DIB bilayer is made thanks to the presence in the bilayer of TG molecules whose level is decreased by the addition of PC[35]. To get insight into the amount of TG present in a PE-DIB bilayer, we measured the thickness of the hydrophobic region of the bilayer by capacitance measurements[36] (Fig. 1c, Supplementary Fig. S1a). The thickness measured in PE-DIBs was 2.68 nm, only ~7% above the thickness of a PC/PE (1:1) DIB, 2.52 nm (Fig. 1c). Importantly, these values are comparable to the thickness of the hydrophobic region in phospholipid vesicles devoid of oil, between 2.3–2.7 nm[18]. Additionally, all-atom molecular dynamics simulations indicate that adding PE to a PC bilayer devoid of oil is sufficient to increase bilayer thickness up to 10% (Supplementary Fig. S1b). Altogether, these data indicate that the thickness of the DIBs made here is similar to that of phospholipid bilayer vesicles and is not significantly affected by the presence of oil.

Since the PC/PE mixture was added to the oil phase, we wanted to know whether this bulk ratio reflects the monolayer composition. We had previously measured the surface tension of monolayers made of PC/PE and found a linear decrease as this ratio increased in bulk oil[35] (from ~2 mN m^{-1} for PE at 100% to ~0.6 mN m^{-1} for 100% PC). This supports that the bulk PC/PE composition reflects the one at the monolayer, as otherwise a plateau of surface tension against PC/PE should be observed. We next asked whether the PC/PE ratio in the monolayer and in the DIB bilayer are identical. To address this, we measured the partitioning of Rhodamine-PE (Rh-PE) between the DIB monolayers and bilayer, in the case of PE and PC/PE (1:1) DIBs. In pure PE-DIB, Rh-PE was uniformly distributed, indicating that the distributions of Rh-PE and PE are similar. In PC/PE, Rh-PE was also almost uniformly distributed, suggesting that the monolayer and the bilayer have a similar PC/PE composition (Fig. 1d, Supplementary Fig. S1c). To confirm this finding, we investigated lipid distribution in model nascent LDs using molecular dynamics simulations, in three systems containing TG and (a) pure DOPC or DOPC/DOPE mixtures, (b) 80/20 and (c) 60/40. We found that DOPE and DOPC mix ideally and their distribution was approximately homogeneous (Supplementary Fig. S1d). DOPC was only slightly enriched in the monolayer compared to the bilayer, while DOPE was slightly enriched in the bilayer compared to the monolayer—the differences being minor in both cases (Supplementary Fig. S1d). Overall, the data confirm that PE/PC mixtures are ideal mixtures, with an approximately even distribution of both lipids between the monolayer and bilayer interfaces.

The above characterizations indicate that DIBs recapitulate sufficiently well conditions of a bilayer containing an oil droplet, as previously shown[35]. We subsequently use DIBs to study protein partitioning.

Monotopic proteins strongly bind to TG-covering monolayers. We screened the monolayer-bilayer partitioning of two classes of proteins or peptides: soluble proteins, targeting to LDs from the cytosol, and monotopic integral membrane proteins (moving from the ER bilayer to LD surface. Soluble peptides were directly added to the buffer droplets. Monotopic membrane proteins were added to the buffer droplets from purified LDs or from proteoliposomes (Fig. 2a); mixing relocalized the proteins from LDs, or proteoliposomes, to the interface between the buffer droplet and the oil phase. Phospholipids from liposomes or LDs also relocalized to this new interface, but their total amount was always much less than the amount of the exogenous phospholipids we added; the latter would control the interfacial lipid composition in all of our systems. In practice, buffer-in-oil droplets containing the proteins at the interface were prepared before adding phospholipids to the oil phase (Fig. 2a). Two droplets were then brought together to form a DIB. The protein partition coefficient was determined 10 min after contact, at equilibrium, by quantifying the enrichment level of the protein in the bilayer relative to the monolayers (Fig. 2b).

For monotopic membrane proteins, we tested Plin1, ACSL3, CG2254, CG9186, oleosin 1, Hpos, and caveolin1 (Fig. 2d, Supplementary Fig. S2a, c), most of which contain helical hairpin and hydrophobic or amphipathic helix motifs responsible for their localization to LDs[37–40]. These proteins were tagged with fluorescent proteins and expressed in cells that were subsequently loaded with oleate to induce LDs. LDs bound by the proteins were purified and added to the DIB system. One limitation of this approach is that other proteins contained in the LDs would also relocalize to the DIB interfaces, although not visible. Furthermore, proteins with single transmembrane domains, not fully crossing the ER bilayer, could target to the LD surface, but this has never been shown clearly so far. To test this hypothesis, we prepared proteoliposomes containing some of the SNARE components bearing a transmembrane helix, but not crossing the bilayer. Finally, we studied a group of soluble proteins, including Plin2–3, Plin1 AH-containing domains[19,24], and the lipid packing sensors ArfGap1-AH[41] and GMAP-210-AH[42] (Fig. 2c, Supplementary Fig. S2c, d).

For all of the tested proteins, we found a stronger partitioning to the monolayers than to the bilayer, independently of PC/PE ratio (Fig. 2e, f, Supplementary Fig. S2e, f). Additionally, HD-containing proteins showed on average a higher LD enrichment than AH proteins (Fig. 2e, Supplementary Fig. S2f), supporting that proteins coming from the ER bilayer better associate with LDs than soluble proteins. For a subgroup of HD proteins, we measured the partitioning in both PC/PE (1:1) and a more biologically relevant composition (DOPE/DOPC/liverPI/cholesterol, 5:3:1:1), and found very similar results (Supplementary Fig. S2a, b).

Finally, the bilayer localization of AH-containing proteins was increased by addition of PE in most cases, but partitioning to the

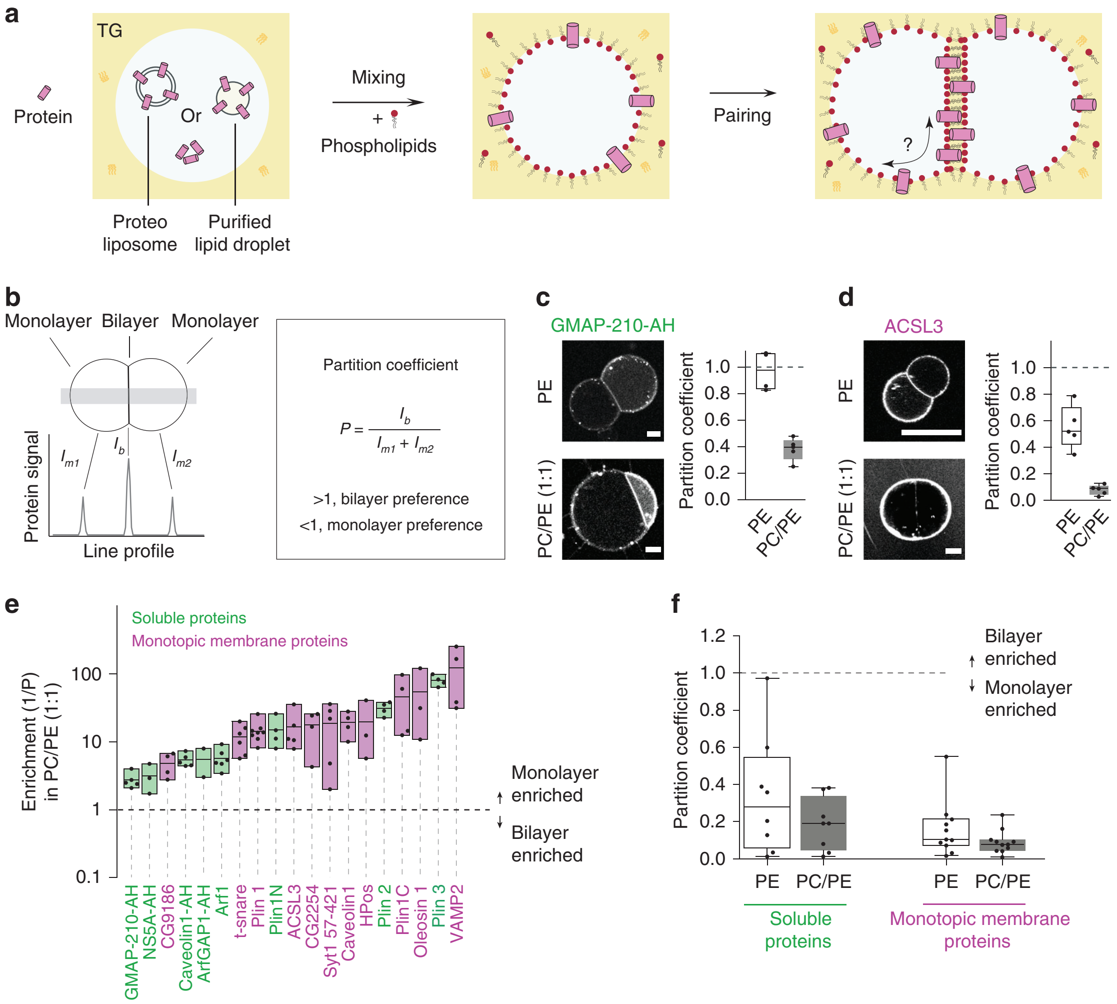


Fig. 2 Partitioning of hydrophobic and amphipathic helix-containing proteins to droplet interface bilayers. a Formation of protein-containing DIBs: soluble protein, proteoliposome, or purified LDs are added to the buffer droplet (left). Mixing the droplet in a TG-containing phospholipids solution allows the relocalization of proteins to the buffer droplet interface (middle). When two drops come close together, their monolayers zip to form a bilayer. The proteins are thus offered to relocalize to the bilayer (right). **b** Protein distribution between the bilayer and the monolayer is determined by the partition coefficient P, which is the protein signal at the bilayer divided by sum of its signals at the monolayers. When a protein partitions preferentially to the bilayer, $p > 1$; if preference is for the monolayer then $p < 1$. **c**, **d** Distribution of GMAP-210-AH (soluble protein) and ACSL3 (monotopic membrane protein) respectively, in DOPE or DOPC/DOPE (1:1) DIBs. Scale bar: 20 µm. The partition coefficient is represented for each condition as box-plots from $n = 5$ independent measurements (excepted for GMAP-210-AH in PE, $n = 4$). **e** Enrichment parameter in DOPC/DOPE (1:1) membranes for AH- (green) and HD- (pink) containing proteins, shown as floating bars (bar limits, min to max values; central line, mean), $2 \leq n \leq 8$ independent measurements were done for each protein. HD-containing proteins coming from LDs are Plin 1, Plin 1C, ACSL3, CG2254, CG9186, Oleosin 1, Caveolin 1, HPos; those coming from proteoliposomes are Syt1 57-421, t-snare, Vamp2; AH-containing proteins coming from LDs are Plin3, Plin2, Plin1N; the other AHs are added soluble. **f** Average partition coefficient of the groups of soluble or monotopic membrane proteins in DOPE or DOPC/DOPE (1:1). Results presented in (Supplementary Fig. S2e) were used to generate (**f**). Box-plots are defined as follow: box limits, upper and lower quartiles; middle line, median; whiskers, minimum and maximum value. In **c**, **d**, **e**, **f** each point is represented as a black dot. Source data are provided as a Source Data file.

monolayer region was still more favorable (Fig. 2f, Supplementary Fig. S2e, f). The negative spontaneous curvature of PE is known to cause lipid packing defects, which can be sensed by AHs[43]. This is well illustrated by the highest partition coefficient (close to 1) obtained in PE with the AH domains of GMAP-210 and ArfGap1, which are lipid packing sensors[43] (Fig. 2c, Supplementary Fig. S2e). In contrast to AHs, the dependence of HD-containing protein on PC/PE was less clear (Supplementary Fig. S2e, f).

In summary, both AH- and HD-containing proteins localized preferentially to the monolayer interface over the bilayer. HD-proteins more strongly partitioned to the monolayer and barely relocated to the bilayer.

KWALP peptides recap the global behavior of HD proteins. At this stage, it is difficult to explain the partitioning trend of the full-length HD proteins. This is in part because most of the proteins, coming from purified LDs, may interact with other unidentified proteins in the system. Also, LD proteins can bear multiple HDs and/or AHs; this is the case for Oleosin1, Caveolin1, HPos, and ACSL3[37,38,40], which possess an AH motif adjacent to their HD motif. To better understand the determinants of partitioning for pure HD domains, we focused on model peptides of the KWALP family. KWALP peptides consist of a repeated leucine-alanine motif (Fig. 3a), bounded by two tryptophan residues at the C-terminus and three lysine residues at the N-terminus; to this N-terminus we added a glycine linked to a

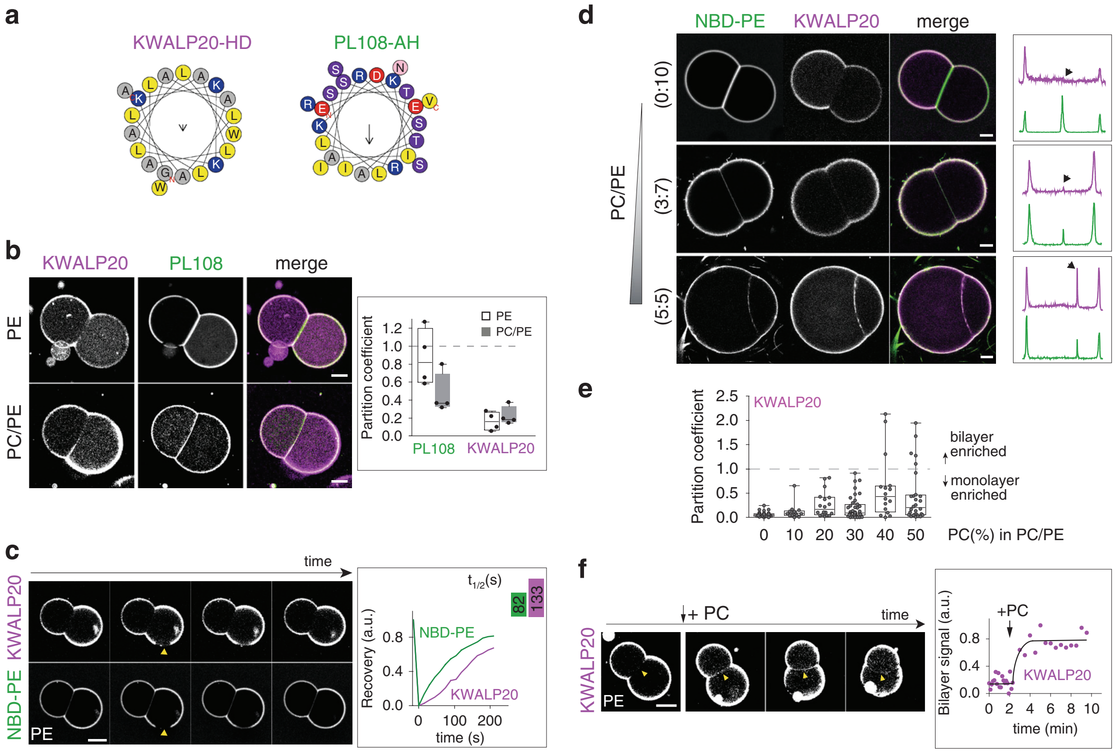


Fig. 3 PC/PE ratio modulates the partitioning of model AH and HD peptides. **a** Helical wheel representation of KWALP20 and PL108-AH, generated from HeliQuest[69]. **b** Distribution of KWALP20 and PL108 in DOPE or DOPC/DOPE (1:1) DIBs. KWALP20 is labeled with and PL108 with NBD. Scale bar: 50 µm. The partition coefficient is plotted for both peptides as box-plots (box limits, upper and lower quartiles; middle line, median; whiskers, minimum and maximum value), from $n = 4$ independent measurements for each condition. Individual data are shown as black dots. **c** FRAP experiment shows that KWALP20 (purple) and lipid (green) signals are mobile at the monolayer. NBD-PE reports for phospholipids. Yellow arrows indicate the area bleached. Scale bar: 50 µm. Recovery half-time was obtained using one-phase association fitting in GraphPad software and is shown in the upper right box. **d** KWALP20 distribution in DIBs of different PC/PE ratio. Line profiles (not displayed) are drawn perpendicular to the bilayer and monolayers (as described in Fig. 2b); the thickness of the line is 30-40% of the bilayer size. The corresponding signals are shown in the box (right); black arrows depict the bilayer signal. Scale bar: 20 µm. **e** Partition coefficient of KWALP20 in DIBs of different DOPC/DOPE ratios represented as box-plots (box limits, upper and lower quartiles; middle line, median; whiskers, minimum and maximum value). Sample size was $n = 31$ for 0% PC, $n = 14$ for 10% PC, $n = 18$ for 20% PC, $n = 41$ for 30% PC, $n = 16$ for 40% PC and $n = 28$ for 50% PC. Each data point is plotted. **f** Relocalization of KWALP20 from the monolayer to the bilayer after addition of DOPC to DOPE DIBs. The bilayer signal is plotted over time. Image brightness is enhanced to improve bilayer viewing. Scale bar: 50 µm. Source data are provided as a Source Data file.

rhodamine-B dye. This peptide features a strong tendency towards helical conformation and transmembrane partitioning, as reviewed from numerous previous studies[28]; therefore they represent a valid model for transmembrane domains of proteins, including those localizing to the Golgi and plasma membranes[17,28,44]. Moreover, KWALP represents an excellent model for the minimal basic hydrophobic sequences, commonly found in HD domains of LD proteins (Supplementary Fig. S3). We used KWALP20, with 16 hydrophobic amino acids (Fig. 3a) and a length (when folded in an α-helix) close to the ER bilayer thickness (requiring ~ 20 hydrophobic amino acids[28]). For comparison, we also studied the partitioning of an AH motif derived from the 11-mer repeat of Perilipin1—termed here PL108 (Fig. 3a)[24]. KWALP20 and PL108 represent useful models for the two classes of HD and AH proteins tested above.

We prepared DIBs containing both KWALP20 and PL108 to compare their distribution under identical experimental conditions. When the DIB contained PE exclusively, PL108 partitioned almost equally between the monolayer and the bilayer (Fig. 3b), while KWALP20 was surprisingly absent from the bilayer (Fig. 3b). When PE/PC (1:1) was used, KWALP20 signal in the bilayer increased, while PL108 bilayer concentration significantly decreased (Fig. 3b). These observations are consistent with the behavior of most HD- and AH-containing proteins (Fig. 2e, Supplementary Fig. S2f): both peptides partition more favorably to the monolayer, especially for KWALP20; in PE lipids, PL108 partitions more evenly, like GM210-AH, and KWALP20 is barely detectable to the bilayer, like Oleosin1, Caveolin1 or Hpos (Supplementary Fig. S2c).

Since KWALP recapitulated the partitioning of most of the full-length HD-proteins (Fig. 2e, f, Supplementary Fig. S2c, e, f), we further investigated the driving forces for its distribution to establish general principles underpinning the enrichment of HDs to LD surface.

PC/PE ratio regulates the partitioning of KWALP. KWALP20 was almost totally absent from the bilayer in PE, while it was well folded in the monolayer (Supplementary Fig. S4a). Importantly, the peptide was laterally mobile, as shown by the rapid recovery of fluorescence following photobleaching (Fig. 3c, Supplementary Fig. S4c); this recovery was indeed due to in-plane diffusion, because recovery from bulk did not happen within 10 minutes (Supplementary Fig. S4b). Since the hydrophobic thickness of the PE DIB bilayer is comparable to the peptide length, it is unlikely that KWALP20 localization to the bilayer was prevented by hydrophobic mismatch. We next increased further the PC/PE ratio, and observed that the concentration of the peptide in the bilayer increased with PC level, but it still remained lower than in the monolayer (Fig. 3d, e). In the bilayer, the KWALP peptide was also mobile but showed a significant tendency to cluster as the PC level was increased (Fig. 3f, Supplementary Fig. S4f). To verify if the bilayer localization was dynamic, and if diffusion was not prevented by peptide aggregation, we followed the peptide signal while changing the phospholipid composition. Starting from a pure PE DIB, where KWALP20 was absent from the bilayer, we added a TG solution containing PC to the oil phase surrounding the droplets (Fig. 3f). The recruitment of PC to the interface of the droplets was demonstrated by the increase in the contact angle between the droplets[35] (Supplementary Fig. S4e) and it was concomitant with an increase of KWALP20 signal in the bilayer (Fig. 3f). We also noticed the appearance of KWALP clusters after PC recruitment (Fig. 3f, Supplementary Fig. S4f), suggesting that clustering is an inducible equilibrium state. For comparison, PL108 followed the opposite trend, as it was excluded from both the bilayer and the monolayer by PC recruitment (Supplementary Fig. S4g). These results suggest that the system has lower free energy when the peptide is at the monolayer, and the free energy gap between configurations where KWALP is at the monolayer or at the bilayer is decreased by PC. This energy gap does not come from a hydrophobic mismatch since when we repeated the above experiments with a longer KWALP version, namely KWALP28 (Supplementary Fig. S3a), which should be longer than the bilayer thickness, the peptide behave almost exactly as KWALP20, within the resolution limits of our measurements (Supplementary Fig. S4c, d).

In conclusion, our results show that KWALP20 partitions dynamically between the monolayer and the bilayer, but it has a clear preference for the monolayer, especially in high PE levels. In cells, if there is no regulation of ER-to-LD partitioning, HDs would be favorably adsorbed at the LD surface as a result of free energy minimization.

Phospholipid shape defines KWALP partitioning. Since changing PC/PE ratio varied partitioning, we hypothesized that the affinity of HDs for lipids may be a driving force for partitioning. An HD protein can interact with phospholipid acyl chains, TG, and water, although interactions with the latter are unfavorable. We wanted to know which interactions would be responsible for the accumulation of the peptides to the monolayer.

We prepared phospholipid-free buffer-in-TG droplets containing KWALP20. The protein signal at the interface was uniform (Fig. 4a, b). When the interface was lined by PE, the signal was also uniform in most cases (Fig. 4a, b). Instead, when PC alone lined the interface, the protein formed clusters (Fig. 4a, b) in which the peptide was mobile (Supplementary Fig. S5a). Similar clustering was observed when PC was added to DIBs (Fig. 3f, Supplementary Fig. S4f), and never observed for AHs. Apparently, protein-protein interactions become more favorable at the PC monolayer interface, suggesting that KWALP has more affinity for TG than for phospholipid acyl chains. We thus hypothesize that the relative contact of an HD with TG and phospholipids determines HD monolayer-bilayer partitioning.

The monolayer of a droplet contiguous with a bilayer is less packed with phospholipids than the bilayer leaflets[9]. In a bilayer, HD peptides would be in contact with phospholipid acyl chains along their full length while, in a monolayer, a significant fraction of the peptide would be in contact with TG. Thus, the higher affinity of HD proteins for TG over phospholipids can explain why HD proteins partition more favorably to monolayers. Why would the PC/PE ratio matter in this picture? Very likely because PC/PE can modulate the probability for HDs to contact with TG. PC and PE do not differ in their acyl chain composition (two oleoyl chains in both cases) but they differ in their average shape: PC has a cylindrical shape while PE is conical (Fig. 4c). Therefore, PC proffers a higher phospholipid monolayer packing which, in turn, reduces the probability of contact of an HD with TGs, and increases the probability of contact with phospholipid acyl chains (Fig. 4c). As a consequence, HD would less efficiently partition to the monolayer when the PC/PE ratio is increased.

Our model suggests that phospholipid shape modulates HD monolayer-bilayer partitioning. To test this, we used dioleoyl phosphatidic acid (PA), which has a negative spontaneous curvature, like PE. We found that, in a PA DIB, KWALP20 was almost excluded from the bilayer (Supplementary Fig. S5b), as observed in PE. To further challenge our hypothesis, we repeated the KWALP20 partitioning experiments in *N*-methyl-PE and in *N*,*N*-dimethyl-PE phospholipids; these are, from a structural standpoint, intermediates between PE and PC (PC is *N*,*N*,*N*-trimethyl PE) (Fig. 4d, Supplementary Fig. S5c). Increasing methylation increased KWALP signal in the bilayer (Fig. 4d, e, Supplementary Fig. S5c, d), an effect similar to increasing PC/PE ratio, in agreement with our prediction.

In conclusion, our data indicate that HDs prefer mixing with TG instead of being in contact with membrane phospholipids. Because phospholipid packing is less compact in a monolayer compared with a bilayer[9], and because a monolayer thickness is half the thickness of a bilayer, partitioning toward monolayers is favored as they expose HD proteins to TG (Fig. 4c). Increasing the monolayer phospholipid packing, for example by increasing the PC/PE ratio, increases HD-phospholipid interaction at the expense of HD-TG interaction. In this case, the peptide less efficiently discriminates the monolayer of the bilayer from the monolayer covering TG and therefore its monolayer accumulation is dampened.

TG is responsible for the accumulation of HDs in monolayers. Our model postulates that KWALP accumulates to LD monolayers because it mixes with TG more favorably than with membrane phospholipids (Fig. 4a, b). To challenge this model, we altered the relative affinity of the peptide for phospholipids over the oil phase by changing the chemical nature of the oil. We chose an oil phase very different from TG, namely silicone oil, in an attempt to trigger major changes in oil-protein affinity. Silicone oil is chemically very different from TG but they both have a high surface tension with water[9].

In PE DIBs made in TG, KWALP20 was absent from the bilayer, as described above (Fig. 5a, c). In contrast, by replacing TG with silicone oil, we systematically observed that KWALP20 was in the bilayer (Fig. 5b, d). Moreover, phospholipid clusters regularly appeared at the monolayer interface and they were enriched in the peptide (Fig. 5b, d). Outside these areas, the peptide signal was weaker at the oil-water monolayer interface. Our interpretation is that the peptide has a higher affinity for phospholipids than for silicone oil, and therefore it preferentially distributes to phospholipid-rich regions, i.e., the bilayer and the phospholipid clusters.

To further validate our findings, we repeated the same experiment with VAMP2, one of the SNARE components that binds to membranes with a transmembrane domain. In TG, VAMP2 was completely absent from the bilayer (Supplementary Fig. S5e, g) while in silicone oil it was in the bilayer and clustered with phospholipids at the oil-water interface (Supplementary Fig. S5f, h), exactly like KWALP20.

Altogether, our results suggest that HD-containing proteins partition to regions where they find the highest molecular affinity. They have more affinity for TG than for phospholipids, and therefore get enriched in sites offering easier access to TG.

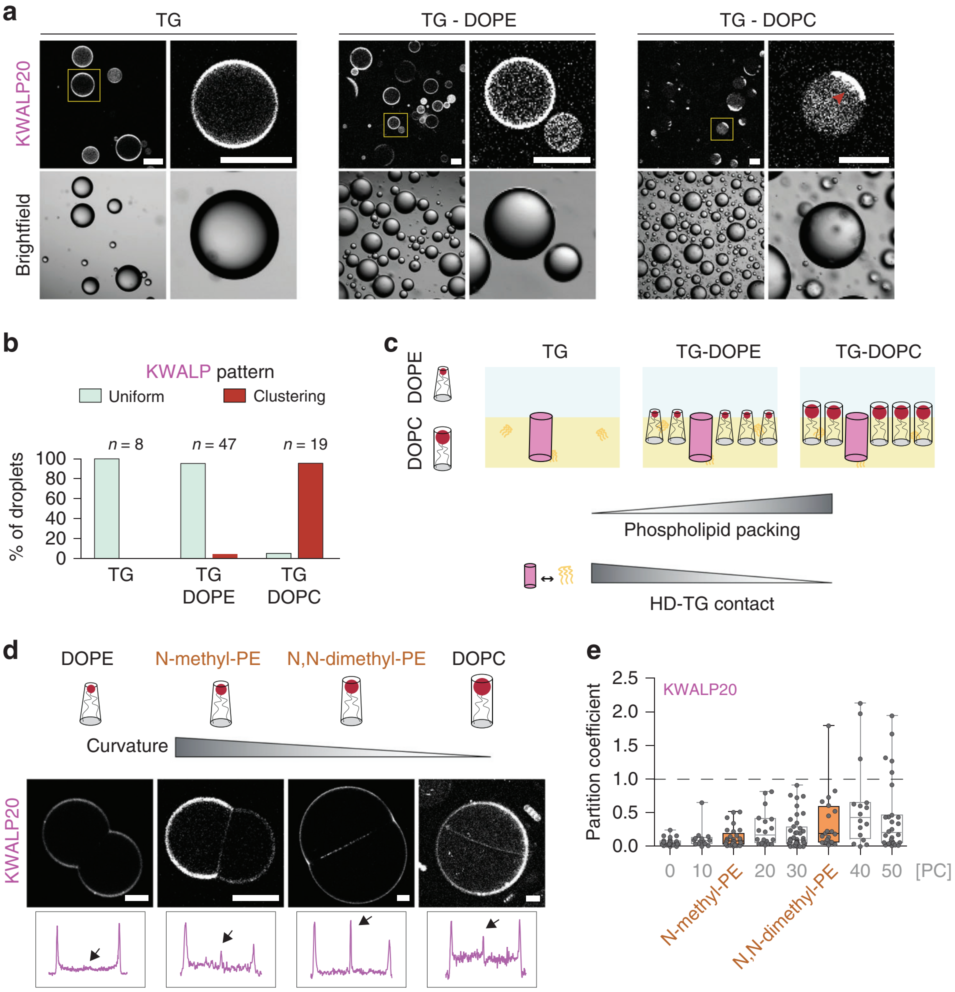


Fig. 4 The partitioning of KWALP is altered by phospholipid shape. a KWALP20 surface distribution in bare TG-buffer droplets or in TG-buffer droplets covered by DOPE or DOPC. Red arrow highlights peptide clustering (in DOPC condition). The yellow square regions are enlarged on the right side of each image. Scale bar: 100 µm. **b** Quantification of KWALP pattern, i.e., uniform (light green) or clustering (red) signal, in TG ($n = 8$), TG + DOPE ($n = 47$) and TG + DOPC ($n = 19$), from n independent measurements. **c** Schematic representation of the difference in phospholipid packing, and thus in HD-TG contact, when DOPE (cone shape) or DOPC (cylinder shape) are present. Increasing DOPC concentration in a DOPC/DOPE monolayer increases the lipid packing and decreases the contact between HDs and TG. **d** Distribution of KWALP20 in DOPE, *N*-methyl-PE, *N*,*N*-dimethyl-PE, and DOPC DIBs. These phospholipids have incremental curvature between that of DOPE and DOPC. KWALP20 is labeled with Rh-B. Line profiles (not displayed) are drawn perpendicular to the bilayer and monolayers (as described in Fig. 2b); the thickness of the line is 30–40% of the bilayer size. Arrows indicate the bilayer signal. Scale bar: 20 µm. **e** Partition coefficient of KWALP20 in DIBs of different compositions is represented as box-plots (box limits, upper and lower quartiles; middle line, median; whiskers, minimum and maximum value). Sample size is $n = 29$ and 20 for KWALP20 in *N*-methyl-PE and *N,N*-dimethyl-PE respectively. Previous results of varying PC/PE ratios (Fig. 3e) are reported in light gray. Individual data points are indicated. Source data are provided as a Source Data file.

KWALP egresses membranes to accumulate in model LDs. The DIB system revealed the existence of an energy gap favoring the higher enrichment of HDs to monolayers, due to their preferential mixing with TG over bilayer phospholipids. Thus, when a nascent LD is formed in a bilayer, as during the early step of LD biogenesis (Fig. 1a), HDs would all preferentially relocalize to the nascent LD. Such behavior has been reported for many HD-containing proteins, including HPos, LiveDrop, or Oleosins[27,37,39,40,45]. We tested this hypothesis.

To mimic the situation of a forming LD, we used the droplet-embedded vesicle system (DEV), which is a giant unilamellar vesicle (GUV) with TG droplets incorporated in between the bilayer leaflets[8] (Fig. 6a). We incorporated KWALP into PC/PE (7/3) GUVs, during GUV electroformation or after, by mixing the GUVs with the peptide, (Fig. 6a, Supplementary Fig. S6a). Next, the KWALP-containing GUVs were mixed with TG-in-water droplets in order to generate DEVs (Fig. 6a). We found that the peptide was massively enriched onto the monolayer side, consistent with our predictions and with results obtained in DIBs (Fig. 6a, b, Supplementary Fig. S6b, c). Next, we used molecular dynamics simulations to account for possible size and curvature effects which are not recapitulated in DEVs or DIBs. We generated bilayers in which 16 or 32 copies of the KWALP20 peptides were all incorporated from one side of a DOPC bilayer. In the absence of TG, the peptides were randomly distributed (Fig. 6c). When TG was incorporated into the bilayer, first it spontaneously nucleated a lens, then almost all the peptides moved to the surface of the lens, as quantified by the peptide distribution profile (Fig. 6d). The peptides remained mobile and were able to transiently move to the bilayer region, indicating that the equilibrium is dynamic and no major kinetic barrier traps the peptides in the monolayer. The same result was obtained for

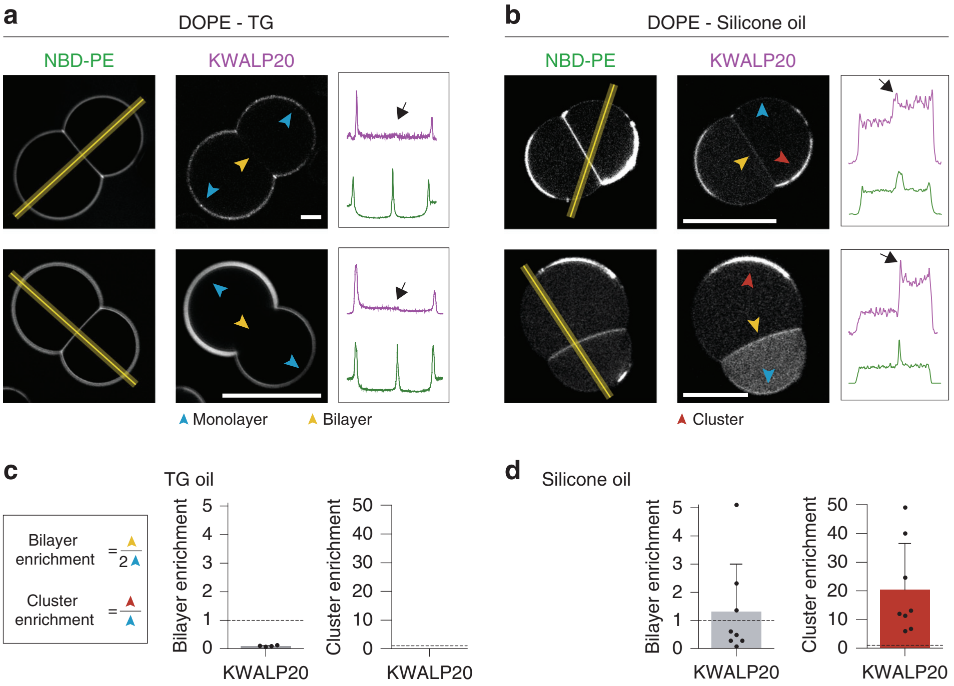


Fig. 5 The distribution of KWALP HD depends on the oil chemistry. a, **b** Distribution of KWALP20 in DOPE DIBs. KWALP20 is labeled with Rh-B. Oil phase is TG (**a**) or silicone oil (**b**). Blue arrows indicate the monolayers, yellow ones indicate the bilayer, and red ones indicate cluster areas. Plot profiles are determined using the yellow lines. The bilayer signal is indicated by a black arrow. Scale bar: 20 µm. **c**, **d** Partition coefficient is reported in gray in TG (**c**) and silicone oil (**d**), as mean ± sd ($n = 2$ and 8 independent measurements respectively). A cluster enrichment coefficient (red) is determined for the experiment in silicone oil, and is shown as mean ± SD ($n = 8$ independent measurements). Each data point is shown as a black dot. Source data are provided as a Source Data file.

KWALP28 peptides (Supplementary Fig. S6d). These results are consistent with the previous ones in the DEV and DIB systems.

Overall, these two sets of data indicate that the free energy of the system is lower when HD proteins are at the monolayer. These results reinforce the idea that HD-containing proteins can sense TG and accumulate at TG hotspots.

Discussion

Protein-lipid interactions has a key role in membrane biology by controlling protein localization and functionality[14,28,31]. In particular, a variety of protein-phospholipid interactions are responsible for the localization of many proteins to specific organelles or membranes regions[28]. TG is not a membrane lipid but a bulk lipid. Our findings support that most HD proteins have a higher affinity for TG over phospholipids in a membrane environment (Fig. 6e). Consequently, under LD biogenesis conditions, HD-containing proteins would more favorably be recruited to sites of TG accumulation, instead of remaining in the ER bilayer (Fig. 6e). Helical hairpins, hydrophobic helices, and transmembrane domains not fully crossing the ER bilayer, would accumulate to nascent LDs. Even AH-containing proteins would do so, but to a lesser extent (Fig. 2f). In short, emerging nascent LDs in the ER would be hotspots that attract nearby HD-containing proteins. Controlling these stages of LD formation will be critical for defining the proteome of the emerging LDs[8,45,46,47] and for keeping ER homeostasis.

While our results predict that HD proteins preferentially accumulate to LD monolayers, clearly not all ER HD-containing proteins target to LDs. Hence, there must be counteracting mechanisms that reduce and prevent unspecific HD-protein targeting to emerging LDs[5,6]. Amino acid composition of an HD might determine the HD-TG affinity and hence the ER-to-LD partitioning extent. For instance, the presence of charged residues in a HD may hamper HD-TG interaction, since embedding charges in a low dielectric milieu such as of a TG phase is unfavorable, and generally requires conformational changes to the protein or interaction with a protein of opposite charge. More generally, as we recently proposed for AHs[9], there could exist sequence motifs tailored with an optimal affinity with TG.

The presence of HDs can perturb the structure of lipid bilayers, generating stresses that tend to reduce protein-phospholipid interactions, for example by clustering proteins, as predicted by theoretical studies and confirmed by molecular simulations[17,28,31]. LD monolayers have more phospholipid packing defects than bilayers[9], and allow exposing TG to water molecules. During LD formation, the relocalization of HDs from the bilayer to a forming LD would reduce the stresses caused by the protein to the bilayer and possibly mask phospholipid packing defects at the LD monolayer. Such partitioning would be highly favorable as it would minimize energy on both bilayer and monolayer interfaces. Actually, even prior to LD assembly, transient TG clusters appearing in the ER bilayer[2] may attract HD-proteins or, inversely, HD-proteins can trigger the clustering of TG molecules around them[48], thereupon promoting LD nucleation and alleviating ER stress.

Finally, LD formation is stimulated by diverse physiological conditions such as excess nutrients or ER stress[11]. During ER stress, the formation of LDs may be stimulated in order to sequester damaged HD-containing proteins to be degraded, by macrolipophagy for example[49]. Indeed, in this case, proteins tend to expose hydrophobic sequences that would be favorably adsorbed to LDs. In this context, LD formation would serve as a protein quality control pathway maintaining ER proteostasis, a function different from the primary role of LDs in metabolism[50]. Accordingly, another mechanism triggering ER bilayer stress is the alteration of ER phospholipid composition[51], especially when

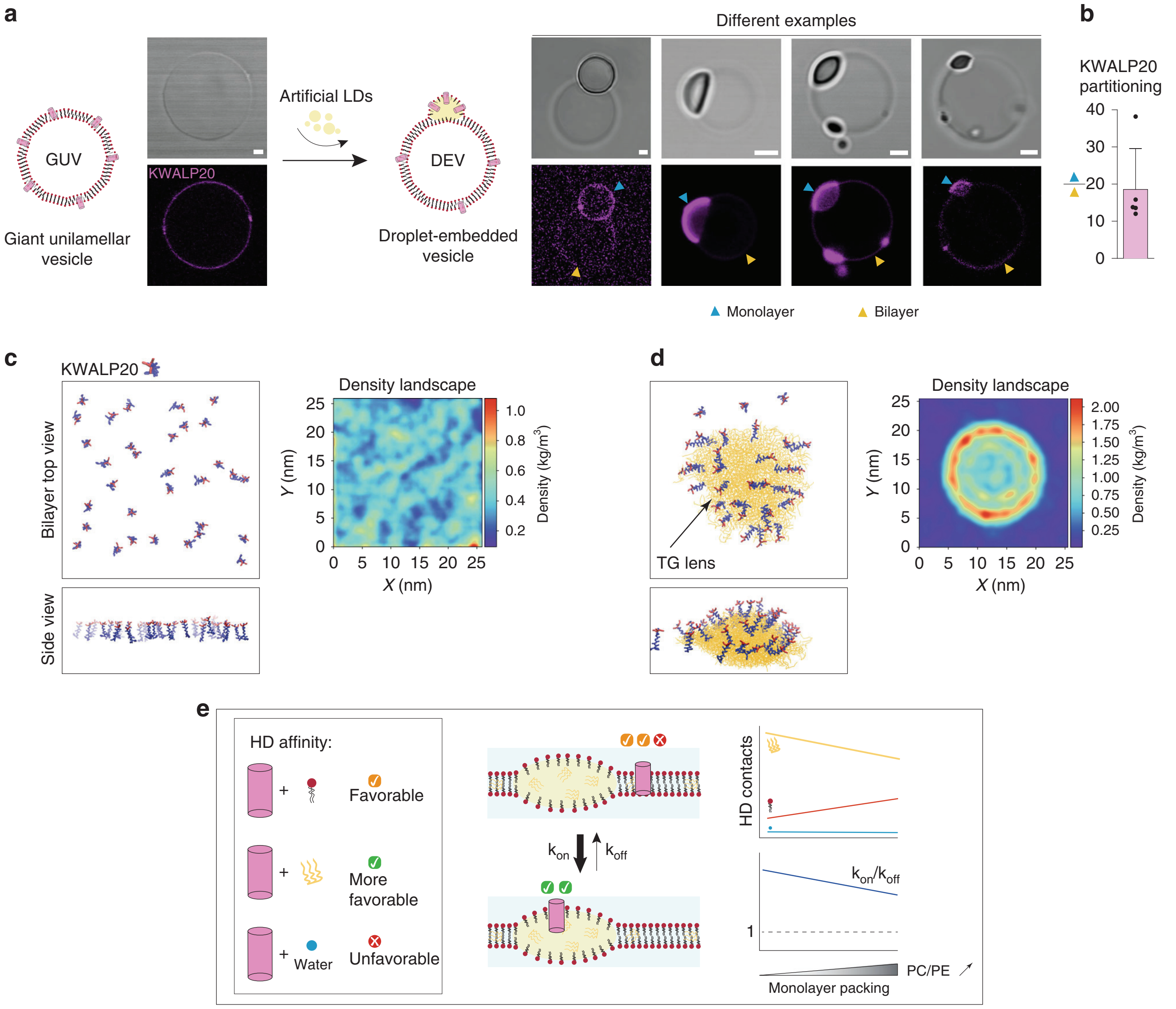


Fig. 6 Nascent LDs are attractive to monotopic proteins. **a** Formation of DEVs-containing KWALP20 labeled with Rh-B: (left) KWALP20 is inserted in GUVs, during GUV electroformation or by mixing GUVs with the peptide; an example of the resulting KWALP20-interted GUV is shown. (right) KWALP20-containing GUVs are mixed with a TG-in-buffer droplet to generate droplet-embedded vesicles; several examples of the resulting DEVs are displayed: The peptide is labeled with Rh-B. The KWALP signals on the monolayer and the bilayer are respectively depicted by blue and yellow arrows. Strong accumulation happens at the monolayer. Image brightness is enhanced to improve monolayer viewing. Scale bar: 2 µm. **b** The ratio between monolayer and bilayer signals is plotted as mean ± SD ($n = 5$ independent measurements). Individual points are indicated. Source data are provided as a Source Data file. **c**, **d** Left: snapshot of molecular dynamics simulations of a bilayer with 32 KWALP20 peptides in parallel orientation, in the absence (**c**) and the presence (**d**) of a TG lens. Hydrophobic amino-acids are represented in blue and charged ones (lysines) in red. On the right of each simulation is displayed the average protein density profiles in the bilayer plane, averaged over the entire MD simulation (20 µs), in the absence (**d**) and in the presence (**e**) of the TG lens. **e** Basic model of how the different interactions of an HD favor its LD monolayer accumulation. HD interacts with TG, phospholipids or water; the monolayer packing regulates the contact between these species and tunes the reaction constant k_{on}/k_{off}.

PC/PE ratio is decreased[10,52,53]. Here, our data bring important insights on how this ratio can modulate the partitioning of HD-containing proteins between ER and LDs: decreased PC/PE favors HD targeting and retention to the LD monolayer. Thus, by tuning PC/PE ratio, cells may be able to shift more HD-containing proteins from the ER bilayer to LDs, or vice versa, for degradation for instance. Along this analysis, increased PE/PC levels in liver is caused by dysfunctions of the phosphatidylethanolamine N-methyltransferase and associated with steatohepatitis[54,55], a condition linked to LD formation. Based on our data, such PE/PC-induced steatosis may be related to abnormal ER-to-LD protein trafficking. In contrast to mammalain cells, Drosophila cells present high PE/PC levels under normal conditions[56]; therefore, the ER-to-LD partitioning extent of HD-proteins in these cell lines may strongly differ from mammalian cells.

In conclusion, our data connect various fields involving protein-lipid interactions, from basic membrane biophysics to membrane biology, lipid metabolism, and cellular proteostasis. Our findings highlight the attractiveness of LD surface for HD-containing proteins. Accumulating neutral lipids would be rapidly detected by proteins bearing these domains. Such non-selective detection is clearly prevented by cells by means to be discovered.

Methods

Material. HEPES, Kacetate, $MgCl_2$, sodium phosphate monobasic, sodium phosphate dibasic, choloroform, trifluoroethanol, Octyl-ß-D-glucopyranoside were bought from Sigma Aldrich. DOPC (1,2dioleyl-sn-glycero-3-phosphocholine), DOPE (1,2dioleyl-sn-glycero-3-phosphoethanolamine), *N*-methyl-PE, *N*,*N*-dimethyl-PE, liver PI, Rhodamine-DOPE and NBD-DOPE were purchased from Avanti Polar Lipids, Inc. CAV1-GFP plasmid was purchased from Sino Biological (catalog no. HG11440-ACG). The following plasmids were gifts: YFP-CG2254 and YFP-CG9186 from Dr. Mathias Beller; GFP-Plin 1, GFP-Plin 1N, GFP-Plin 1C, mcherry-Plin 2 and mcherry-Plin 3 from Dr. David Savage; EGFP-ACSL3 plasmid from Dr. Joachim Füllekrug[57]; GFP-HPos from Prof. Albert Pol[40]. Cells were obtained from American Type Culture Collection and no contamination for mycoplasma was detected.

Peptides and proteins preparation. RhB-KWALP peptides, RhB-ArfGAP1, RhB-GMAP-210 and NBD-CAV1-AH were synthesized by peptide 2.0 Inc., NBD-PL108 was made by Proteogenix SAS, and RhB-NS5A was synthesized by Eric Diesis. All the peptides were chemically synthesized and purified by reverse phase high-performance liquid chromatography (HPLC). Their purity was higher than 95%, as determined by analytical HPLC and their mass was confirmed by mass spectrometry. The amino-acid sequences of the peptides are:

KWALP20: RhB-GKKKLALALALALALALWWA-Amide
KWALP28: RhB-GKKKLALALALALALALALALALALWWA-Amide
ArfGAP1-AH: RhB-FLNSAMSSLYSGWSSFTTRAKKFAK
GMAP-210-AH: RhB-MSSWLGGLGSGLGQSLGQVGGSLASLTGQISNFTKDML
CAV1-AH: NBD-LFEAVGKIFSNVRINLQKEI
PL108: NBD-PPEKIASELKDTISTRLRSARNSISVPIAS
NS5A: RhB-SGSWLRDVWDWVCTILTDFKNWLTSKLFPKL-Amide

Plin proteins, CG2254, CG9186, CAV1, HPos, NS5A and ACSL3 were obtained from purified lipid droplets be using the following protocol. LD purification from Huh7 cells expressing fluorescently tagged LD proteins: cells from five 15 cm dishes were harvested, washed once in ice-cold PBS, and lysed using a 30 G needle in 1 ml of homogenization buffer containing 20 mM Tris and complete^{TM} protease and phosphatase inhibitors, at pH 7.5; for LD isolation, 1 ml of cell lysates was mixed with 1 ml of 60% sucrose in Tris-EDTA buffer supplemented with protease inhibitors, successively overlaid with 20, 10, and 0% buffered sucrose in an 5 ml Ultra-Clear centrifuge tubes (Beckman). The tube was centrifuged for 16 h at 100,000 G and 4 °C, using an SW60 rotor in a Beckman L8-70 centrifuge. The upper 300 µl fraction was collected from as the LD fraction.

Fluorescently labeled Arf1 was generated using an Arf1-variant in which the single cysteine residue of Arf1 was replaced with serine, and the C-terminal lysine was replaced with cysteine, yielding Arf1- C159S-K181C. In short, human Arf1-C159S-K181C and yeast *N*-myristoyltransferase were coexpressed in Escherichia coli supplied with BSA-loaded myristate. Cell lysates were subjected to 35% ammonium sulfate, and the precipitate, enriched in myristoylated Arf1, was further purified by DEAE-ion exchange. Eluted fractions of interest were concentrated in spin-column filters with a 10-kD cutoff (Millipore), and fluorescently labeled using Cy3-maleimide (GE Healtcare) according to the manufacturer's protocol. To remove excess dye, samples were purified by gel filtration using a Superdex 75 column (GE Healthcare).

Oleosin1 lipid droplets were obtained from Arabidopsis seeds, provided by Dr. Martine Miquel.

Vamp2, Tsnare (complex of syntaxin1a and SNAP25) and Synaptotagmin 1 57–421 C277A, E265C (Syt1) were produced and purified by the team of Frédéric Pincet. The proteins Vamp2, Tsnare and Syt1 (solubilized in Octyl-ß-D-glucopyranoside (OG) micelles) were labeled with a fluorescent probe Atto-565 maleimide (Atto-tec, GmbH), according to the manufacturer's instructions. Free-dye was removed by gel-filtration, using a Sephadex G25 column (PD-minitrap G25, GE Healthcare). Labeled-proteins were then purified in DOPC/DOPE 1:1 proteoliposomes (P/L 1:1000): DOPE and DOPC were mixed in a glass tube, then the chloroform was removed under an argon flow and the glass tube was left under vacuum for at least one hour. The resulting lipid film was rehydrated with the Atto565-protein solution during 30 minutes. The sample was then diluted 3 times to decrease OG concentration below its cmc and a dialysis was performed overnight in a 10 kDa Slide-A-Lyzer dialysis cassette (Thermo Scientific) in order to remove OG and keep the protein into liposomes. Final buffer was the following: 25 mM HEPES pH 7.4, 120 mM KCl, 1 mM DTT (with 0.5 mM $CaCl_2$ for Syt1).

Droplet interface bilayer formation. Unless mentioned, in vitro experiments were performed in HKM buffer: 50 mM HEPES, 120 mM Kacetate, 1 mM $MgCl_2$ at pH 7.4. KWALP peptides were dissolved in trifluorethanol at 200 µM, and then diluted in HKM to get a final concentration of 10 µM. PL108 was solubilized at 50 µM in HKM, CAV1-AH at 10 µM in HKM (with 0.1% DMSO), GMAP-210 at 2 µM (with 0.1% DMSO, 16 mM urea, 80 µM DTT), ArfGAP1 at 8 µM in HKM (with 0.1% DMSO). NS5A was diluted in HKM to obtain a final concentration of 1 µM and 10% of trifluoroethanol was added to ensure of its folding. The proteins in LD (Plin 1, Plin 1 C, ACSL3, CG2254, CG9186, Oleosin 1, CAV1, HPos) or proteoliposomes (Syt1 57–421, t-snare, VAMP2) were used directly. DTT was added at a final concentration of 2.5 mM in case of aggregation.

Phospholipids (eventually with 0.2% of labeled-PE) were evaporated under a stream of argon to remove the chloroform. The resulting lipid film was then re-solubilized to the desired concentration (0.2% w/w) in trioctanoate (or silicone oil), strongly vortexed and sonicated for 10 min to ensure a complete dissolution. To form DIBs, buffer-in-oil emulsions were made using 10 µl HKM (or peptide/protein solution) dispersed in 100 µl of trioctanoate. This emulsion was strongly vortexed in order to let the protein relocalize to the surface of the droplets. Then, the same volumes of peptide/protein emulsion and lipids in oil phase were put together and the resulting emulsion was placed on a hydrophobic coverslip (glass coverslip #0 from Menzel Glaser, Braunschweig, Germany, which was covered by PDMS). The sample was let to equilibrate for 10 min and was then observed by confocal fluorescence microscopy (LSM 800, Carl Zeiss, Oberkochen, Germany), with a ×10 or oil-immersed ×63 objective depending of the size of the droplets. When the emulsion is poured onto the observation glass, droplets which are closer to each other spontaneously adhere, because of the poor solubility of the phospholipids in the oil phase, and form a bilayer. The final lipid concentration in the oil phase is then 0.1% w/w and the interfacial lipid composition is determined by lipid composition in this oil phase.

To study the effect of PC on the localization of KWALP or PL108 peptide in a dynamic way, the two peptides were both used at the same concentration of 50 µM in DOPE DIBs. Then, 5 µl of DOPC 0.2% in trioctanoate (containing 10% $CHCl_3$) was added to the sample.

Giant unilamellar vesicles formation. Phospholipids (DOPC/DOPE (7:3 or 6:4)) in chloroform at 2.5 mM were spread on an indium tin oxide (ITO)-coated glass plate. After chloroform evaporation, the resulting lipid film was then placed under vacuum for 1 h. The chamber was sealed with another ITO-coated glass plate. GUVs were grown by electroformation in a sucrose solution (0.1 g ml^{-1}, ≈280 mosmol kg^{-1}) with the following settings: 100 Hz, 1.25 V, for 1.5 to 2 h. They were then collected carefully with a Pasteur pipette, placed in a Eppendorf® tube and stored at 4 °C.

Droplets embedded vesicles formation. Droplets were made using an oil-in-water emulsion: 20 µl of trioctanoate were mixed with 100 µl of HKM buffer. The solution is then sonicated to form small droplets. 10 µl of 20 µM KWALP peptide solution was added to 40 µl of a GUV solution, which were then incubated with 20 µl of droplets for 5 min. We also added KWALP to dried phospholipids prior the electroformation and this also led to the incorporation of KWLAP to GUVs, which were subsequently used to make DEV (Fig. 6). With both approaches, the DEV/KWALP sample was then placed on a glass coverslip pretreated with 10% (w/w) BSA and washed three times with buffer, and it was then observed by confocal fluorescence microscopy (LSM 800, Carl Zeiss, Oberkochen, Germany), with an oil-immersed ×63 objective.

Electrical measurement. Aqueous droplets in oil were blown at the tip of micropipettes containing Ag/AgCl electrodes (connected to an Axopatch 200B amplifier—Molecular Device) and filled with electrolyte buffer. Micropipettes are made from borosilicate capillary (Harvard Apparatus, 1.0 mm OD×0.50 mm, ID×150mm) pulled with a micropipette puller (Sutter Instrument) to obtain tip with inner diameter of 2 µm. Before any use, tip of the micropipettes was treated dipping in a dimethyldichlorosilane solution to avoid capillary wetting by the aqueous droplets. Micropipettes were manipulated through MP225 and MP285 micromanipulators (Sutter Instrument). After blowing droplets at each micropipettes tip, 5 min are waited to allow monolayer formation, then micropipettes are moved to put into contact the two droplets to allow formation of the bilayer. Once DIB is stable, the electrical measurement was performed. This consisted in repeatedly imposing a 20 mV voltage step for 300 ms between the two sides of the DIB and measuring the resulting current. At the same time as capacitance measurements, images of droplets were acquired using a IDS camera mounted on an Olympus IX71 inverted microscope with a 20x objective to measure bilayer area of the DIB.

Thickness calculation. The capacitance value C was obtained from the fitting by an exponential of the transient capacitive current at the beginning of the voltage step to determine its time constant. The thickness of the bilayer was then calculated assuming that the bilayer can be assimilated to a dielectric material using the relation: $e = \frac{\varepsilon_r.\varepsilon_0.S}{C}$ where ε_0 is the permittivity of vacuum, ε_r the dielectric constant of the material ($\varepsilon_r = 2.8$)[58] and S the surface of the bilayer calculated from images treated on ImageJ.

Molecular dynamics simulations. To study protein distribution in nascent LDs, we carried out molecular dynamics (MD) simulations at the coarse-grained level using the MARTINI force field[50,59,60] (version 2.2). First, we generated a system containing 2016 DOPC lipids, 625 triolein (TO) molecules, and approximately 83,000 water particles; the approximate size was ca. 27 × 27 × 18 nm. TO molecules were initially dispersed homogeneously in the DOPC bilayer, and phase-separated spontaneously to form an oil lens in the bilayer. The system was simulated for 20 µs, and its shape and properties did not change during the last 10 µs. Then, protein-containing systems were generated from the equilibrated lens system, inserting 16 or 32 copies of different transmembrane peptides. Peptides were always inserted in the bilayer region of the system. We used 2 similar peptide sequences, KWALP20 and KWALP28, both in parallel and anti-parallel orientation (i.e., half of the peptides pointing up and half pointing down). In total, we built 8 different protein-containing lens systems. For each system, MD simulations were carried out for 20 µs, and the last 10 µs were used for analysis.

To study the distribution of phospholipids in LDs, we carried out simulations of large nascent LDs, containing 18144 phospholipids (DOPC and/or DOPE), 7500 TO molecules, and ca. 1.9 million water particles; the system size was approximately 78 × 78 × 40 nm. We carried out 3 simulations: one with 100% DOPC, one with DOPC:DOPE 80:20, and one with DOPC:DOPE 60:40. Each simulation was carried out for 30 µs, and the last 20 µs were used for analysis.

All coarse-grained MD simulations were carried out with GROMACS (v2016.4) software[61], using the leap-frog integrator and a time step of 20 fs. Non-bonded interactions were calculated with the Verlet neighborlist algorithm, with a Verlet buffer tolerance of 10^{-6} kJ mol^{-1} ps^{-1} and a cutoff of 1.1 nm; electrostatic interactions were shifted to zero from 0 nm, long-range electrostatics were calculated with the reaction-field method ($\varepsilon R = 15$, $\varepsilon RF = \infty$); Lennard-Jones potential was shifted to zero at the cutoff. The stochastic velocity rescaling thermostat[62] with a time constant of 1 ps was used to maintain the temperature of the membrane (lipids and proteins) and the solvent separately at 300 K. Pressure was controlled semi-isotropically using the Parrinello–Rahman barostat[63] with a reference pressure of 1 bar, compressibility of 4×10^{-4} bar^{-1}, and a time constant of 12 ps.

Analysis of protein density was carried out with in-house software[17] after re-centering the trajectory, using the center of mass of the largest TO cluster as the center of the simulation box.

Analyses of TO content in the bilayer, DOPC:DOPE contact fraction and mixing, and phospholipid distribution between bilayer and monolayer region were carried out with in-house software, freely available on our web site (https://mmsb.cnrs.fr/en/team/mobi-en/softwares/).

To study the effect of DOPE lipids on membrane thickness, we carried out all-atom simulations of pure DOPC and DOPC:DOPE 1:1 mixtures, using the CHARMM36 force field[64] and the TIP3P water model[65]. Simulation boxes contained 100 lipids (50 per leaflet) and 5000 water molecules, and simulation time was 400 ns.

Simulations were carried out with the GROMACS 2020 software, using the leap-frog integrator and a time step of 2 fs. Non-bonded interactions were calculated with the Verlet neighborlist algorithm, with a Verlet buffer tolerance of 10^{-6} kJ mol^{-1} ps^{-1} and a cutoff of 1.2 nm. The PME algorithm[66,67] was used for long-range electrostatics. The temperature was maintained at 298 K using the stochastic velocity rescaling thermostat[62] with a time constant of 1 ps. Pressure was controlled with the semi-isotropic Parrinello–Rahman barostat[63], with a reference pressure of 1 bar, compressibility of 4.5×10^{-4} bar^{-1}, and a time constant of 10 ps. Analysis of mass density was carried out over the last 300 ns of the trajectories, with standard GROMACS tools.

Circular dichroism. CD spectra were recorded over the wavelength range 185–250 nm, at 0.2 nm intervals and 20 nm min^{-1} scan speed, on a Jasco 815 spectropolarimeter (Jasco Inc., Easton, MD). Temperature was kept at 20 °C. Spectra measurements were performed in a 1 mm path length quartz cells from Hellma GmbH (internal volume 200 µl). Experiments were done either in TFE, 10 mM phosphate buffer pH 7.4, DOPC small unilamellar vesicles or TG emulsions with or without phospholipids (DOPC, DOPE, DOPC/DOPE 1:1). To prepare small unilamellar vesicles, DOPC in chloroform was put in a glass tube and the chloroform was removed using a stream of argon. Then the resulting lipid film was dried under vacuum for at least 30 min, and it was rehydrated with phosphate buffer and vortexed strongly. Finally, the lipid solution was sonicated to reduce the size of the vesicles. The oil-in-buffer emulsions were done by mixing 30 µl trioctanoate (eventually 0.2% w/w phospholipids) with 500 µl of buffer, and then sonicating the mixture. KWALP concentration was 20 µM and DOPC liposomes concentration was 1 mM. Data obtained were collected and processed using the software Spectra Manager®, and are then reported as molar ellipticity per residue (degree dmol^{-1} cm^2 residue^{-1}), given by:

$$[\theta]_{\mathrm{molar}} = \frac{100 \times \theta}{c \times l \times N}$$

where θ is the recorded ellipticity in degrees, c is the peptide concentration in mol l^{-1}, l is the cell path-length in cm and N is the number of residues of the peptide. In order to estimate the peptide secondary structure content, an analysis of CD spectra was done using CDPro software[68].

Statistics. Data analysis and representation were performed in Prism 7 (GraphPad Software, US). Information about sample size, errors bars and statistical tests are reported in each figure legend.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Data supporting the findings of this manuscript are available from the corresponding author upon reasonable request. A reporting summary for this Article is available as a Supplementary Information file. Source data are provided with this paper.

Received: 8 November 2019; Accepted: 8 July 2020;
Published online: 07 August 2020

References

1. Olzmann, J. A. & Carvalho, P. Dynamics and functions of lipid droplets. *Nat. Rev. Mol. Cell Biol.* **20**, 137–155 (2019).
2. Santinho, A. et al. Membrane Curvature Catalyzes Lipid Droplet Assembly. *Curr. Biol.* **30**, 2481–2494.e6 (2020).
3. Choudhary, V., Ojha, N., Golden, A. & Prinz, W. A. A conserved family of proteins facilitates nascent lipid droplet budding from the ER. *J. Cell Biol.* **211**, 261–271 (2015).
4. Pol, A., Gross, S. P. & Parton, R. G. Review: biogenesis of the multifunctional lipid droplet: lipids, proteins, and sites. *J. Cell Biol.* **204**, 635–646 (2014).
5. Bersuker, K. & Olzmann, J. A. Establishing the lipid droplet proteome: Mechanisms of lipid droplet protein targeting and degradation. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **1862**, 1166–1177 (2017).
6. Dhiman, R., Caesar, S., Thiam, A. R. & Schrul, B. Mechanisms of protein targeting to lipid droplets: a unified cell biological and biophysical perspective. In *Seminars in Cell & Developmental Biology* (Elsevier, 2020).
7. Kory, N., Farese, R. V. Jr & Walther, T. C. Targeting fat: mechanisms of protein localization to lipid droplets. *Trends Cell Biol.* **26**, 535–546 (2016).
8. Chorlay, A. et al. Membrane asymmetry imposes directionality on lipid droplet emergence from the ER. *Dev. Cell* **50**, 25–42 (2019).
9. Chorlay, A. & Thiam, A. R. Neutral lipids regulate amphipathic helix affinity for model lipid droplets. *J. Cell Biol.* **219**, e201907099 (2020).
10. Ho, N., Xu, C. & Thibault, G. From the unfolded protein response to metabolic diseases–lipids under the spotlight. *J. Cell Sci.* **131**, jcs199307 (2018).
11. Thiam, A. R. & Dugail, I. Lipid droplet–membrane contact sites–from protein binding to function. *J. Cell Sci.* **132**, jcs230169 (2019).
12. Ernst, R., Ballweg, S. & Levental, I. Cellular mechanisms of physicochemical membrane homeostasis. *Curr. Opin. Cell Biol.* **53**, 44–51 (2018).
13. Sezgin, E., Levental, I., Mayor, S. & Eggeling, C. The mystery of membrane organization: composition, regulation and roles of lipid rafts. *Nat. Rev. Mol. Cell Biol.* **18**, 361–374 (2017).
14. Marsh, D. Lateral pressure profile, spontaneous curvature frustration, and the incorporation and conformation of proteins in membranes. *Biophys. J.* **93**, 3884–3899 (2007).
15. Bacle, A., Gautier, R., Jackson, C. L., Fuchs, P. F. & Vanni, S. Interdigitation between triglycerides and lipids modulates surface properties of lipid droplets. *Biophys. J.* **112**, 1417–1430 (2017).
16. Prévost, C. et al. Mechanism and determinants of amphipathic helix-containing protein targeting to lipid droplets. *Dev. Cell* **44**, 73–86 (2018).
17. Castillo, N., Monticelli, L., Barnoud, J. & Tieleman, D. P. Free energy of WALP23 dimer association in DMPC, DPPC, and DOPC bilayers. *Chem. Phys. Lipids* **169**, 95–105 (2013).
18. Lewis, B. A. & Engelman, D. M. Lipid bilayer thickness varies linearly with acyl chain length in fluid phosphatidylcholine vesicles. *J. Mol. Biol.* **166**, 211–217 (1983).
19. Ajjaji, D. et al. Dual binding motifs underpin the hierarchical association of perilipins1–3 with lipid droplets. *Mol. Biol. Cell* **30**, 703–716 (2019).
20. Exner, T. et al. An alternative membrane topology permits lipid droplet localization of peroxisomal fatty acyl-CoA reductase 1. *J. Cell Sci.* **132**, jcs223016 (2019).
21. Pataki, C. I. et al. Proteomic analysis of monolayer-integrated proteins on lipid droplets identifies amphipathic interfacial α-helical membrane anchors. *Proc. Natl. Acad. Sci. USA* **115**, E8172–E8180 (2018).
22. Čopič, A. et al. A giant amphipathic helix from a perilipin that is adapted for coating lipid droplets. *Nat. Commun.* **9**, 1332 (2018).
23. Thiam, A. R. et al. COPI buds 60-nm lipid droplets from reconstituted water–phospholipid–triacylglyceride interfaces, suggesting a tension clamp function. *Proc. Natl. Acad. Sci. USA* **110**, 13244–13249 (2013).
24. Rowe, E. R. et al. Conserved amphipathic helices mediate lipid droplet targeting of perilipins 1–3. *J. Biol. Chem.* **291**, 6664–6678 (2016).
25. Jacquier, N. et al. Lipid droplets are functionally connected to the endoplasmic reticulum in Saccharomyces cerevisiae. *J. Cell Sci.* **124**, 2424–2437 (2011).
26. Markgraf, D. F. et al. An ER protein functionally couples neutral lipid metabolism on lipid droplets to membrane lipid synthesis in the ER. *Cell Rep.* **6**, 44–55 (2014).
27. Wilfling, F. et al. Arf1/COPI machinery acts directly on lipid droplets and enables their connection to the ER for protein targeting. *Elife* **3**, e01607 (2014).
28. Andersen, O. S. & Koeppe, R. E. Bilayer thickness and membrane protein function: an energetic perspective. *Annu Rev. Biophys. Biomol. Struct.* **36**, 107–130 (2007).
29. Bowie, J. U. Solving the membrane protein folding problem. *Nature* **438**, 581 (2005).
30. Phillips, R., Ursell, T., Wiggins, P. & Sens, P. Emerging roles for lipids in shaping membrane-protein function. *Nature* **459**, 379 (2009).

31. Lee, A. G. How lipids affect the activities of integral membrane proteins. *Biochim. Biophys. Acta* **1666**, 62–87 (2004).
32. Killian, J. A. Hydrophobic mismatch between proteins and lipids in membranes. *Biochim. Biophys. Acta* **1376**, 401–416 (1998).
33. Booth, M. J., Schild, V. R., Downs, F. G. & Bayley, H. Functional aqueous droplet networks. *Mol. Biosyst.* **13**, 1658–1691 (2017).
34. Thiam, A. R., Bremond, N. & Bibette, J. From stability to permeability of adhesive emulsion bilayers. *Langmuir* **28**, 6291–6298 (2012).
35. Ben M'barek, K. et al. ER membrane phospholipids and surface tension control cellular lipid droplet formation. *Dev. Cell* **41**, 591–604.e7 (2017).
36. Gross, L. C., Heron, A. J., Baca, S. C. & Wallace, M. I. Determining membrane capacitance by dynamic control of droplet interface bilayer area. *Langmuir* **27**, 14335–14342 (2011).
37. Huang, C.-Y. & Huang, A. H. Unique motifs and length of hairpin in oleosin target the cytosolic side of endoplasmic reticulum and budding lipid droplet. *Plant Physiol.* **174**, 2248–2260 (2017).
38. Ostermeyer, A. G. et al. Accumulation of caveolin in the endoplasmic reticulum redirects the protein to lipid storage droplets. *J. Cell Biol.* **152**, 1071–1078 (2001).
39. Thul, P. J. et al. Targeting of the Drosophila protein CG2254/Ldsdh1 to a subset of lipid droplets. *J. Cell Sci.* **130**, 3141–3157 (2017).
40. Kassan, A. et al. Acyl-CoA synthetase 3 promotes lipid droplet biogenesis in ER microdomains. *J. Cell Biol.* **203**, 985–1001 (2013).
41. Bigay, J., Casella, J.-F., Drin, G., Mesmin, B. & Antonny, B. ArfGAP1 responds to membrane curvature through the folding of a lipid packing sensor motif. *EMBO J.* **24**, 2244–2253 (2005).
42. Magdeleine, M. et al. A filter at the entrance of the Golgi that selects vesicles according to size and bulk lipid composition. *Elife* **5**, e16988 (2016).
43. Bigay, J. & Antonny, B. Curvature, lipid packing, and electrostatics of membrane organelles: defining cellular territories in determining specificity. *Dev. Cell* **23**, 886–895 (2012).
44. Killian, J. A. & Nyholm, T. K. Peptides in lipid bilayers: the power of simple models. *Curr. Opin. Struct. Biol.* **16**, 473–479 (2006).
45. Wang, H. et al. Seipin is required for converting nascent to mature lipid droplets. *Elife* **5**, e16582 (2016).
46. Grippa, A. et al. The seipin complex Fld1/Ldb16 stabilizes ER–lipid droplet contact sites. *J. Cell Biol.* **211**, 829–844 (2015).
47. Salo, V. T. et al. Seipin regulates –lipid droplet contacts and cargo delivery. *The EMBO J.* **35**, 2699–2716 (2016).
48. Chung, J. et al. LDAF1 and Seipin Form a Lipid Droplet Assembly Complex. *Dev. Cell* **51**, 551–563.e7 (2019).
49. Vevea, J. D. et al. Role for lipid droplet biogenesis and microlipophagy in adaptation to lipid imbalance in yeast. *Dev. Cell* **35**, 584–599 (2015).
50. de Jong, D. H. et al. Improved parameters for the martini coarse-grained protein force field. *J. Chem. Theory Comput.* **9**, 687–697 (2012).
51. Halbleib, K. et al. Activation of the unfolded protein response by lipid bilayer stress. *Mol. Cell* **67**, 673–684 (2017).
52. van der Veen, J. N. et al. The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. (Elsevier, 2017).
53. Thibault, G. et al. The membrane stress response buffers lethal effects of lipid disequilibrium by reprogramming the protein homeostasis network. *Mol. Cell* **48**, 16–27 (2012).
54. Li, Z. et al. The ratio of phosphatidylcholine to phosphatidylethanolamine influences membrane integrity and steatohepatitis. *Cell Metabolism* **3**, 321–331 (2006).
55. Vance, D. E. Physiological roles of phosphatidylethanolamine N-methyltransferase. *Biochim. Biophys. Acta.* **1831**, 626–632 (2013).
56. Guan, X. L. et al. Biochemical membrane lipidomics during Drosophila development. *Dev. Cell* **24**, 98–111 (2013).
57. Poppelreuther, M. et al. The N-terminal region of acyl-CoA synthetase 3 is essential for both the localization on lipid droplets and the function in fatty acid uptake. *J. Lipid Res.* **53**, 888–900 (2012).
58. Valincius, G. et al. Soluble amyloid β-oligomers affect dielectric membrane properties by bilayer insertion and domain formation: implications for cell toxicity. *Biophys. J.* **95**, 4845–4861 (2008).
59. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & De Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **111**, 7812–7824 (2007).
60. Monticelli, L. et al. The MARTINI coarse-grained force field: extension to proteins. *J. Chem. Theory Comput* **4**, 819–834 (2008).
61. Abraham, M. J. et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25 (2015).
62. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
63. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
64. Klauda, J. B. et al. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).
65. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
66. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
67. Essmann, U. et al. A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
68. Sreerama, N. & Woody, R. W. A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Anal. Biochem.* **209**, 32–44 (1993).
69. Gautier, R., Douguet, D., Antonny, B. & Drin, G. HELIQUEST: a web server to screen sequences with specific α-helical properties. *Bioinformatics* **24**, 2101–2102 (2008).

Acknowledgements

We are thankful to all the group members for their valuable comments and critical discussions. We are thankful to Dr. Frederic Pincet and Dr. Bruno Antonny for helpful discussions, and for providing materials. We are also thankful to Dr. Joachim Füllekrug, Dr. Mathias Beller, Dr. Martine Miquel, Dr. Rainer Beck, Dr. Paul Heo, Dr. David Savage, Dr. François Penin, for providing plasmids and materials. This work was supported the ANR-NANODROP, ANR-17-CE11-0003 to ART and LM, ANR-MOBIL, ANR-18-CE11-0012-01, and Paris Sciences et Lettres to ART. L.M. acknowledged funding from INSERM.

Author contributions

A.R.T. and L.C. designed project. L.C. performed in vitro experiments with the help of M.O. V.N. and L.M. performed and analyzed molecular dynamics simulations, P.G. and N.R. performed D.I.B. thickness measurements. A.R.T. and L.C. analyzed data and wrote the manuscript edited by co-authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-17585-8.

Correspondence and requests for materials should be addressed to A.R.T.

Peer review information *Nature Communications* thanks Rumiana Dimova, Ziliang Zhao, and other, anonymous, reviewers for their contributions to the peer review of this work.

Appendix C

Martini 3: a general purpose force field for coarse-grained molecular dynamics

ARTICLES
https://doi.org/10.1038/s41592-021-01098-3

nature methods

Check for updates

Martini 3: a general purpose force field for coarse-grained molecular dynamics

Paulo C. T. Souza[1,2] ✉, Riccardo Alessandri[1], Jonathan Barnoud[1,3], Sebastian Thallmair[1,4], Ignacio Faustino[1], Fabian Grünewald[1], Ilias Patmanidis[1], Haleh Abdizadeh[1], Bart M. H. Bruininks[1], Tsjerk A. Wassenaar[1], Peter C. Kroon[1], Josef Melcr[1], Vincent Nieto[2], Valentina Corradi[5], Hanif M. Khan[5,6], Jan Domański[7,8], Matti Javanainen[9,10], Hector Martinez-Seara[9], Nathalie Reuter[6], Robert B. Best[8], Ilpo Vattulainen[10,11], Luca Monticelli[2], Xavier Periole[12], D. Peter Tieleman[5], Alex H. de Vries[1] and Siewert J. Marrink[1] ✉

The coarse-grained Martini force field is widely used in biomolecular simulations. Here we present the refined model, Martini 3 (http://cgmartini.nl), with an improved interaction balance, new bead types and expanded ability to include specific interactions representing, for example, hydrogen bonding and electronic polarizability. The updated model allows more accurate predictions of molecular packing and interactions in general, which is exemplified with a vast and diverse set of applications, ranging from oil/water partitioning and miscibility data to complex molecular systems, involving protein–protein and protein–lipid interactions and material science applications as ionic liquids and aedamers.

The molecular dynamics simulation technique has become an indispensable tool in natural sciences, offering a spatio-temporal resolution unmatched by any experimental technique[1]. A major bottleneck of molecular dynamics is the limited time and length scales that are accessible. To overcome this limitation, coarse-grained (CG) models representing groups of atoms by effective beads, have achieved widespread use[2]. The Martini model is among the most popular CG models in the field of biomolecular simulation, due to its easy-to-use building block principle. Martini relies on a four-to-one mapping scheme (that is, on average four heavy atoms and associated hydrogens are mapped into one CG bead), and has been parametrized using a top-down approach with thermodynamic partitioning data as the main target[3,4]. Nonbonded interactions between neutral beads of Martini are solely described by Lennard-Jones potentials, while charged beads also include Coulombic interactions. The interaction strength of the Lennard-Jones potential (that is, its well depth) is used to discriminate between different levels of polarity of the CG beads. The model features four main classes of CG bead types, denoted C, N, P and Q representing nonpolar, intermediately polar, polar and charged chemical groups, respectively[4]. Sublabels are used to make a further distinction within a class in terms of degree of polarity or hydrogen donor/acceptor capabilities. In principle, all beads are the same size, denoted regular (R) beads. By way of exception, special small (S) beads were introduced to model ring-like compounds for which a four-to-one mapping scheme is inadequate[4]. To reproduce correct stacking and hydrogen-bonding distances between nucleotides, even smaller tiny (T) beads were found necessary[5]. Parametrization of the cross-interactions between S and T beads with R beads, however, was done on an ad hoc basis.

The Martini force field is used in a wide range of applications in diverse fields including structural biology[6–8], biophysics[9,10], biomedicine[11], nanotechnology[12,13] and materials design[14,15]. With its growing use, however, a number of shortcomings of the Martini model have recently been identified. One of the most important problems is the observation that certain molecules tend to interact too strongly. This has been reported for proteins and carbohydrates in solution, as well as for membrane embedded proteins[16–18]. The origin lies among others in small but systematic deviations in packing and intermolecular interactions[19]. Besides, the coverage of chemical space for broader applications was uneven, and in some cases, such as selectivity of nucleobase pairing[5,20], consistency was difficult to obtain given the limited bead types and sizes. To alleviate these problems, we undertook a rebalancing of all nonbonded interaction terms of the Martini model, including the addition of new beads and labels. The new version, called Martini 3, enables more accurate simulations of molecular systems in general. In this paper, we present the key features of Martini 3 combined with examples of new applications and improvements in relation to the previous Martini model.

Results

Reparametrization of the beads. In Martini 3, the parametrization strategy was based on the construction of prototype

---

[1]Groningen Biomolecular Sciences and Biotechnology Institute and Zernike Institute for Advanced Material, University of Groningen, Groningen, the Netherlands. [2]Molecular Microbiology and Structural Biochemistry, UMR 5086 CNRS and University of Lyon, Lyon, France. [3]Intangible Realities Laboratory, University of Bristol, School of Chemistry, Bristol, UK. [4]Frankfurt Institute for Advanced Studies, Frankfurt am Main, Germany. [5]Centre for Molecular Simulation and Department of Biological Sciences, University of Calgary, Calgary, Alberta, Canada. [6]Department of Chemistry and Computational Biology Unit, University of Bergen, Bergen, Norway. [7]Department of Biochemistry, University of Oxford, Oxford, UK. [8]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD, USA. [9]Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague, Czech Republic. [10]Computational Physics Laboratory, Tampere University, Tampere, Finland. [11]Department of Physics, University of Helsinki, Helsinki, Finland. [12]Department of Chemistry, Aarhus University, Aarhus C, Denmark.
✉e-mail: paulocts@gmail.com; s.j.marrink@rug.nl

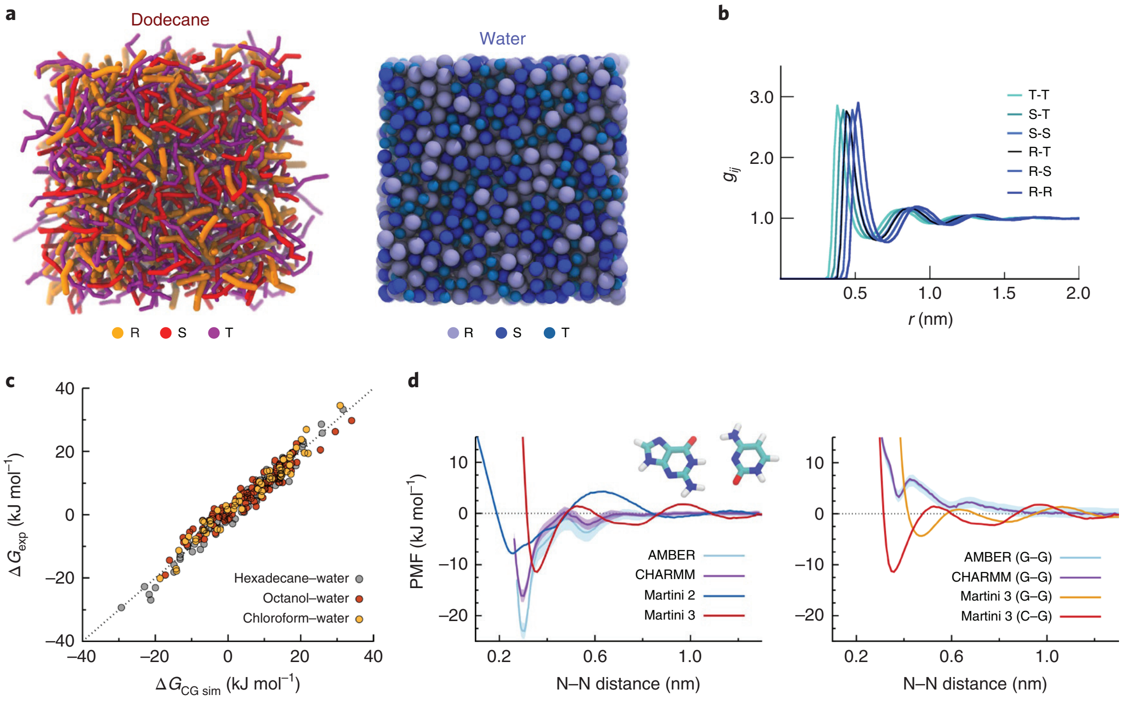


Fig. 1 | Rebalancing R, S and T beads. a, Snapshots of simulation boxes containing mixtures of dodecane and water in three resolutions. **b**, Radial distribution functions (g_{ij}) for all bead combinations in the multi-resolution mixture of water. **c**, Water–oil transfer free energies (ΔG) computed for around 260 data points using Martini 3. **d**, Hydrogen-bonding PMF between nucleobases. On the left, comparison between Martini 2 and 3 for the cytosine–guanine base pair. On the right, comparison of the cytosine–guanine (C–G) and guanine–guanine (G–G) base pairs using Martini 3. In both plots, CHARMM and AMBER atomistic data are also reported[5] for comparison. Errors are estimated with bootstrapping and displayed as transparent shades. In the case of Martini, errors are smaller than 0.1 kJ mol^{-1}, and hence are not visible on the graphs.

models of polar and nonpolar molecules in all three Martini resolutions. Self- (R-R, S-S, and T-T) and cross-interactions (R-S, R-T and S-T) of the different bead sizes were optimized to be well-balanced (Supplementary Notes). In terms of chemical types, the beads were separated into three blocks: organic, ions and water (Supplementary Notes). The organic (containing P, N and C beads) and ion (Q beads) blocks have been subjected to independent parametrizations, where different trends in self-interaction, solvation and transfer free energy on bead size change were included (Supplementary Notes). In contrast to the previous version, water is defined as a separate bead type (called W), which enables optimization of water properties independently from other targets; for example, the freezing of water at room temperature (a problem sometimes encountered with the previous water model) no longer occurs. In addition, it is available in three different sizes as well (Supplementary Notes). Together with this optimization strategy, the new Martini 3 model also features a fully revised interaction matrix (Supplementary Notes) and new intermediate interaction levels, added to smoothen the transition between chemical types (Supplementary Notes). Bead assignment and validation of the models were not only based on experimental transfer free energies, but also included solvent miscibility data (Supplementary Notes and Supplementary Results) and a series of benchmark tests, ranging from structural properties of bilayers to dimerization potentials of mean force (PMF) of proteins (Methods, Supplementary Notes and Supplementary Results).

The improved interaction balance between regular and smaller bead types is illustrated by the close to ideal mixing behavior of pure solvents composed of molecules mapped at different resolutions (Fig. 1a). Integration of radial distribution functions, defined as Kirkwood–Buff integrals (G_{ij}), are used here to quantify the degree of miscibility of the multi-resolution liquid water model (Fig. 1b). Theoretically, pair differences in Kirkwood–Buff integrals (ΔG_{ij}) should be equal to zero for all i,j pairs in ideal mixtures[21,22] while real mixtures that closely approach ideal behavior (such as benzene–toluene) show values around ±1 cm^3 mol^{-1} (ref. [22]). Our multi-resolution water model shows Δ$G_{ij} \approx -0.3$ cm^3 mol^{-1}, indicating that the balance achieved with the new parametrization closely captures an ideal mixing behavior.

The accuracy of CG models containing ring or branched fragments, which rely heavily on smaller bead types, is also greatly increased in Martini 3. For example, the binary mixing behavior of various solvents (Supplementary Results) and the transfer free energies of many linear, branched and ring-like compounds (Fig. 1c and Supplementary Results) are now in very good agreement with experimental data. The mean absolute error of transfer free energies compared to the experimental data is 2 kJ mol^{-1}, with 86% of the molecules presenting errors lower than 3 kJ mol^{-1}.

Another benefit of the recalibrated interactions is the disappearance of the artificially large desolvation free energy barriers that contribute to slow dissociation processes of the previous Martini 2 models. The problem that was initially observed in dimerization of nucleobases[5,19] is thus solved, as highlighted by the comparison of Martini 2 and Martini 3 PMFs between cytosine and guanine (left panel of Fig. 1d). Note that there is room for further improvement, as the free energy minima of the CG PMF profiles with Martini 3 are shifted relative to the all-atom profiles because the bead sizes representing nitrogen-containing groups are not optimal to reproduce hydrogen-bonding distances. In addition, the difference between C–G and G–G base pairs is not as large as in the atomistic case (roughly 20 kJ mol^{-1}). However, it is still large enough (roughly 8 kJ mol^{-1}) to provide specificity.

The proper balancing of R, S and T beads in Martini 3 also indicates that the mapping choice of an arbitrary molecule to its Martini representation is now better defined. S and T beads are not only suited to represent ring-like compounds, but also used for cases involving 3-to-1 and 2-to-1 mapping of linear and branched chemical groups (Supplementary Notes).

Covering the chemical space with new beads and labels. Together with a thorough revision of the interaction strengths, in Martini 3, we extend the number of chemical bead types and the ability to modify the bead properties depending on the chemical details of the underlying moieties. Each P, N and C class now has six bead types with different degrees of polarity, which enables a more precise definition of different chemical groups by assigning them to certain bead types. Additionally, we introduce a new X-class of beads to model halo-compounds (Supplementary Notes). In the previous version of Martini, some of the bead types were already subclassified according to their ability to act as hydrogen bond donor, acceptor or both. This property can be now attributed to all bead types of an intermediate or polar nature (N or P class). The effective interaction strength between donor and acceptor pairs is increased, whereas donor–donor and acceptor–acceptor pairs are weakened (Supplementary Notes). For example, Martini 3 correctly reproduces the trends in hydrogen bond-based pairing of nucleobases[5,20] without the use of special-purpose beads specifically for nucleobases (right panel of Fig. 1d). Note that chemical groups that can act as both donor and acceptor at the same time are always represented by the pure beads of the P and N class in Martini 3.

Next to the fine-tuning based on hydrogen-bonding capabilities, we introduce the possibility to change the interactions based on the electronic polarizability. Depending on inductive or conjugate effects caused by chemical functionalization, nonpolar molecules can be polarized; that is, they can acquire an electron-donor (or 'enriched', label 'e') or electron-acceptor (or 'vacancy', label 'v') character, which can promote preferential interactions. Polarizable groups in Martini 3 can be distinguished through the label 'e/v' that can only be applied to the C- and X-class. A nice example of their application is the strong and specific interaction between electron-donor and electron-acceptor aromatic rings in aedamers, a class of molecules that have been studied extensively in the context of biomimetic folding and self-assembly[23,24]. The use of 'e/v' allows Martini 3 to capture the preferential interaction between 1,5-dialkoxynaphthalene (DAN) and naphthalene diimide (NDI) (left panel of Fig. 2a) experimentally observed via nuclear magnetic resonance (NMR) titration[23] and atomistic simulation data. Self-assembly of amide-linked tetramers shows preferential formation of alternating stacks of DAN and NDI, which is also measured by NMR and isothermal titration calorimetry investigations[24]. On top of hydrogen bonding and electron polarization labels, all beads can have their self-interaction fine-tuned by other sublabels (as further described in the Supplementary Notes).

Chemical groups carrying monovalent charges +1/−1 are represented in Martini by the class of Q beads (Supplementary Notes). The original Martini model only considers monovalent ions, and was solely optimized for regular bead sizes that represented small ions and their first hydration shell. In Martini 3, charged groups can have R, S or T sizes. The tiny size category allows modeling of small, bare ions, enabling applications that involve ion binding where (part of) the hydration shell is lost. This feature is exemplified by the binding of sodium ions (represented by a charged tiny bead) to a buried small cavity localized in the core of the adenosine A_{2A} receptor (Fig. 2b). X-ray crystallographic[25] and ligand binding assays[26] confirm the importance of sodium ions for the structure and for the allosteric modulation of the A_{2A} receptor. Note that an extensive validation of the lipid models in Martini 3 was performed to allow simulations of transmembrane and peripheral membrane proteins (Supplementary Results).

In addition to the smaller sizes, the Q class was also expanded to five bead types, following the classical Hofmeister series trend[27,28] (Supplementary Notes and Supplementary Results). At one extreme, the Q5 bead may be used to represent hard monovalent ions with the smallest polarizability, for example inorganic ions such as $R_2PO_4^-$. At the other end of the Martini–Hofmeister series, the Q1 type models polarizable soft monovalent ions, such as $N(CH_3)_4^+$, and includes ion–π interactions via the Lennard-Jones potential. Such differences in behavior of the different Q-bead types are exemplified by molecular dynamics simulations of the anion transfer between aqueous solutions and organophosphonium-based ionic liquids (Fig. 2c and Supplementary Results). Harder ions such as Cl^- (modeled as TQ5 with −1 charge) tend to stay in the water phase, together with Na^+ ions (TQ5$^+$ bead). In contrast, softer ions such as ClO_4^- (Q2$^-$ bead) can exchange with Br^- (SQ4$^-$ bead) or (partially) PF_6^- (Q1$^-$ bead) from the ionic liquid phase. In the case of the biphasic system using trihexyltetradecylphosphonium bromide ([P_{66614}][Br]), direct comparison to experimental data shows good agreement for the anion transfer trends[28,29]. The new Q-bead types also impact biologically relevant systems, as exemplified by preferential cation–π interaction between choline groups (Q1$^+$ bead) of phosphatidylcholine lipids and aromatic residues of the *Bacillus thuringiensis* phosphatidylinositol-specific phospholipase C (*Bt*PI-PLC). In the previous version, such specific interactions between soft ions and aromatic molecules were solely included in the recently updated polarizable Martini implementation[30]. However, in Martini 3, the different Q-bead types allow easier (implicit) incorporation of such interactions without the need for additional partial charges.

On top of the new chemical types, all Q beads can use the hydrogen-bonding labels (called in this case 'p/n'). They represent organic charged molecules or fragments, such as R-CH$_2$-COO$^-$ and R-CH$_2$-NH$_3^+$, and also introduce modifications in the Hofmeister trends of the pure Q beads (Supplementary Notes). Positively charged hydrogen donors ('p' label) interact more strongly with nonpolar beads, as expected in cation–π interactions. On the other hand, negatively charged hydrogen acceptors ('n' label), have stronger interactions with neutral polar beads, mimicking the stronger hydrogen bonds with anions. To complete the ion block, we explicitly include a new D bead for divalent ions (such as Ca^{2+}), which are typically hard ions.

Improving packing and protein–protein interactions. Another change in philosophy with respect to the previous Martini models is the definition of bonded interactions. Instead of using the center of mass of the mapped chemical groups to define the geometry of the molecule, we now use a size-shape concept aimed at preserving the volume of molecules in comparison to all-atom reference structures. This choice and the proper use of Martini 3 bead sizes (Supplementary Notes) lead to more realistic molecular packing. As a consequence, the hydration of protein pockets and channels is improved, as illustrated by the Fragaceatoxin C (FraC) nanopore inserted in a lipid bilayer (Fig. 3a). The pore of FraC remains open over the whole trajectory in Martini 3, as indicated by X-ray crystallography[31] and electro-osmotic flow assays[32], while in Martini 2 it is closed.

Another example of accurate packing is the stacking predictions of thiophene derivatives in bulk heterojunction solar cells composed of poly(3-hexyl-thiophene) (P3HT) and phenyl-C61-butyric acid methyl ester (PCBM) (Fig. 3b). The morphology of these organic solar cells is a determinant for high-efficiency devices[33]. The scattering profiles computed with Martini 3 show improved agreement with Martini 3 in relation to P3HT lamellar (peak around $q \cong 0.45$ Å^{-1}) and stacking ($q \cong 1.65$ Å^{-1}) experimental distances[33,34].

The use of bonds based on molecular volume and the appropriate choice of chemical bead types, sizes and labels also controls the interaction density of the model, which has an important impact

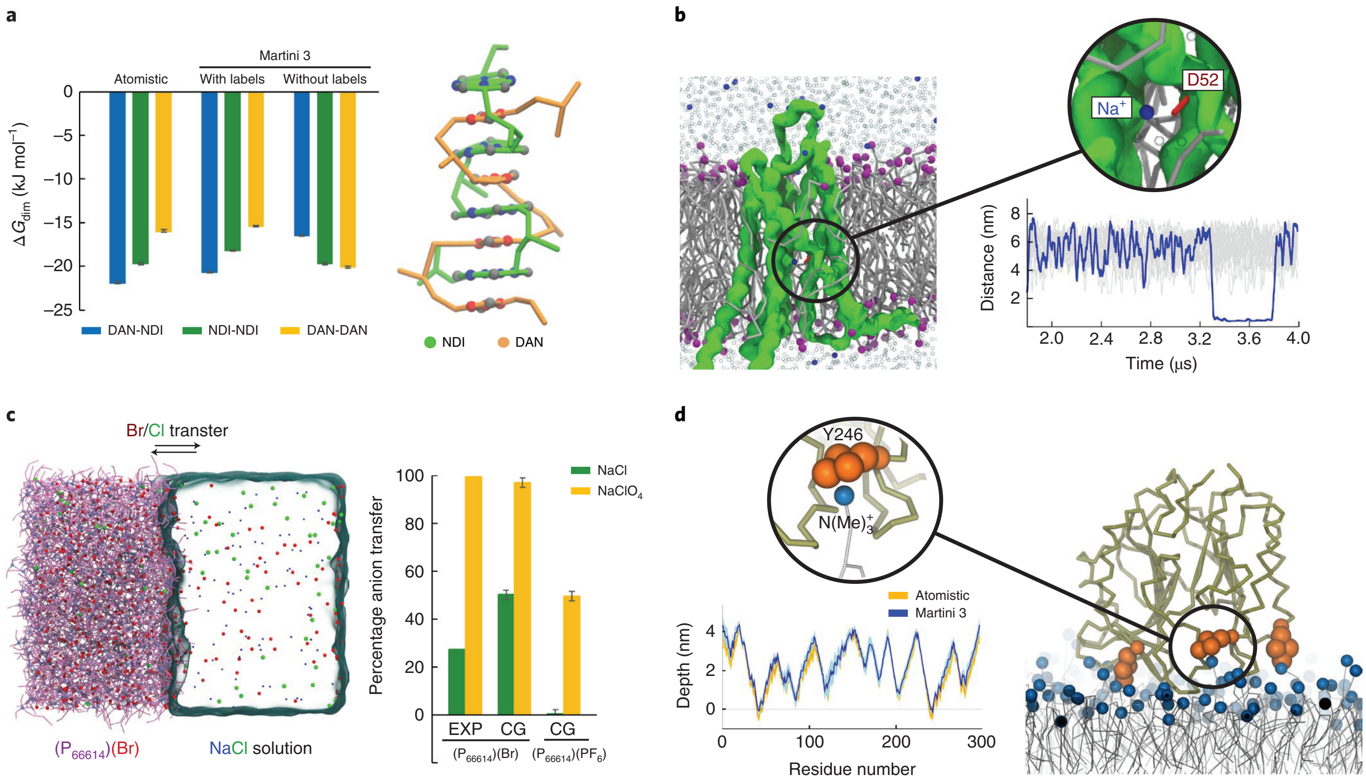


Fig. 2 | New chemical bead types, sublabels and applications. a, Self-assembly of aedamers. The left panel shows the dimerization free energies (ΔG_{dim}) of pegylated monomers of DAN and NDI. Errors are estimated with bootstrapping. The right panel shows the self-assembled duplex dimer formed by amide-linked tetramers of NDI (green) and DAN (orange). **b**, As indicated by X-ray crystallography[25], sodium ions (charged TQ5 bead) can bind to a buried small cavity in the core of the adenosine A_{2A} receptor. **c**, Charged Q beads in Martini 3 follow the classical Hofmeister series, as exemplified by the anion transfer between salt aqueous solutions and organophosphonium-based ionic liquids (right panel). Errors in the average anion transfer percentage are estimated by block averaging. **d**, Preferential cation–π interaction between choline groups (Q1 bead) of phosphatidylcholine lipids and aromatic residues of the *B. thuringiensis* phosphatidylinositol-specific phospholipase C (*Bt*PI-PLC). The depth of insertion of each amino acid of *Bt*PI-PLC is in very good agreement with the insertion obtained from an atomistic molecular dynamics simulation[30].

on the strength of collective interactions between molecules[19]. To test to what extent the changes in nonbonded and bonded interactions reduce the over-estimated aggregation of proteins, we performed extensive simulations comprising solutions of soluble proteins as well as membrane embedded proteins. These systems were simulated under conditions in which proteins do not aggregate and, preferentially, stay as monomers. For soluble proteins (Fig. 3c), qualitative tests were performed with the headpiece domain of chicken villin[35], and the modified and mutated cellulose-binding domain from *Cellulomonas fimi* (EXG-CBM), which is an example of a protein completely free of charged side chains that can maintain solubility, stability and function[36]. Trends are well captured in Martini 3, with both proteins mainly staying as monomers in pure water (with only counter-ions to neutralize the system in the case of villin). The villin headpiece showed salting-in behavior (that is, less aggregation) under addition of 0.4 M of NaCl, which was also observed for certain soluble proteins at low ionic strengths[37]. On the other hand, EXG-CBM only showed salting-out behavior (that is, more aggregation), which was expected based on experimental data[36]. In contrast, both proteins aggregate in Martini 2, forming a single and stable aggregate during the simulation.

Polyleucine (K_2-L_{26}-K_2) was selected to evaluate the aggregation propensity in membranes. Experimental evidence with this transmembrane protein model indicates a preference for the monomeric state in a bilayer environment[38–40]. Both Martini 2 and 3 show that the hydrophobic mismatch between transmembrane length and membrane thickness can play a role in the aggregation, with Martini 3 showing a higher percentage of the monomeric state (Fig. 3d). To quantitatively evaluate the strength of protein–protein interactions in a membrane environment, we also considered the dimerization of four selected transmembrane helices: the transmembrane domains of the receptor tyrosine kinases EphA1 and ErbB1; the red blood cell protein glycophorin A (GpA); as well as the well-known model peptide WALP23 (left panel of Fig. 3e). For EphA1 and ErbB1, experimental dimerization free energies in a membrane environment have been estimated using Förster resonance energy transfer (FRET)[41,42]. For GpA, dimerization free energies range from around −15 kJ mol^{-1} (in various cell membrane environments)[43,44] to −31.5 kJ mol^{-1} (GALLEX assay in *E. coli* inner membranes)[45,46] or −50.6 kJ mol^{-1} (steric trap in 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) bilayers)[47]. WALP peptides have been characterized thoroughly during the past two decades, including their self-association[48]. For each one of the four peptide dimers, we compared experimental dimerization free energies with the free energies predicted by the Martini 2 and Martini 3 models. Martini 3 shows not only to be able to capture the correct trends, but also to quantitatively reproduce the experimental affinities. The binding mode also becomes improved as highlighted for GpA (right panel of Fig. 3e). The GpA homodimer structure with Martini 3 closely resembles experimental results obtained with NMR spectroscopy and crystallography[49–51].

In summary, for both soluble and transmembrane proteins, we observed that the Martini 3 models are in much better agreement with experimental data than before. Another advantage of the current

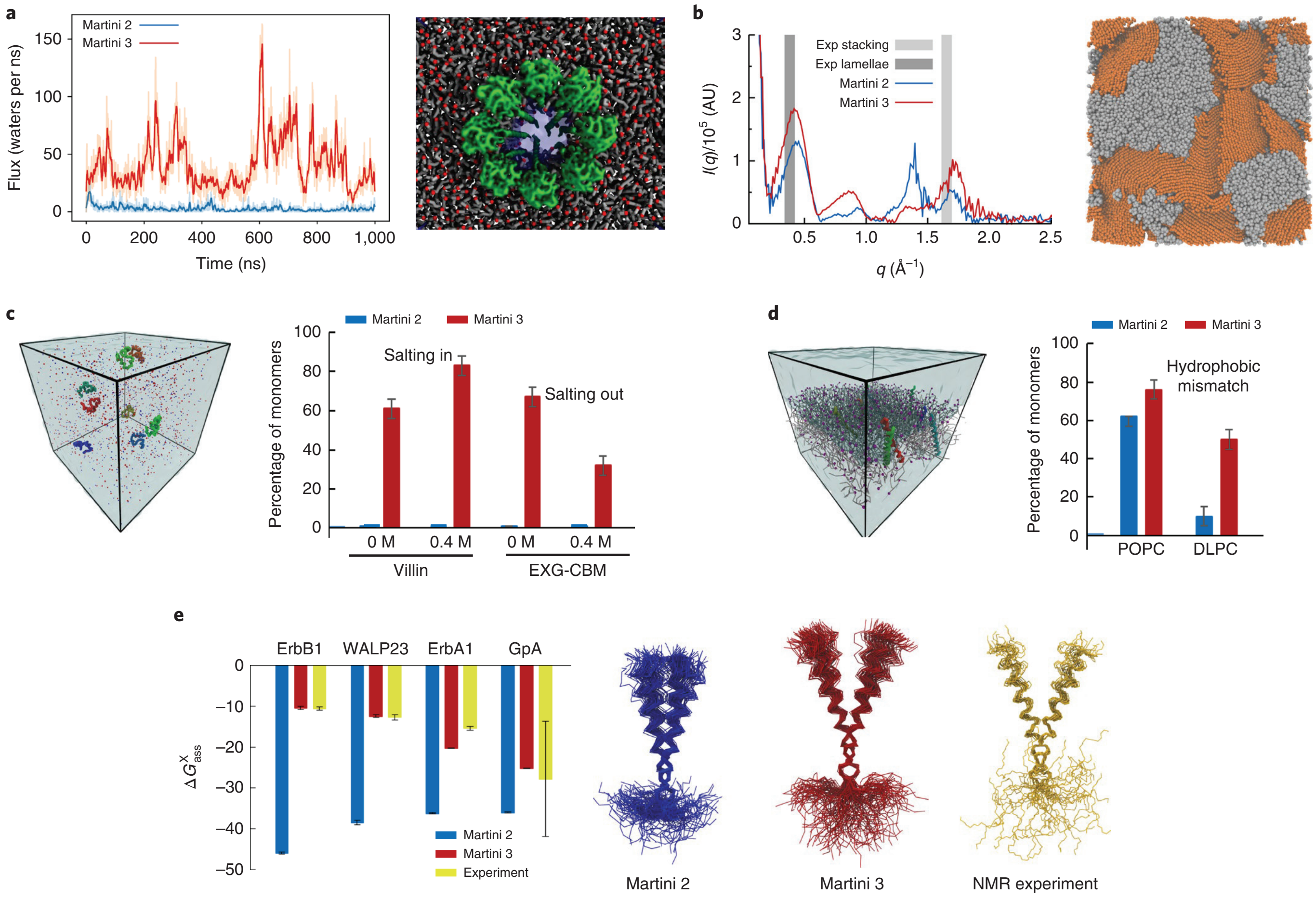


Fig. 3 | Improving packing, cavities and reducing protein stickiness. **a**, Hydration of FraC nanopore inserted in a lipid bilayer. **b**, Scattering profiles and a Martini 3 snapshot of a bulk heterojunction morphology of P3HT (in orange) and PCBM (in gray) formed after solvent evaporation and annealing simulations. *I(q)* corresponds to scattering intensity and *q* is the reciprocal space vector. **c**, Aggregation levels of the soluble proteins villin headpiece and the modified EXG-CBM in different salt concentrations. **d**, Aggregation levels of polyleucine helices in POPC and 1,2-dilauroyl-sn-glycero-3-phosphocholine (DLPC) bilayers. Errors in the average monomer percentage of **c** and **d** are estimated by block averaging. **e**, Dimerization of transmembrane helices. The left panel shows a comparison between experimental and calculated values for the mole fraction standard Gibbs free energy of dimerization ($\Delta G^{X}_{\mathrm{ass}}$) of the following transmembrane protein domains: ErbB1, EphA1, WALP23 and GpA. Simulation errors are estimated with bootstrapping while experimental data were obtained from the literature[41–48]. In the case of GpA, error was estimated by the mean absolute error of four independent experimental data[43–47]. A comparison between experimental and simulated binding modes of GpA is highlighted in the right panel. The experimental structure was taken from solution NMR in micelles (Protein Data Bank accession code 1AFO)[49]. Near identical experimental structures were obtained by solid-state NMR in nanodiscs and X-ray crystallography in a lipid cubic phase[49–51].

Martini 3 protein model is that the default bead type representing the protein backbone in Martini 3 (a regular P2 bead) no longer depends on the secondary structure. In addition, the representation of protein flexibility can now be improved by the use of Gō-like models[52].

Discussion

In this paper, we have described the new version of the Martini force field, which shows numerous improvements in relation to the previous version. However, inherent limitations to the process of coarse graining, related to transferability and representability problems[53–55] are still part of the model. An important drawback is the limited structural detail that is a consequence of representing multiple atoms with isotropic interaction sites. This is most noticeable for the Martini water model, which represents four water molecules with a single Lennard-Jones site and will certainly not capture any of the higher order structural correlations of real water. The role of explicit water in a CG model such as Martini is mostly to provide a good solvent for polar compounds resulting in realistic partitioning. For applications requiring finer details, structure-based CG models are more suitable[56,57]. Another fundamental limitation is the entropy–enthalpy compensation. The loss of internal degrees of freedom for groups of atoms represented by a CG bead inevitably reduces the entropy of the system. Since the Martini force field is based on reproducing free energies, this requires a concomitant reduction in the enthalpy. As consequence, inaccurate entropy–enthalpy balance affects the temperature dependence of several properties and reduces the transferability to different state points. To probe transferability, we performed temperature-dependent calculations for a number of solvent systems as well as lipid membranes (Supplementary Results). Temperature-dependent properties, such as the heat expansion coefficient and heat capacity of water and n-octane, are very well captured by Martini 3, but this is not true for the hydrophobic effect that shows the opposite trend with respect to atomistic models, in line with previous findings[58]. Note that the use of higher-resolution S or T particles does not remedy this problem, as these bead types were parameterized mainly to be compatible with the regular (R type) beads and should be used primarily to represent parts of the system that cannot be adequately mapped with R particles. Potential improvements with respect to the temperature transferability of our model could make use of environment dependent potentials[59] or CG beads with embedded sites, such as the polarizable water models[60,61] where incorporation of quadrupole moment might be required[58]. Bottom-up CG models that are derived with minimization of the information loss[54] as parameterization target might also perform better. For a more in depth discussion of these and related issues with respect to the Martini coarse-graining philosophy, we refer to previous papers[62,63].

Keeping these limitations in mind, Martini 3 offers a versatile and easy-to-use generic force field to simulate molecular processes at a semi-quantitative level of accuracy. In relation to the previous model, the excessive over-estimated aggregation[19] is substantially reduced. We expect that Martini 3 will allow for more realistic predictions of protein interactions, as well as more accurate simulations of molecular systems in general. The increased number of bead types and interaction levels makes the model even more versatile, covering a larger part of chemical space with appropriate building blocks. Based on this foundation, further optimizations are currently ongoing, including the use of Gō-potentials to alleviate limitations of protein conformational flexibility, a reoptimization of the bonded potentials of lipids and other biomolecular classes, as well as a compatible polarizable water model for applications requiring more realistic screening of electrostatic interactions. Finally, we foresee new application horizons for the Martini model in the field of materials science[64,65] and high-throughput drug design[66].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests and statements of data and code availability are available at https://doi.org/10.1038/s41592-021-01098-3.

Received: 16 June 2020; Accepted: 22 February 2021;
Published online: 29 March 2021

References

1. Bottaro, S. & Lindorff-Larsen, K. Biophysical experiments and biomolecular simulations: a perfect match? *Science* **361**, 355–360 (2018).
2. Ingólfsson, H. I. et al. The power of coarse graining in biomolecular simulations. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **4**, 225–248 (2014).
3. Marrink, S. J., De Vries, A. H. & Mark, A. E. Coarse grained model for semiquantitative lipid simulations. *J. Phys. Chem. B* **108**, 750–760 (2004).
4. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & de Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **111**, 7812–7824 (2007).
5. Uusitalo, J. J., Ingólfsson, H. I., Akhshi, P., Tieleman, D. P. & Marrink, S. J. Martini coarse-grained force field: extension to DNA. *J. Chem. Theory Comput.* **11**, 3932–3945 (2015).
6. Abellón-Ruiz, J. et al. Structural basis for maintenance of bacterial outer membrane lipid asymmetry. *Nat. Microbiol.* **2**, 1616–1623 (2017).
7. Yen, H. Y. et al. PtdIns(4,5)P2 stabilizes active states of GPCRs and enhances selectivity of G-protein coupling. *Nature* **559**, 423–427 (2018).
8. Van Eerden, F. J., Melo, M. N., Frederix, P. W. J. M., Periole, X. & Marrink, S. J. Exchange pathways of plastoquinone and plastoquinol in the photosystem II complex. *Nat. Commun.* **8**, 15214 (2017).
9. Vögele, M., Köfinger, J. & Hummer, G. Hydrodynamics of diffusion in lipid membrane simulations. *Phys. Rev. Lett.* **120**, 268104 (2018).
10. Agostino, M. D., Risselada, H. J., Lürick, A., Ungermann, C. & Mayer, A. A tethering complex drives the terminal stage of SNARE-dependent membrane fusion. *Nature* **551**, 634–638 (2017).
11. Jeena, M. T. et al. Mitochondria localization induced self-assembly of peptide amphiphiles for cellular dysfunction. *Nat. Commun.* **8**, 26 (2017).
12. Jiang, Z. et al. Subnanometre ligand-shell asymmetry leads to Janus-like nanoparticle membranes. *Nat. Mater.* **14**, 912–917 (2015).
13. Maingi, V. et al. Stability and dynamics of membrane-spanning DNA nanopores. *Nat. Commun.* **8**, 14784 (2017).
14. Frederix, P. W. J. M. et al. Exploring the sequence space for (tri-)peptide self-assembly to design and discover new hydrogels. *Nat. Chem.* **7**, 30–37 (2015).
15. Bochicchio, D., Salvalaglio, M. & Pavan, G. M. Into the dynamics of a supramolecular polymer at submolecular resolution. *Nat. Commun.* **8**, 147 (2017).
16. Stark, A. C., Andrews, C. T. & Elcock, A. H. Toward optimized potential functions for protein-protein interactions in aqueous solutions: osmotic second virial coefficient calculations using the MARTINI coarse-grained force field. *J. Chem. Theory Comput.* **9**, 4176–4185 (2013).
17. Javanainen, M., Martinez-Seara, H. & Vattulainen, I. Excessive aggregation of membrane proteins in the Martini model. *PLoS ONE* **12**, e0187936 (2017).
18. Schmalhorst, P. S., Deluweit, F., Scherrers, R., Heisenberg, C.-P. & Sikora, M. Overcoming the limitations of the MARTINI force field in simulations of polysaccharides. *J. Chem. Theory Comput.* **13**, 5039–5053 (2017).
19. Alessandri, R. et al. Pitfalls of the Martini model. *J. Chem. Theory Comput.* **15**, 5448–5460 (2019).
20. Uusitalo, J. J., Ingólfsson, H. I., Marrink, S. J. & Faustino, I. Martini coarse-grained force field: extension to RNA. *Biophys. J.* **113**, 246–256 (2017).
21. Ben-Naim, A. *Molecular Theory of Solutions* (Oxford Univ. Press, 2006).
22. Ploetz, E. A., Bentenitis, N. & Smith, P. E. Kirkwood–Buff integrals for ideal solutions. *J. Chem. Phys.* **132**, 164501 (2010).
23. Zych, A. J. & Iverson, B. L. Synthesis and conformational characterization of tethered, self-complexing 1,5-dialkoxynaphthalene/1,4,5,8-naphthalenetetracarboxylic diimide systems. *J. Am. Chem. Soc.* **122**, 8898–8909 (2000).
24. Gabriel, G. J. & Iverson, B. L. Aromatic oligomers that form hetero duplexes in aqueous solution. *J. Am. Chem. Soc.* **124**, 15174–15175 (2002).
25. Liu, W. et al. Structural basis for allosteric regulation of GPCRs by sodium ions. *Science* **337**, 232–236 (2012).
26. Gao, Z. G. & Ijzerman, A. P. Allosteric modulation of A(2A) adenosine receptors by amiloride analogues and sodium ions. *Biochem. Pharmacol.* **60**, 669–676 (2000).
27. Okur, H. I. et al. Beyond the Hofmeister series: ion-specific effects on proteins and their biological functions. *J. Phys. Chem. B* **121**, 1997–2014 (2017).
28. Dupont, D., Depuydt, D. & Binnemans, K. Overview of the effect of salts on biphasic ionic liquid/water solvent extraction systems: anion exchange, mutual solubility, and thermomorphic properties. *J. Phys. Chem. B* **119**, 6747–6757 (2015).
29. Naert, P., Rabaey, K. & Stevens, C. V. Ionic liquid ion exchange: exclusion from strong interactions condemns cations to the most weakly interacting anions and dictates reaction equilibrium. *Green Chem.* **20**, 4277–4286 (2018).
30. Khan, H. M. et al. Capturing choline-aromatics cation–π interactions in the MARTINI force field. *J. Chem. Theory Comput.* **16**, 2550–2560 (2020).
31. Tanaka, K., Caaveiro, J. M. M., Morante, K., González-Mañas, J. M. & Tsumoto, K. Structural basis for self-assembly of a cytolytic pore lined by protein and lipid. *Nat. Commun.* **6**, 6337 (2015).
32. Huang, G., Willems, K., Soskine, M., Wloka, C. & Maglia, G. Electro-osmotic capture and ionic discrimination of peptide and protein biomarkers with FraC nanopores. *Nat. Commun.* **8**, 935 (2017).
33. Alessandri, R., Uusitalo, J. J., De Vries, A. H., Havenith, R. W. A. & Marrink, S. J. Bulk heterojunction morphologies with atomistic resolution from coarse-grain solvent evaporation simulations. *J. Am. Chem. Soc.* **139**, 3697–3705 (2017).
34. Chiu, M. Y., Jeng, U. S., Su, C. H., Liang, K. S. & Wei, K. H. Simultaneous use of small- and wide-angle X-ray techniques to analyze nanometerscale phase separation in polymer heterojunction solar cells. *Adv. Mater.* **20**, 2573–2578 (2008).
35. Petrov, D. & Zagrovic, B. Are current atomistic force fields accurate enough to study proteins in crowded environments? *PLoS Comput. Biol.* **10**, e1003638 (2014).
36. Højgaard, C. et al. A soluble, folded protein without charged amino acid residues. *Biochemistry* **55**, 3949–3956 (2016).
37. Ruckenstein, E. & Shulgin, I. L. Effect of salts and organic additives on the solubility of proteins in aqueous solutions. *Adv. Colloid Interface Sci.* **123–126**, 97–103 (2006).
38. Zhou, F. X., Cocco, M. J., Russ, W. P., Brunger, A. T. & Engelman, D. M. Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nat. Struct. Biol.* **7**, 154–160 (2000).
39. Zhou, F. X., Merianos, H. J., Brunger, A. T. & Engelman, D. M. Polar residues drive association of polyleucine transmembrane helices. *Proc. Natl Acad. Sci. USA* **98**, 2250–2255 (2001).
40. Grau, B. et al. The role of hydrophobic matching on transmembrane helix packing in cells. *Cell Stress* **1**, 90–106 (2017).
41. Chen, L., Merzlyakov, M., Cohen, T., Shai, Y. & Hristova, K. Energetics of ErbB1 transmembrane domain dimerization in lipid bilayers. *Biophys. J.* **96**, 4622–4630 (2009).
42. Artemenko, E. O., Egorova, N. S., Arseniev, A. S. & Feofanov, A. V. Transmembrane domain of EphA1 receptor forms dimers in membrane-like environment. *Biochim. Biophys. Acta* **1778**, 2361–2367 (2008).
43. Sarabipour, S. & Hristova, K. Glycophorin A transmembrane domain dimerization in plasma membrane vesicles derived from CHO, HEK 293T, and A431 cells. *Biochim. Biophys. Acta - Biomembr.* **1828**, 1829–1833 (2013).

44. Chen, L., Novicky, L., Merzlyakov, M., Hristov, T. & Hristova, K. Measuring the energetics of membrane protein dimerization in mammalian membranes. *J. Am. Chem. Soc.* **132**, 3628–3635 (2010).
45. Nash, A., Notman, R. & Dixon, A. M. De novo design of transmembrane helix–helix interactions and measurement of stability in a biological membrane. *Biochim. Biophys. Acta - Biomembr.* **1848**, 1248–1257 (2015).
46. Finger, C. et al. The stability of transmembrane helix interactions measured in a biological membrane. *J. Mol. Biol.* **358**, 1221–1228 (2006).
47. Hong, H., Blois, T. M., Cao, Z. & Bowie, J. U. Method to measure strong protein–protein interactions in lipid bilayers using a steric trap. *Proc. Natl Acad. Sci. USA* **107**, 19802–19807 (2010).
48. Sparr, E. et al. Self-association of transmembrane α-helices in model membranes: importance of helix orientation and role of hydrophobic mismatch. *J. Biol. Chem.* **280**, 39324–39331 (2005).
49. MacKenzie, K. R., Prestegard, J. H. & Engelman, D. M. Transmembrane helix dimer: structure and implications. *Science* **276**, 131–133 (1997).
50. Trenker, R., Call, M. E. & Call, M. J. Crystal structure of the glycophorin A transmembrane dimer in lipidic cubic phase. *J. Am. Chem. Soc.* **137**, 15676–15679 (2015).
51. Domański, J., Sansom, M. S. P., Stansfeld, P. J. & Best, R. B. Balancing force field protein–lipid interactions to capture transmembrane helix–helix association. *J. Chem. Theory Comput.* **14**, 1706–1715 (2018).
52. Souza, P. C. T., Thallmair, S., Marrink, S. J. & Mera-Adasme, R. An allosteric pathway in copper, zinc superoxide dismutase unravels the molecular mechanism of the G93A amyotrophic lateral sclerosis-linked mutation. *J. Phys. Chem. Lett.* **10**, 7740–7744 (2019).
53. Brini, E. et al. Systematic coarse-graining methods for soft matter simulations-a review. *Soft Matter* **9**, 2108–2119 (2013).
54. Foley, T. T., Shell, M. S. & Noid, W. G. The impact of resolution upon entropy and information in coarse-grained models. *J. Chem. Phys.* **143**, 243104 (2015).
55. Wagner, J. W., Dama, J. F., Durumeric, A. E. P. & Voth, G. A. On the representability problem and the physical meaning of coarse-grained models. *J. Chem. Phys.* **145**, 044108 (2016).
56. Wörner, S. J., Bereau, T., Kremer, K. & Rudzinski, J. F. Direct route to reproducing pair distribution functions with coarse-grained models via transformed atomistic cross correlations. *J. Chem. Phys.* **151**, 244110 (2019).
57. Noid, W. G., Chu, J. W., Ayton, G. S. & Voth, G. A. Multiscale coarse-graining and structural correlations: connections to liquid-state theory. *J. Phys. Chem. B.* **111**, 4116–4127 (2007).
58. Wu, Z., Cui, Q. & Yethiraj, A. Driving force for the association of hydrophobic peptides: the importance of electrostatic interactions in coarse-grained water models. *J. Phys. Chem. Lett.* **2**, 1794–1798 (2011).
59. Jin, J., Yu, A. & Voth, G. A. Temperature and phase transferable bottom-up coarse-grained models. *J. Chem. Theory Comput.* **16**, 6823–6842 (2020).
60. Yesylevskyy, S. O., Schäfer, L. V., Sengupta, D. & Marrink, S. J. Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comput. Biol.* **6**, e1000810 (2010).
61. Michalowsky, J., Schäfer, L. V., Holm, C. & Smiatek, J. A refined polarizable water model for the coarse-grained MARTINI force field with long-range electrostatic interactions. *J. Chem. Phys.* **146**, 054501 (2017).
62. Marrink, S. J. & Tieleman, D. P. Perspective on the Martini model. *Chem. Soc. Rev.* **42**, 6801–22 (2013).
63. Bruininks, B. M. H., Souza, P. C. T. & Marrink, S. J. in *Biomolecular Simulations: Methods and Protocols* (eds Bonomi, M. & Camilloni, C.) 105–127 (Humana Press Inc., 2019).
64. Liu, J. et al. Enhancing molecular n-type doping of donor-acceptor copolymers by tailoring side chains. *Adv. Mater.* **30**, 1704630 (2018).
65. Vazquez-Salazar, L. I., Selle, M., de Vries, A., Marrink, S. J. & Souza, P. C. T. Martini coarse-grained models of imidazolium-based ionic liquids: from nanostructural organization to liquid–liquid extraction. *Green Chem.* **22**, 7376–7386 (2020).
66. Souza, P. C. T. et al. Protein–ligand binding with the coarse-grained Martini model. *Nat. Commun.* **11**, 3714 (2020).

Methods

CG models. CG mappings of small molecules were initially inspired by the standard Martini 2 models, when they were available. Due to the well-balanced properties of the regular (R), small (S) and tiny (T) beads in Martini 3, the CG models now follow more specific rules for mapping. For instance, over-representing 3-to-1 or 2-to-1 fragments by the usage of R beads is always avoided. Aromatic rings without substituents are composed of T beads and, in case of substituents, S beads are used. Aliphatic rings without substituents are usually based on S beads, which better reproduce their molecular shape. More technical details about the mapping rules and bead types used are given in the Supplementary Notes. As in the previous version of Martini[5,20,67–69], bonded parameters are based on atomistic simulations or high-resolution experimental data. The main difference in Martini 3 lies in the protocol to derive bond lengths, which are now based on matching overall volume and shape of the molecules (Supplementary Notes). In this spirit, the bonded parameters of the protein models were also slightly modified from the standard Martini 2.2 values[68,70], including the addition of side chain corrections[71], based on experimental reference structures. Backbone bead types do not depend on the secondary structure anymore, but are now represented by P2 beads, except for proline (SP1a), alanine (SP2, with an additional bead for the side chain) and glycine (SP1). Adapted versions of Gō-like models[72] or Elastic Networks[73] were used to maintain the tertiary protein structure. All CG protein models were constructed using Martinize2, described in Supplementary Codes. Lipid mapping was inspired by the previous Martini model[4,74–76], but now following the Martini 3 rules for mapping and also with adaptations in the bonded parameters inspired by the 'extensible model' of Carpenter et al.[77].

General setup for CG molecular dynamics simulations and analysis. Settings for the CG simulations followed, in general, the 'new' Martini set of simulation parameters[78] using the leap-frog algorithm[79] for integrating the equations of motion. The Verlet neighbor search algorithm[80] is used to update the neighbor list every 20 steps with a buffer tolerance of $0.005\,\mathrm{kJ\,mol^{-1}\,ps^{-1}}$. For the Lennard-Jones terms, we used a cutoff scheme with a value of 1.1 nm and the Verlet cutoff scheme[81] for the potential-shift. Long range electrostatic interactions were treated with reaction field[82] or particle mesh Ewald[83], with relative permittivity set to $\varepsilon_r = 15$ and a cutoff value of 1.1 nm. Reaction field was used for most of the systems, except the ones explicitly mentioning particle mesh Ewald. Periodic boundary conditions were used in all three dimensions. For the production simulations, the velocity rescaling thermostat[84] (coupling time constant of 1.0 ps) and the Parrinello–Rahman barostat[85] (coupling time constant of 12.0 ps) were employed to maintain temperature and pressure, respectively. Except for equilibration runs, a time step of 20 fs was used for all systems. CG simulation settings are available as input files for GROMACS on the Martini portal http://cgmartini.nl. GROMACS 2016.x and 2018.x were used to run all the molecular dynamics simulations[86,87]. For automated running and managing the Martini 3 simulations, we provide an updated version of Martinate[88,89], described in Supplementary Codes. All the analysis were performed using gmx analysis tools (GROMACS 2016 and 2018)[86,87], VMD v.1.9.4a12 (ref. [90]), xmgrace (v.5.1.25) and MDAnalysis[91]. The graphs were plotted using Excel 2016, xmgrace (v.5.1.25) and gnuplot (v.5.2). Figures were compiled using VMD v.1.9.4a12 and Inkscape v.1.1.

Parameter calibration, tests and validation. To parametrize the Lennard-Jones parameters of single beads and also test the Martini 3 CG models, many molecular systems and methods were used in this work. The overall iterative approach was not based in rigorous separation of calibration and validation groups. As Martini is based on pair interactions, it is hard to find simple systems that cover enough points in the interaction matrix for all bead size combinations. So, complex systems are not only used for validation but can also be part of the calibration. The tests performed were separated in 'tiers', which represent systems with different level of complexity. In tier 0, isolate beads and simple molecules are mainly used for calibration of Lennard-Jones parameters, with balance of different bead sizes and thermodynamics data (for example, liquid–liquid partitioning and miscibility) used as the main targets. In the intermediate tier 1, bilayer properties are checked, together with qualitative tests, and applied to systems such as soluble and transmembrane proteins. These qualitative tests are designed as 'yes-or-no' questions to evaluate the overall quality of the force field. At the same time, some points in the interaction matrix were also tested and fine-tuned here. In the final tier 2, quantitative tests involving complex systems are performed, including comparisons with experimental or atomistic simulation data. Here most of the systems are considered for validation. For a complete overview of the parametrization strategy used, see the Supplementary Notes.
The Supplementary Notes provide details of the specific systems and methods used in the tests performed to parametrize and validate the new Martini 3 Lennard-Jones parameters.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Force-field parameters and procedures (for example, tutorials) are publicly available at http://cgmartini.nl. Simulation data (for example, trajectories) supporting the results of this paper are available from the corresponding authors upon reasonable request.

Code availability

Martinize2 (for which the manuscript is in preparation) and Martinate codes used in this work are publicly available at https://github.com/marrink-lab/. For more detailed information, see Supplementary Codes.

References

67. López, C. A. et al. Martini coarse-grained force field: extension to carbohydrates. *J. Chem. Theory Comput.* **5**, 3195–3210 (2009).
68. Monticelli, L. et al. The MARTINI coarse-grained force field: extension to proteins. *J. Chem. Theory Comput.* **4**, 819–834 (2008).
69. Grunewald, F., Rossi, G., de Vries, A. H., Marrink, S. J. & Monticelli, L. Transferable MARTINI model of poly(ethylene oxide). *J. Phys. Chem. B* **122**, 7436–7449 (2018).
70. de Jong, D. H. et al. Improved parameters for the martini coarse-grained protein force field. *J. Chem. Theory Comput.* **9**, 687–97 (2013).
71. Herzog, F. A., Braun, L., Schoen, I. & Vogel, V. Improved side chain dynamics in MARTINI simulations of protein–lipid interfaces. *J. Chem. Theory Comput.* **12**, 2446–2458 (2016).
72. Poma, A. B., Cieplak, M. & Theodorakis, P. E. Combining the MARTINI and structure-based coarse-grained approaches for the molecular dynamics studies of conformational transitions in proteins. *J. Chem. Theory Comput.* **13**, 1366–1374 (2017).
73. Periole, X., Cavalli, M., Marrink, S.-J. & Ceruso, M. A. Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *J. Chem. Theory Comput.* **5**, 2531–2543 (2009).
74. Wassenaar, T. A., Ingólfsson, H. I., Böckmann, R. A., Tieleman, D. P. & Marrink, S. J. Computational lipidomics with insane: a versatile tool for generating Custom membranes for molecular simulations. *J. Chem. Theory Comput.* **11**, 2144–2155 (2015).
75. Melo, M. N., Ingólfsson, H. I. & Marrink, S. J. Parameters for Martini sterols and hopanoids based on a virtual-site description. *J. Chem. Phys.* **143**, 243152 (2015).
76. López, C. A., Sovova, Z., van Eerden, F. J., de Vries, A. H. & Marrink, S. J. Martini force field parameters for glycolipids. *J. Chem. Theory Comput.* **9**, 1694–1708 (2013).
77. Carpenter, T. S. et al. Capturing phase behavior of ternary lipid mixtures with a refined Martini coarse-grained force field. *J. Chem. Theory Comput.* **14**, 6050–6062 (2018).
78. de Jong, D. H., Baoukina, S., Ingólfsson, H. I. & Marrink, S. J. Martini straight: boosting performance using a shorter cutoff and GPUs. *Comput. Phys. Commun.* **199**, 1–7 (2016).
79. Hockney, R. W., Goel, S. P. & Eastwood, J. W. Quiet high-resolution computer models of a plasma. *J. Comput. Phys.* **14**, 148–158 (1974).
80. Páll, S. & Hess, B. A flexible algorithm for calculating pair interactions on SIMD architectures. *Comput. Phys. Commun.* **184**, 2641–2650 (2013).
81. Verlet, L. Computer 'experiments' on classical fluids. I. Thermodynamical properties of Lennard–Jones molecules. *Phys. Rev.* **159**, 98–103 (1967).
82. Tironi, I. G., Sperb, R., Smith, P. E. & Van Gunsteren, W. F. A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **102**, 5451–5459 (1995).
83. Essmann, U. et al. A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
84. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
85. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
86. Abraham, M. J. et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
87. Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
88. Wassenaar, T. A., Ingólfsson, H. I., Prieß, M., Marrink, S. J. & Schäfer, L. V. Mixing MARTINI: electrostatic coupling in hybrid atomistic-coarse-grained biomolecular simulations. *J. Phys. Chem. B.* **117**, 3516–3530 (2013).
89. Wassenaar, T. A. et al. High-throughput simulations of dimer and trimer assembly of membrane proteins. The DAFT approach. *J. Chem. Theory Comput.* **11**, 2278–91 (2015).
90. Humphrey, W., Dalke, A. & Schulten, K. VMD—visual molecular dynamics. *J. Molec. Graph.* **14**, 33–38 (1996).
91. Gowers, R. J. et al. MDAnalysis: a Python package for the rapid analysis of molecular dynamics simulations. in *Proc. 15th Python Sci. Conference* 98–105 (2016).

Acknowledgements

We thank all members of the S.J.M. group and also external users for testing Martini 3 in its open-beta version. In particular, we thank C. F. E. Schroer, P. W. J. M. Frederix, W. Pezeshkian, M. N. Melo, H. I. Ingólfsson, M. Tsanai, M. König, P. A. Vainikka, T. Zijp, L. Gaifas, J. H. van der Woude, M. Espinoza Cangahuala, M. Scharte, J. Cruiming, L. M. van der Sleen, V. Verduijn, A. H. Beck Frederiksen, B. Schiøtt, M. Sikora, P. Schmalhorst, R. A. Moreira, A. B. Poma, K. Pluhackova, C. Arnarez, C. A. López, E. Jefferys and M. S. P. Sansom for their preliminary tests with a lot of different systems including aedamers, sugars, amino acids, deep eutectic solvents, lipids, peptides and proteins. We also thank the Center for Information Technology of the University of Groningen for providing access to the Peregrine high-performance computing cluster. We acknowledge the National Computing Facilities Foundation of The Netherlands Organization for Scientific Research (NWO), CSC–IT Center for Science Ltd (Espoo, Finland) and CINES (France) for providing computing time. Work in the S.J.M. group was supported by an European Research Council advanced grant no. 'COMP-MICR-CROW-MEM'. R.A. thanks the NWO (Graduate Programme Advanced Materials, no. 022.005.006) for financial support. L.M. acknowledges the Institut National de la Santé et de la Recherche Medicale and the Agence Nationale de la Recherche for funding (grant no. ANR-17-CE11-0003) and GENCI-CINES for computing time (grant no. A0060710138). S.T. acknowledges the support from the European Commission via a Marie Skłodowska-Curie Actions individual fellowship (MicroMod-PSII, grant agreement no. 748895). M.J. thanks the Emil Aaltonen foundation for financial support. I.V. thanks the Academy of Finland (Center of Excellence program (grant no. 307415)), Sigrid Juselius Foundation, the Helsinki Institute of Life Science fellow program and the HFSP (research grant no. RGP0059/2019). R.B.B. and J.D. were supported by the intramural research program of the NIDDK, NIH. Their work used the computational resources of the NIH HPC Biowulf cluster (http://hpc.nih.gov). H.M.-S. acknowledges the Czech Science Foundation (grant no. 19-19561S). J.B. acknowledges funding from the TOP grant from S.J.M. (NWO) and the EPSRC program grant no. EP/P021123/1. Work in D.P.T.'s group is supported by the Natural Sciences and Engineering Research Council (Canada) and Compute Canada, funded by the Canada Foundation for Innovation. D.P.T. acknowledges further support from the Canada Research Chairs program. N.R. acknowledges funding from the Norwegian Research Council (FRIMEDBIO nos. 251247 and 288008) and computational resources from UNINETT SIGMA2 AS (grant no. NN4700K). H.M.K. acknowledges funding from the University of Calgary through the 'Eyes High Postdoctoral Fellowship' program.

Author contributions

P.C.T.S. and S.J.M. conceived the project with suggestions from R.A., A.H.V., J.B. and S.T. P.C.T.S. generated and optimized all bead parameters. P.C.T.S., R.A. and J.B. generated the topology and bonded parameters of all CG models with suggestions from S.T. and I.F. P.C.T.S., R.A., A.H.V. and F.G. performed the simulations and analysis involving transfer free energies, solvent and polymer properties. P.C.T.S., S.T., J.B. and J.M. performed the simulations and analysis involving lipid bilayers. P.C.T.S., I.F. and R.A. performed the simulations and analysis involving nucleobases. P.C.T.S., I.P. and A.H.V. generated the models and performed the simulations and analysis involving aedamers. P.C.T.S. and F.G. generated the models and performed the simulations and analysis involving ionic liquids and ionic water solutions. R.A. generated the models and performed the simulations and analysis involving bulk heterojunctions, with suggestions from L.M. regarding the fullerene model. P.C.T.S., J.B., H.A., R.A., B.M.H.B., S.T., J.M., V.N., X.P., M.J., H.M.K., J.D., V.C. and H.M.-S. performed the simulations and analysis involving amino acids, peptides and proteins. J.B., T.A.W., P.C.K. and S.T. developed some tools and scripts used to generate the CG models and to run the molecular dynamics simulations. L.M., R.B.B., P.T., N.R., I.V., A.H.V. and S.J.M. provided guidance and supervision in the studies performed by their respective group members and collaborators. P.C.T.S. and S.J.M. wrote the main manuscript, with contributions from all the authors. P.C.T.S. prepared the figures with contributions from R.A., B.M.H.B., H.M.K. and A.H.V. P.C.T.S. wrote the Methods with contributions from all the authors. P.C.T.S. wrote the Supplementary Information, with contributions from all the authors. All the authors revised and approved the final version of the manuscript, Methods and Supplementary Information.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41592-021-01098-3.

Correspondence and requests for materials should be addressed to P.C.T.S. or S.J.M.

Peer review information *Nature Methods* thanks the anonymous reviewers for their contribution for the peer review of this work. Arunima Singh was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.

Appendix D

Facilitating CG Simulations with MAD: The MArtini Database Server

JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim

Application Note

Facilitating CG Simulations with MAD: The MArtini Database Server

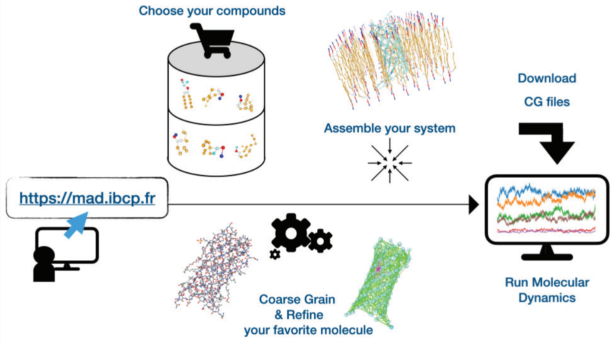
Cécile Hilpert,[#] Louis Beranger,[#] Paulo C. T. Souza,[#] Petteri A. Vainikka, Vincent Nieto, Siewert J. Marrink, Luca Monticelli, and Guillaume Launay*

Cite This: *J. Chem. Inf. Model.* 2023, 63, 702−710

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: The MArtini Database (MAD - https://mad.ibcp.fr) is a web server designed for the sharing of structures and topologies of molecules parametrized with the Martini coarse-grained (CG) force field. MAD can also convert atomistic structures into CG structures and prepare complex systems (including proteins, lipids, etc.) for molecular dynamics (MD) simulations at the CG level. It is dedicated to the generation of input files for Martini 3, the most recent version of this popular CG force field. Specifically, the MAD server currently includes tools to submit or retrieve CG models of a wide range of molecules (lipids, carbohydrates, nanoparticles, etc.), transform atomistic protein structures into CG structures and topologies, with fine control on the process and assemble biomolecules into large systems, and deliver all files necessary to start simulations in the GROMACS MD engine.



INTRODUCTION

Coarse-grained (CG) force fields allow simulations of macromolecular systems on time and length scales beyond reach for atomistic descriptions. During the past two decades coarse-graining has become a popular solution for the study of a large variety of biological problems[1,2] as well as in materials science.[3] This created the need for tools to facilitate the preparation and analysis of CG system. These tools are often provided as web services with graphical interfaces that require no installation by the user, while guiding him/her in the choice of the parameters and the validation of the results. Currently available web services are meeting various needs such as supporting the generation of multiple or specific CG force field parameters, preparing the necessary files for molecular dynamics (MD) simulations, performing quick equilibration of systems, running MD simulations for simple CG models or MD simulation analysis and files of membrane proteins in lipid bilayers.[4−16]

The Martini force field is one of the most popular choices among the coarse-grained force fields available[17,18] and offers the possibility to describe molecular interactions in systems containing lipids,[17,19,20] proteins,[21,22] carbohydrates,[23−25] nucleic acids,[26,27] polymers,[28−30] nanoparticles,[31,32] and other molecular systems, recently reviewed in refs 33 and 34. With the release of Martini 3,[18] many new systems and applications were also in reach of the model, including drug-like small-molecules,[35,36] ionic liquids,[37] deep eutectic solvents[38] and poly electrolyte coacervates.[39] Specific tools for the Martini community already exist,[20,22,40−42] but they are only partially covering the needs of the users, in particular the available web services. The CGMD/MERMAID Web server[14,15] is designed for the simulation of membrane or soluble proteins but is currently limited to version 2 of the Martini force field. The CHARMM-GUI server[43,44] allows for the preparation of MD input files for versions 2 and 3 of Martini but offers limited possibility to edit the models.

The CG representation of a molecule in Martini is obtained by combining molecular fragments representing well-defined chemical moieties, modeled by particles or beads. A total of 843 particle types are currently available in Martini 3.[18] Such a large number of different particle types can represent with high specificity the polarity, size, and hydrogen bonding capabilities of the building blocks they represent. Determining the most appropriate set of beads and their associated bonded parameters is a critical step in the preparation of a CG molecule. This so-called “parametrization” step often requires expert knowledge of the molecule chemistry associated with a validation protocol to ensure that thermodynamic properties can still be reproduced under the CG model. Thus, the creation of accurate CG representation of molecules, while more than ever accessible, remains a daunting crafting task for

Received: October 31, 2022
Published: January 19, 2023

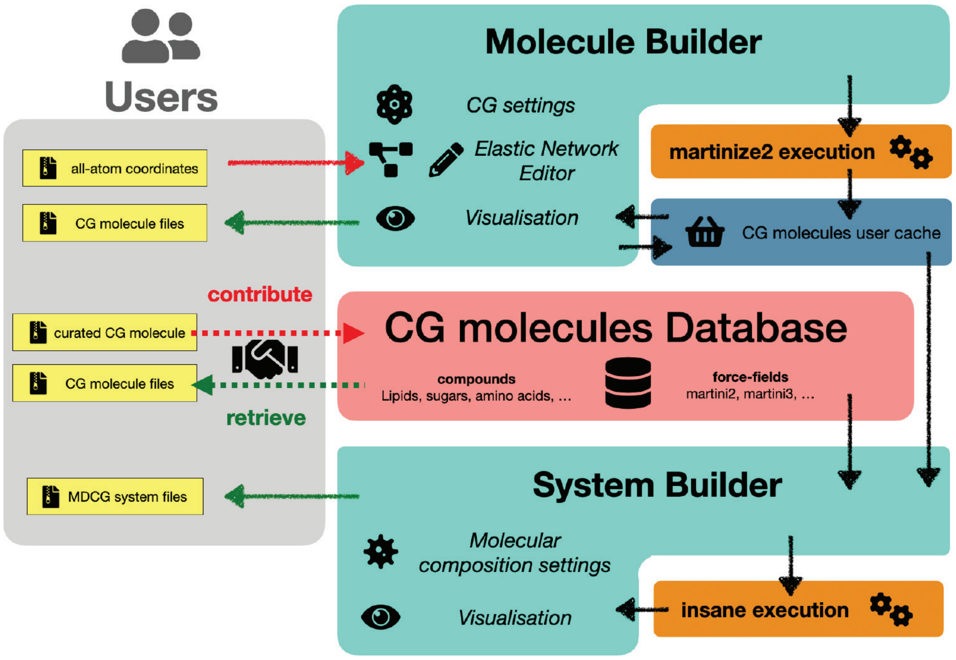


Figure 1. Workflow of the web server. CG models of molecules can be retrieved from or contributed to the MAD:Database, which covers a range of molecule types and force field versions. Alternatively, all-atom structures can be submitted to the MAD:Molecule Builder with control over the CG process by the *martinize2* program.[49] Every user is granted a private storage which holds a copy of all user generated models. The MAD:System Builder interface can use files from the main database and from the user private storage. Submitted structures are processed by *Insane* to obtain the full set of requires files to run MD.

most users. Although many automatic parametrization protocols have been recently developed,[45−48] there is still a need for an extended and curated database of models that can be used as reference to assess protocols accuracy and help at their calibration. Furthermore, as the set of Martini models is growing and diversifying, the need of CG molecule repositories is becoming critical for the long-term availability of molecular models and to make simulations reproducible.

In this context, the MAD server aims at making the setup of a CG simulation with Martini accessible to the wide Martini users community, by providing the resources to obtain the CG molecules and preparing the entire system for MD simulation. To achieve this goal, the MAD server is extending the capabilities of the original Martini molecule repository (http://cgmartini.nl/) in three directions. The MAD:Database is storing a large collection of CG molecules readily available for MD simulations. The database is organized in a user-friendly way with modern content browsing and viewing capabilities. For macromolecules such as proteins, not present in the database, the MAD:Molecule Builder tool can produce coarse-grained models based on its uploaded all-atom coordinates. Finally, the MAD:System Builder assembles many CG molecules in a simulation box and delivers all the files necessary to start the MD simulation.

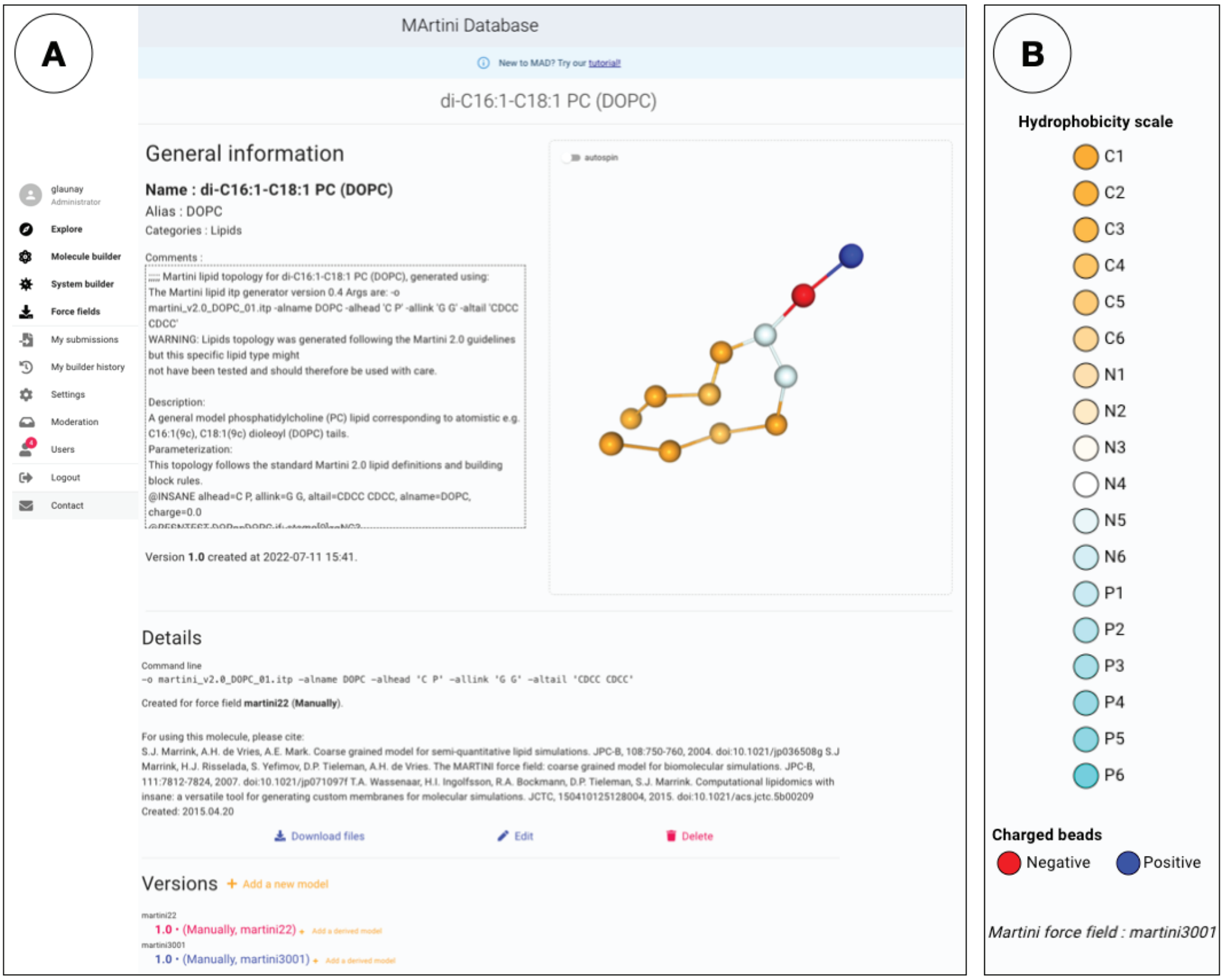
Because both MAD:Molecule Builder and MAD:System Builder services make heavy usage of the server computational resources, a login is required to use them. Logged-in users are also granted access to a private storage and a job history with resuming capabilities.

MATERIAL AND METHODS

Overall Organization. The MAD ecosystem can be depicted as a database of CG molecules and CG tools connected together with computational resources and private storage for the user (see Figure 1). The main point of access to the MAD web server is its welcome page (https://mad.ibcp.fr), where a left menu hosts links to directly explore the database, access the tools (Molecule Builder and System Builder), or download the supported force field files. The lower-half of the menu is for users to manage their account: profile, molecule contributed to the database, and submitted jobs.

In fact, every registered MAD user is granted a private history module, accessible from this menu under the “my builder history” icon. Here, all his/her previous CG operations performed on the MAD server are reported to the user. Each job is labeled by its input, date of creation, the version of the targeted force field, and the type of the operation. From the history panel, the user can visualize a model and eventually resume its modifications. Links to download the corresponding CG files are also available. Models can be deleted from this panel; in any case, their data will not be conserved over 15 days on the MAD file system.

Database. The database is designed to contain a wide variety of CG molecules which can be expanded through uploaded contributions by users. Each entry of the database corresponds to a particular association of molecule and force field version. The information on an entry is stored as a specific collection of CG files: topology files (.top and itp extension) and coordinate files (gro extension). Currently, GROMACS[50] is the main supported MD engine, but with possible future implementation of Martini in codes as OpenMM,[51] NAMD,[52] or ddcMD,[53] more file formats may be supported by MAD. Force field conversion tools[54−56] between different MD engines may be considered by users, but features of Martini models as virtual sites may not be simply adapted in other codes.



di-C16:1-C18:1 PC (DOPC)

General information

Name : di-C16:1-C18:1 PC (DOPC)

Alias : DOPC

Categories : Lipids

Comments :

```
;;;;; Martini lipid topology for di-C16:1-C18:1 PC (DOPC), generated using:
The Martini lipid itp generator version 0.4 Args are: -o
martini_v2.0_DOPC_01.itp -alname DOPC -alhead 'C P' -allink 'G G' -altail 'CDCC
CDCC'
WARNING: Lipids topology was generated following the Martini 2.0 guidelines
but this specific lipid type might
not have been tested and should therefore be used with care.

Description:
A general model phosphatidylcholine (PC) lipid corresponding to atomistic e.g.
C16:1(9c), C18:1(9c) dioleoyl (DOPC) tails.
Parameterization:
This topology follows the standard Martini 2.0 lipid definitions and building
block rules.
@INSANE alhead=C P, allink=G G, altail=CDCC CDCC, alname=DOPC,
charge=0.0
```

Version **1.0** created at 2022-07-11 15:41.

Details

Command line

```
-o martini_v2.0_DOPC_01.itp -alname DOPC -alhead 'C P' -allink 'G G' -altail 'CDCC CDCC'
```

Created for force field **martini22** (**Manually**).

For using this molecule, please cite:
S.J. Marrink, A.H. de Vries, A.E. Mark. Coarse grained model for semi-quantitative lipid simulations. JPC-B, 108:750-760, 2004. doi:10.1021/jp036508g S.J Marrink, H.J. Risselada, S. Yefimov, D.P. Tieleman, A.H. de Vries. The MARTINI force field: coarse grained model for biomolecular simulations. JPC-B, 111:7812-7824, 2007. doi:10.1021/jp071097f T.A. Wassenaar, H.I. Ingolfsson, R.A. Bockmann, D.P. Tieleman, S.J. Marrink. Computational lipidomics with insane: a versatile tool for generating custom membranes for molecular simulations. JCTC, 150410125128004, 2015. doi:10.1021/acs.jctc.5b00209
Created: 2015.04.20

Versions

martini22
1.0 • (Manually, martini22)
martini3001
1.0 • (Manually, martini3001)

Figure 2. (A) MAD page of dioleoylphosphatidylcholine (DOPC). (B) Particle color code used in MAD. The color scale broadly follows particle polarity, with red and blue colors for charged particles. The different force field models are accessible in the Versions section.

The welcome page of the database (mad.ibcp.fr/explore) displays a top section dedicated to the custom search and a bottom section which gives direct access to the currently stored molecule in table form. Each table row links to a single molecule description page. Alternatively, molecules can be searched in the database with the top section formula of the welcome page. Here, molecules can be searched by force field name (different Martini versions are available), creation methods (manually or by different tools), and biochemical category (e.g., carbohydrates, lipids, etc.) and by free text searches within name, alias, or whole entry.

The search and the table browsing methods will both link to the MAD description page (Figure 2) of a molecule. The MAD molecule pages are made of four sections. The first one is the General information section which displays the alias, name, and category of the molecule along with all the comments section extracted from the corresponding itp file. The comments can provide useful information such as the command line arguments used to generate the CG files. The top-right section of a MAD molecule page is an interactive CG molecular view of the compound, where each sphere is a Martini bead with color and size representing the bead types (Figure 2B). The Details section displays references for the molecules and a download link to bundled topology and structure files.

At the bottom of the page, the Versions section displays the version tracking trees of all the available models of the molecule. Each of these trees corresponds to a specific version of the Martini force field: only force field trees with at least one model in the database are shown. Within a tree, a node represents the specific model (and files) of a molecule stored in the database. The node of the currently displayed database entry is colored in red, and the others node in the tree are links to the web page of different models of the same molecule. The children of a molecular model are models which were declared as being derived from this parent model. By clicking on the edit button, the user can effectively submit a new "child" version of the model currently displayed.

The database is open, i.e., any registered user can submit a new CG model. Submission requires the Martini files (itp, top) and at least one reference to a publication describing the derivation of the parameters. Submitted molecules will go through a quick curation process by Martini developers. If the molecule is new to the database, a new entry will be created. In cases where previous versions would exist, the newly submitted version will be added to the corresponding position (molecule type and force field number) in the version tracking tree.

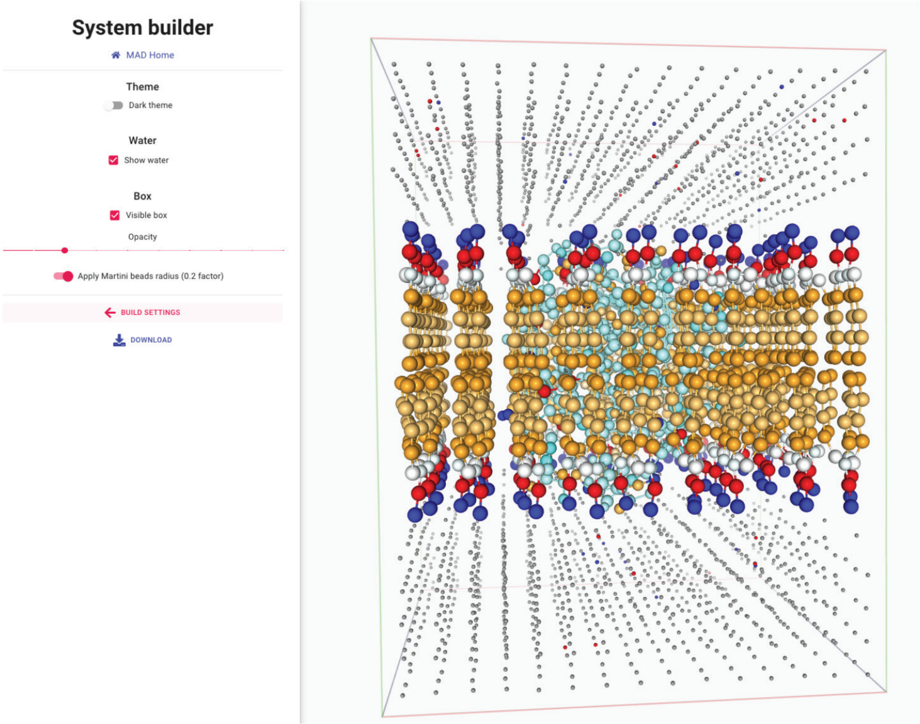


Figure 3. A view on the System Builder results: phospholipid bilayer and embedded protein(cyan) displayed in a water-filled simulation box.

Molecule Builder. The MAD:Molecule Builder tool generates the CG structure and topology from an all-atom structure. It is currently used to coarse-grain proteins, but other types of polymers will be accepted in the future (i.e., polynucleotides). The MAD:Molecule Builder is built on top of the *martinize2* program. It guides the user in the choice of the input parameters and provides handy edition/postprocessing capabilities of the output CG structures. The interface of the MAD:Molecule Builder is centered around an interactive molecular representation of the user molecule. This interactive viewer is accompanied by a left-ended panel which provides the set of control commands appropriate to each molecular building stage.

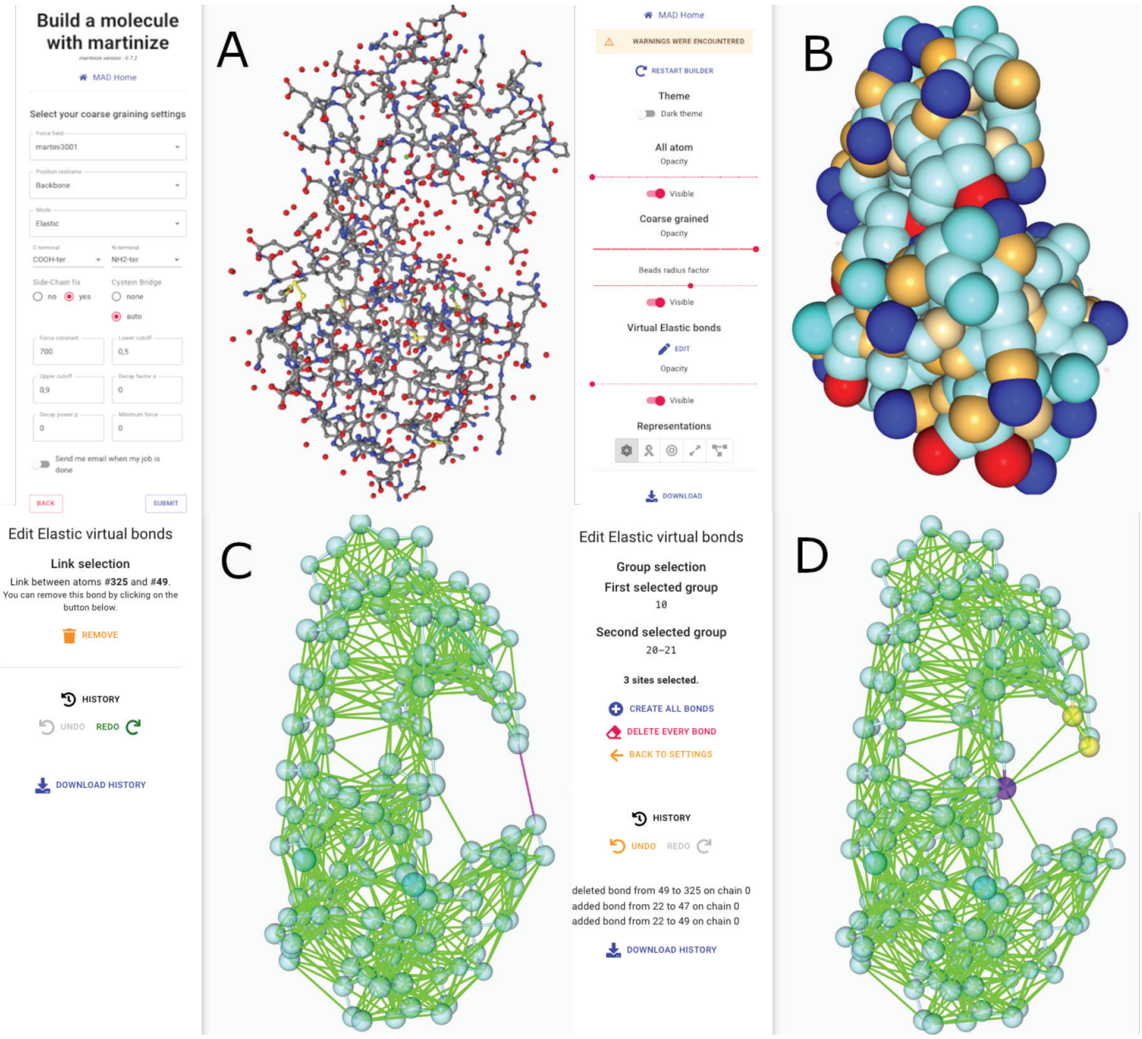
As a first step, the user uploads the all-atom structure (in PDB format) to be processed. The uploaded structure is displayed in the molecular viewer, along with CG settings in the left panel. These settings control the execution of the *martinize2* program. The force fields drop-down menu features the different versions of the Martini force field available. Currently supported versions are Martini 2.2, Martini 2.2 polarized, and Martini 3.0.

The Mode option controls the setting of distance restraints (DR), which can be based on all-beads or backbone positions only. DR are generally useful during the equilibration simulation of the molecule, in order to maintain the protein fold. Three values of Mode are available: classic, elastic,[57] and GOMartini.[58] The classic mode does not add any DR, while elastic and Go model options respectively apply additional harmonic and Lennard-Jones potentials. The Go model contact map follows the approach of Wołek et al.,[59] as implemented by Moreira et al.[60]

The N-terminal and C-terminal fixes modify the protein terminal particles, to improve representations of functional groups charges and geometries and are activated by default. The user may activate the "side chain fix", which promotes protein stability and increases the reliability of the structures during MD by the addition of dihedral angles between SC1-BB-BB-SC1 beads; this provides more realistic side chain orientations.[61] The "cystein bridges" options controls the automatic detection of cysteine residues and the application of covalent bonds between cysteine side chains when the distance between the sulfur atoms is below a threshold of 0.216 nm.

In the case of long-running computations, the email toggle option makes it possible to be contacted by email upon MAD:Molecule Builder job completion. The email will enclose a link to access and visualize the data, which are privately stored on the server for a period of 15 days. In the case of a job failure, the user would have access to the appropriate logs. Most failures are caused by improper values in the uploaded PDB file, which can easily be traced and fixed using the provided failure logging information.

Once the structure has been coarse-grained, the user can inspect it and, most importantly, can interactively modify the set of DR that *martinize2* produced. According to the chosen value of the Mode option, DR may be modeled by elastic or Go potentials. This defines a network of DR meant to preserve the protein structure (intradomain) while promoting molecular motions of interests (domain−domain interactions). By default, the *martinize2* software will deduce an initial network of DR from the all-atom structure. This network often needs to be edited based on (experimental) information on protein structure and dynamics. The DR network editor of the MAD:Molecule Builder greatly facilitates this editing process. The creation/deletion of a DR requires a simple click on two beads to add or delete an elastic bond; more sophisticated modifications are achieved by the usage of a selection language to build all DR at once between two selected groups of beads. The MAD:Molecule Builder automatically encodes the DR network in the appropriate CG files. All modifications to the



Build a molecule with martinize

martinize version : 0.7.2

Select your coarse graining settings

Force field: martini3001

Position restrains: Backbone

Mode: Elastic

C-terminal: COOH-ter

N-terminal: NH2-ter

Side-Chain fix: ○ no ● yes

Cystein Bridge: ○ none ● auto

Force constant: 700

Lower cutoff: 0,5

Upper cutoff: 0,9

Decay factor a: 0

Decay power p: 0

Minimum force: 0

Send me email when my job is done

Edit Elastic virtual bonds

Link selection

Link between atoms **#325** and **#49**. You can remove this bond by clicking on the button below.

Edit Elastic virtual bonds

Group selection

First selected group

`10`

Second selected group

`20-21`

3 sites selected.

```
deleted bond from 49 to 325 on chain 0
added bond from 22 to 47 on chain 0
added bond from 22 to 49 on chain 0
```

Figure 4. Sequential views of the MAD:Molecule Builder during the process of coarse-graining the bacteriophage T4 lysozyme (pdb code 181L): from the initial settings of parameters on the all-atom structure (A) to the visualization of the beads at a 0.6 scale on the CG structure (B), the direct (C) selection of an ER to be removed, and (D) the query based selection of amino acids to connect through newly created ER. Green links on panels C and D do not represent covalent bonds but ER between CG beads.

network can be reverted through the history section of the MAD:Molecule Builder.

Any stage of molecular editing in the MAD:Molecule Builder can be privately stashed as an update of the current model or as a new model into the user private history; the user can access them later to resume the modifications or download the structure and topology files.

System Builder. The MAD:System Builder is built on top of the *Insane* software,[20] a powerful tool for the setup of large macromolecular systems in a simulation box. The MAD:System Builder can combine several CG structures to create large systems ready for use in MD simulations. The following types of structures are allowed: models from the MAD:Database, CG molecules from the user's private stash, and topology and structure files uploaded by the user.

The first type of system that the MAD:System Builder can produce is phospholipid bilayers in water solution. The builder allows us to configure the lipid types and ratio, including the possibility of different compositions for the upper and lower leaflets. Salt concentration and total charge can be set by the user, and polarizable water can be used.

A macromolecule, usually a protein, in water solution constitutes the second type of system available within the MAD:System Builder. The macromolecule files (GRO, TOP, ITP) can be uploaded or imported from the database of compounds or from the user's private stash. If no lipids are added during the setup, the resulting simulation box will comprise one instance of the provided macromolecule in a box of water molecules.

The MAD:System Builder can also setup CG MD files for a system made of a protein embedded in a lipid bilayer surrounded by water molecules. Similarly to the previous cases, the user is guided through dedicated interface panels to define the protein, lipids, and solvent of the system.

The MAD:System Builder will assist the user with the setting of *Insane* to deliver the appropriate simulation box. For all three types of systems, the MAD:System Builder features common and advanced controls with default values applicable to most cases. Advanced users may override several parameters if so needed: a useful common setting is the geometry of the simulation box. Caution should be exercised when modifications are made to the advanced settings as they could impair the simulation.

Following the computation of the system on the MAD cluster, a view of the computed simulation box is provided. Here, the left-hand panel features visualization options and a download link to the files required to start the simulation (Figure 3).

RESULTS AND DISCUSSION

Database Content. The MAD:Database currently comprises a total of 383 CG molecules belonging to the following categories: carbohydrates, polymers, amino acids, lipids, ions, phytochemical, solvents, surfactants, synthetic nanoparticles, and small molecules. The largest category is the lipids with currently 218 entries. Molecular entries in the database can be available for the following force fields: Martini 3.01, Martini 2.2, and polarized Martini 2 versions (called 2.2P). The specific Martini 2 version developed by Monticelli and collaborators that is dedicated to nanoparticles and certain polymers is also included (called Martini 2.2 with CNP).[31,32,62] The MAD:Database currently holds the CG models of 291 and 202 molecules for Martini 2 and Martini 3, respectively. Because the database is open to submission by users, these numbers are expected to change over time, with Martini 3 CG counts progressively exceeding Martini 2.

Case Study. The elastic network approach consists of a set of harmonic potentials added on top of the Martini model to conserve the tertiary structure of proteins.[57] The network is fully dependent on the PDB structure used as reference, with the number of elastic restraints (ER) defined by the upper and lower distance cutoff MAD:Molecule Builder parameters. The rigidity of the protein model is defined by the number of ER and by the force constant used. Optimal parameters for the elastic network depend of the studied protein system. It is recommended to use experimental or atomistic simulation data to calibrate the parameters of the elastic network. To illustrate the interest of the elastic network tool in MAD:Molecule Builder, here we show how to build models of T4 lysozyme, a protein from the bacteriophage T4 (pdb code 181L).[63] Once the structure file is open in the MAD:Molecule Builder, the force field is set to "martini301". With "Elastic" activated, the default setup provides for a force constant of 700 kJ/(mol nm^2), with the lower and upper cutoffs at 0.5 and 0.9 nm, respectively. Finally, neutral termini, auto assignment of cysteine bridges, and side-chain fix are applied (Figure 4A). Upon completion of the CG process, the resulting model is displayed (Figure 4B) using the same beads color scales as the database of compounds (Figure 2B). The automatic construction of elastic network from the all-atom structure can lead to artifact ER between CG beads. Within the MAD:Molecule Builder, visual inspection of the restraints network of T4 lysozyme protein superimposed onto its CG model facilitates the identification of such a case. Indeed, an abnormal ER is found to be present between T21 and Q141 (Figure 4C). As the two amino acids are distant in structure, their ER may greatly impair the flexibility of the CG model. To better prepare the model for subsequent MD simulations, it is recommended to delete the restraint by directly clicking to select and remove (Figure 4C). Alternatively, rigidity can be added to the model in the N-terminal subdomain as exemplified in Figure 4D where all possible ER between two amino acids selections, D10 (violet color) and D20-T21 (yellow color), are created in a single click.

CONCLUSION AND PERSPECTIVES

We presented here the MAD server, a new web resource dedicated to the preparation of MD systems with the Martini coarse-grained force field. The MAD server provides a large collection of CG molecules ready to be used. Newly parametrized and published molecules can be uploaded to the MAD:Database. For molecules not yet published, topology and structure files can be provided from the user's computer. To the community, the MAD:Database could provide the storing and version tracking of large collections of models, for instance to other CG web services like the CHARMM-GUI and CGMD platform, keeping these initiatives updated with the new releases of Martini force field. In addition to the database, all-atom structures can be coarse-grained by the MAD:Molecule Builder. In a final step, CG molecules from any of these sources can be combined by the MAD:System Builder to produce topology and structure files MD-ready. We strongly encourage users to contribute to the repository by uploading their favorite models.

For the new users, a tutorial is available at https://mad.ibcp.fr/tutorial/. We plan on expanding the capabilities of the MAD server through the future integration of algorithms useful to the MARTINI users community, such as Polyply[42] and Martini-sour.[64]

ASSOCIATED CONTENT

Data Availability Statement

The MAD server is built on a front-end to back-end architecture. The front-end, which is based on the version 16.9 of the React web component framework, carries most of the steps for the submission, validation, visualization, and edition of structure. The back-end is a NodeJS/Express platform which interacts with a slurm scheduler to launch GROMACS,[50] *martinize2*, and *Insane*[20] jobs on the MAD cluster. MAD uses the public version of the *Insane* software (https://github.com/Tsjerk/Insane) and the version 0.73 of the *martinize2*[49] and its companion library, the vermouth package (https://github.com/marrink-lab/vermouth-martinize). The public database of molecules is operated by the NoSQL Couch SGDB. All the molecular visualizations are performed by the NGL[65] JavaScript library. All scripts are publicly available for download from the separated repositories of the front-end and back-end code bases, respectively located at https://github.com/MMSB-MOBI/martinize-db-client and https://github.com/MMSB-MOBI/martinize-db.

AUTHOR INFORMATION

Corresponding Author

Guillaume Launay − *Microbiologie Moléculaire et Biochimie Structurale (MMSB), 69367 Lyon, France;* orcid.org/0000-0003-0177-8706; Email: guillaume.launay@ibcp.fr

Authors

Cécile Hilpert − *Microbiologie Moléculaire et Biochimie Structurale (MMSB), 69367 Lyon, France;* orcid.org/0000-0002-7825-9376

Louis Beranger − *Microbiologie Moléculaire et Biochimie Structurale (MMSB), 69367 Lyon, France;* orcid.org/0000-0001-5206-2164

Paulo C. T. Souza − *Microbiologie Moléculaire et Biochimie Structurale (MMSB), 69367 Lyon, France;* orcid.org/0000-0003-0660-1301

Petteri A. Vainikka − *Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, The Netherlands;* orcid.org/0000-0002-3570-0977

Vincent Nieto − *Microbiologie Moléculaire et Biochimie Structurale (MMSB), 69367 Lyon, France;* orcid.org/0000-0002-5633-2114

Siewert J. Marrink − *Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, The Netherlands;* orcid.org/0000-0001-8423-5277

Luca Monticelli − *Microbiologie Moléculaire et Biochimie Structurale (MMSB), 69367 Lyon, France;* orcid.org/0000-0002-6352-4595

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jcim.2c01375

Author Contributions

[#]C.H., L.B., and P.C.T.S. contributed equally to this work.

Author Contributions

Cécile Hilpert, Louis Beranger, and Paulo C. T. Souza performed the development and data acquisition, with contributions from Petteri A. Vainikka and Vincent Nieto; Luca Monticelli and Guillaume Launay designed the study; Siewert J. Marrink, Luca Monticelli, Paulo C. T. Souza, and Guillaume Launay wrote the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank all external users for testing MAD during its development and the technical support from Alexis Michon and Samuel Bosquin (CNRS:UAR3760). Luca Monticelli thanks the French National Institute of Health and Medical Research (INSERM) for the support. Paulo C. T. Souza acknowledges the supported by French National Center for Scientific Research (CNRS). Further funding of Paulo C. T. Souza and Luca Monticelli came from a research collaboration with PharmCADD. Siewert J. Marrink acknowledges funding by the European Union (ERC Advanced, COMP-O-CELL, 101053661).

■ REFERENCES

(1) Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A. E.; Kolinski, A. Coarse-Grained Protein Models and Their Applications. *Chem. Rev.* **2016**, *116*, 7898−7936.

(2) Marrink, S. J.; Corradi, V.; Souza, P. C. T.; Ingólfsson, H. I.; Tieleman, D. P.; Sansom, M. S. Computational Modeling of Realistic Cell Membranes. *Chem. Rev.* **2019**, *119*, 6184−6226.

(3) Chen, Z.; Huo, J.; Hao, L.; Zhou, J. Multiscale modeling and simulations of responsive polymers. *Curr. Opin. Chem. Eng.* **2019**, *23*, 21−33.

(4) Hospital, A.; Andrio, P.; Fenollosa, C.; Cicin-Sain, D.; Orozco, M.; Gelpí, J. L. MDWeb and MDMoby: an integrated web-based platform for molecular dynamics simulations. *Bioinformatics* **2012**, *28*, 1278−1279.

(5) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859−1865.

(6) Kuriata, A.; Gierut, A. M.; Oleniecki, T.; Ciemny, M. P.; Kolinski, A.; Kurcinski, M.; Kmiecik, S. CABS-flex 2.0: a web server for fast simulations of flexibility of protein structures. *Nucleic Acids Res.* **2018**, *46* (W1), W338−W343.

(7) Feig, M. Local Protein Structure Refinement via Molecular Dynamics Simulations with locPREFMD. *J. Chem. Inf. Model.* **2016**, *56*, 1304−1312.

(8) Heo, L.; Feig, M. PREFMD: a web server for protein structure refinement via molecular dynamics simulations. *Bioinformatics* **2018**, *34*, 1063−1065.

(9) Czaplewski, C.; Karczynska, A.; Sieradzan, A. K.; Liwo, A. UNRES server for physics-based coarse-grained simulations and prediction of protein structure, dynamics and thermodynamics. *Nucleic Acids Res.* **2018**, *46*, W304−W309.

(10) Noel, J. K.; Levi, M.; Raghunathan, M.; Lammert, H.; Hayes, R. L.; Onuchic, J. N.; Whitford, P. C. SMOG 2: A Versatile Software Package for Generating Structure-Based Models. *PLoS Comput. Biol.* **2016**, *12*, No. e1004794.

(11) Kimmett, T.; Smith, N.; Witham, S.; Petukh, M.; Sarkar, S.; Alexov, E. ProBLM web server: protein and membrane placement and orientation package. *Computational and Mathematical Methods in Medicine* **2014**, *2014*, 838259.

(12) Schneider, J.; Ribeiro, R.; Alfonso-Prieto, M.; Carloni, P.; Giorgetti, A. Hybrid MM/CG Webserver: Automatic Set Up of Molecular Mechanics/Coarse-Grained Simulations for Human G Protein-Coupled Receptor/Ligand Complexes. *Front. Mol. Biosci.* **2020**, DOI: 10.3389/fmolb.2020.576689.

(13) Margreitter, C.; Petrov, D.; Zagrovic, B. Vienna-PTM web server: a toolkit for MD simulations of protein post-translational modifications. *Nucleic Acids Res.* **2013**, *41*, W422−W426.

(14) Damre, M.; Marchetto, A.; Giorgetti, A. MERMAID: dedicated web server to prepare and run coarse-grained membrane protein dynamics. *Nucleic Acids Res.* **2019**, *47*, W456−W461.

(15) Marchetto, A.; Chaib, Z. S.; Rossi, C. A.; Ribeiro, R.; Pantano, S.; Rossetti, G.; Giorgetti, A. CGMD Platform: Integrated Web Servers for the Preparation, Running, and Analysis of Coarse-Grained Molecular Dynamics Simulations. *Molecules* **2020**, *25*, 5934.

(16) Newport, T. D.; Sansom, M. S.; Stansfeld, P. J. The MemProtMD database: a resource for membrane-embedded protein structures and their lipid interactions. *Nucleic Acids Res.* **2019**, *47*, D390−D397.

(17) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **2007**, *111*, 7812−7824.

(18) Souza, P. C. T.; Alessandri, R.; Barnoud, J.; Thallmair, S.; Faustino, I.; Grünewald, F.; Patmanidis, I.; Abdizadeh, H.; Bruininks, B. M. H.; Wassenaar, T. A.; Kroon, P. C.; Melcr, J.; Nieto, V.; Corradi, V.; Khan, H. M.; Domański, J.; Javanainen, M.; Martinez-Seara, H.; Reuter, N.; Best, R. B.; Vattulainen, I.; Monticelli, L.; Periole, X.; Tieleman, D. P.; de Vries, A. H.; Marrink, S. J. Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nat. Methods* **2021**, *18*, 382−388.

(19) López, C. A.; Sovova, Z.; van Eerden, F. J.; de Vries, A. H.; Marrink, S. J. Martini Force Field Parameters for Glycolipids. *J. Chem. Theory Comput.* **2013**, *9*, 1694−1708.

(20) Wassenaar, T. A.; Ingólfsson, H. I.; Böckmann, R. A.; Tieleman, D. P.; Marrink, S. J. Computational Lipidomics with insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *J. Chem. Theory Comput.* **2015**, *11*, 2144−2155.

(21) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S. J. The MARTINI coarse-grained force field: extension to proteins. *J. Chem. Theory Comput.* **2008**, *4*, 819−834.

(22) de Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schäfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved Parameters for the Martini Coarse-Grained Protein Force Field. *J. Chem. Theory Comput.* **2013**, *9*, 687−697.

(23) López, C. A.; Rzepiela, A. J.; De Vries, A. H.; Dijkhuizen, L.; Hünenberger, P. H.; Marrink, S. J. Martini Coarse-Grained Force Field: Extension to Carbohydrates. *J. Chem. Theory Comput.* **2009**, *5*, 3195−3210.

(24) Shivgan, A. T.; Marzinek, J. K.; Huber, R. G.; Krah, A.; Henchman, R. H.; Matsudaira, P.; Verma, C. S.; Bond, P. J. Extending

the Martini Coarse-Grained Force Field to N-Glycans. *J. Chem. Inf. Model.* **2020**, *60*, 3864−3883.

(25) Grünewald, F.; Punt, M. H.; Jefferys, E. E.; Vainikka, P. A.; König, M.; Virtanen, V.; Meyer, T. A.; Pezeshkian, W.; Gormley, A. J.; Karonen, M.; Sansom, M. S. P.; Souza, P. C. T.; Marrink, S. J. Martini 3 Coarse-Grained Force Field for Carbohydrates. *J. Chem. Theory Comput.* **2022**, *18*, 7555−7569.

(26) Uusitalo, J. J.; Ingólfsson, H. I.; Akhshi, P.; Tieleman, D. P.; Marrink, S. J. Martini Coarse-Grained Force Field: Extension to DNA. *J. Chem. Theory Comput.* **2015**, *11*, 3932−3945.

(27) Uusitalo, J. J.; Marrink, S. J.; Faustino, I. Martini Coarse-Grained Force Field: Extension to RNA. *Biophys. J.* **2017**, *113*, 246−256.

(28) Rossi, G.; Giannakopoulos, I.; Monticelli, L.; Rostedt, N. K. J.; Puisto, S. R.; Lowe, C.; Taylor, A. C.; Vattulainen, I.; Ala-Nissila, T. A MARTINI Coarse-Grained Model of a Thermoset Polyester Coating. *Macromolecules* **2011**, *44*, 6198−6208.

(29) Rossi, G.; Monticelli, L.; Puisto, S. R.; Vattulainen, I.; Ala-Nissila, T. Coarse-graining polymers with the MARTINI force-field: polystyrene as a benchmark case. *Soft Matter* **2011**, *7*, 698−708.

(30) Grunewald, F.; Rossi, G.; de Vries, A. H.; Marrink, S. J.; Monticelli, L. Transferable MARTINI Model of Poly(ethylene Oxide). *J. Phys. Chem. B* **2018**, *122*, 7436−7449.

(31) Wong-Ekkabut, J.; Baoukina, S.; Triampo, W.; Tang, I.-M.; Tieleman, D. P.; Monticelli, L. Computer simulation study of fullerene translocation through lipid membranes. *Nat. Nanotechnol.* **2008**, *3*, 363−368.

(32) Baoukina, S.; Monticelli, L.; Tieleman, D. P. Interaction of Pristine and Functionalized Carbon Nanotubes with Lipid Membranes. *J. Phys. Chem. B* **2013**, *117*, 12113−12123.

(33) Marrink, S. J.; Monticelli, L.; Melo, M. N.; abd, D.; Peter Tieleman, R. A.; Souza, P. C. T. Two decades of Martini: Better beads, broader scope. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2022**, No. e1620.

(34) Alessandri, R.; Grünewald, F.; Marrink, S. J. The Martini Model in Materials Science. *Adv. Mater.* **2021**, *33*, No. e2008635.

(35) Souza, P. C. T.; Thallmair, S.; Conflitti, P.; Ramírez-Palacios, C.; Alessandri, R.; Raniolo, S.; Limongelli, V.; Marrink, S. J. Protein-ligand binding with the coarse-grained Martini model. *Nat. Commun.* **2020**, *11*, 3714.

(36) Alessandri, R.; Barnoud, J.; Gertsen, A. S.; Patmanidis, I.; de Vries, A. H.; Souza, P. C. T.; Marrink, S. J. Martini 3 Coarse-Grained Force Field: Small Molecules. *Adv. Theory Simul.* **2022**, *5*, 2100391.

(37) Vazquez-Salazar, L. I.; Selle, M.; de Vries, A. H.; Marrink, S. J.; Souza, P. C. T. Martini coarse-grained models of imidazolium-based ionic liquids: from nanostructural organization to liquid−liquid extraction. *Green Chem.* **2020**, *22*, 7376−7386.

(38) Vainikka, P.; Thallmair, S.; Souza, P. C. T.; Marrink, S. J. Martini 3 Coarse-Grained Model for Type III Deep Eutectic Solvents: Thermodynamic, Structural, and Extraction Properties. *ACS Sustainable Chem. Eng.* **2021**, *9*, 17338−17350.

(39) Tsanai, M.; Frederix, P. J. M.; Schroer, C. F. E.; Souza, P. C. T.; Marrink, S. J. Coacervate formation studied by explicit solvent coarse-grain molecular dynamics with the Martini model. *Chem. Sci.* **2021**, *12*, 8521−8530.

(40) Pezeshkian, W.; König, M.; Wassenaar, T. A.; Marrink, S. J. Backmapping triangulated surfaces to coarse-grained membrane models. *Nat. Commun.* **2020**, *11*, 2296.

(41) Vickery, O. N.; Stansfeld, P. J. CG2AT2: an Enhanced Fragment-Based Approach for Serial Multi-scale Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2021**, *17*, 6472−6482.

(42) Grünewald, F.; Alessandri, R.; Kroon, P. C.; Monticelli, L.; Souza, P. C. T.; Marrink, S. J. Polyply; a python suite for facilitating simulations of macromolecules and nanomaterials. *Nat. Commun.* **2022**, *13*, 68.

(43) Qi, Y.; Ingólfsson, H. I.; Cheng, X.; Lee, J.; Marrink, S. J.; Im, W. CHARMM-GUI Martini Maker for Coarse-Grained Simulations with the Martini Force Field. *J. Chem. Theory Comput.* **2015**, *11*, 4486−4494.

(44) Hsu, P.-C.; Bruininks, B. M. H.; Jefferies, D.; de Souza, P. C. T.; Lee, J.; Patel, D. S.; Marrink, S. J.; Qi, Y.; Khalid, S.; Im, W. CHARMM-GUI Martini Maker for modeling and simulation of complex bacterial membranes with lipopolysaccharides. *J. Comput. Chem.* **2017**, *38*, 2354−2363.

(45) Bereau, T.; Kremer, K. Automated Parametrization of the Coarse-Grained Martini Force Field for Small Organic Molecules. *J. Chem. Theory Comput.* **2015**, *11*, 2783−2791.

(46) Graham, J. A.; Essex, J. W.; Khalid, S. PyCGTOOL: Automated Generation of Coarse-Grained Molecular Dynamics Models from Atomistic Trajectories. *J. Chem. Inf. Model.* **2017**, *57*, 650−656.

(47) Empereur-Mot, C.; Pesce, L.; Doni, G.; Bochicchio, D.; Capelli, R.; Perego, C.; Pavan, G. M. Swarm-CG: Automatic Parametrization of Bonded Terms in MARTINI-Based Coarse-Grained Models of Simple to Complex Molecules via Fuzzy Self-Tuning Particle Swarm Optimization. *ACS Omega* **2020**, *5*, 32823−32843.

(48) Potter, T. D.; Barrett, E. L.; Miller, M. A. Automated Coarse-Grained Mapping Algorithm for the Martini Force Field and Benchmarks for Membrane−Water Partitioning. *J. Chem. Theory Comput.* **2021**, *17*, 5777−5791.

(49) Kroon, P. C.; Grünewald, F.; Barnoud, J.; van Tilburg, M.; Souza, P. C. T.; Wassenaar, T. A.; Marrink, S.-J. Martinize2 and Vermouth: Unified Framework for Topology Generation. *arXiv* **2022**, DOI: 10.48550/arXiv.2212.01191.

(50) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.

(51) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput. Biol.* **2017**, *13*, No. e1005659.

(52) Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V.; Schulten, K.; Chipot, C.; Tajkhorshid, E. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **2020**, *153*, 044130.

(53) Zhang, X.; Sundram, S.; Oppelstrup, T.; Kokkila-Schumacher, S. I. L.; Carpenter, T. S.; Ingólfsson, H. I.; Streitz, F. H.; Lightstone, F. C.; Glosli, J. N. ddcMD: A fully GPU-accelerated molecular dynamics program for the Martini force field. *J. Chem. Phys.* **2020**, *153*, 045103.

(54) Rusu, V. H.; Horta, V. A. C.; Horta, B. A. C.; Lins, R. D.; Baron, R. MDWiZ: a platform for the automated translation of molecular dynamics simulations. *J. Mol. Graphics Modell.* **2014**, *48*, 80−86.

(55) Shirts, M. R.; Klein, C.; Swails, J. M.; Yin, J.; Gilson, M. K.; Mobley, D. L.; Case, D. A.; Zhong, E. D. Lessons learned from comparing molecular dynamics engines on the SAMPL5 dataset. *J. Comput.-Aided Mol. Des.* **2017**, *31*, 147−161.

(56) Lesch, V.; Diddens, D.; Bernardes, C. E. S.; Golub, B.; Dequidt, A.; Zeindlhofer, V.; Sega, M.; Schröder, C. ForConX A forcefield conversion tool based on XML. *J. Comput. Chem.* **2017**, *38*, 629−638.

(57) Periole, X.; Cavalli, M.; Marrink, S.-J.; Ceruso, M. A. Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition. *J. Chem. Theory Comput.* **2009**, *5*, 2531−2543.

(58) Poma, A. B.; Cieplak, M.; Theodorakis, P. E. Combining the MARTINI and Structure-Based Coarse-Grained Approaches for the Molecular Dynamics Studies of Conformational Transitions in Proteins. *J. Chem. Theory Comput.* **2017**, *13*, 1366−1374.

(59) Wołek, K.; Gómez-Sicilia, Àngel; Cieplak, M. Determination of contact maps in proteins: A combination of structural and chemical approaches. *J. Chem. Phys.* **2015**, *143*, 243105.

(60) Moreira, R. A.; Chwastyk, M.; Baker, J. L.; Guzman, H. V.; Poma, A. B. All atom simulations snapshots and contact maps analysis

scripts for SARS-CoV-2002 and SARS-CoV-2 spike proteins with and without ACE2 enzyme. (https://zenodo.org/record/3817447)

(61) Herzog, F. A.; Braun, L.; Schoen, I.; Vogel, V. Improved Side Chain Dynamics in MARTINI Simulations of Protein−Lipid Interfaces. *J. Chem. Theory Comput.* **2016**, *12*, 2446−2458.

(62) Monticelli, L. On Atomistic and Coarse-Grained Models for C60 Fullerene. *J. Chem. Theory Comput.* **2012**, *8*, 1370−1378.

(63) Morton, A.; Matthews, B. W. Specificity of ligand binding in a buried nonpolar cavity of T4 lysozyme: Linkage of dynamics and structural plasticity. *Biochemistry* **1995**, *34*, 8576−8588.

(64) Grünewald, F.; Souza, P. C. T.; Abdizadeh, H.; Barnoud, J.; de Vries, A. H.; Marrink, S. J. Titratable Martini model for constant pH simulations. *J. Chem. Phys.* **2020**, *153*, 024118.

(65) Rose, A. S.; Hildebrand, P. W. NGL Viewer: a web application for molecular visualization. *Nucleic Acids Res.* **2015**, *43*, W576−579.