Martini 3 building blocks for Lipid Nanoparticle design

Lisbeth Ravnkilde Kjølbye, Mariana Valério, Markéta Paloncýová, Luís Borges-Araújo, Roberto Pestana-Nobles, Fabian Grünewald, Bart Bruininks, Rocío Araya-Osorio, Martin Šrejber, Raul Mera-Adasme, et al.

Martini 3 building blocks for Lipid Nanoparticle design

Lisbeth R. Kjølbye[1], Mariana Valério[1,2,3†], Markéta Paloncýová[4†], Luís Borges-Araújo[1,2,3†], Roberto Pestana-Nobles[1,5], Fabian Grünewald[6], Bart M. H. Bruininks[7], Rocío Araya-Osorio[8], Martin Šrejber[4], Raul Mera-Adasme[8], Luca Monticelli[1], Siewert J. Marrink[9], Michal Otyepka[4,10], Sangwook Wu[11,12*], Paulo C.T. Souza[1,2,3*]

[1]*Molecular Microbiology and Structural Biochemistry (MMSB, UMR 5086), CNRS and University of Lyon, Lyon, France.*
[2] *Laboratoire de Biologie et Modélisation de la Cellule, CNRS, UMR 5239, Inserm, U1293, Universite Claude Bernard Lyon 1, Ecole Normale Supérieure de Lyon, 46 allée d'Italie, 69364, Lyon, France.*
[3] *Centre Blaise Pascal de Simulation et de Modélisation Numérique, Ecole Normale Supérieure de Lyon, 47 allée d'Italie, 69364, Lyon, France.*
[4]*Regional Centre of Advanced Technologies and Materials, Czech Advanced Technology and Research Institute (CATRIN), Palacký University Olomouc, Šlechtitelů 27, Olomouc 779 00, Czech Republic.*
[5]*Laboratorio de Simulación Molecular y Bioinformática, Centro de Investigación en Ciencias de la Vida (CICV), Universidad Simón Bolívar, Carrera 59 59-65, Barranquilla 080002, Colombia.*
[6]*Heidelberg Institute for Theoretical Studies (HITS), Heidelberg, Germany.*
[7]*Department of Physics, University of Helsinki, P.O. Box 64, FI-00014 Helsinki, Finland.*
[8]*Departamento de Quimica, Facultad de Ciencias, Universidad de Tarapacá, Av. Gral. Velasquez 1775, Arica, Chile.*
[9]*Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, the Netherlands.*
[10]*IT4Innovations, VŠB - Technical University of Ostrava, 17. listopadu 2172/15, Ostrava-Poruba 708 00, Czech Republic.*
[11]*PharmCADD, 1102 ho, 60, Centum jungang-ro, Haeundae-gu, Busan, 48059, Republic of Korea.*
[12]*Department of Physics, Pukyong National University 48513, Busan, Republic of Korea.*

[*]*Corresponding authors: s.wu@pharmmcadd.com, paulo.telles_de_souza@ens-lyon.fr*
[†]*These authors contributed equally.*

Abstract: Lipid nanoparticles (LNPs) represent a promising platform for advanced drug and gene delivery, yet optimizing these particles for specific cargos and cell targets poses a complex, multifaceted challenge. Furthermore, there is a pressing need for a more comprehensive understanding of the underlying technology. Experimental studies are costly and often provide low-resolution information. Molecular dynamics (MD) simulations allow us to study these particles at a higher resolution, enhancing our understanding. However, studying these systems at atomic resolutions is both challenging and computationally expensive, as well as time-consuming. Coarse-grained (CG) models, such as Martini 3, are positioned as promising tools for studying LNPs. To enable CG-MD studies of LNPs, accurate and validated models of their components are needed. Here, we present a substantial extension of the Martini 3 library of lipids, covering the most important LNP components, including over a hundred of ionizable lipid (IL) models, along with natural occurring sterol models and PEGylated lipid models. We furthermore present initial protocols for screening fusion efficacy across different lipid formulations and for constructing whole LNPs at CG resolution, enabling future studies of these nanoparticles.

1. Introduction

The interest in lipid nanoparticles (LNPs) has increased tremendously because of their large potential in vaccines as a delivery platform. However, the understanding of this technology is still substantially lacking and remains a multivariate problem, making it difficult to investigate and decouple specific variables. An LNP typically consists of five major components, an ionizable

lipid (IL), a sterol, a helper phospholipid, polyethylene glycol conjugated (or PEGylated) lipids, and cargo. In addition to optimizing each component, the relative ratio between them can also be optimized depending on the cargo, target cell specificity or stability, not to mention the assembly process [1, 2]. Although experimental studies can explore certain variations under certain conditions, extensive screening of ratios and components is neither easy nor cheap, especially when non-commercially available lipids need to be synthesized and tested [3]. Additionally, experiments often lack high-resolution structural and dynamical information about the architecture and phase changes of LNPs under different pH values [4, 5].

The currently accepted mode of action for LNP delivery is that, upon injection into the bloodstream, a corona shell is formed around the particle, where the APO E peptide is an essential component in addition to albumin. APO E is responsible for LNP endocytosis into the cell. The resulting endosome formed with the LNP inside will then mature, resulting in a pH drop from 7 to around 6-5. Upon the pH drop, several of the ILs in the LNP become protonated which is believed to induce structural changes in the LNP that enable particle fusion with the endosomal membrane and the eventual release of its cargo into the cytosol. In the case of mRNA cargo, as in the COVID vaccines, the mRNA is then transcribed by the ribosome [6, 7].

Several LNP properties are optimized to improve their potency, such as particle and cargo stability in storage and/or blood, fusion efficacy in relation to endosomal escape, or even target-cell specificity [2, 6, 8]. Regarding LNP stability, PEGylated lipids are often used to improve particle 'stealthiness' in terms of stability in the blood, and to help maintain a certain particle size while in storage [9]. The main instability, however, has been identified as originating from the cargo, whether RNA or DNA [10]. Given the extensive experimental data readily available on RNA degradation rates, AI-based approaches are already a reality for the computational design of new RNA sequences [11]. However, the situation is different for LNP lipid composition and its influence on vaccine efficacy. Likely due to the lack of large, curated databases, there are currently no well-established AI-based protocols for optimizing lipid composition, despite its clear significance in enhancing LNP properties. Only initial attempts at *in silico* screening lipid compositions have been made [12].

An additional issue with current LNP technology is that, only a very small percentage of the cargo, in the case of RNA, actually reaches the cytosol of target cells. One of the biggest hurdles is the cargo endosomal escape. To optimize this crucial process, it is necessary to enhance LNP fusion with the endosome while avoiding excessive fusogenicity with other lipid compartments [13]. Understanding the pH-dependent phase behavior of LNPs is essential for optimizing endosomal escape [1]. Despite several studies that attempt to shed light on the matter, a comprehensive understanding is still lacking due to the multivariate nature of LNPs [5, 14–18]. Overall, the literature still lacks a deep molecular and functional understanding of LNP technology. Although some small details have been revealed, the puzzle remains incomplete and requires further integration of experimental and computational approaches to achieve a clear understanding of the whole picture.

Molecular dynamics (MD) simulations are a valuable tool for studying, in detail, the packing behavior and organization of LNPs, as well as their endosomal release process. Combining MD studies of LNPs with available experimental data could provide valuable high resolution structural insight that would be otherwise inaccessible [15, 19, 20]. For example, MD simulations have suggested that optimizing the packing of the Moderna ILs around the RNA nucleobases via π-π stacking may enhance cargo stability [21]. Such computational strategies could be used to further optimize and advance LNP technology. However, the sheer complexity and size of LNPs puts them beyond the reach of traditional spatial and time scales currently accessible to atomistic MD simulations, posing challenges in terms of computational resources. This is where coarse grained (CG) MD simulations can play a major role, alleviating these limitations [1, 22, 23].

The Martini CG force field (FF) [24, 25] is a widely used CG model, in particular for MD

simulations involving lipids in complex biological environment [26] and their interplay with other biomolecules and synthetic materials [25, 27]. Martini enhances sampling efficiency and reduces system complexity by representing two to four non-hydrogen atoms, along with their associated hydrogens, as a single CG bead. This model employs a building block approach and integrates both bottom-up and top-down methods to derive bonded and non-bonded interactions, respectively. A sub-selection of previous works have highlighted the possibilities for delivery systems being studied with with Martini force field, examples ranging from simulating the escape of DNA from a lipoplex [23] to studying the lipid nano-emulation for vitamin delivery [28]. Some wok has also been on LNPs, studying the core and pH dependency [22, 29]. The latest version, Martini 3, significantly improves interaction accuracy and greatly expands chemical space coverage [30] making it a promising new tool for designing and optimizing LNPs, which was not as feasible with earlier models [1]. In this work, we introduce an extensive and new Martini 3 library that includes ILs, sterols, and PEGylated lipids. Additionally, we provide detailed protocols for exploring fusion processes, investigating pH-dependent phase behavior of LNPs, and constructing and analyzing whole LNPs in solution or during interactions with target membranes. This work aims to lay the foundation for computational design and screening of LNP formulations using Martini CG MD simulations.

2. Results

2.1. LNP lipid models

2.1.1. Ionizable lipid parameterization - a fragment based approach

To build Martini 3 CG models of ionizable lipids, a building block approach was employed (Figure 1). This approach consists of modeling several IL fragments – representing various regions of the IL, such as the headgroup, linker region or acyl chain – which can then be assembled to model entire ILs (Figure 1). Not only does this simplify the parameterization process, but it also allows researchers to use these fragment models to build other IL models or get inspiration for new ones, making it easier to study different LNPs by changing specific chemical groups within the lipids. The naming convention adopted for the two-tailed IL models consists of the two-letter combination for the tails (typically used for phospholipids), followed by a two-letter combination for the head/linker region. For the head, the first letter remains the same across protonation states, while the second letter changes to "P" for the charged protonated state (see Figure 1A).

To obtain the necessary parameters for the IL CG models, atomistic MD simulations were conducted for each fragment. These simulations provided the bonded parameters and solvent accessible surface area (SASA) values essential for accurate model construction. In the absence of experimental data for the fragments, we also used COSMO-RS, a quantum chemistry-based method, to predict the thermodynamic properties [31]. This approach has consistently been recognized as one of the best prediction tools for oil/water partitioning data [32, 33]. An additional advantage of COSMO-RS is its ability to predict logP for not only the octanol/water but also for other partitioning systems, which is crucial for the chemical transferability of the Martini CG models. In this work, we used all three gold standard partitioning systems employed in Martini 3 parametrization: octanol/water, hexadecane/water, and chloroform/water. We also ensured consistency across both the fragment library and the existing phospholipid library. For some fragments, compromises on the CG partitioning free energy were necessary to maintain consistency and transferability within the library. It is worth noting that COSMO-RS typically has an mean absolute error (MAE) range of 2-3 kJ/mol [32, 33], which justifies our preference for maintaining internal consistency with the Martini bead assignment rather than perfectly matching the free energies.

Figure 1B shows the matching of CG and atomistic SASA, indicating an excellent agreement, which reflect not only the accurate fitting of bond distance but also the choices of bead sizes. The

overall comparison between COSMO-RS and Martini 3 CG partitioning free energies showed a reasonable agreement with a R^2 = 0.80 and MAE = 6.4 kJ/mol (Figure 1C). The matching between the approaches was worse for larger, branched hydrophobic fragments, such as some of the linker fragments containing several carbons and one hydrophilic group such as an ester or ether. These discrepancies are expected as experimentally it is impossible to accurately estimate these values, with any predictions becoming less trustworthy [34]. Additionally, as already mentioned, there is an inherent error of the COSMO-RS estimates. Therefore, the main focus for selecting the types of beads was to maintain consistency between fragments, chemical groups, and other Martini 3 lipid models. Hence, the difference in partitioning free energy between the CG models and predicted values, is not an issue for the final IL models. The behavior of the ILs was later compared with atomistic models to further validate the choices.

In total 15 head and 10 linker fragments were made along with 2 fragments for constructing PEGylated lipids. When combined with the available tail parameters, this setup allows for the generation of hundred of different lipids. As an example, we provided a total of 117 unique coarse-grained (CG) models of ILs, each in both neutral and charged states (total of 234 parameter files). Furthermore, additional IL models featuring branched tails were constructed, including the well-known ALC-0315 and SM-102, which are used in Pfizer and Moderna vaccines, respectively [35, 36].

2.1.2. Validation of bonded terms from fragments when used as part of ILs

The bonded parameters obtained for the fragments were further validated with four ILs from the literature, for which CHARMM36 parameters are available [20]. We focus on the two widely known ILs, DLMC3 and DLKC2, which will henceforth be referred to as DUM3 / DUMP and DUK2 / DUKP, depending on the protonation state, or simply MC3 and KC2. Using the CHARMM36 FF [20, 37], these two IL types were simulated in a IL:POPC bilayer (30:70 mol%) during 1 μs for both the fully charged and neutral state. The corresponding CG systems were simulated for 1 μs as well, and the bonded distributions were compared between the atomistic and CG simulations. See Figure S2-11 for the SASA and bonded distributions for DUM3/MP and DUK2/KP, the rest can be found on the GitHub (https://github.com/Martini-Force-Field-Initiative/M3-Ionizable-Lipids/). This comparison aimed to verify that the bonded parameters derived from the OPLS fragment simulations match well with those obtained from whole ILs, albeit obtained using a different atomistic force field. In addition to these ILs, Amber FF simulations of Pfizer (ALC-0315) and Moderna (SM-102) from the work of Paloncová *et al*. [16] were also used to cross-validate their CG equivalents. The bonded and SASA distributions showed a good overlap between the atomistic lipid, fragment and coarse grained lipid model.

2.1.3. Increasing the Martini 3 LNP toolbox with sterols and PEGylated lipids

Optimizing LNPs presents a multi-variable challenge. Typically, the composition of LNPs consists of a base mixture comprising an ionizable lipid, a sterol, a helper lipid like DSPC or PE, and a PEGylated lipid. [4] Several studies have focused on optimizing the ionizable lipid, with the emphasis on apparent pKa, fusion efficacy and biodegradability [38–40]. Other studies have investigated the role of the helper lipid in LNP structure and fusion [41]. Recent studies are now beginning to reveal the impact of replacing commonly used cholesterol with alternative sterols, such as plant sterol β -sitosterol. These investigations demonstrate that such substitutions not only alter LNP properties like shape and structure but can also markedly improve fusion efficacy [42–44].

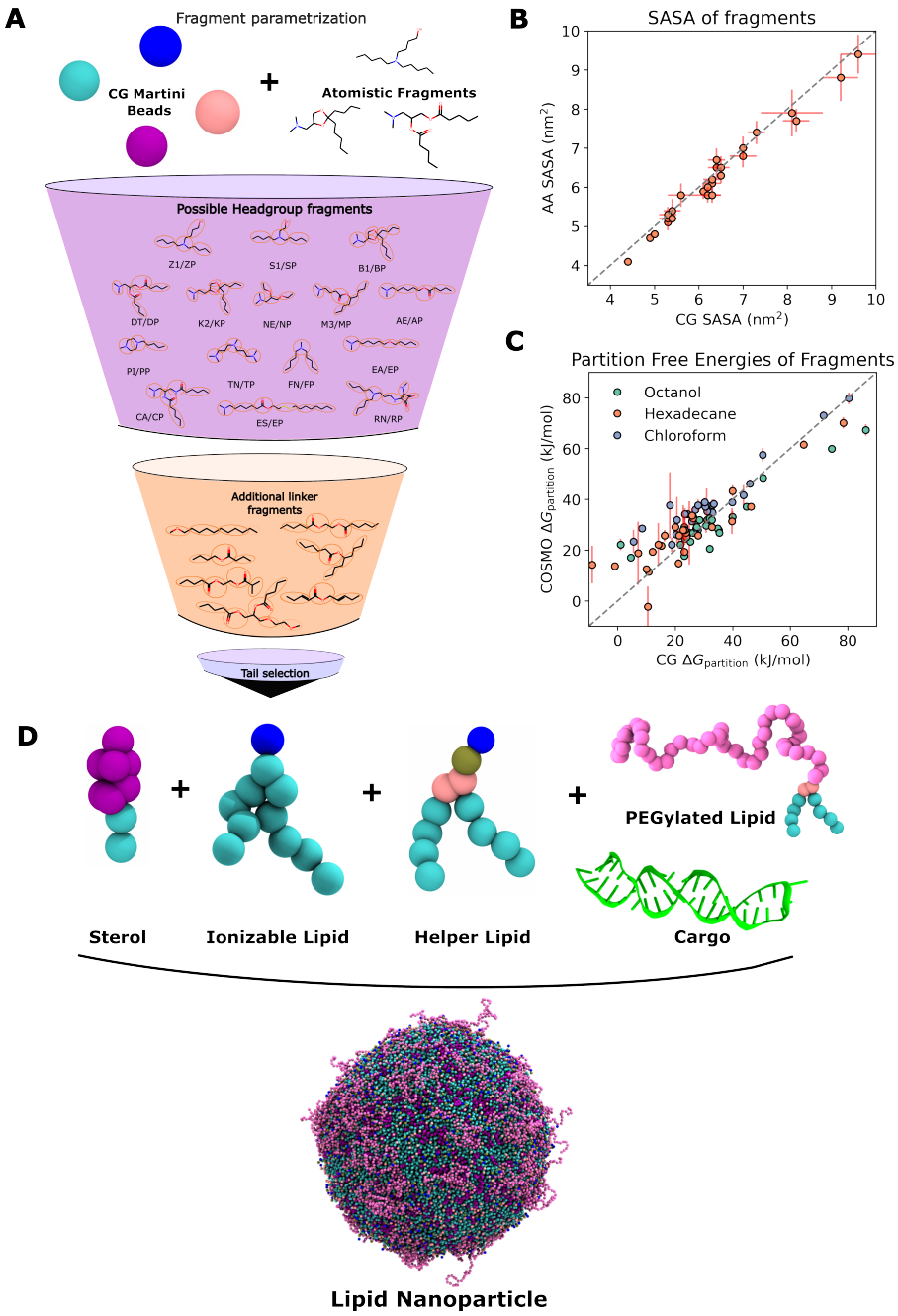

Fig. 1. Representing the fragment-based approach to construct several ionizable lipid models. A) The funnel illustrates the fragment-based approach and the step-by-step process for constructing an ionizable lipid. It begins with headgroup fragments and their corresponding names in the larger section of the funnel, followed by additional linker fragments that can be incorporated. Finally, the tails are attached to complete the lipid structure. B) Solvent Accessible Surface Area (SASA) for atomistic vs coarse grained fragments. C) COSMO-RS vs CG calculated partitioning free energies for the coarse grained fragments. D) The chemical structure of the ALC-0315 ionizable lipid is depicted, with the head, linker, and tails regions highlighted to correspond with the colors shown in the funnel above. Additionally, examples of other key components of a LNP are presented, including sterol, helper lipid, PEGylated lipid, and RNA cargo.

We expanded our library with four naturally occurring sterols (Ergosterol, Campesterol, β-sitosterol, and Stigmasterol), modeled based on the new cholesterol model for Martini 3 [45]. The sterol models were simulated in DSPC bilayers and compared with cholesterol, ensuring a correct trend in terms of packing, insertion, and melting temperature relative to the experimental data. See Figure S12-24 in the supporting information for the bonded distributions along with these properties mentioned above. In addition to sterols, we have also included PEGylated lipid models in the Martini 3 LNP toolbox. Leveraging the versatility of Polyply, [46] adjustable PEG head lengths are available, which could be combined with the Martini 3 tail library and with fragments exclusively generated to facilitate the modeling of DMG-PEG and ALC-0159. As an example, we provide a model of DMG-PEG 2000 and ALC-0159. It is important to note, however, that the use of PEG in Martini 3 currently requires some adjustments in the interactions between SN3r and charged (Q) beads [47].

2.2. Validation of the ionizable lipid models

2.2.1. Reproducing IL charge-dependent lamellar/oil lens behavior

Generally, in bilayer simulations involving helper lipids and different combinations of head/linker and tail compositions from ILs, it is expected that neutral species will accumulate in the middle, forming an oil lens between the two leaflets and leading to a significant increase in membrane thickness. To demonstrate that our model can also capture this effect, we performed Martini simulations of DLMC3 and DLKC2 (DUM3 / DUMP and DUK2 / DUKP) when mixed with POPC (10:90, 20:80, and 30:70 IL:POPC mol%). CHARMM36 atomistic simulations were performed of bilayers at the same concentrations of ILs for comparison. These analyses were performed in either the fully charged or neutral state, representing low- and high-pH situations, respectively.

Figure 2A and S25 shows the density across the bilayer normal for the DUM3/MP and DUK2/KP headgroups in both the atomistic and corresponding CG systems. At low pH the IL is assumed fully charged, while at high pH it is neutral. As confirmed by the snapshots of the simulations highlighted in Figure 2B, the IL maintains a lamellar topology at low pH (charged headgroup), while at high pH (neutral headgroup) it accumulates in between the POPC leaflets, generating the expected oil lens. This behaviour aligns well with the knowledge in the literature, where neutral MC3 and cholesterol are concluded to form an oil core within LNPs, while multilamellar and inverse hexagonal phases are proposed to form under low pH when ILs are charged [14, 18]. Additional results from atomistic simulations also indicate the same behavior [19, 20].

Notably, the CG models of neutral MC3 (DUM3) and KC2 (DUK2) tend to aggregate slightly more in the interleaflet region compared to the corresponding atomistic simulations. Several parameterizations of MC3 (DUM3) with the CHARMM and Slipids force field have been presented in the literature and compared to experimental data such as Neutron Reflectrometry [48]. For the CHARMM parameters the accumulation of the neutral specie in the interleaflet is underestimated. [48] The bilayer thickness of neutral MC3 in DOPC bilayers at 10 and 15 mol% IL (50 mM NaCl, 288 K) was experimentally determined to be 41 ± 3 and 40 ± 3 Å [48]. For comparison we simulated the same compositions of MC3:DOPC under the same conditions and measured a bilayer thickness of 38.20 ± 0.27 and 39.61 ± 0.31 Å, respectively, aligning well with the experimentally *in vitro* determined values. Hence the increased accumulation of the neutral species observed in the CG simulations may not match entirely with the atomistic simulations but rather appears to agree well with the *in vitro* experimental data. These findings validate our model's ability to replicate the expected behavior across IL protonation states.

2.2.2. Expanding IL Model Validation: Effects of Headgroup, Linker, and Tail Variations

To further validate our IL models, we took advantage of our building block approach for constructing ILs and extended our tests to a larger sub-selection of headgroups, linkers, and tails in 15×15 nm^2 IL:DSPC (50:50 mol%) bilayer patches. From these simulations, we calculated the area per lipid (APL) and the hydrophobic thickness, specifically the distances between the two leaflets excluding the polar headgroups. See Figure S26 for the head/linkers and tails under study along with the APL and thickness values. The APL was only calculated for the charged species that maintained the lamellar topology. A trend was observed across the lipid tails, with the DU tail (CDDC CDDC) having the highest APL. The fully saturated tails DL and DS consistently showed lower and similar APL across the lipid types (see Figure S26 for more details). This trend aligns with observations in common phospholipids, where polyunsaturated tails exhibit higher APL compared to saturated ones [49]. The choice between saturated and unsaturated tails for LNP usage is a topic of discussion in the literature; unsaturated tails, despite being prone to oxidation and presenting an instability issue, have demonstrated a significant improvement in LNP transfection efficacy [50, 51].

2.2.3. pH-dependent phase behaviour reproduced

It has been suggested that LNP organization depends not only on their lipid composition and cargo but also on pH of the environment. [5, 14, 43, 52]. LNPs containing siRNA or mRNA are proposed to have a multilamellar [18] or inverse hexagonal / worm-like internal structure [5, 14], respectively. The lamellar topology has been shown to only be maintained with charged MC3 (DUMP) based on SAXS (Small Angle X-ray Scattering) data [18]. The following question is whether the same ILs can form an inverse hexagonal phase, given the right conditions, such as pH [5].

Experimental studies [5, 14] have found that both MC3 and KC2 form an inverse hexagonal phase at low pH with a hydration of around 7–12 water molecules per lipid. They also propose that the core of LNPs consists of around 75 mol% IL and 25 mol% cholesterol. The presence of cholesterol has previously been suggested to help maintain the inverted hexagonal phase of the LNPs based on atomistic MD simulations [17]. The tests for formation of inverse hexagonal phase were therefore performed with KC2 and MC3 in the presence of either DSPC or cholesterol.

Double bilayer systems were constructed of both MC3 (DUM3/DUMP) and KC2 (DUK2/DUKP) with 10 mol% DSPC or with 25 mol% cholesterol and low hydration (around 7 water molecules per lipid) and neutralizing ions, representing low and high pH, assuming fully charged or neutral lipids, based on the experimental conditions in the work by Philipp *et al.* [5].

Indeed, both MC3 and KC2 form the inverse hexagonal phase in with both DSPC and cholesterol, however only at low pH, assuming a plus one charge for all the ILs. The inverse hexagonal phase does however not seem to depend on the presence or absence of either DSPC nor cholesterol, see Figure 2C and D. An additional test was performed with only MC3 and KC2 at low pH, and in fact the inverse hexagonal phase was formed as well. The simulations reach the same conclusion as obtained experimentally [5], that the phase is determined by the ionizable lipid and not the helper lipid DSPC nor cholesterol.

At high pH, the neutral MC3 or KC2 does not form any topology resembling an inverse hexagonal phase. In the presence of DSPC small inverse micelle pockets are formed, with the same being observed for the cholesterol-containing systems, though there, the neutral IL mixes with cholesterol and separates from the water phase completely, see Figure 2C and D. This aligns well with experimental SAXS data [5, 14], suggesting that the parameters can indeed reproduce the known pH-dependent phase behaviour of these ILs.

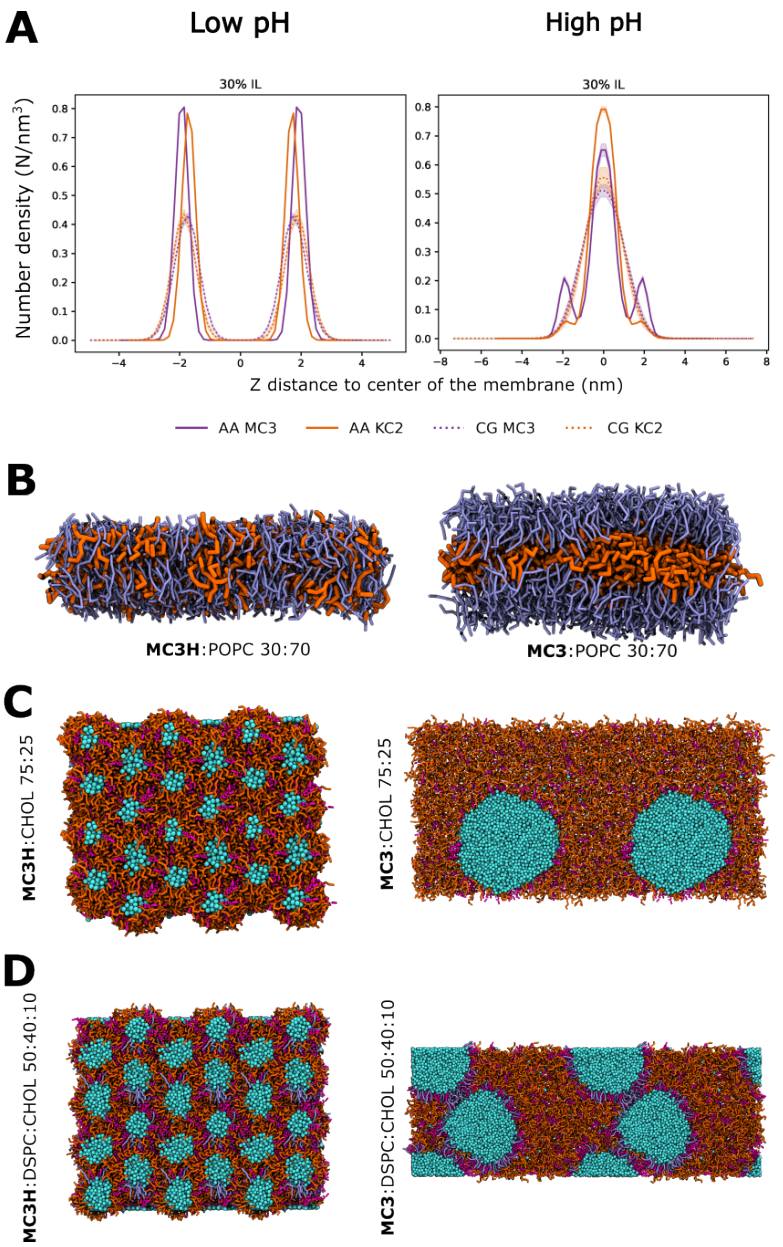

Fig. 2. A) Density of headgroups of MC3 (DUM3) and KC2 (DUK2) across the membrane normal at low and high pH. B) Snapshot of the bilayers containing MC3 (in orange) and POPC lipids (in purple) at high and low pH. C) Snapshots of systems containing MC3:Cholesterol (75:25) forming an inverted hexagonal phase and inverse micellar phase at low and high pH, respectively. D) Snapshots of systems containing MC3:DSPC:Cholesterol (50:40:10) forming an inverted hexagonal phase and inverse micellar phase at low and high pH, respectively. MC3 and MC3H lipids are shown in orange, DSPC in purple, cholesterol in pink and the water bead in cyan.

2.3. *Towards in silico screening of LNP formulations: fusion assays*

A large library of IL models can enable various studies, including the screening and characterization of different LNP lipid formulations, in terms of their packing, phase behaviour, and fusiogenic behaviour against different membrane models. To illustrate the potential of this IL model library and to propose an approach for assessing LNP efficiency, we focus here on quantifying the fusiogenic properties of different formulations. For this purpose, we adapted the protocol by Hub *et al.* [53] to perform Potential of Mean Force (PMF) fusion assays of stalk formation between different LNP formulations and a plasma membrane model [54]. The LNP interacting leaflet of this membrane is composed of (in mol%) 24 % POPC, 12 % PAPC, 2 % POPE, 6 % DIPE, 24 % DPSM and 31 % CHOL while the non-interacting leaflet contains 14 % POPC, 8 % PAPC, 5 % POPE, 16 % DIPE, 11 % DPSM, 28 % CHOL, 25 % PAPS, 2 % PAP6. Additionally, we also studied the stalk formation between the LNP formulations and a membrane model enriched with 10 mol % PAPS lipids in the LNP-interacting leaflet to better mimic the compositional differences observed in the endosome. PAPS was also included in this leaflet based on the hypothesis that anionic lipids, such as PS, could play an important role in the transfection mechanism, and to better represent the increase in negatively charged lipid found in late endosomes [55, 56].

The initial barrier to membrane fusion is the formation of a hemifusion stalk, which should be followed by the formation of a fusion pore [23, 57, 58], process not covered by our fusion assays. Here we use the PMF method to calculate the free energy landscape of the fusion intermediates, providing insights into the energetic barriers and stability of the stalk. Figure 3A presents a schematic overview of the process by which the LNP enters the cell and leads to protein expression, with stalk formation highlighted in Figure 3B as the proposed initial barrier for endosomal escape.

2.3.1. Acyl-chain dependent fusion efficacy

The experimental work by Sabnis *et al.* [36] marked a pivotal point in the optimization of the IL chemistry, focusing on enhancing protein expression and subsequent fusion efficacy. From the well-established MC3 lipid, they derived nine additional ILs. Compared to MC3, one derivative performed worse (Lipid 2), while others exhibited significantly enhanced performance (such as Lipid 10) in terms of protein expression. The distinctions between Lipid 2, 10, and MC3 lie in increasing saturation and branching of the tails, coupled with the introduction of an additional ester group, which makes Lipid 10 more biodegradable [36]. In an attempt to reproduce these experimental trends, we performed a fusion PMF assay with bilayers mimicking LNP formulations (IL:CHOL:DSPC at 50:40:10 mol% ratio), were MC3, Lipid 2, or Lipid 10 where the ILs studied, and a plasma membrane [54], as mentioned above.

These fusion PMFs (Figure3D) revealed a substantial decrease in the free energy of stalk formation for Lipid 10 compared to MC3, and a slight increase for Lipid 2 compared to MC3, consistent with the protein expression trend reported by Sabnis *et al.* [36]. Other studies have also shown that replacing alkenes with alkanes and introducing esters and branching of the lipids tails enhances delivery, since it helps in lipid protrusion, enabling the lipid to more easily overcome the dehydration barrier between the two membranes, and thereby forming the stalk [59, 60]. This in turn helps to explain why Lipid 10 outperforms MC3 in the PMF assays, due to the increased branching and introduction of an additional ester.

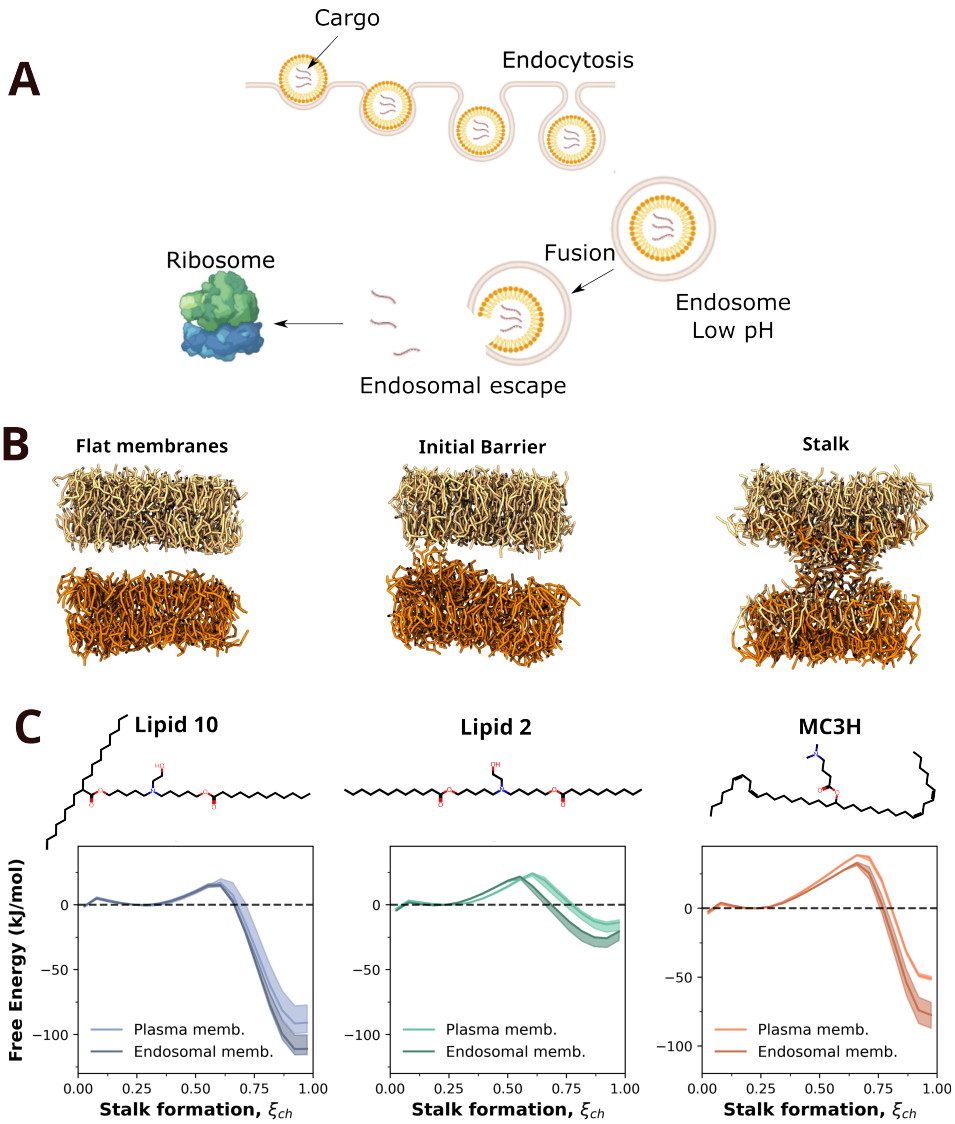

Fig. 3. Understanding the first stages of endosomal escape through PMFs of stalk formation between LNP bilayer containing the MC3 lipid, Lipid 2 or Lipid 10, and an plasma or endosomal bilayer. A) Currently accepted mode of action for LNP delivery. Once the LNP is in the bloodstream a protein corona is formed around the particle enabling endocytosis (not shown in the schematic representation for clarity), resulting in the LNP being encapsulated into the endosome within the cell. The endosome maturates resulting in a pH drop from around 7 to 6-5. Upon the pH change the ionizable lipids become protonated, enabling interaction with the negatively charged lipids in the endosome. B) Representative frames of Martini simulation system of the flat LNP and plasma/endosomal membranes, the transition state, and the stalk (from left to right, respectively). The LNP model membrane is shown in orange while the target plasma membrane is shown in yellow. C) Comparison of stalk formation PMFs between LNPs bilayers containing MC3 lipid, Lipid 2 or Lipid 10, and the plasma membrane or the endosomal membrane models. The error is shown as shading in the same color as the corresponding line.

These fusion assays were conducted between a membrane mimicking the LNP formulations and a simplified plasma membrane, with negatively charged lipids predominantly present in the inner leaflet. However, considering the possibility that in the endosome, negatively charged lipids might migrate to the LNP-interacting leaflet, we sought to evaluate the impact of this potential redistribution. To this end, we performed additional PMF fusion assays against a membrane model enriched with 10 mol % PAPS lipids in the LNP-interacting leaflet to better mimic the compositional differences observed in the endosome. PS was used as the negative lipid for simplicity and comparison to an earlier similar study [23].

The results for these fusion assays demonstrated a slight reduction in both the initial energetic barrier for stalk formation and in the actual stalk across all three lipids (Figure 3C), though for Lipid 10 most of the PMF curves overlap the difference in free energy for the stalk formation at x=1 show a statistically meaningful difference, confirming the trend. This reduction in the barrier reinforces the idea that anionic lipids, such as PS, may play an important role in the endosomal membrane during the transfection mechanism. Morover, the overall findings align with experimental data indicating that LNPs containing Lipid 10 lead to higher protein expression [36], thereby enhancing transfection efficiency.

2.3.2. Sterol-dependent fusion efficacy

In two other significant studies [42, 43], the influence of sterols on transfection has also been emphasized. Various naturally occurring cholesterol analogs (phytosterols) have been shown to markedly enhance transfection efficiency for different ILs, including the widely recognized MC3 lipid. To test this, we conducted additional PMF fusion assays using LNP bilayers containing MC3:STEROL:DSPC (50:40:10 mol%), incorporating cholesterol, β-sitosterol, ergosterol, and stigmasterol to the LNP bilayer. These were tested against the plasma membrane model [54], as mentioned previously.

The resulting fusion PMFs show a more stable stalk formation with β-sitosterol, ergosterol, or stigmasterol compared to cholesterol (Figure 4). This finding correlates with the increased order observed for MC3 with double unsaturated tails (DU) and suggests a potential mechanism by which substituting cholesterol with phytosterols in LNPs enhances fusion efficacy.

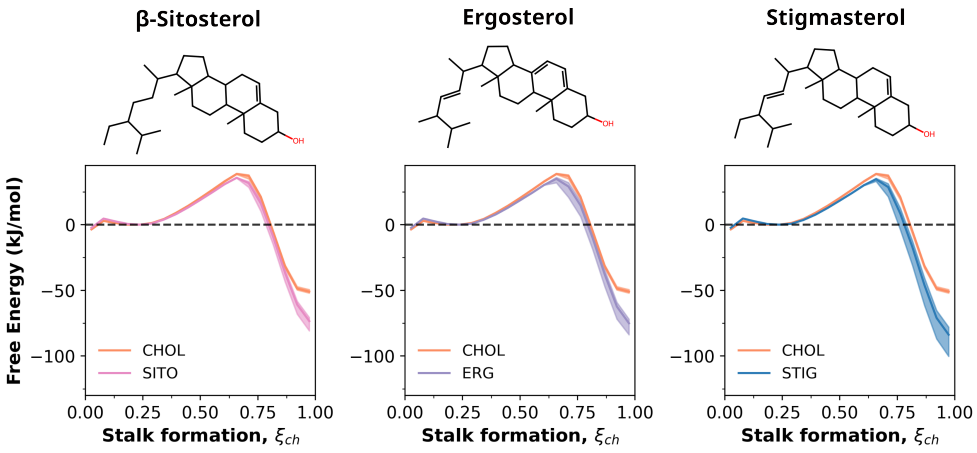

Fig. 4. Potential of mean force calculations of stalk formation with LNP membranes containing different sterols. PMFs of stalk formation between a plasma membrane model [54] and an LNP membrane containing MC3:DSPC:STEROL, with three different sterols.

2.4. Building LNPs at Coarse-Grained Resolution

Currently, the understanding of LNP structure and organization is limited, primarily based on experimental techniques such as Small-Angle X-ray, Neutron Scattering (SAXS, SANS) and low-resolution Cryo-Transmission Electron Microscopy data [5, 14]. While these methods provide valuable insights, they often lack the resolution needed to fully elucidate the complex architecture of LNPs. Simulation studies attempting to model and further understand LNP organization have also been limited [15, 16, 19, 22, 23]. These studies typically focus on smaller bilayer patches or particles, examining how pH variations affect lipid packing and distribution. Notably, only one successful study has simulated the transfection of DNA cargo from a lipoplex [23].

Despite these efforts, our understanding of LNP structure remains incomplete, with various factors such as cargo type, lipid formulation, manufacturing technique, pH, and buffer composition still needing comprehensive exploration [5, 14, 18, 52, 61, 62]. Proposed models for LNP organization range widely—from multilamellar structures and nanostructured cores to homogeneous core-shells, worm-like models, inverse hexagonal phases, bicontinuous cubic phases, and even "bleb" compartments for LNPs containing mRNA [5, 14, 18, 52, 61, 62]. Among these, the inverse hexagonal model is frequently mentioned for smaller cargoes such as siRNA or polyA RNA, while the 'bleb' compartment model is often proposed for LNPs carrying mRNA [5, 52].

Understanding these structures in greater detail is crucial for advancing the field, particularly in the context of improving LNP efficacy and stability. To address these challenges, we now turn our attention to the protocols and tools available for building coarse-grained LNP models, followed by a case study analyzing the structure of an LNP composed of Lipid 2, Lipid 10, and MC3.

2.4.1. Protocols and Tools for Building Coarse-Grained LNPs

To enable further studies of whole particles, we developed a protocol for constructing LNP models. As an initial test case, we used an inverse hexagonal core at low pH, as observed experimentally [5, 14]. This example demonstrates the protocol's capability to handle various LNP structures.

When constructing an LNP, several considerations must be taken into account. LNPs are potentially metastable structures that can significantly depend on the assembly process. In this context, an integrative modeling approach—where LNPs are prebuilt based on available experimental evidence and computational predictions—can be a valuable strategy for achieving accurate representations of these complex particles.

In this work, we built a protocol to generate LNPs with the following parameters: a core hydration level of approximately 12 water molecules per lipid, and a predefined lipid species distribution with a total IL:DSPC:CHOL (for cholesterol) ratio of about 50:10:40 mol%. The protocol constructs a low pH LNP with an inverse hexagonal core, encapsulated 7 strands of RNA with water and neutralizing ions, and a coating consisting of a monolayer along the core channels and a bilayer at the channel ends, closing off the channels. The core comprises IL and cholesterol in a 3:1 ratio, while the coating is made of the remaining IL, cholesterol, and DSPC. These considerations are based on experimental data for an MC3 LNP [5, 14].

The protocol allows for scaling the size of the cargo and the LNP and also adapt the core to other LNP structures and shapes. Factors such as the protonation level of the ILs, the N/P (nitrogen to phosphate) ratio, core hydration level, and ion concentration can be customized by the user, guided by existing experimental data. Additionally, PEGylated lipids can be incorporated into the LNP coating as observed experimentally [14]. The protocol utilizes several tools, including TS2CG [63], Polyply [46], Packmol [64], Blender [65], MDVWhole [66], MDVContainment [67], the Insane script [49], and VMD [68], combined into four steps (detailed in the Methods section). Figure 5 provides an overview of the four main steps involved in the

building protocol, along with two test cases demonstrating different possible organizations for larger particles. The protocol and scripts are available on GitHub (https://github.com/Martini-Force-Field-Initiative/M3-Ionizable-Lipids/). Since the protocol is based on highly versatile tools such as TS2CG [63] and Polyply [46], it can be used as a starting point for constructing other LNP structures, such as the 'bleb' compartment LNP see in 5.

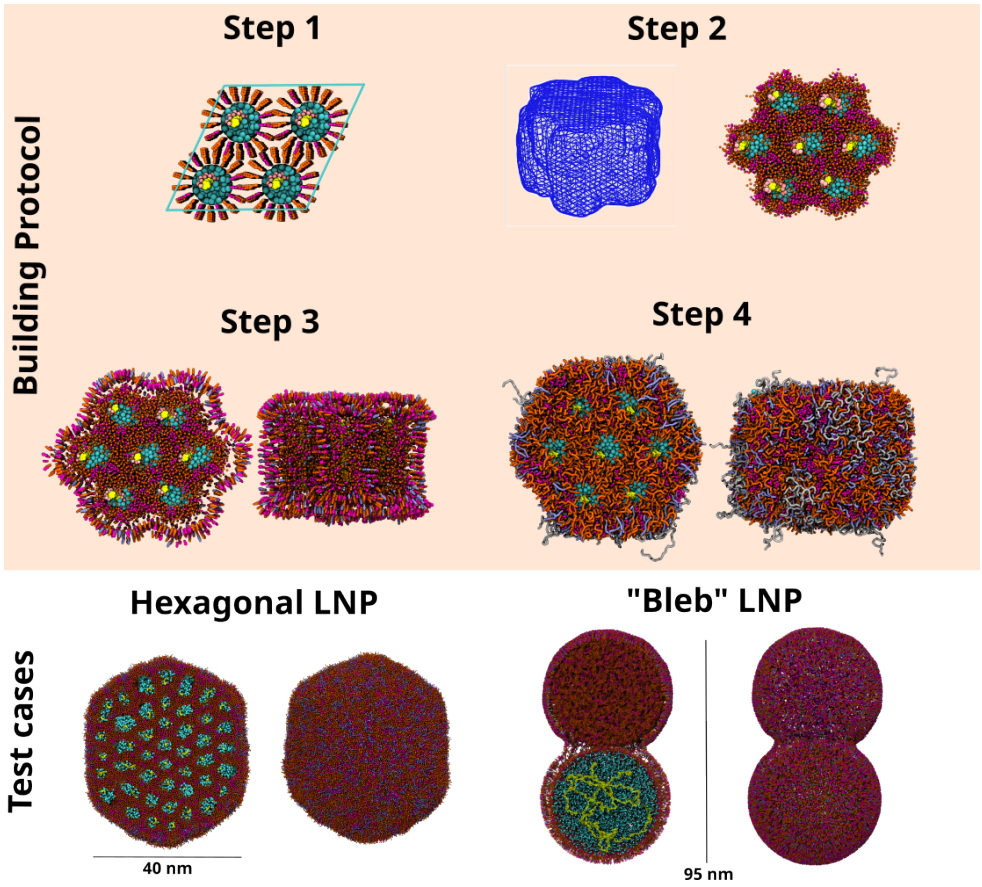

Fig. 5. Representation of the building protocol for LNPs starting at low pH. Step 1) Build and simulate a periodic H_{II} phase with possible cargo and the desired water and ion content. Step 2) Extract the core by multiplying the periodic H_{II} phase in x, y, and z, depending on the desired size. Use VMD to generate a triangulated surface of the core. Step 3) Coat the core using TS2CG with the triangulated surface. A bilayer is placed on top of the otherwise open channels, and a monolayer along the sides of the channels. Step 4) If desired, PEGylated lipids can be added to the surface of the LNP using the PolyPly tool. Two test cases are shown below: one with an inverse hexagonal core and the other phase-separated into two compartments—a 'bleb' compartment containing water, ions, and cargo, and a compartment containing neutral ionizable lipids and cholesterol.

2.4.2. Case study: Analysis of LNP structure composed of Lipid 2, Lipid 10, and MC3

The protocol described previously was used to construct three LNPs with IL:DSPC:CHOL in a 50:10:40 ratio, incorporating the three ILs tested in the PMF fusion assays: MC3, Lipid 2, and Lipid 10. PEGylated lipids were excluded, as they typically constitute only 1–2% of LNP formulations and because these particles were later used to study interactions with an

endosomal membrane model. In all three LNPs, the ILs were protonated. Seven RNA strands, comprising AU duplexes in a helical conformation, were used as cargo and encapsulated in an inverse hexagonal phase. During the 10 μs simulations, all LNPs maintained a compact structure (Figure 6), preserving the encapsulated RNA and a generally spherical shape with an approximate diameter of ~20 nm. However, noticeable variations in internal morphology were observed when comparing LNPs formulated with Lipid 10 and MC3 to those formulated with Lipid 2.

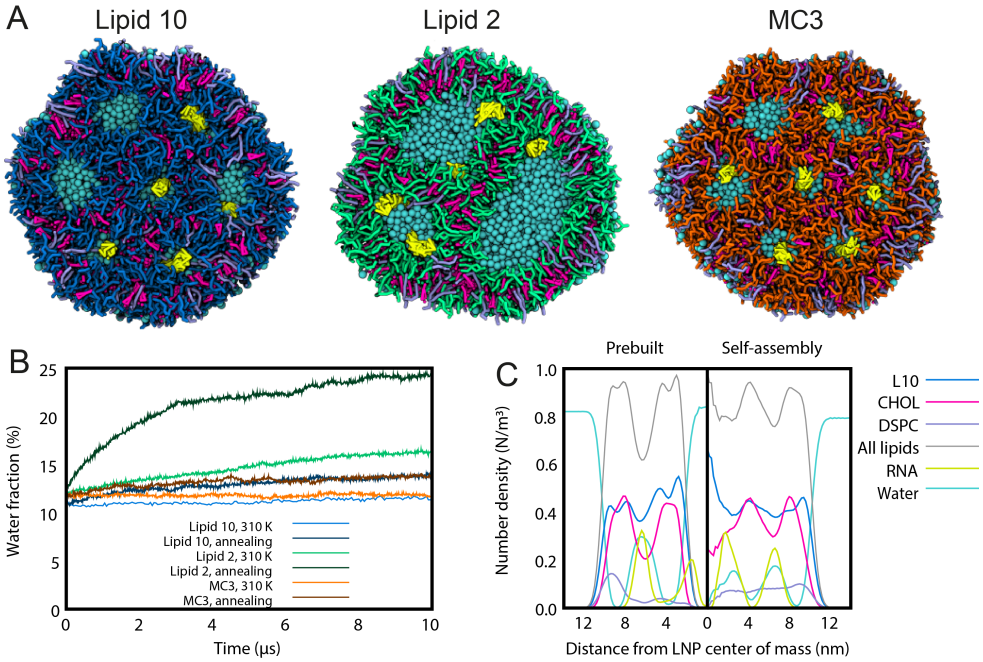

Fig. 6. Structure of LNP models, all containing 7 strands of cargo: (A) Cross-section of LNP structures after 10 μs of simulations at 310 K. Note, because it's a cross-section not all RNA strands are visible, however all are present within the water compartments in the LNP. Water inside LNPs is shown in cyan balls, RNA in yellow, CHOL in magenta, DSPC in violet, Lipid 10 in blue, Lipid 2 in green and MC3 in orange. (B) Time evolution of LNP hydration. (C) Radial number density profile of Lipid 10 containing LNP resulting from pre-built and self-assembly simulations. Water density is scaled by the factor of 0.1 to fit into plot.

While Lipid 10 and MC3 maintained an inverse hexagonal phase, Lipid 2 formed a multicompartmental vesicle with lipids preferring a lamellar organization (Figure 6). Intermolecular radial distribution function (RDF) between RNA surfaces of MC3 LNP did not evolve significantly during the simulation, with two peaks, at 4.7 and 9.3 nm (Figure S28). Lipid 10 was less organized, with RDF peaks at 4.3, 8.4 and 10.7 nm. Visually, both these LNPs preserved the inverse hexagonal phase for at least part of the LNP, in agreement with experimental observation of MC3 [5, 14]. On the other hand, Lipid 2 reshaped its internal organization, formed multicompartmental vesicle with RNA located on the water/lipid interface (Figure 6 Panel A), seemingly similar to the reported 'bleb' compartments observed with Cryo-EM pictures [52].

To investigate the impact of equilibration process, we performed additional simulated annealing of the LNPs. In all of the cases, final LNPs were less regular and more hydrated, see Figure 6B. At 310 K, the LNPs in inverse hexagonal phase (Lipid 10 and MC3) contained ~12 mass% of water (Table S1) while Lipid 2-based LNP consisted ~16 % of water (Figure 6). Heating and cooling the systems led to LNPs with ~14 % of water in inverse hexagonal phase and 24 % in

case of multicompartmental vesicle of Lipid 2. The original ~690 Cl^- beads stayed mostly inside the LNP, without significant differences between the systems (Table S1).

To further assess the robustness of the prebuilt structures, we conducted self-assembly simulations using the same lipid compositions as in the pre-equilibrated structures. Within the first tens of nanoseconds, lipids assembled around the RNA duplexes. In all of the systems (including Lipid 2), lipids adopted an inverse hexagonal phase. As expected from an initial random distribution of lipids, DSPC was not present only on the LNP surface, but equally distributed in its interior (Figure 6C). Furthermore, LNPs were less hydrated (9–10 mass% of water inside LNP, Table S1) and with a significant reduction of the amount of Cl^- beads trapped within the core of LNP: 112 Cl^- for Lipid 10, 49 Cl^- for Lipid 2 and 128 Cl^- for MC3.

The findings suggest that the building protocol is reasonable, with MC3 LNP serving as a control due to its known formation of an inverse hexagonal phase with polyA RNA at pH 6–4.5 [5, 14]. For Lipid 10 and 2 there is no experimental data on their organization at low pH, however, we do observe Lipid 2 adapting a completely different organization rather quickly. This shows, that even though we pre-built with an inverse hexagonal phase, other organizations can easily be observed and that the LNP, at least in this case, is not stuck in the inverse hexagonal phase.

Self-assembly simulations demonstrate that, for these smaller particles, ILs can rapidly organize around RNA cargo to form structures similar to inverse hexagonal phases. However, a potential limitation is the discrepancy in ion and hydration levels between prebuilt and self-assembly simulations (Table S1). Additionally, the self-assembly simulations did not achieve a neutral core in terms of neutralizing ions, likely due to the limitations of the reaction-field method used for electrostatics. Ion translocation through the LNP can also be a slow process. Water molecules in the core equilibrate relatively quickly (Figure S29), whereas ions require a longer time to reach equilibrium, highlighting the importance of considering ion dynamics in the initial structure. Overall, the protocol provides a viable approach for prebuilding LNPs with experimentally guided parameters, offering a foundation for further studies and refinements.

2.4.3. Unbiased simulations of LNP-membrane fusion

Besides demonstrating the capabilities of the models for the three LNPs, we also inserted the LNPs into a lipid vesicle representing the same membranes studied in the PMF fusion assays of Section 2.3. Two different membrane models were used: one plasma membrane model without charged lipids in the interacting leaflet, and another, called endosomal membrane model, with 10 mol% of PAPC lipids replaced by negatively charged PAPS.

During our simulations with LNPs embedded in lipid vesicles lacking negatively charged lipid species in the interacting leaflet, no fusion phenomenon was observed. For each LNP, we extended one of the simulations to over 20 µs, simulating more than 120 µs in total (Table S2). Consistent with the high energy barrier observed in the PMF fusion assays, no LNP-vesicle fusion occurred. This result also aligns with the kinetic barrier described by Leikin et al. [69], which suggests that dehydration of the contact region or another initiation mechanism is required for fusion to occur.

Incorporating negatively charged lipids into the interacting leaflet of the lipid vesicle may accelerate the initiation of fusion between positively charged LNPs and the endosomal membrane [55, 56]. To test this hypothesis, we modified the vesicles to contain 10 mol% of negatively charged PAPS in the inner leaflet and performed three 10 µs LNP-vesicle simulations for Lipid 10, Lipid 2, and MC3. For each IL system, LNP-vesicle fusion was observed in one replica (Figure 7), while in the other three replicas, the LNP remained inside the vesicle without any vesicle-LNP contacts. Although the unbiased approach did not clearly distinguish between the three IL types, it confirmed that LNP and vesicle lipid fusion is not a barrierless process, even with a fraction of negatively charged lipid in the vesicle. However, this barrier is reduced when

negatively charged lipids are present in the target membrane.

The unbiased simulations offer insights into the initial stages of the RNA transfection, in particular the fusion of LNP and ensosomal membrane. The LNP diffuses inside the vesicle and approaches the vesicle membrane, separated by a thin layer of water (Figure 7A). The lipid surfaces come close enough that individual lipids can interact. Negatively charged PAPS interacts with the positively charged IL, and one of them is pulled out of the lipid layer, lying with its tail on the lipid surface (lipid handshake, Figure 7B). If the lipid adopts an extended conformation and embeds its other tail into the lipid vesicle (in the case of IL) or LNP (in the case of PAPS), fusion begins (Figure 7C). Once the two lipid compartments are connected by this single lipid, fusion is initiated and progresses rapidly (Figure 7,D).

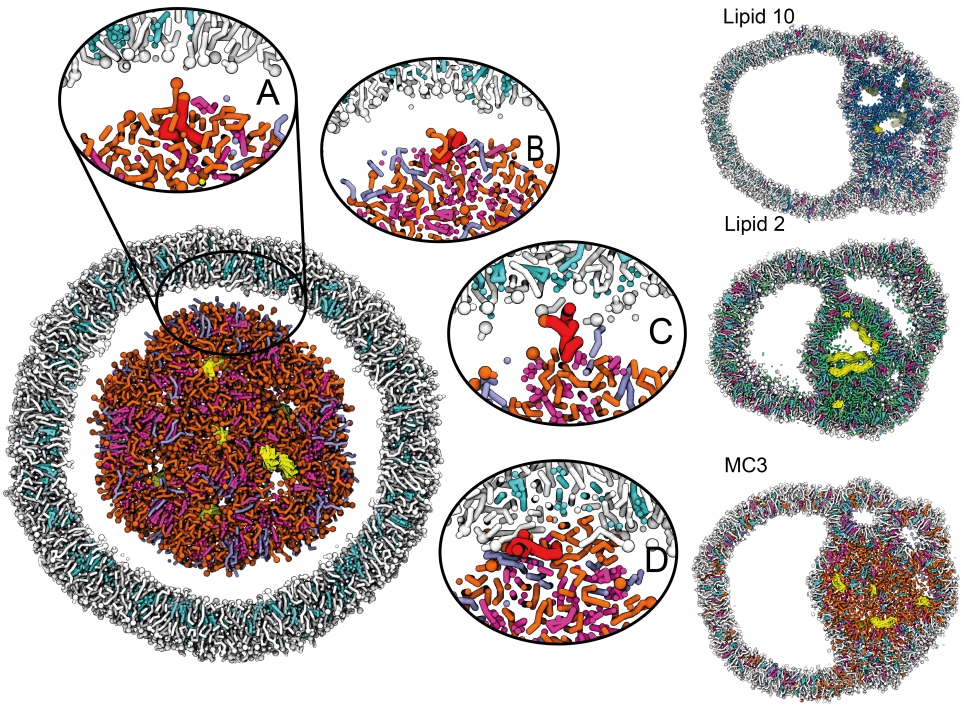

Fig. 7. Left: Cross-section of the initial structure of MC3-LNP embedded in a lipid vesicle (left) and a detailed insight into the initiation of the LNP-vesicle fusion (A–D). The initial fused MC3 is shown in red. Right: Cross-section of fused structures after 7 μs (Lipid 10) and 10 μs (Lipid 2, MC3) of simulations with vesicle CHOL in cyan, other vesicle lipids in grey, RNA in yellow, LNP CHOL in pink, DSPC in purple, Lipid 10 in blue, Lipid 2 in green and MC3 in orange. Water and ions are omitted for clarity.

LNP fusion with the lipid vesicle induces deformation of the vesicle and mixing of LNP and vesicle lipids. Initially, the intact lipid vesicle maintained a spherical shape, but the fusion process disrupted this geometry, leading to elongations and other deformations of the vesicle (Figure 7). Cholesterol from both the LNP and vesicle intermingled and became evenly distributed throughout the final structure. DSPC diffused into the inner leaflet of the vesicle, while a small fraction of ILs flipped to the outer leaflet. However, the majority of ILs remained in contact with RNA (Figure 7). In the final configuration, the RNA was predominantly surrounded by cholesterol (originating from both the vesicle and the LNP) and some residual DSPC. Notably, very few other vesicle lipids, apart from cholesterol, interacted with RNA by the end of the simulation. The RNA-rich compartment was encased in a lipid monolayer primarily composed

of vesicle lipids (Figure 7). Although the induced deformations and RNA-lipid clustering might impact the stability of the resulting particle, we did not observe any RNA translocation outside the vesicle. This phenomenon has been reported for dsDNA in lipoplexes [23], but for RNA in LNPs, it remains an open question.

While the ultimate goal of efficiently studying transfection remains a challenge, it lies beyond the scope of this paper. Nevertheless, we have developed and presented the necessary models and protocols to facilitate these studies in future research.

3. Discussions

Modeling LNPs and designing new formulations present unique challenges for physics-based computational approaches due to several critical requirements. Firstly, achieving high chemical specificity and accuracy in lipid interactions, while incorporating diverse system components – such as solvents, ions, the target endosomal membrane lipids and LNP components, including the encapsulated cargos like nucleic acids – can be demanding. Secondly, accurately simulating the spatial and temporal scales necessary for LNP behavior and drug delivery poses significant computational challenges. The complexity of these systems requires extensive computational resources to model interactions over long timescales and large spatial domains. Additionally, capturing the conformational flexibility of molecules and the precise thermodynamics of LNP assembly and drug delivery is essential. Despite advances in *in silico* methods, many approaches struggle to fully meet these demands, including coarse-grained (CG) methods.

Recent advancements in the Martini model, particularly Martini 3 [30], offer notable improvements. The latest version features enhanced bead packing through optimized cross-interactions between bead sizes, especially for charged groups, and re-parameterized bonded distances based on molecular volume and shape. These upgrades result in more accurate representations of lipid assemblies and improved packing of complex molecules. Furthermore, the expanded range of chemical bead types in Martini 3 enhances coverage of the chemical space, making it especially useful for modeling intricate LNP compositions used in vaccine delivery and gene therapy applications. The model's application is increasingly shifting towards more complex biological systems [26], including attempts to model entire cells [70]. In support of this direction, several tools have been developed to aid these efforts [46, 63, 71, 72].

While our work highlights the significant advancements brought by Martini 3 and associated tools such as TS2CG [63] and Polyply [46], it is important to acknowledge that these models and approaches also have limitations and potential issues. One notable limitation is related to the atomistic models used to parameterize bonded terms of lipids. Although lipid atomistic force fields have undergone extensive benchmarking and evaluations over the years [73, 74], most efforts have focused on standard biological phospholipids, particularly PC lipids. A careful evaluation of models for ILs is still necessary. CHARMM36 [20, 37] with its extensive libraries provided by CHARMM-GUI [75], is one of the most promising atomistic force fields for this purpose. Another significant concern arises from the parametrization of non-bonded terms in Martini. Given the lack of experimental data, the choice of bead chemical types in this work partially relied on COSMO-RS [31] predictions. While COSMO-RS has demonstrated accuracy in logP predictions [32, 33] and has been used effectively in other studies for Martini parametrization [76], the adoption of this approach means that the Martini 3 lipid models developed here are based on a fully bottom-up strategy. This kind of approach should still be considered with caution, as it may not fully align with the Martini parametrization philosophy. This is one of the reasons why consistency with experimentally determined bead assignments was also considered. Further experimental validation of atomistic force fields and thermodynamic properties will be crucial for future refinements of the Martini 3 IL models presented here.

Regardless of the quality of the Martini 3 LNP models, which will benefit from community evaluation with their release, an important aspect of this work is the library of IL models.

Although this study parametrizes one of the most extensive sets of Martini lipids in the literature, the library is still relatively small compared to the diversity of ILs explored in the literature. The building-block approach using overlapping fragments is a promising idea that can expand the library in the near future while maintaining consistency across different lipids. Automatic tools for modeling small-molecule fragments and lipids, as Automartini [77], CGCompiler [78] and Bartender [79], should further facilitate the expansion of lipid libraries. Further development of AI-based tools could be particularly beneficial in this regard. Another more specific consideration involves lipid tails. The Martini model has traditionally used a simplified approach, where the same chain represents two tail lengths. For example, the classic DPPC lipid model represents both di-C16:0 and di-C18:0 lengths. Given that Martini 3 can represent up to three resolutions via beads of different sizes, current developments [80] may allow for more precise distinction of tail lengths, which would significantly enhance the accuracy and expand the IL library models.

It is also important to note that the RNA Martini 3 models used in this work are still under development. While these prototype models have shown reasonable results in stabilizing LNP organization in the 'bleb' or inverted hexagonal assemblies (6A), they currently have limitations in accurately representing RNA structure and possible conformational changes induced by the lipid environment [81]. Further development of the RNA models could provide better insights into how secondary structure and overall nucleobase content might influence LNP assembly and transfection efficiency. Presently, optimization efforts for LNP lipid components have been conducted separately from RNA sequence optimization. This separation, although necessary due to current limitations, underscores the need for more integrated approaches that can simultaneously account for both factors, ultimately leading to more effective LNP designs for drug delivery applications.

Beyond the Martini models themselves, there are important considerations regarding the protocols we suggest. The degree of protonation of ILs at different pH levels was not explored in this work. We approximated IL lipids with pH 7 or higher as fully neutral and lower pH (5-6) as fully protonated. Better consideration of the expected degree of protonation based on apparent pKas, the influence of the environment (especially nucleic acid cargo), or the use of constant pH approaches [82, 83] could improve the assignment of protonation states. Further development of Martini constant pH approaches, such as the Titratable variant [84], would also be beneficial.

Another assumption, now made in building protocols and unbiased fusion assays, concerns the variability of LNP morphologies, which are possibly metastable assemblies that can vary in size and structure depending on the experimental process used in their production. Integrative modeling approaches for prebuilding models seem ideal in this scenario, as precisely modeling microfluidic experimental processes [3] used to synthesize LNPs is currently not feasible with computational methods. However, the prebuild approach requires experimental information, particularly some low-resolution data about the expected organization of the LNP core. Lack of information on the concentration of water-soluble ions inside the internal water compartments may be a concern in this approach. The self-assembly of small LNPs, as discussed in Section 2.4, highlights potential issues with ions. More experimental and theoretical studies should be conducted to clarify the amount of ions and each LNP component in different subcompartments. Further experimental validation of the standard Martini 3 model in reproducing LNP structures would also be desirable, in particular concerning the electrostatic treatment used.

The PMF fusion assays described in Section 2.3 have proven extremely valuable for quantifying the efficacy of stalk formation, the initial step in fusion, across different LNP formulations and target membrane compositions. However, this approach has limitations: it is performed in the absence of curvature in both the LNP and target membrane, and it does not account for cargo. Additionally, the fusion protocol only quantifies stalk formation and does not encompass other steps involved in complete transfection. While curvature may be incorporated by new approaches [85] and additional collective variables could be applied to explore subsequent steps in

transfection, such as pore formation [86], PMF fusion assays rely on one fundamental assumption: it assumes that the fusion process occurs in an equilibrium regime, where lipids of LNP and endosomal membrane fully mix, which may not reflect the real endosomal escape process. Unbiased fusion simulations presented in Section 2.4.3 provide a natural alternative to PMF fusion assays, as they naturally account for curvature and cargo and reached similar conclusions regarding the impact of negatively charged lipids on fusion. However, unbiased simulations require significantly more sampling to distinguish between formulations and to fully capture the transfection process [23], which remains challenging. Previous work has demonstrated that observing the entire process is possible with unbiased simulations [23], although not observed in our work. The lack of observed phenomena may be related to limitations in Martini, even in its latest version, particularly in forming pores with potentially still show too larger barriers [87].

Finally, although our work has highlighted the importance of target endosomal models in understanding transfection trends, larger benchmarks with more extensive sets of LNP formulations are needed. Also comparing against *in vitro* experimental fusion trends, and not only with *in vivo* protein expression [88], may help to isolate the impact of fusion from other factors affecting RNA delivery efficiency, such as LNP structure and the interaction strength of ILs with nucleic acids. Detailed lipidomics studies of endosomal membrane composition in different cell types could also improve the accuracy of observed trends and aid in developing cell-specific LNP formulations, which is a significant goal in the field, especially for promising cancer vaccines [89, 90].

4. Conclusion

We have developed and presented a comprehensive toolbox for *in silico* studies of Martini 3 LNPs, offering a robust platform for advancing nucleic acid delivery systems. This toolbox includes an extensive library of ionizable lipids and naturally occurring sterol models, specifically designed for Martini 3, which can serve as essential building blocks in the study of LNPs. Additionally, we provide a detailed protocol for constructing full-scale LNPs, guided by experimental data such as cargo contents, PEGylation, lipid formulation, hydration levels, internal ion concentration, and N/P ratio. Moreover, we have introduced a protocol for screening and quantifying the fusion efficacy of various LNP formulations.

These models and protocols not only enable the detailed study of LNP organization, structure, and fusion but also pave the way for routine screening of lipid mixtures, prediction of LNP stability, and optimization of fusion efficacy. Through our showcase of Martini 3 LNP applications in fusion assays, we have demonstrated the critical role of endosomal negatively charged lipids in enhancing transfection efficiency and identified trends in fusion efficiency across different LNP compositions. This underscores the importance of accurately modeling the endosomal membrane to improve LNP predictions. Simulating whole LNPs with our toolkit can possibly provide, in the future, additional insights into how lipid formulation influences corona formation in the bloodstream, which is also a critical factor for LNP internalization into cells.

In summary, the toolbox we provide here empowers researchers to perform *in silico* optimization of LNP compositions, potentially enhancing the efficacy of nucleic acid delivery, gene therapies, and beyond. This work represents a significant step forward in the design and development of more effective LNP-based delivery systems for a wide range of therapeutic applications.

5. Data availability

All the protocols, parameters, test systems, and plots for bonded distributions for each fragment and lipid tested can be found on this GitHub page: https://github.com/Martini-Force-Field-Initiative/M3-Ionizable-Lipids/

6. Methods

All simulations were run using the GROMACS package versions 2019 and 2018 [91] and analyzed using the python packages, MDAnalysis [92,93], Numpy [94], Scipy, scikit-learn, matplotlib [95], Fatslim [96], MDVWhole [66], or GROMACS [91]. Visualization was performed using VMD, Inkscape, or the Python RDkit package. [97–99] All CG simulations were modelled with the Martini 3 CG model for biomolecular simulations [30]. For packing and constructing the systems the packages Blender [65], insane [49], Packmol [64], TS2CG [63], PolyPly [46], and MDVContainment [67] were used.

6.1. Fragment modeling

The design of the fragments was based on relevant ILs and divided into head and linker fragments. The fragments were ensured to overlap with each other so that they could later be combined. The mapping of the head or linker fragments was designed such that they were compatible with the phospholipid tails and consistent across fragments, thereby ensuring the building block approach. The bond distances were based on the center of geometry rule, with the mapping chosen to get an optimal molecular volume and shape while keeping chemical groups within one bead, as described in the parametrization guidelines of Martini 3 [30, 100] and tutorials [101].

6.1.1. Atomistic modelling of fragments

Once the fragments were identified and mapped, their atomistic parameters were obtained using the software LigPargen [102], with the OPLS-AA/1.14*CM1A(-LBCC) force field [103]. Each fragment was simulated for 1 μs in water with 0.15 mM NaCl. For very hydrophobic linker fragments, the fragment was simulated in bulk as well as in water.

Fragment systems were minimized using steepest descent with 5000 steps followed by a relaxation of 250 ps using 1 fs timestep, and a production run of 1 μs using 2 fs time step. The temperature and pressure were held constant at 298 K and 1 bar using the Berendsen thermostat and barostat [104] for relaxation and the V-rescale thermostat [105] and Parrinello-Rahman barostat [106] for the production run.

The pressure coupling was isotropic with a compressibility of 4.5×10^{-5} bar^{-1}. The temperature was coupled using τ = 1 ps during both relaxation and production run. For pressure coupling, τ = 1 ps was set during relaxation and 5 ps during the production. The electrostatics were treated with the PME algorithm [107] using a cutoff of 1.2 nm, while the van der Waals interactions were truncated, smoothly switching the force to zero between 1.0 and 1.2 nm [108]. Bonds involving hydrogen atoms were restricted using the LINCS algorithm [109].

6.1.2. Parametrization of bonded terms

Atomistic fragment simulations were mapped into coarse grained (CG) resolution using MDAnalysis [92, 93]. The bonded terms were then measured and the distributions plotted. The fragment was then modeled in CG and simulated for 1 μs in water with 0.15 mM NaCl. The bonded distributions were fitted onto the corresponding atomistic distributions. For the highly hydrophobic fragment, a bulk simulation was performed as well for 1 μs.

For the CG fragment simulations, the system was minimized using 500 steps using the deepest descent algorithm, followed by a relaxation of 10 ns, using 10 fs timestep, and final a 1 μs production run using 20 fs time step.

Temperature and pressure were kept constant during both the relaxation and production runs at 310 K and 1 bar. The coupling for the pressure was isotropic, using the Berendsen [104] and Parinello-Rahman barostats [106] with τ = 12.0 ps during respectively the relaxation and production run, having the compressibility set to 3×10^{-4} bar^{-1}. The Berendsen [104] and Velocity-rescale thermostat [105] were used respectively for relaxation and production run with τ

= 1.0 ps for maintaining the temperature. The reaction field method was used to treat electrostatics while van der Waals interactions were truncated after 1.1 nm (New-RF" Martini settings [110]) using the potential-shift Verlet scheme [108].

6.1.3. Reference partitioning free energies

Given the lack of experimental data, the reference partitioning free energies were estimated using COSMO-RS approach [31]. Three partitioning systems were considered: octanol/water, hexadecane/water and chloroform/water. For each molecule, 500 ps of semiempirical Born-Oppenheimer molecular dynamics were performed at the GFN2-xtb [111] level, in gas phase, with the xtb program [112]. Samples were taken from the resulting trajectory in such a way that the spacing between samples corresponded to a value for the trajectory self-correlation function less or equal to 0.05. Each sample was optimized at the BP86-D3(BJ)/def2-TZVP [113–116] density-functional level of theory, and frequencies were obtained at the same level. Gibbs free energies were obtained in gas-phase from those frequencies. The systems were then optimized at the pre-defined BP-TZVP-COSMO level with a single-point calculation at the BP-TZVPD-FINE-COSMO-SP level [31] for each sample. All the density-functional calculations used the RI approximation [117] and were performed in Turbomole v7.7 [118]. From those calculations we obtained the data required by COSMO-RS [31], as implemented in the COSMOtherm program. Finally, the values obtained were averaged, weighted by their relative abundance in gas phase, calculated from their Gibbs free energies, and the logarithm of the average was obtained. The error bars correspond to the inter-quartile range for the non-weighted logP of all the employed conformers.

6.1.4. Bead assignment, estimated CG partitioning free energies and SASA

Bead types and sizes were selected according to Martini 3 parametrization guidelines, ensuring consistency with default bead assignments for specific chemical groups. These choices were validated by estimating CG partitioning free energies. The calculated partitioning free energies were obtained by Thermodynamic Integration [119]. The fragment was solvated in a $5\times5\times5$ nm^2 box with either water or 25 % hydrated octanol [120,121]. A set of 11 simulations were performed with equally spaced points from 0 to 1, where fragment-solvent interactions were scaled down from full ($\lambda = 0$) to zero ($\lambda = 1$). A stochastic integrator [122] was applied and each simulation was minimized with 10,000 deep descent steps followed by a 50 ns production run. A soft-core potential was applied with the parameters of $\alpha = 0.5$ and p = 1. The free energy was then calculated using the Multistate Bennett Acceptance Ratio (MBAR) [123].

To ensure correct packing further validate the choices for bead sizes, the CG Solvent Accessible Surface Area (SASA) was fitted against the atomistic mapped SASA as well, having less then 10% difference. The SASA was calculated using the GROMACS package [91] as described in the Supporting Information of the Martini 3 paper [30].

6.2. Lipid modeling

Once the fragments were done, they were combined with known ILs from the literature and tested against CHARMM36 [20, 37]. Bilayer simulations in both atomistic and CG resolution were performed with a mixture of IL:POPC of 30:70 mol%.

6.2.1. CG bilayer simulations

Bilayer patches were constructed with a modified version of the insane software [49], which is available in the GitHub repository associated with this work. The systems were neutralized and ions added to obtained a 0.15 mM NaCl concentration using the GROMACS software [91]. The systems were simulated with the Martini 3 force field [30]. All systems were subjected to the same relaxation protocol. First a minimization of 500 steps using steepest descent

followed by 10 ns of relaxation and finally 1–2 μs production run. The target temperature was 310 K during both relaxation and production runs using, respectively, the Berendsen [104] and Velocity-rescale thermostat [105] with τ = 1 ps, coupling lipids separately from the solvent. The pressure was maintained with a semiisotropic pressure coupling during relaxation and production run using respectively the Berendsen [104] and Parrinello-Rahman [106] pressure coupled with τ = 12 ps and the compressibility at 3×10^{-4} bar^{-1}. The verlet buffer tolerance was set to 0.002 kJ mol^{-1} ps^{-1}. The reaction field method was used to treat electrostatics and van der Waals interactions were truncated after 1.1 nm ("New-RF" Martini settings [110]) using the potential shift Verlet scheme [108].

6.2.2. Atomistic bilayer simulations

The atomistic bilayer patches were constructed using the CHARMM-GUI Membrane Builder tool [75]. The CHARMM36 [20, 37] force field (FF) and the TIP3P water model were used, and the systems were neutralized with 0.15 M NaCl. The systems were first minimized using the steepest descent algorithm for 5000 steps, followed by 6 steps of relaxation, each 1 ns long with a timestep of 1 fs, where position restraints on headgroups and dihedrals were gradually lowered from 1000 to 0 kJ mol^{-1}nm^{-2}. For the production run, the timestep was increased to 2 fs.

For both the relaxation and production runs, the temperature and pressure were held constant at 310 K and 1 bar, respectively. The pressure was held constant using a semiisotropic coupling. In the relaxation the Berendsen barostat and thermostat [104] was applied with τ = 1.0 ps, while in the production run the Velocity-rescale thermostat [105] with τ = 1.0 ps and Parrinello-Rahman barostat were used with τ = 5.0 ps and a compressibility of 4.5×10^{-5} bar^{-1}. The lipids and solvent were coupled independently.

Electrostatics were treated with the PME method [107] using a cutoff point of 1.2 nm while van der Waals interactions were truncated, smoothly switching the force to zero between 1.0 and 1.2 nm. The bonds involving hydrogen atoms were constrained using the LINCS algorithm [109].

6.3. Sterol modelling

Sterol modelling followed the same parameterization strategies as those employed to model cholesterol for Martini 3 [45]. Most sterol molecules parameterized in this work share their sterol core with cholesterol, only differing on their aliphatic chains (β-sitosterol, stigmasterol and campesterol). To model each of these a building block approach was used, where the cholesterol sterol core was attached to the corresponding aliphatic chain. Aliphatic chain mapping and bead typing followed standard Martini 3 parameterization guidelines [30]. Chain dynamics and bonded parameters were modeled using reference CHARMM36 simulations of the corresponding sterols [124], both inserted in POPC membranes and in water. To match atomistic reference dynamics, 1 to 2 dihedral potentials and 2 to 4 angle potentials were employed. This is not unlike the strategy used to model the Martini 3 cholesterol model [45], albeit complicated by the fact that the sterols parameterized in this work have aliphatic chains made up of a larger number of beads.

Ergosterol required us to model their sterol core. The same single-frame virtual site topology approach that was used to model the cholesterol sterol core was applied to these molecules [45]. Additionally, the mapping was kept as close as possible to that used for cholesterol. In short, three beads (typically R1, R2 and C1) were used to define a rigid triangular frame assembled using linear constraints. The beads that make up this triangular frame were then used as the constructing particles from which the remaining beads (other than aliphatic chain beads) were defined as out-of-plane virtual sites. Parameters for virtual site construction were obtained as described in detail elsewhere [45, 125, 126], from CHARMM36 simulations of each sterol inserted in POPC bilayers. Briefly, the atomistic references were CG-mapped and the positions of the particles were then averaged after rotationally and translationally fitting the entire molecule

to the virtual-site frame. The placement of the virtual site beads relative to the $frame_{ijk}$ was then defined as a function of ij, ik and of their cross product. Nonbonded interactions between virtual site particles were excluded. Martini bead masses were then redistributed from the virtual site beads to the beads making up the virtual site frame. The corresponding aliphatic chains were then attached to each core as described above.

Each sterol molecule was then validated to ensure correct behaviour and numerical stability. For stability validation, large sterol containing bilayers $30 \times 30\,nm^2$ were simulated at 20 fs for over 5 microseconds. For sterol behavior validation, $10{\times}10\,nm^2$ DPPC lipid bilayers containing increasing concentrations of sterols were simulated at 323 K. Several membrane biophysical properties were then probed, such as area per lipid, membrane thickness and acyl chain order. Trends were compared between each sterol and against a cholesterol reference.

6.3.1. Bilayer melting temperature protocol

As part of the validation of the sterols, the melting temperature was calculated using a simulated annealing approach, where the simulations starts at 373 K, followed by a quenching of the temperature at a rate of 1 K/100 ns, until reaching 273 K, followed by a heating of 1 K/100 ns returning to 373 K. The Lindemann index was then calculated for all lipid beads in blocks of 100 ns, such that each block corresponds to a change of approximately 1 K (Eq. 1, 2). Overheating (T+) and under-cooling (T-) temperatures were determined using the largest gradient in the Lindemann index. The mean T+ and T- were calculated from 5 repeats each of 2 μs. The simulation Tm was the obtained through Equation 3, known ad the "hysteresis method" [127].

$$q_i = \frac{1}{(N-1)} \sum_{i \neq j} (\sqrt{\langle r_{ij}^2 \rangle - \langle r_{ij} \rangle^2}) \tag{1}$$

$$Lindemann = \frac{1}{(N)} \sum_{i=i}^{N} (q_i) \tag{2}$$

$$T_m = T_+ + T_- - \sqrt{T_+ T_-} \tag{3}$$

The pressure, temperature, van der Waals and electrostatic interactions were treated as described in section 5.2.1.

6.4. Nucleic acid models

A preliminary Martini 3 model for nucleic acids was used in this work for the simulations with RNA. For simulations performed in LNP with an inverted hexagonal phase, 21 nucleotide-long polyA dsRNA sequences were modeled using the same preliminary model employed in recent biomolecular condensate studies [128, 129]. In this model, bead mapping and bonded parameters were based on the phosphate and ribose of the Martini 2 nucleic acid models [130, 131], with adaptations to the bead types. Phosphate groups are represented by Q5 beads, while ribose is represented by SN3a-SP1 fragments. Aromatic nucleobase rings were already available with Martini 3 [30]. Since this model was still a prototype, the π-stacking between nucleobases was reinforced via an elastic networks. For the long 1,000 nucleotide polyA ssRNA used in the bleb LNP, a more advanced, but still preliminary, model was utilized. This model did not include elastic networks, with mapping and bonded terms aligned with Martini 3 guidelines. Specifically, ribose fragments followed the recent guidelines suggested for carbohydrates [132]. Although promising, this new prototype model presents some numerical instability, with all production simulations performed using a 10 fs time step instead of the 20 fs time step used in the rest of the CG simulations performed in this work. Coordinate and parameter files are provided on the GitHub repository associated with this work.

6.5. LNP - Endosomal Bilayer Fusion assays

Bilayer fusion assays between LNP membranes with different ionizable lipids (MC3H, Lipid 2 or Lipid 10) and a target plasma membrane were performed using the method described in [53]. We used a plasma membrane model called A8 [54], which has reduced complexity, but captures reasonably well the properties of more complex compositions. The LNP interacting leaflet of this membrane is composed of (in mol%) 24 % POPC, 12 % PAPC, 2 % POPE, 6 % DIPE, 24 % DPSM and 31 % CHOL while the non-interacting leaflet contains 14 % POPC, 8 % PAPC, 5 % POPE, 16 % DIPE, 11 % DPSM, 28 % CHOL, 25 % PAPS, 2 % PAP6. An additional membrane model, enriched with 10 mol % PAPS lipids in the LNP-interacting leaflet, was also considered to better model the differences in composition in the endosome. All bilayer fusion systems were built with an automated protocol written in python, and the same protocol was used for all studies of these systems, differing only in the composition of the membranes. Initially, two independent membrane bilayers corresponding to the target LNP and endosomal target membrane are built in a $10 \times 10 \times 8\,\mathrm{nm}^3$ box using the insane software [49]. Each system was then hydrated and neutralized using sodium ions. This was followed by energy minimization and equilibration for 20 ns to relax the box dimensions completely.

Subsequently, the two membranes were stacked on top of each other (as shown in Figure 4A). To fully hydrate the distal leaflets, the box was extended along the z-axis, and additional water was added to until there were 5 water beads per lipid between the proximal leaflets. Sodium ions were introduced to neutralize the system's charge. The double membrane system was then equilibrated for an additional 20 ns.

To avoid waters from permeating the leaflets and changing the hydration of the inner layer, we applied flat-bottomed position restraints to all water beads. The reference positions of all flat-bottomed position restraints were taken as the center of the box along z. The force constant for the quadratic potential was set to $100\,\mathrm{kJ\ mol^{-1}\ nm^{-2}}$. Together, these flat-bottomed position restraints potentials allowed normal diffusion and fluctuation of water in both compartments and applied only if water beads deeply penetrated the hydrophobic membrane.

The initial frames for the umbrella sampling (US) simulations were extracted from constant-velocity pulling simulations. In these simulations, the systems were pulled from reaction coordinate $\xi\mathrm{ch} = 0.25$ to $\xi\mathrm{ch} = 1$ over 1000 ns using a force constant of $5000\,\mathrm{kJ\ mol^{-1}}$, for five replicates. Nineteen umbrella windows were employed, with reference positions ranging from $\xi\mathrm{ch} = 0.1$ to 1 in increments of 0.05, where 1 is the fully formed stalk. The force constant was set to $5000\,\mathrm{kJ\ mol^{-1}}$. Each window was simulated for 1 μs, with the initial 500 ns excluded for equilibration. An integration time step of 20 fs was utilized. The potential of mean force (PMF) was computed using the weighted histogram analysis method (WHAM) [133].

6.6. Constructing whole LNPs

6.6.1. Building protocol

The construction of whole LNPs with inverse hexagonal core utilizes the tool TS2CG [63], Packmol [64], Polyply [46], the insane script [49], VMD [68] and the python packages fatslim [96], MDAnalysis [92], MDVContainment [67], and MDVwhole [66]. The procedure is an evolution of a protocol previously described for building cationic lipid/DNA lipoplex models studied in the context of gene therapy delivery [23, 134]. The different steps for constructing the particles are described in Jupyter notebooks and made available on the GitHub repository associated with this work. Within the nookbooks are the parameters for construction set as well.

Step 1 involves constructing a periodic system of the inverse hexagonal phase of a given lipid composition using TS2CG in combination with MDAnalysis and fatslim. A single cylinder is constructed with TS2CG, with the principal axis aligned with the z-axis. The cylinder is then filled with possible cargo, water and ions using Packmol. The cylinder is then replicated 3 times,

such that there is 4 cylinders in a hexagonal periodic box. This system is then relaxed for 10 ns.

Step 2 is the extraction of a hexagon from the periodic inverse hexagonal system. This is done through duplicating the relaxed inverse hexagonal system using GROMACS, after which MDVContainment along with MDVWhole are used to identify the different channels and extracting a whole hexagon with 7 channels in total.

Step 3 is the coating of the core hexagonal system. VMD is used in combination with Blender to obtain a triangulated surface of the hexagon, which is used as input for TS2CG to generate a monolayer and a bilayer of the desired lipid composition for the coating of the hexagon. At the top of the hexagon the channels are open, hence a bilayer is placed at either end of the channels such that the hydrophilic headgroups of the lipids are facing both the open channels as well as the solvent. Along the channels the monolayer is placed, with the lipid headgroups facing the solvent.

Step 4 involves possible PEGylation of the coating using Polyply, resulting in a LNP with 20 nm in diameter.

A similar procedure was performed to build a larger LNP with a diameter of approximately 40 nm. In this case, after step 1, the relaxed inverse hexagonal system was replicated in step 2 to generate an LNP with 44 channels in total, followed by the same procedures described in steps 3 and 4. For the 'bleb' structure, steps 1 and 2 were replaced by the construction of two vesicle-like compartments using TS2CG and Polyply. One compartment contained water, ions, and a long 1,000 ssRNA chain, while the other contained neutral ionizable lipids and cholesterol. After construction, both compartments were relaxed and the outer coating was generated, as described in Steps 3 and 4. After construction, the resulting LNP particles are then solvated and the desired ions are added using insane and GROMACS, followed by 5 steps of relaxation, as detailed in the Jupyter notebooks.

6.6.2. Assessment of the pre-build structures

The pre-built LNPs were equilibrated for 10 μs at a) constant temperature of 310 K and b) via rapid periodic simulated annealing with linear heating and cooling the system between 290 and 390 K in 40 ns cycles. The remaining simulation parameters were the same as in other production simulations in this manuscript, except for the use of isotropic pressure coupling.

To also compare different strategies for generating LNP models, we performed also self-assembly simulations of the same composition of lipids and RNA molecules as in the prebuilt structures. The molecules were randomly inserted into a cubic simulation box (18 nm side). The box was enlarged (35 nm side), hydrated, neutralized with Cl^- ions and then Cl^- and Na^+ ions were added in physiological conditions. The system was equilibrated and simulated for 10 μs with annealing, as in the case of pre-assembled LNPs and in previous work [15]. The structures of the LNPs were then characterized in terms of their size, shape, density distribution and hydration level using the same protocols provided in ref [15], with average values calculated for the last 2 μs of the simulation.

In order to evaluate the role of initial hydration and ionization level of the LNP, we modified the structure of LNP with Lipid 2. We reduced the number of water beads up to a full dehydration of the initial structure (90, 80, 50 and 0% of original water level). Further, we reduced the number of Cl^- ions to 90, 50 and 0% of original level at the fully hydrated structure and to 50 and 0% at the fully and 50 % dehydrated LNP structure (Figure S29). The LNPs were simulated for 10 μs with annealing as other LNPs in this paper and their level of their hydration was evaluated as in ref [15].

Further, we tested the changing LNP morphology dependent on the protonation level of Lipid 2. We used the initial structure of fully protonated LNP with Lipid 2 and deprotonated a fraction of Lipid 2 (Figure S30), keeping the system net-neutral (replacing the required amount of Cl^- ions by water beads) and simulated them for 10 μs with annealing. The protonation or deprotonation of 10% of ILs did not make a significant difference in the LNP structure. Deprotonation of more

molecules of Lipid 2 resulted in RNA cluster surrounded by a bulk lipid phase (Figure S30).

6.7. Unbiased fusion simulations of whole LNPs

For unbiased fusion simulations we used the approach of simulating an LNP encapsulated in a lipid vesicle. Three lipid vesicles with outer diameters ranging from 15–16 nm were prepared by TS2CG tool [63] with water pores of 2 nm in diameter. The lipid composition of the lipid vesicle corresponded to the plasma membrane model used in the fusion assays in section 2.3 with negatively charged lipids (PAPS) in the outer non-interacting membrane leaflet. Lipid tails were excluded from the region of water pore with cylindrical restraints and a short vacuum energy minimization was performed (as recommended in the original vesicle formation protocol). The equilibrated structure of the LNP was inserted inside the lipid vesicle to align the LNP and vesicle centers of gyration. The LNP was restrained at its original position for the whole time of the equilibration (with force constant of 1,000 kJ mol^{-1} nm^{-2}). Afterwards, the system was solvated and neutralized with Na^+ ions and minimized again. During system equilibration, neutral and negatively charged vesicle lipids were treated differently. Neutral lipids were only restrained by keeping their tails out of the water pore (with 1,000 kJ mol^{-1} nm^{-2}), that was gradually closing (with 2.0, 1.5, 1.0 nm in diameter). Charged lipids were kept out of the pore with higher force (with 10,000 kJ mol^{-1} nm^{-2}) and their headgroups were restrained spherically to stay in the outer leaflet. This process did not prevent all flip-flops or vesicle deformations, therefore after pore reduction to 1.0 nm in diameter, charged PAPS in the internal leaflet were changed to PAPC, and PAP6 was removed. Further, lipids in vesicle deformations were removed. The equilibration with 1.0 nm pore was repeated, followed by steps with 0.5 and 0.2 nm of pore diameter. Finally, all restraints were removed and the LNP-vesicle system was simulated for at least 10 μs with the same simulation protocol as bare LNPs at 310 K (see the Table S2). For each LNP, three different vesicles were prepared and simulated independently.

To evaluate the role of negatively charged lipids in the fusion process, we prepared another set of lipid vesicles with 10 % of lipids of the internal leaflet exchanged for PAPS, mimicking the endosomal membrane. We used the same lipid vesicles as in the previous case (after the equilibration), exchanged the required amount of PAPC with PAPS and performed 10 μs of simulations.

7. Acknowledgments

The authors would like to thank Dr Ignacio Faustino, Linus Grünewald and Jan A. Stevens for their contributions in the preliminary Martini 3 models of nucleic acids and their implementation in Martinize2 and Polyply codes. P.C.T.S. also acknowledge Riccardo Pellarin and Carlos M. Marques for the valuable discussions. L.R.K., L.B.A and P.C.T.S would like to thank the support of the French National Center for Scientific Research (CNRS) and the funding from research collaboration agreements with PharmCADD. This work was granted access to the HPC resources of IDRIS, CINES and TGCC under the allocations 2022-A0120713456 and 2023-A0140713456 made by GENCI. M.V., L.B.A and P.C.T.S also acknowledge the support of the PSMN (Pôle Scientifique de Modélisation Numérique) and the Centre Blaise Pascal's IT test platform at ENS de Lyon (Lyon, France) for the computer facilities. The platform operates the SIDUS solution developed by Emmanuel Quemener [135]. R.M-A. and R.A-O. thank ANID-Chile for support under FONDECYT N. 1240992. S.J.M. acknowledges funding from the European Research Council with the Advanced grant "COMP-O-CELL" (101053661). L.M. acknowledges funding by the Institut National de la Santé et de la Recherche Médicale (INSERM). M.P., M.Š. and M.O. acknowledge VSB-Technical University of Ostrava, IT4Innovations National Supercomputing Center, Czech Republic for granting this project access to the LUMI supercomputer owned by the EuroHPC Joint Undertaking, hosted by CSC (Finland) and the LUMI consortium with the support of the Ministry of Education, Youth and Sports of the Czech Republic through

e-INFRA CZ (grant 90254). COST Action CA21101 is also acknowledged. The work was also supported by the ERDF/ESF project TECHSCALE (No. CZ.02.01.02/00/22_008/0004578) and by the European Union under the REFRESH - Research Excellence for Region Sustainability and High-tech Industries project (No. CZ.10.03.01/00/22_003/0000048) via the Operational Program Just Transition.

References

1. L. R. Kjolbye, G. P. Pereira, A. Bartocci, M. Pannuzzo, S. Albani, A. Marchetto, B. Jiménez-García, J. Martin, G. Rossetti, M. Cecchini, S. Wu, L. Monticelli, and P. C. T. Souza, "Towards design of drugs and delivery systems with the Martini coarse-grained model," QRB Discov. **3** (2022).
2. Y. Jia, X. Wang, L. Li, F. Li, J. Zhang, and X.-J. Liang, "Lipid nanoparticles optimized for targeting and release of nucleic acid," Adv. Mater. **36**, 2305300 (2024).
3. M. Mehta, T. A. Bui, X. Yang, Y. Aksoy, E. M. Goldys, and W. Deng, "Lipid-based nanoparticles for drug/gene delivery: An overview of the production techniques and difficulties encountered in their industrial development," ACS Mater. Au **3**, 600–619 (2023).
4. M. Cárdenas, R. A. Campbell, M. Yanez Arteta, M. J. Lawrence, and F. Sebastiani, "Review of structural design guiding the development of lipid nanoparticles for nucleic acid delivery," Curr. Opin. Colloid Interface Sci. **66** (2023).
5. J. Philipp, A. Dabkowska, A. Reiser, K. Frank, R. Krzysztoń, C. Brummer, B. Nickel, C. E. Blanchet, A. Sudarsan, M. Ibrahim, S. Johansson, P. Skantze, U. Skantze, S. Östman, M. Johansson, N. Henderson, K. Elvevold, B. Smedsrød, N. Schwierz, L. Lindfors, and J. O. Rädler, "pH-dependent structural transitions in cationic ionizable lipid mesophases are critical for lipid nanoparticle function," Proc. National Acad. Sci. United States Am. **120**, e2310491120 (2023).
6. M. Schlich, R. Palomba, G. Costabile, S. Mizrahy, M. Pannuzzo, D. Peer, and P. Decuzzi, "Cytosolic delivery of nucleic acids: The case of ionizable lipid nanoparticles," Bioeng. Transl. Med. **6** (2021).
7. L. Digiacomo, S. Renzi, E. Quagliarini, D. Pozzi, H. Amenitsch, G. Ferri, L. Pesce, V. De Lorenzi, G. Matteoli, F. Cardarelli, and G. Caracciolo, "Investigating the mechanism of action of dna-loaded pegylated lipid nanoparticles," Nanomedicine: Nanotechnology, Biol. Med. **53**, 102697 (2023).
8. C. Hald Albertsen, J. A. Kulkarni, D. Witzigmann, M. Lind, K. Petersson, and J. B. Simonsen, "The role of lipid components in lipid nanoparticles for vaccines and gene therapy," Adv. Drug Deliv. Rev. **188**, 114416 (2022).
9. W. Li and F. C. Szoka, "Lipid-based nanoparticles for nucleic acid delivery," Pharm. Res. **24**, 438–449 (2007).
10. S. H. Kiaie, N. Majidi Zolbanin, A. Ahmadi, R. Bagherifar, H. Valizadeh, F. Kashanchi, and R. Jafari, "Recent advances in mRNA-LNP therapeutics: immunological and pharmacological aspects," J. Nanobiotechnology **20** (2022).
11. V. Agarwal and D. R. Kelley, "The genetic and biochemical determinants of mRNA degradation rates in mammals," Genome Biol. **23** (2022).
12. S. Moayedpour, J. Broadbent, S. Riahi, M. Bailey, H. V. Thu, D. Dobchev, A. Balsubramani, R. N.D. Santos, L. Kogler-Anele, A. Corrochano-Navarro, S. Li, F. U. Montoya, V. Agarwal, Z. Bar-Joseph, and S. Jager, "Representations of lipid nanoparticles using large language models for transfection efficiency prediction," Bioinformatics **40**, btae342 (2024).
13. S. Chatterjee, E. Kon, P. Sharma, and D. Peer, "Endosomal escape: A bottleneck for LNP-mediated therapeutics," Proc. National Acad. Sci. United States Am. **121**, e2307800120 (2024).
14. M. Y. Arteta, T. Kjellman, S. Bartesaghi, S. Wallin, X. Wu, A. J. Kvist, A. Dabkowska, N. Székely, A. Radulescu, J. Bergenholtz, and L. Lindfors, "Successful reprogramming of cellular protein production through mRNA delivered by functionalized lipid nanoparticles," Proc. National Acad. Sci. United States Am. **115**, E3351–E3360 (2018).
15. M. Paloncýová, M. Šrejber, P. Čechová, P. Kührová, F. Zaoral, and M. Otyepka, "Atomistic Insights into Organization of RNA-Loaded Lipid Nanoparticles," J. Phys. Chem. B **127**, 1158–1166 (2023).
16. M. Paloncýová, P. Čechová, M. Šrejber, P. Kührová, and M. Otyepka, "Role of Ionizable Lipids in SARS-CoV-2 Vaccines As Revealed by Molecular Dynamics Simulations: From Membrane Structure to Interaction with mRNA Fragments," J. Phys. Chem. Lett. **12**, 11199–11205 (2021).
17. M. Ramezanpour and D. P. Tieleman, "Computational Insights into the Role of Cholesterol in Inverted Hexagonal Phase Stabilization and Endosomal Drug Release," Langmuir **38**, 7462–7471 (2022).
18. J. A. Kulkarni, M. M. Darjuan, J. E. Mercer, S. Chen, R. Van Der Meel, J. L. Thewalt, Y. Y. C. Tam, and P. R. Cullis, "On the Formation and Morphology of Lipid Nanoparticles Containing Ionizable Cationic Lipids and siRNA," ACS Nano **12**, 4787–4795 (2018).
19. M. F. Trollmann and R. A. Böckmann, "mRNA lipid nanoparticle phase transition," Biophys. journal **121**, 3927–3939 (2022).
20. S. Park, Y. K. Choi, S. Kim, J. Lee, and W. Im, "CHARMM-GUI Membrane Builder for Lipid Nanoparticles with Ionizable Cationic Lipids and PEGylated Lipids," J. Chem. Inf. Model. **61**, 5192–5202 (2021).
21. M. Cornebise, E. Narayanan, Y. Xia, E. Acosta, L. Ci, H. Koch, J. Milton, S. Sabnis, T. Salerno, and K. E. Benenato, "Discovery of a Novel Amino Lipid That Improves Lipid Nanoparticle Performance through Specific Interactions with mRNA," Adv. Funct. Mater. **32** (2022).

22. A. K. Leung, I. M. Hafez, S. Baoukina, N. M. Belliveau, I. V. Zhigaltsev, E. Afshinmanesh, D. P. Tieleman, C. L. Hansen, M. J. Hope, and P. R. Cullis, "Lipid nanoparticles containing siRNA synthesized by microfluidic mixing exhibit an electron-dense nanostructured core," J. Phys. Chem. C **116**, 18440–18450 (2012).
23. B. M. Bruininks, P. C. T. Souza, H. I. Ingólfsson, and S. J. Marrink, "A molecular view on the escape of lipoplexed dna from the endosome," eLife **9** (2020).
24. S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries, "The MARTINI force field: Coarse grained model for biomolecular simulations," J. Phys. Chem. B **111**, 7812–7824 (2007).
25. S. J. Marrink, L. Monticelli, M. N. Melo, R. Alessandri, D. P. Tieleman, and P. C. T. Souza, "Two decades of martini: Better beads, broader scope," WIREs Comput. Mol. Sci. **13**, e1620 (2023).
26. S. J. Marrink, V. Corradi, P. C. T. Souza, H. I. Ingólfsson, D. P. Tieleman, and M. S. Sansom, "Computational modeling of realistic cell membranes," Chem. Rev. **119**, 6184–6226 (2019). PMID: 30623647.
27. R. Alessandri, F. Grünewald, and S. J. Marrink, "The martini model in materials science," Adv. Mater. **33**, 2008635 (2021).
28. N. Machado, B. M. Bruininks, P. Singh, L. Dos Santos, C. Dal Pizzol, G. d. C. Dieamant, O. Kruger, A. A. Martin, S. J. Marrink, P. C. Souza *et al.*, "Complex nanoemulsion for vitamin delivery: droplet organization and interaction with skin membranes," Nanoscale **14**, 506–514 (2022).
29. Y. Cao, J. Zhu, J. Kou, D. P. Tieleman, and Q. Liang, "Unveiling interactions of tumor-targeting nanoparticles with lipid bilayers using a titratable martini model," J. Chem. Theory Comput. **20**, 4045–4053 (2024).
30. P. C. T. Souza, R. Alessandri, J. Barnoud, S. Thallmair, I. Faustino, F. Grünewald, I. Patmanidis, H. Abdizadeh, B. M. Bruininks, T. A. Wassenaar, P. C. Kroon, J. Melcr, V. Nieto, V. Corradi, H. M. Khan, J. Domański, 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," Nat. Methods 2021 18:4 **18**, 382–388 (2021).
31. A. Klamt, "The cosmo and cosmo-rs solvation models," WIREs Comput. Mol. Sci. **1**, 699–709 (2011).
32. M. Işık, T. D. Bergazin, T. Fox, A. Rizzi, J. D. Chodera, and D. L. Mobley, "Assessing the accuracy of octanol–water partition coefficient predictions in the sampl6 part ii log p challenge," J. Comput. Mol. Des. **34**, 335–370 (2020).
33. T. D. Bergazin, N. Tielker, Y. Zhang, J. Mao, M. R. Gunner, K. Francisco, C. Ballatore, S. M. Kast, and D. L. Mobley, "Evaluation of log P, pK a, and log D predictions from the SAMPL7 blind challenge," J. Comput. Mol. Des. **35**, 771–802 (2021).
34. J. Sangster, "Octanol-Water Partition Coefficients of Simple Organic Compounds," J. Phys. Chem. Ref. Data **18**, 1111–1229 (1989).
35. K. J. Hassett, K. E. Benenato, E. Jacquinet, A. Lee, A. Woods, O. Yuzhakov, S. Himansu, J. Deterling, B. M. Geilich, T. Ketova, C. Mihai, A. Lynn, I. McFadyen, M. J. Moore, J. J. Senn, M. G. Stanton, Almarsson, G. Ciaramella, and L. A. Brito, "Optimization of Lipid Nanoparticles for Intramuscular Administration of mRNA Vaccines," Mol. therapy. Nucleic acids **15**, 1–11 (2019).
36. S. Sabnis, E. S. Kumarasinghe, T. Salerno, C. Mihai, T. Ketova, J. J. Senn, A. Lynn, A. Bulychev, I. McFadyen, J. Chan, Almarsson, M. G. Stanton, and K. E. Benenato, "A Novel Amino Lipid Series for mRNA Delivery: Improved Endosomal Escape and Sustained Pharmacology and Safety in Non-human Primates," Mol. therapy : journal Am. Soc. Gene Ther. **26**, 1509–1519 (2018).
37. J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. J. MacKerell, and R. W. Pastor, "Update of the charmm all-atom additive force field for lipids: Validation on six lipid types," The J. Phys. Chem. B **114**, 7830–7843 (2010). PMID: 20496934.
38. S. C. Semple, A. Akinc, J. Chen, A. P. Sandhu, B. L. Mui, C. K. Cho, D. W. Sah, D. Stebbing, E. J. Crosley, E. Yaworski, I. M. Hafez, J. R. Dorkin, J. Qin, K. Lam, K. G. Rajeev, K. F. Wong, L. B. Jeffs, L. Nechev, M. L. Eisenhardt, M. Jayaraman, M. Kazem, M. A. Maier, M. Srinivasulu, M. J. Weinstein, Q. Chen, R. Alvarez, S. A. Barros, S. De, S. K. Klimuk, T. Borland, V. Kosovrasti, W. L. Cantley, Y. K. Tam, M. Manoharan, M. A. Ciufolini, M. A. Tracy, A. De Fougerolles, I. MacLachlan, P. R. Cullis, T. D. Madden, and M. J. Hope, "Rational design of cationic lipids for siRNA delivery," Nat. Biotechnol. 2010 28:2 **28**, 172–176 (2010).
39. M. J. Carrasco, S. Alishetty, M. G. Alameh, H. Said, L. Wright, M. Paige, O. Soliman, D. Weissman, T. E. Cleveland, A. Grishaev, and M. D. Buschmann, "Ionization and structural properties of mRNA lipid nanoparticles influence expression in intramuscular and intravascular administration," Commun. Biol. 2021 4:1 **4**, 1–15 (2021).
40. F. Ding, H. Zhang, J. Cui, Q. Li, and C. Yang, "Boosting ionizable lipid nanoparticle-mediated in vivo mRNA delivery through optimization of lipid amine-head groups," Biomater. Sci. **9**, 7534–7546 (2021).
41. X. Cheng and R. J. Lee, "The role of helper lipids in lipid nanoparticles (LNPs) designed for oligonucleotide delivery," Adv. drug delivery reviews **99**, 129–137 (2016).
42. S. Patel, N. Ashwanikumar, E. Robinson, Y. Xia, C. Mihai, J. P. Griffith, S. Hou, A. A. Esposito, T. Ketova, K. Welsher, J. L. Joyal, Almarsson, and G. Sahay, "Naturally-occurring cholesterol analogues in lipid nanoparticles induce polymorphic shape and enhance intracellular delivery of mRNA," Nat. Commun. 2020 11:1 **11**, 1–13 (2020).
43. Y. Eygeris, S. Patel, A. Jozic, G. Sahay, and G. Sahay, "Deconvoluting Lipid Nanoparticle Structure for Messenger RNA Delivery," Nano Lett. **20**, 4543–4549 (2020).
44. J. Zhang, H. Fan, D. A. Levorse, and L. S. Crocker, "Interaction of cholesterol-conjugated ionizable amino lipids with biomembranes: Lipid polymorphism, structure-activity relationship, and implications for siRNA delivery," Langmuir **27**, 9473–9483 (2011).

45. L. Borges-Araújo, A. C. Borges-Araújo, T. N. Ozturk, D. P. Ramirez-Echemendia, B. Fábián, T. S. Carpenter, S. Thallmair, J. Barnoud, H. I. Ingólfsson, G. Hummer, D. P. Tieleman, S. J. Marrink, P. C. T. Souza, and M. N. Melo, "Martini 3 coarse-grained force field for cholesterol," J. Chem. Theory Comput. **19**, 7387–7404 (2023).
46. F. Grünewald, R. Alessandri, P. C. Kroon, L. Monticelli, P. C. T. Souza, and S. J. Marrink, "Polyply; a python suite for facilitating simulations of macromolecules and nanomaterials," Nat. Commun. 2022 13:1 **13**, 1–12 (2022).
47. F. Grünewald, "Material design using martini: Accelerating discovery through coarse-grained simulations," Ph.D. thesis, University of Groningen (2023).
48. M. Ibrahim, J. Gilbert, M. Heinz, T. Nylander, and N. Schwierz, "Structural insights on ionizable Dlin-MC3-DMA lipids in DOPC layers by combining accurate atomistic force fields, molecular dynamics simulations and neutron reflectivity," Nanoscale **15**, 11647–11656 (2023).
49. T. A. Wassenaar, H. I. Ingólfsson, R. A. Böckmann, D. P. Tieleman, and S. J. Marrink, "Computational lipidomics with insane: A versatile tool for generating custom membranes for molecular simulations," J. Chem. Theory Comput. **11**, 2144–2155 (2015).
50. K. Hashiba, Y. Sato, M. Taguchi, S. Sakamoto, A. Otsu, Y. Maeda, T. Shishido, M. Murakawa, A. Okazaki, and H. Harashima, "Branching Ionizable Lipids Can Enhance the Stability, Fusogenicity, and Functional Delivery of mRNA," Small Sci. **3**, 2200071 (2023).
51. Y. Li, Z. Ye, H. Yang, and Q. Xu, "Tailoring combinatorial lipid nanoparticles for intracellular delivery of nucleic acids, proteins, and drugs," Acta Pharm. Sinica B **12**, 2624–2639 (2022).
52. M. L. Brader, S. J. Williams, J. M. Banks, W. H. Hui, Z. H. Zhou, and L. Jin, "Encapsulation state of messenger RNA inside lipid nanoparticles," Biophys. J. **120**, 2766–2770 (2021).
53. C. S. Poojari, K. C. Scherer, and J. S. Hub, "Free energies of membrane stalk formation from a lipidomics perspective," Nat. Commun. 2021 12:1 **12**, 1–10 (2021).
54. H. I. Ingólfsson, H. Bhatia, T. Zeppelin, W. F. Bennett, K. A. Carpenter, P. C. Hsu, G. Dharuman, P. T. Bremer, B. Schiøtt, F. C. Lightstone, and T. S. Carpenter, "Capturing Biologically Complex Tissue-Specific Membranes at Different Levels of Compositional Complexity," J. Phys. Chem. B **124**, 7819–7829 (2020).
55. Y. S. Tarahovsky, R. Koynova, and R. C. MacDonald, "Dna release from lipoplexes by anionic lipids: Correlation with lipid mesomorphism, interfacial curvature, and membrane fusion," Biophys. J. **87**, 1054–1064 (2004).
56. T. Kobayashi, M.-H. Beuchat, J. Chevallier, A. Makino, N. Mayran, J.-M. Escola, C. Lebrand, P. Cosson, T. Kobayashi, and J. Gruenberg, "Separation and characterization of late endosomal membrane domains," J. biological chemistry **277**, 32157–32164 (2002).
57. V. S. Markin, M. M. Kozlov, and V. L. Borovjagin, "On the theory of membrane fusion. The stalk mechanism." Gen. Physiol. Biophys. **3**, 361–377 (1984).
58. L. V. Chernomordik and M. M. Kozlov, "Mechanics of membrane fusion," Nat. structural & molecular biology **15**, 675–683 (2008).
59. Y. Yan, X. Liu, L. Wang, C. Wu, Q. Shuai, Y. Zhang, and S. Liu, "Branched hydrophobic tails in lipid nanoparticles enhance mRNA delivery for cancer immunotherapy," Biomaterials **301** (2023).
60. L. Miao, J. Lin, Y. Huang, L. Li, D. Delcassian, Y. Ge, Y. Shi, and D. G. Anderson, "Synergistic lipid compositions for albumin receptor mediated delivery of mRNA to the liver," Nat. Commun. 2020 11:1 **11**, 1–13 (2020).
61. J. Viger-Gravel, A. Schantz, A. C. Pinon, A. J. Rossini, S. Schantz, and L. Emsley, "Structure of Lipid Nanoparticles Containing siRNA or mRNA by Dynamic Nuclear Polarization-Enhanced NMR Spectroscopy," The journal physical chemistry. B **122**, 2073–2081 (2018).
62. L. Schoenmaker, D. Witzigmann, J. A. Kulkarni, R. Verbeke, G. Kersten, W. Jiskoot, and D. J. Crommelin, "mRNA-lipid nanoparticle COVID-19 vaccines: Structure and stability," Int. J. Pharm. **601**, 120586 (2021).
63. W. Pezeshkian, M. König, T. A. Wassenaar, and S. J. Marrink, "Backmapping triangulated surfaces to coarse-grained membrane models," Nat. Commun. 2020 11:1 **11**, 1–9 (2020).
64. L. Martinez, R. Andrade, E. G. Birgin, and J. M. Martínez, "PACKMOL: A package for building initial configurations for molecular dynamics simulations," J. Comput. Chem. **30**, 2157–2164 (2009).
65. B. R. Kent, *3D Scientific Visualization with Blender®*, 2053-2571 (Morgan Claypool Publishers, 2015).
66. B. M. H. Bruininks, T. A. Wassenaar, and I. Vattulainen, "Unbreaking Assemblies in Molecular Simulations with Periodic Boundaries," J. Chem. Inf. Model. **63**, 3448–3452 (2023).
67. B. M. H. Bruininks, "GitHub - BartBruininks/mdvcontainment: Robust characterization of inside and outside in periodic spaces. — github.com," https://github.com/BartBruininks/mdvcontainment. [Accessed 01-08-2024].
68. W. Humphrey, A. Dalke, and K. Schulten, "VMD: Visual molecular dynamics," J. Mol. Graph. **14**, 33–38 (1996).
69. S. L. Leikin, M. M. Kozlov, L. V. Chernomordik, V. S. Markin, and Y. A. Chizmadzhev, "Membrane fusion: Overcoming of the hydration barrier and local restructuring," J. Theor. Biol. **129**, 411–425 (1987).
70. J. Stevens, F. Grünewald, P. van Tilburg, M. König, B. Gilbert, T. Brier, Z. Thornburg, Z. Luthey-Schulten, and S. Marrink, "Molecular dynamics simulation of an entire cell," Front. Chem. **11** (2023).
71. M. D. Andreasen, P. C. T. Souza, B. Schiøtt, and L. Zuzic, "Creating coarse-grained systems with coby: Towards higher accuracy in membrane complexity," bioRxiv (2024).
72. B. M. H. Bruininks, A. S. Thie, P. C. T. Souza, T. A. Wassenaar, S. Faraji, and S. J. Marrink, "Sequential Voxel-Based Leaflet Segmentation of Complex Lipid Morphologies," J. Chem. Theory Comput. **17**, 7885 (2021).
73. O. S. Ollila and G. Pabst, "Atomistic resolution structure and dynamics of lipid bilayers in simulations and experiments,"

Biochimica et Biophys. Acta (BBA) - Biomembr. **1858**, 2512–2528 (2016).
74. M. Javanainen, P. Heftberger, J. J. Madsen, M. S. Miettinen, G. Pabst, and O. H. S. Ollila, "Quantitative comparison against experiments reveals imperfections in force fields' descriptions of popc–cholesterol interactions," J. Chem. Theory Comput. **19**, 6342–6352 (2023). PMID: 37616238.
75. J. Lee, D. S. Patel, J. Ståhle, S. J. Park, N. R. Kern, S. Kim, J. Lee, X. Cheng, M. A. Valvano, O. Holst, Y. A. Knirel, Y. Qi, S. Jo, J. B. Klauda, G. Widmalm, and W. Im, "CHARMM-GUI Membrane Builder for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans," J. Chem. Theory Comput. **15**, 775–786 (2019).
76. L. F. Dettmann, O. Kühn, and A. A. Ahmed, "Martini-based coarse-grained soil organic matter model derived from atomistic simulations," J. Chem. Theory Comput. **20**, 5291–5305 (2024). PMID: 38831535.
77. T. Bereau and K. Kremer, "Automated parametrization of the coarse-grained martini force field for small organic molecules," J Chem Theory Comput. **11**, 2783–2791 (2015).
78. K. S. Stroh, P. C. T. Souza, L. Monticelli, and H. J. Risselada, "Cgcompiler: Automated coarse-grained molecule parametrization via noise-resistant mixed-variable optimization," J. Chem. Theory Comput. **19**, 8384–8400 (2023). PMID: 37971301.
79. G. P. Pereira, R. Alessandri, M. Domínguez, R. Araya-Osorio, L. Grünewald, L. Borges-Araújo, S. Wu, S. J. Marrink, P. C. T. Souza, and R. Mera-Adasme, "Bartender: Martini 3 bonded terms via quantum mechanics-based molecular dynamics," J. Chem. Theory Comput. **20**, 5763–5773 (2024). PMID: 38924075.
80. C. Empereur-mot, K. B. Pedersen, R. Capelli, M. Crippa, C. Caruso, M. Perrone, P. C. T. Souza, S. J. Marrink, and G. M. Pavan, "Automatic optimization of lipid models in the martini force field using swarmcg," J. Chem. Inf. Model. **63**, 3827–3838 (2023). PMID: 37279107.
81. M. Paloncỳová, M. Pykal, P. Kührová, P. Banáš, J. Šponer, and M. Otyepka, "Computer aided development of nucleic acid applications in nanotechnologies," Small **18**, 2204408 (2022).
82. N. Aho, P. Buslaev, A. Jansen, P. Bauer, G. Groenhof, and B. Hess, "Scalable constant ph molecular dynamics in gromacs," J. Chem. Theory Comput. **18**, 6148–6160 (2022). PMID: 36128977.
83. H. A. F. Santos, D. Vila-Viçosa, V. H. Teixeira, A. M. Baptista, and M. Machuqueiro, "Constant-ph md simulations of dmpa/dmpc lipid bilayers," J. Chem. Theory Comput. **11**, 5973–5979 (2015).
84. F. Grünewald, P. C. T. Souza, H. Abdizadeh, J. Barnoud, A. H. de Vries, and S. J. Marrink, "Titratable Martini model for constant pH simulations," The J. Chem. Phys. **153**, 024118 (2020).
85. M. König, R. de Vries, F. Grünewald, S. Marrink, and W. Pezeshkian, "Curvature-induced lipid sorting beyond the critical packing parameter," bioRxiv (2023).
86. J. S. Hub, "Joint reaction coordinate for computing the free-energy landscape of pore nucleation and pore expansion in lipid membranes," J. Chem. Theory Comput. **17**, 1229–1239 (2021).
87. W. D. Bennett and D. P. Tieleman, "Water defect and pore formation in atomistic and coarse-grained lipid membranes: Pushing the limits of coarse graining," J. Chem. Theory Comput. **7**, 2981–2988 (2011).
88. S. M. Lee, Q. Cheng, X. Yu, S. Liu, L. T. Johnson, and D. J. Siegwart, "A systematic study of unsaturation in lipid nanoparticles leads to improved mrna transfection in vivo," Angewandte Chemie Int. Ed. **60**, 5848–5853 (2021).
89. X. Hou, T. Zaks, R. Langer, and Y. Dong, "Lipid nanoparticles for mrna delivery," Nat. Rev. Mater. **6**, 1078–1094 (2021).
90. E. Kon, N. Ad-El, I. Hazan-Halevy, L. Stotsky-Oterin, and D. Peer, "Targeting cancer with mrna–lipid nanoparticles: key considerations and future prospects," Nat. Rev. Clin. Oncol. **20**, 739–754 (2023).
91. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindah, "Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," SoftwareX **1-2**, 19–25 (2015).
92. R. J. Gowers, M. Linke, J. Barnoud, T. J. E. Reddy, M. N. Melo, S. L. Seyler, J. Domanski, D. L. Dotson, S. Buchoux, I. M. Kenney, and O. Beckstein, "MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations," Proc. 15th Python Sci. Conf. pp. 98–105 (2019).
93. N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, and O. Beckstein, "MDAnalysis: A toolkit for the analysis of molecular dynamics simulations," J. Comput. Chem. **32**, 2319–2327 (2011).
94. C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, "Array programming with NumPy," Nat. 2020 585:7825 **585**, 357–362 (2020).
95. J. D. Hunter, "Matplotlib: A 2D graphics environment," Comput. Sci. Eng. **9**, 90–95 (2007).
96. S. Buchoux, "FATSLiM: a fast and robust software to analyze MD simulations of membranes," Bioinformatics **33**, 133–134 (2017).
97. W. Humphrey, A. Dalke, and K. Schulten, "VMD – Visual Molecular Dynamics," J. Mol. Graph. **14**, 33–38 (1996).
98. J. Stone, J. Gullingsrud, P. Grayson, and K. Schulten, "A system for interactive molecular dynamics simulation," in *2001 ACM Symposium on Interactive 3D Graphics,* J. F. Hughes and C. H. Séquin, eds. (ACM SIGGRAPH, New York, 2001), pp. 191–194.
99. "Rdkit: open-source cheminformatics," https://www.rdkit.org.
100. R. Alessandri, J. Barnoud, A. S. Gertsen, I. Patmanidis, A. H. de Vries, P. C. T. Souza, and S. J. Marrink, "Martini 3 coarse-grained force field: Small molecules," Adv. Theory Simulations **5**, 2100391 (2022).
101. R. Alessandri, S. Thallmair, C. G. Herrero, R. Mera-Adasme, S. J. Marrink, and P. C. T. Souza, "A Practical

Introduction to Martini 3 and its Application to Protein-Ligand Binding Simulations," in *A Practical Guide to Recent Advances in Multiscale Modeling and Simulation of Biomolecules,* (AIP Publishing LLC).
102. L. S. Dodda, I. C. De Vaca, J. Tirado-Rives, and W. L. Jorgensen, "LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands," Nucleic Acids Res. **45**, W331–W336 (2017).
103. L. S. Dodda, J. Z. Vilseck, J. Tirado-Rives, and W. L. Jorgensen, "1.14*CM1A-LBCC: Localized Bond Charge Corrected CM1A Charges for Condensed-Phase Simulations," The journal physical chemistry. B **121**, 3864 (2017).
104. H. J. C. Berendsen, J. P. M. Postma, W. F. V. Gunsteren, A. Dinola, and J. R. Haak, "Molecular dynamics with coupling to an external bath," The J. Chem. Phys. **81**, 3684–3690 (1984).
105. G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," J. Chem. Phys. **126**, 14101 (2007).
106. M. Parrinello and A. Rahman, "Polymorphic transitions in single crystals: A new molecular dynamics method," J. Appl. Phys. **52**, 7182–7190 (1981).
107. T. Darden, D. York, and L. Pedersen, "Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems," The J. Chem. Phys. **98**, 10089–10092 (1993).
108. S. Páll and B. Hess, "A flexible algorithm for calculating pair interactions on SIMD architectures," Comput. Phys. Commun. **184**, 2641–2650 (2013).
109. B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, "LINCS: A Linear Constraint Solver for molecular simulations," J. Comput. Chem. **18**, 1463–1472 (1997).
110. D. H. de Jong, S. Baoukina, H. I. Ingólfsson, and S. J. Marrink, "Martini straight: Boosting performance using a shorter cutoff and GPUs," Comput. Phys. Commun. **199**, 1–7 (2016).
111. C. Bannwarth, S. Ehlert, and S. Grimme, "Gfn2-xtb—an accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions," J. Chem. Theory Comput. **15**, 1652–1671 (2019).
112. C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher, and S. Grimme, "Extended tight-binding quantum chemistry methods," WIREs Comput. Mol. Sci. **11**, e1493 (2021).
113. A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," Phys. Rev. A **38**, 3098–3100 (1988).
114. J. P. Perdew, "Density-functional approximation for the correlation energy of the inhomogeneous electron gas," Phys. Rev. B **33**, 8822–8824 (1986).
115. F. Weigend and R. Ahlrichs, "Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for h to rn: Design and assessment of accuracy," Phys. Chem. Chem. Phys. **7**, 3297–3305 (2005).
116. S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu," The J. Chem. Phys. **132**, 154104 (2010).
117. K. Eichkorn, F. Weigend, O. Treutler, and R. Ahlrichs, "Auxiliary basis sets for main row atoms and transition metals and their use to approximate coulomb potentials," Theor. Chem. Accounts **97**, 119–124 (1997).
118. S. Balasubramani, G. Chen, S. Coriani, M. Diedenhofen, M. Frank, Y. Franzke, F. Furche, R. Grotjahn, M. Harding, C. Hättig, A. Hellweg, B. Helmich-Paris, C. Holzer, U. Huniar, M. Kaupp, A. Marefat Khah, S. Karbalaei Khani, T. Müller, F. Mack, B. Nguyen, S. Parker, E. Perlt, D. Rappoport, K. Reiter, S. Roy, M. Rückert, G. Schmitz, M. Sierka, E. Tapavicza, D. Tew, C. van Wüllen, V. Voora, F. Weigend, A. Wodyński, and J. Yu, "Turbomole: Modular program suite for ab initio quantum-chemical and condensed-matter simulations," The J. Chem. Phys. **152** (2020).
119. J. G. Kirkwood, "Statistical Mechanics of Fluid Mixtures," The J. Chem. Phys. **3**, 300–313 (1935).
120. J. L. MacCallum and D. P. Tieleman, "Structures of neat and hydrated 1-octanol from computer simulations," J. Am. Chem. Soc. **124**, 15085–15093 (2002).
121. S. E. DeBolt and P. A. Kollman, "Investigation of Structure, Dynamics, and Solvation in 1-Octanol and Its Water-Saturated Solution: Molecular Dynamics and Free-Energy Perturbation Studies," J. Am. Chem. Soc. **117**, 5316–5340 (1995).
122. N. Goga, A. J. Rzepiela, A. H. de Vries, S. J. Marrink, and H. J. Berendsen, "Efficient algorithms for langevin and DPD dynamics," J. Chem. Theory Comput. **8**, 3637–3649 (2012).
123. M. R. Shirts and J. D. Chodera, "Statistically optimal analysis of samples from multiple equilibrium states," J. Chem. Phys. **129** (2008).
124. J. B. Lim, B. Rogaski, and J. B. Klauda, "Update of the cholesterol force field parameters in charmm," The J. Phys. Chem. B **116**, 203–210 (2012).
125. S. J. Marrink, A. H. de Vries, T. A. Harroun, J. Katsaras, and S. R. Wassall, "Cholesterol shows preference for the interior of polyunsaturated lipid membranes," J. Am. Chem. Soc. **130**, 10–11 (2008).
126. M. N. Melo, H. I. Ingólfsson, and S. J. Marrink, "Parameters for Martini sterols and hopanoids based on a virtual-site description," J. Chem. Phys. **143** (2015).
127. Y. Zhang and E. J. Maginn, "A comparison of methods for melting point calculation using molecular dynamics simulations," J. Chem. Phys. **136** (2012).
128. M. Tsanai, P. J. M. Frederix, C. F. E. Schroer, P. C. T. Souza, and S. J. Marrink, "Coacervate formation studied by explicit solvent coarse-grain molecular dynamics with the martini model," Chem. Sci. **12**, 8521–8530 (2021).
129. H. I. Ingólfsson, A. Rizuan, X. Liu, P. Mohanty, P. C. Souza, S. J. Marrink, M. T. Bowers, J. Mittal, and J. Berry, "Multiscale simulations reveal tdp-43 molecular-level interactions driving condensation," Biophys. J. **122**, 4370–4381 (2023).

130. J. J. Uusitalo, H. I. Ingólfsson, P. Akhshi, D. P. Tieleman, and S. J. Marrink, "Martini coarse-grained force field: Extension to dna," J. Chem. Theory Comput. **11**, 3932–3945 (2015).
131. J. J. Uusitalo, H. I. Ingólfsson, S. J. Marrink, and I. Faustino, "Martini coarse-grained force field: Extension to rna," Biophys. J. **113**, 246–256 (2017).
132. F. Grünewald, M. H. Punt, E. E. Jefferys, P. A. Vainikka, M. König, V. Virtanen, T. A. Meyer, W. Pezeshkian, A. J. Gormley, M. Karonen, M. S. P. Sansom, P. C. T. Souza, and S. J. Marrink, "Martini 3 coarse-grained force field for carbohydrates," J. Chem. Theory Comput. **18**, 7555–7569 (2022).
133. S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, "THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method," J. Comput. Chem. **13**, 1011–1021 (1992).
134. B. M. H. Bruininks, P. C. T. Souza, and S. J. Marrink, *A Practical View of the Martini Force Field* (Springer New York, New York, NY, 2019), pp. 105–127.
135. E. Quemener and M. Corvellec, "SIDUS—the solution for extreme deduplication of an operating system," Linux J. **2013**, 3 (2013).