# Lipid shape and packing are key for optimal design of pH-sensitive mRNA lipid nanoparticles

Giulio Tesei[a,b,1] (ID), Ya-Wen Hsiao[c], Aleksandra Dabkowska[d], Gunnar Grönberg[e] (ID), Marianna Yanez Arteta[d] (ID), David Ulkoski[d] (ID), David J. Bray[c] (ID), Martin Trulsson[b] (ID), Johan Ulander[f], Mikael Lund[b,1] (ID), and Lennart Lindfors[d,1] (ID)

The ionizable-lipid component of RNA-containing nanoparticles controls the pH-dependent behavior necessary for an efficient delivery of the cargo—the so-called endosomal escape. However, it is still an empirical exercise to identify optimally performing lipids. Here, we study two well-known ionizable lipids, DLin-MC3-DMA and DLin-DMA using a combination of experiments, multiscale computer simulations, and electrostatic theory. All-atom molecular dynamics simulations, and experimentally measured polar headgroup $pK_a$ values, are used to develop a coarse-grained representation of the lipids, which enables the investigation of the pH-dependent behavior of lipid nanoparticles (LNPs) through Monte Carlo simulations, in the absence and presence of RNA molecules. Our results show that the charge state of the lipids is determined by the interplay between lipid shape and headgroup chemistry, providing an explanation for the similar pH-dependent ionization state observed for lipids with headgroup $pK_a$ values about one-pH-unit apart. The pH dependence of lipid ionization is significantly influenced by the presence of RNA, whereby charge neutrality is achieved by imparting a finite and constant charge per lipid at intermediate pH values. The simulation results are experimentally supported by measurements of $\alpha$-carbon $^{13}$C-NMR chemical shifts for eGFP mRNA LNPs of both DLin-MC3-DMA and DLin-DMA at various pH conditions. Further, we evaluate the applicability of a mean-field Poisson–Boltzmann theory to capture these phenomena.

lipid nanoparticle | RNA encapsulation | lipid ionization | pH controlled release

The intracellular delivery of polynucleotides, which is crucial for gene therapy and messenger RNA (mRNA)-based therapeutics, requires sophisticated delivery systems. This is due to the inherent susceptibility to enzymatic degradation of polynucleotides and to their limited bioavailability, which primarily stems from their large size and net charge. mRNA-based therapies, while not requiring entry to the cell nucleus, necessitate targeted delivery to the appropriate cells and efficient release of mRNA in the cytosol, where it is made available for translation (1). While several different delivery methods exist, lipid nanoparticles (LNPs) have emerged as one of the most attractive and effective delivery systems for mRNA-based therapeutics, as underscored by their important role in the development of two Covid-19 vaccines.

Since the pioneering work of Safinya et al. (2), a correlation has been established between lipid structure in the LNP and transfection efficiency. Experimental evidence indicates a significant uptake of LNPs through ApoE-LDLr receptor-mediated endocytosis (3, 4), whereas the main bottleneck for LNP delivery is attributed to the endosomal escape (5, 6). Nowadays, LNPs for mRNA delivery are designed and formulated with cationic ionizable lipids so that they form inverse hexagonal phases (7) upon endosomal acidification. This phase transition promotes the escape of mRNA from the endosome, thereby ensuring efficient translation and protein expression (7, 8). Accordingly, the most prominent hypothesis ascribes the efficiency of mRNA-LNP delivery systems to the pH-dependent formation and stability of the inverse hexagonal phase in the LNP core (7). In this context, the study of Jayaraman et al. (9) on short interfering RNA (siRNA) LNPs has led to the common assumption that transfection efficiency may be maximized by an optimal apparent $pK_a$ value of the ionizable lipid component. However, data from the same study also highlight that LNPs based on ionizable lipids with similar apparent $pK_a$ values show orders of magnitude differences in biological efficacy. Further, the curvature of the lipid monolayer is also considered to be a key parameter since some of the most efficient lipids have been shown to form inverse phases in LNPs also in the absence of mRNA (7). Therefore, we here focus our attention on inverse hexagonal phases and the effect of both lipid shape and headgroup intrinsic $pK_a$ on phase morphology.

To guide the design of efficient delivery systems, we need molecular-level understanding of the pH dependence of the structural properties of the core of LNPs. However, a detailed

## Significance

Lipid nanoparticles (LNPs) containing siRNA or mRNA represent an exciting approach to combat diseases. However, the design of efficient and reasonably safe lipid nanoparticles is currently a costly and trial-and-error activity. We use advanced simulations and NMR experiments to study key component of LNPs, the so-called ionizable lipid component. We show how the shape of the ionizable lipid and the intrinsic $pK_a$ of its polar headgroup dictate the pH-dependent response of the LNP, which is required for an efficient cytoplasmic delivery of RNA. We argue that the ionization of RNA-containing LNPs is better described by an isoelectric point, rather than the apparent $pK_a$ values. Our results contribute to a rational basis for optimizing RNA delivery via LNPs.

experimental characterization of such systems is challenging. Although biophysical approaches such as small-angle X-ray scattering (SAXS) can provide structural parameters such as the center–center spacing of the inverse hexagonal phase, the spontaneous curvature of the water–lipid interface and elastic moduli of the monolayer are less straightforward to obtain. Molecular simulations have emerged as a natural complement to experiments, and there has been an increased interest in simulations of systems relevant for mRNA-LNP delivery (10, 11). There is a lack of understanding of the role of individual lipids in the formation and pH-dependent stability of LNPs and several AA simulation studies have highlighted the interdependence of the lipid charge state and the morphology of lipid/RNA phases. Cornebise et al. recently presented novel ionizable lipids with improved LNP performance (12) and, through AA simulations, attributed the enhanced efficiency to specific interactions between lipids and RNA. Further, AA simulations have been employed to investigate the associative behavior of RNA at neutral pH for three different ionizable lipids (13) and the pH-dependent interactions between Dlin-KC2-DMA and POPC in the absence of RNA (14). Notably, the AA simulations by Ramezanpour et al. (14) showed segregation of Dlin-KC2-DMA from POPC in a monolayer at neutral pH, supporting the hypothesis, proposed by Kulkarni et al. (15), of the formation of an isotropic water-in-oil microemulsion in the LNP core. A multiscale approach was recently employed to study the pH-dependent behavior of lamellar phases of the ionizable cationic lipids DODMA and DOPC with and without mRNA (16). AA simulations were initiated by backmapping longer coarse-grained (CG) simulations, whereas the pH dependence was modeled using different fixed ratios of charged species. Further, a large-scale AA simulation study was performed for lipid mixtures used in the Moderna and Pfizer-BioNTech Covid-19 vaccines (17). This study investigated the stability and self-assembly of lipid mixtures in the presence mRNA fragments and reported the spontaneous formation of nano-disks as well as cubic and lamellar phases.

The spatiotemporal complexity of LNP delivery poses a considerable challenge for AA simulations. It is for example known from experiments that LNPs require hours to days to equilibrate, which is well above the time scales that can be explored in AA studies. Therefore, CG modeling is a promising approach for simulations of LNPs, especially for studies of phase transitions and large-scale reorganization or escape events (18). CG simulations of the spontaneous formation of inverse hexagonal phases of DNA lipoplexes were performed using the Martini force field (19). These simulations reproduced the center–center spacing obtained from SAXS experiments and recapitulated experimentally observed transitions between lamellar and inverse hexagonal phases induced by varying lipid composition (19). Moreover, CG simulations were employed to investigate fusion of DNA-lipoplexes with the endosomal membrane (20).

As opposed to the above-mentioned simulation studies which used fixed lipid ionization states, in this work, we performed CG reactive canonical Monte Carlo simulations (21) wherein both the number of molecules and protonation states of lipids are free to fluctuate subject to an external chemical potential. For our simulations, we parametrized a CG model using experimental p$K_a$ values of the lipid headgroups and structural features derived from SAXS experiments and AA simulations. Our CG Monte Carlo simulations capture the central aspects of the pH-dependent behavior of ionizable lipid phases in the absence and presence of mRNA molecules. These include the spontaneous formation of inverse hexagonal phases, the pH dependence of the lipid charge state, and the role of lipid shapes and packing (19, 22). We measured α-carbon $^{13}$C-NMR chemical shifts for LNPs of Dlin-DMA

(DD) and Dlin-MC3-DMA (MC3), which are ionizable lipids with different headgroup intrinsic p$K_a$ values. The ionization state of the lipid headgroups in the core of LNPs is influenced by the local environment through electrostatic interactions with surrounding charged species, such as other lipid headgroups, salt ions, and mRNA. Besides these interactions, image–charge repulsive forces induced by the interfacing aqueous and apolar lipid phase (23) contribute to modulate the apparent stoichiometric p$K_a$ of the lipids, which typically deviates significantly from the intrinsic p$K_a$ value of the headgroup. The CG model developed in this study captures these effects and, combined with Poisson–Boltzmann calculations, provides a molecular-level explanation for the similar values of the apparent p$K_a$ values of DD and MC3 in LNPs, which we estimated from the NMR data. Therefore, our work addresses current gaps in the interpretation of experimental data for the characterization of LNPs and contributes to an improved understanding of the role of pH and lipid shape in LNP morphology. In turn, this knowledge may translate into principles for the design of novel lipids that could lead to more efficacious and safer delivery systems.

## Results

**All-Atom Molecular Dynamics Simulations.** To study pH-dependent effects on the ionizable lipid/RNA system in inverse hexagonal phase (H$_{II}$), we developed a pH-sensitive CG model in which the main parameters are the intrinsic p$K_a$ of the lipid headgroups and the shape of the lipids. The latter affects lipid packing and the curvature of the aqueous channel, where RNA, water molecules, and counterions reside. To design and parametrize the coarse-grained model, we first performed AA simulations of two ionizable lipids, MC3 and DD (*SI Appendix*, Fig. S1), in hexagonal periodic cells. Since the inverse hexagonal phase is expected to form under low pH conditions, lipids are modeled as fully ionized. The charged head groups of the lipids are arranged around a cylindrical channel which is filled with water molecules and counterions (Fig. 1A). In the simulated system, the lipid phase is first put in contact with a bulk aqueous phase located at the ends of the hexagonal cell (*Materials and Methods*), which allows the size of the water channel to fluctuate. After reaching convergence in the center–center distance, water and salt contents, the system is made truly periodic in all three dimensions, and equilibrated once again.

To validate the simulation protocol, we use the experimentally well-studied H$_{II}$ phase of dioleylethanolamine (DOPE)-H$_2$O and found our simulation data for DOPE-H$_2$O to be in satisfactory agreement with experimental results from the literature (24) (*SI Appendix*, Fig. S2). Therefore, we applied the same protocol to study the fully ionized MC3 and DD but also including the LNP helper lipid cholesterol (ionizable lipid/cholesterol molar ratio of 3:1) (25). AA simulation results for the center–center distance, the water channel radius, and the area per lipid obtained for MC3 and DD systems are shown in Fig. 1 *B–D*. Although both lipids have the same headgroup and C18:2 tails (differing only in the linker region), our simulations indicate that both the center–center distance and the aqueous channel radius are smaller for H$_{II}$ phases of the MC3 than for the DD system (*SI Appendix*, Fig. S3). Whereas the structure of the H$_{II}$ phase of DD has not yet been characterized experimentally, simulations for MC3 indicate a center–center distance of 7.3 nm, which is significantly larger than the value obtained from SAXS experiments (*d* = 6.2 nm) (7). It should be noted, however, that the experimental data were obtained using an acidic buffer (50 mM citrate, pH 3) containing 25 vol% ethanol so that our all-atom simulations are not fully comparable with the SAXS data.
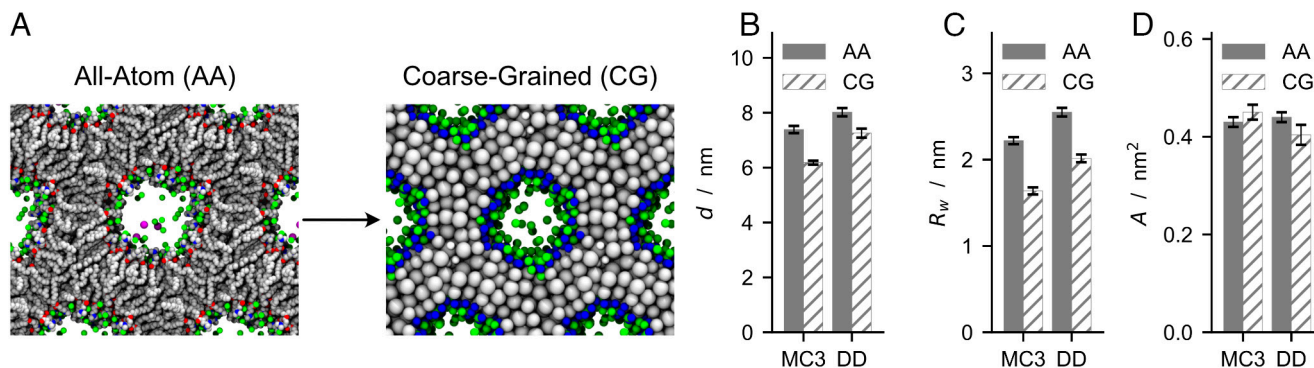
**Fig. 1.** Comparison between all-atom (AA) and coarse-grained (CG) models of the core of empty LNPs. (*A*) Snapshots from all-atom and coarse-grained simulations of fully ionized MC3 at 25 °C. The AA representation shows carbon (gray), oxygen (red), nitrogen (blue), chloride (green), and sodium (purple) atoms. The CG representation shows tail (gray), neutral head (white), protonated head (blue), and chloride (green) beads. (*B–D*) Estimates of the center–center separation (*B*), radius of the water core (*C*), and area per lipid (*D*) from all-atom (gray) and coarse-grained (hatched) simulations of MC3 and DD lipids in inverse hexagonal phase.

**Coarse-Grained Monte Carlo Simulations.** To investigate how lipid shape and pH affect the structure of the core of mRNA-containing LNPs, we developed a pH-dependent, coarse-grained model based our all-atom molecular dynamics simulation data (Fig. 1) as well as experimental headgroup $pK_a$ values of MC3 and DD lipids. To estimate these, we synthesized the polar DD and MC3 head groups and performed potentiometric titrations at 37 °C, resulting in intrinsic $pK_a$ values of 9.41 (±0.04) and 8.46 (±0.07) for MC3 and DD, respectively (26). As a phenomenological model of the ionizable lipids, we use a three-bead lipid model which readily self-assembles into bilayer and micelle phases (27). It has previously been shown that the preferred phase formed by the lipid model can be controlled by varying the size of the bead representing the headgroup (28). Therefore, we use geometrical arguments (*Materials and Methods*) to design a three-bead lipid with a conical aspect ratio, which favors the formation of an $H_{II}$ phase with center–center separation, water channel radius, and area per lipid in reasonable agreement with all-atom simulations for systems without RNA (Fig. 1). The conical lipid model is then simulated in the presence of mRNA and NaCl in the reactive canonical ensemble (21) (Fig. 2 *A* and *B*), in which the number of lipids; ions; the dimensions of the simulation cell; and the protonation state of lipid head groups are free to fluctuate. Within this framework, the lipid protonation state is governed by i) solution pH, ii) intrinsic $pK_a$ of the head group, and iii) molecular interactions in the system.

To isolate the effect of the lipid shape on the structural properties of the system, we simulated ten lipid models of different bead sizes but similar intrinsic $pK_a$ and molecular volume (Fig. 2 and *SI Appendix*, Fig. S4 and Table S1). From these, we selected a small (S), a medium (M), and a large (L) model, which are ranked by head bead size in ascending order and by the size of the second tail bead in descending order (Fig. 2*H*). The lipids have approximately equal molecular volume, whereas the deviation from a cylindrical shape is larger for the S than for the L model (Fig. 2*H*). We quantify this deviation using the critical packing parameter, CPP, which is larger than one for all our models, as expected for lipids forming inverse phases. For systems containing mRNA and lipids with $pK_a$ = 9.41, we observe that both the center–center distance, *d,* and the radius of the water channel, $R_w$, are differently influenced by pH, depending on the aspect ratio of the constituent lipids (Fig. 2 *C* and *D*). While at pH < 6 the three lipid models show similar *d* values, at larger pH values, *d* decreases with increasing head size, as the lipid shape becomes more prominently cylindrical, $d_S > d_M > d_L$ (Fig. 2*C*). The dependence of $R_w$ on pH follows the same trend as *d* at high pH,

whereas larger head beads impart larger $R_w$ values at low pH (Fig. 2*D*). At pH 3 and 4, the large deviation of the $R_w$ value for the L lipids with respect to the other models highlights a difference in phase behavior. Namely, L lipids in the presence of mRNA form a lamellar phase at pH 3 to 4 and a stable $H_{II}$ phase at pH > 4, whereas M and S lipids form a stable $H_{II}$ phase in the explored pH range (Fig. 3). Despite the relatively small differences in $R_w$ across the models, our simulations predict a significant decrease in charge per lipid with decreasing head-bead size. The effective $pK_a$, defined as the pH value at which half of the lipids are positively charged, is 6 for L lipids and ~5 for S lipids (Fig. 2*E*). This corresponds to a down-shift of 3 to 4 units compared to the intrinsic $pK_a$ of the head group. The decrease in charge per lipid with decreasing head-bead size is accompanied by an increase in lipid packing, which can be quantified by the area per lipid (Fig. 2*F*). The decrease in area per lipid with increasing pH and decreasing head-bead size reflects both changes in $R_w$ and in the lipid-to-nucleotide ratio, which increases as the mRNA molecule becomes less extended (*SI Appendix*, Fig. S5). Besides packing, the shape of the lipids also affects their average orientational order, which we quantified by the order parameter $\langle \cos \theta \rangle$, where $\theta$ is the angle between the head-to-tail vector and the radial vector pointing toward the middle tail bead. For the S and M lipids, we observed that $\langle \cos \theta \rangle$ decreases with increasing pH (Fig. 2 *G* and *J*). On the other hand, the $H_{II}$ phase formed by the L model between pH 5 and 8 is characterized by high orientational order (Fig. 2 *G* and *I*). Notably, the trends described for L, M, and S lipids are fully consistent with those observed for the other seven lipid models designed in this study (*SI Appendix*, Fig. S4), which underscores the pivotal role of the critical packing parameter in influencing the internal structure and pH response of LNPs.

In our model, we use a dielectric constant of 49 to account for the nanoconfinement of water in the core of the LNPs (29, 30) Increasing the dielectric constant by 1/3—to match the value of 74 for bulk water at 37 °C—results in weaker headgroup–phosphate interactions. In turn, this effect leads to a shift toward larger center–center distances, lower lipid order, and decreased charge per lipid by a similar extent irrespective of lipid shape (*SI Appendix*, Fig. S6).

DD and MC3 are modeled as conical lipids with different intrinsic $pK_a$ values and different headgroup sizes (Fig. 4 *I* and *J*), as inferred from our AA simulations which show that DD forms $H_{II}$ phases with larger *d*, $R_w$, and *A* values than MC3 in the absence of mRNA (Fig. 1). Namely, we modeled MC3 and DD as S and L lipids, respectively, and used the $pK_a$ values of the lipid head groups determined experimentally in this study.
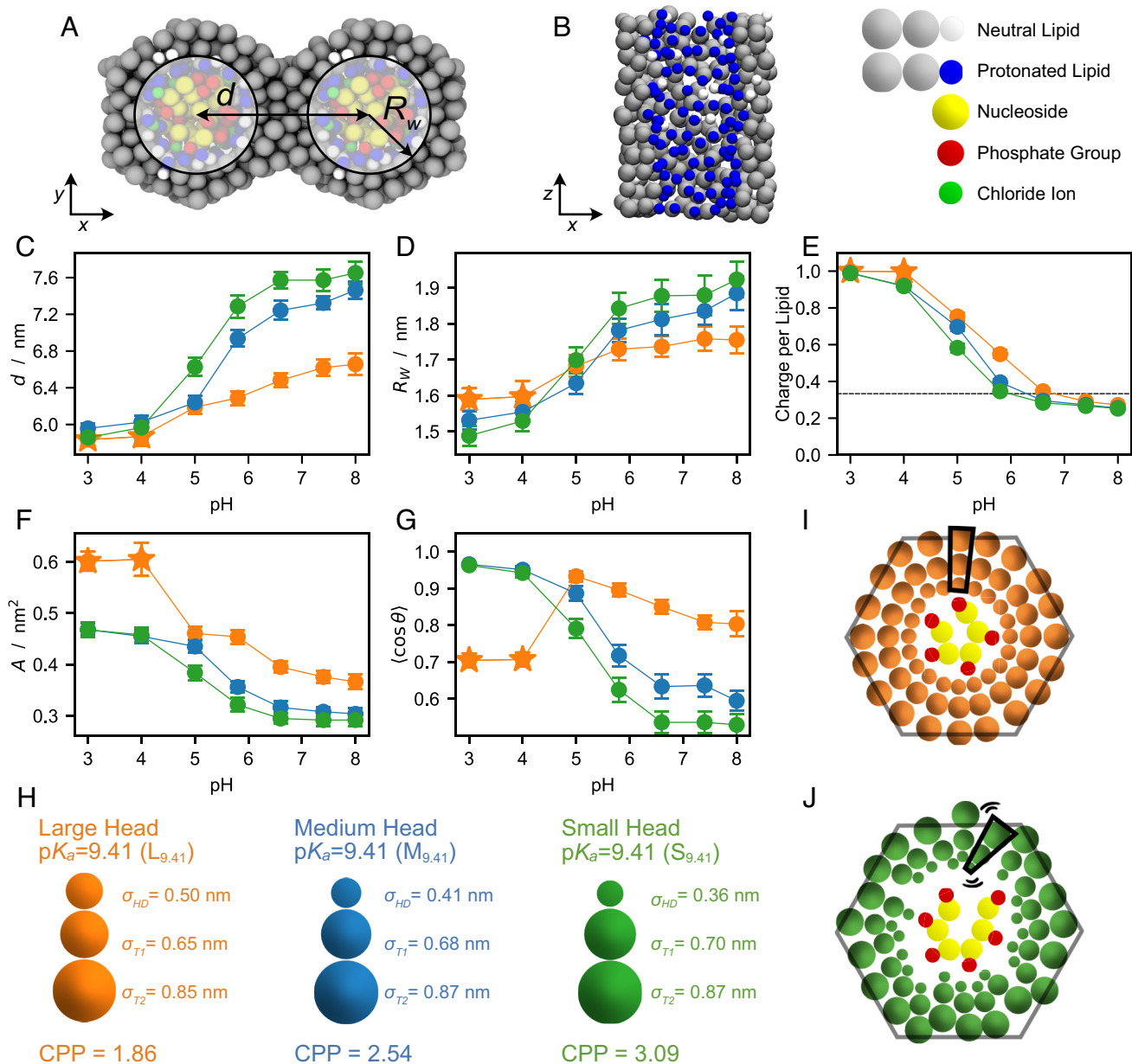
**Fig. 2.** Effect of lipid shape on the core of mRNA-containing LNPs at a fixed intrinsic p$K_a$ value. (*A*) Top-view molecular visualization of the coarse-grained model of the core of mRNA-containing LNPs. (*B*) Lipid monolayer in H$_{\text{II}}$ phase viewed from the center of the water core. (*C*) Center–center separation, (*D*) radius of the water core, (*E*) charge per lipid, and (*F*) area per lipid, (*G*) order parameter as a function of pH calculated from CG simulations of L$_{9.41}$ (orange), M$_{9.41}$ (blue), and S$_{9.41}$ (green) lipid models. Circles (○) and stars (★) indicate systems in H$_{\text{II}}$ and lamellar phase, respectively. (*H*) Schematic illustrations of the L, M, and S coarse-grained models and their critical packing parameters (CPP). (*I* and *J*) Schematic illustration of the packing and lipid orientational order in H$_{\text{II}}$ phases formed by (*I*) L and (*J*) S lipids at pH ≥ 6. Lipids are shown in orange (*I*) and (green), whereas nucleosides and phosphate groups are shown in yellow and red, respectively.

At pH 3 to 5, our CG MC simulations of MC3 and DD in the presence of mRNA predict a lipid-to-nucleotide ratio of ~3 (*SI Appendix*, Fig. S7*E*), which is in good agreement with the experimental composition of the mRNA-containing LNPs (7). At pH < 6, the larger head bead of the DD model imparts a larger $R_w$ than the MC3 model (*SI Appendix*, Fig. S7*B*). However, as opposed to the L model with p$K_a$ = 9.41, the lower charge per lipid of the DD model results in weaker electrostatic repulsion between head groups, which enables tighter lipid packing and a larger lipid-to-nucleotide ratio than for L$_{9.41}$ (*SI Appendix*, Fig. S7*E*). Because of this combination of ionization and shape effects, DD and MC3 show similar values of the apparent lipid p$K_a$ (*SI Appendix*, Fig. S7*C*).

**$^{13}$C-NMR Measurements of LNP Lipid Degree of Ionization.** To estimate the effect of pH on the protonation state of the ionizable-lipid component of LNPs, we measured $^{13}$C-NMR spectra in the chemical shift region corresponding to the α-carbons attached to the nitrogen atom in the MC3 and DD head groups (Fig. 4 *A* and *B* and *SI Appendix*, Fig. S8). At pH 3 to 5, the spectra for empty and eGFP mRNA-containing LNPs are very similar, which indicates that the presence of mRNA has insignificant influence on the local environment of the lipid head groups. However, at neutral pH, the differences between empty and mRNA-containing LNPs are pronounced. For empty LNPs, the spectra follow the expected signature of a pH-induced transition, that is, narrow peaks at low and high pH, with a broadening at intermediate
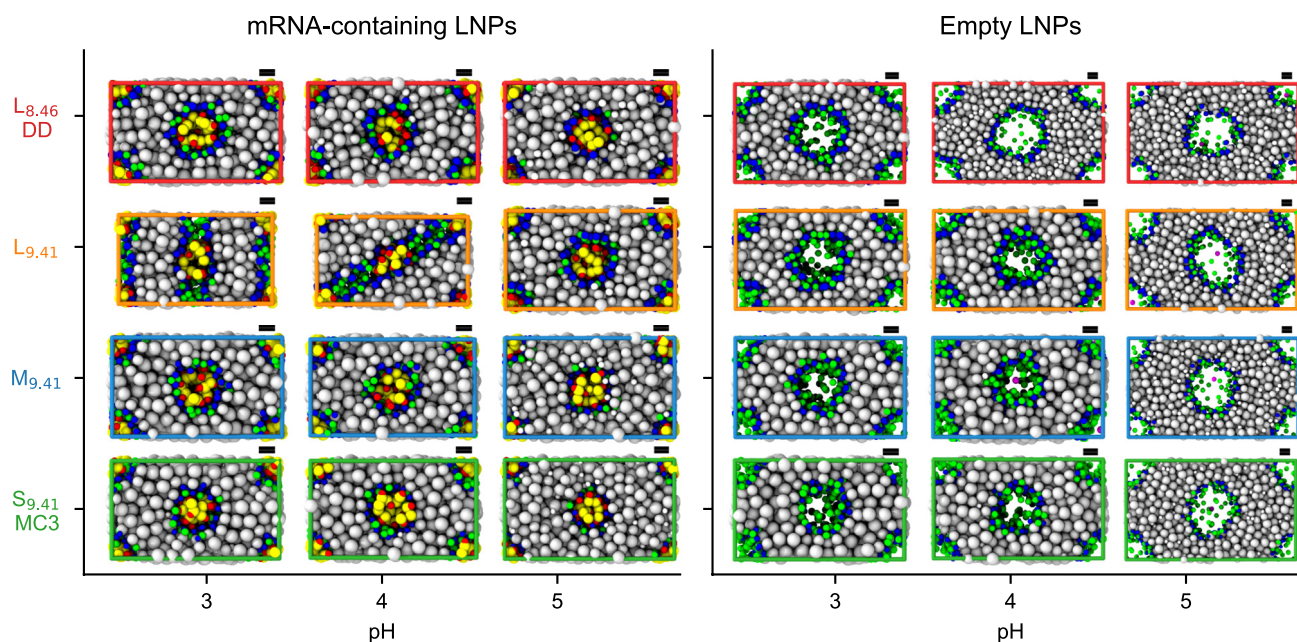
**Fig. 3.** Top-view molecular visualizations of the core of LNPs as a function of pH. Snapshots are from coarse-grained Monte Carlo simulations of the core of mRNA-containing (*Left*) and empty (*Right*) LNPs modeled using the $L_{8.46}$ (red), $L_{9.41}$ (orange), $M_{9.41}$ (blue), and $S_{9.41}$ (green) lipid models. Each rectangle shows a section of area $d \times \sqrt{3}d$ obtained by placing two hexagonal periodic images side by side. Beads are color coded as shown in the legend of Fig. 1*B*. Scale bars in black indicate 1 nm.

pH due to proton exchange. Conversely, for mRNA LNPs, we observe sharp signals only at low pH values and a broadening which persists beyond neutral pH, which suggests a surprising heterogeneity of charge states also at high pH. Similar results are obtained for empty and mRNA LNPs of DD (*SI Appendix*, Fig. S8).

Likewise, the chemical shifts for the α-carbon atoms of DD and MC3 have similar values with and without mRNA at low pH, whereas at pH > 5, the chemical shifts increase more steeply in the absence than in the presence of mRNA (Fig. 4 *C* and *D*). Encouraged by the coarse-grained simulation data, we make the approximation that the chemical shifts observed for the empty LNPs at pH 3 and 8 are truly representative of the fully ionized and the neutral state, respectively, also for the mRNA-containing LNPs. We then use those "end-point" chemical shifts to calculate the degree of ionization for both empty and mRNA LNPs of DD (Fig. 4*E*) and MC3 (Fig. 4*F*). Interestingly, we observe that for mRNA LNPs, the degree of ionization levels out to a value ~1/3, which corresponds to the charge neutralization of mRNA by ionizable lipids since the mRNA LNPs of both MC3 and DD were prepared at a 3:1 lipid-to-nucleotide molar ratio.

We also predicted the degree of ionization using constant-pH CG simulations for the L and S lipid models, using the experimental polar headgroup $pK_a$ values of 8.46 and 9.41 to represent DD and MC3, respectively (Fig. 4 *G* and *H*). The agreement between experiments and simulations is encouraging, and our molecular model suggests that a more prominently conical lipid shape can almost completely compensate for the effect of a higher headgroup $pK_a$ value on the properties of an mRNA-containing $H_{II}$ phase.

**Mean-Field Theory.** We finally used electrostatic mean-field theory at the Poisson–Boltzmann (PB) level to investigate the effect of pH on the lipid ionization state. For our numerical calculations, we used geometrical and thermodynamic parameters obtained from CG simulations, namely, the radius of the water core, the nucleotide-to-lipid ratio, the concentrations of sodium and chloride ions, and the lipid density (Fig. 5*A*). PB calculations

yield ionization curves with similar shapes as the predictions from CG simulations, where the charge per lipid ranges from unity at low pH values, to around 0.3 (≈the nucleotide-to-lipid ratio) at high pH values. However, in contrast to CG simulations, PB calculations predict a monotonous decrease of the charge per lipid with increasing pH and decreasing $pK_a$. Notably, PB calculations lack the plateau in the degree of ionization between pH 6 and 8 (Fig. 5*B*), which we observed in the ionization profiles obtained from CG simulations and NMR experiments (Fig. 4). Further, compared to CG simulations, the effect of the lipid size on the degree of ionization in PB calculations is less pronounced and shows the opposite trend, i.e., a systematic decrease in charge per lipid with increasing headgroup size (Fig. 5*B*).

## Discussion

For constant lipid volume and constant intrinsic $pK_a$ value, increasing pH leads to an increase in the hexagonal center–center separation, $d$, regardless of headgroup size (Fig. 2*C*). This general trend is due to weakened electrostatic attraction between lipids and nucleotides as the lipids gradually neutralize at higher pH. At low pH, the lipids are fully protonated (Fig. 2*E*), and the radius of the water core, $R_w$, increases with increasing headgroup size, as the smaller headgroups of S lipids can get closer to the oppositely charged nucleotides than the larger headgroups of L lipids (Fig. 2*D*). The trend is reversed at high pH where L lipids form an $H_{II}$ phase with smaller $d$ and $R_w$ values compared to S lipids, reflecting how neutral lipids pack in the absence of strong lipid–nucleotide interactions (Fig. 2 *I* and *J*). Of the three lipid models, neutral L lipids form an $H_{II}$ phase with higher lipid order, which is less sensitive to electrostatic interactions and weakly affected by changes in pH (Fig. 2 *G* and *I*). Conversely, S and M lipids rely on electrostatics for tight packing at low pH and form a more disordered phase when deprotonated (Fig. 2 *G* and *J*).

The lipid protonation state strongly depends on the local environment (Fig. 2*E*), and the apparent (or stoichiometric) $pK_a$ is downshifted by 4 to 5 units from the intrinsic $pK_a$ of the isolated
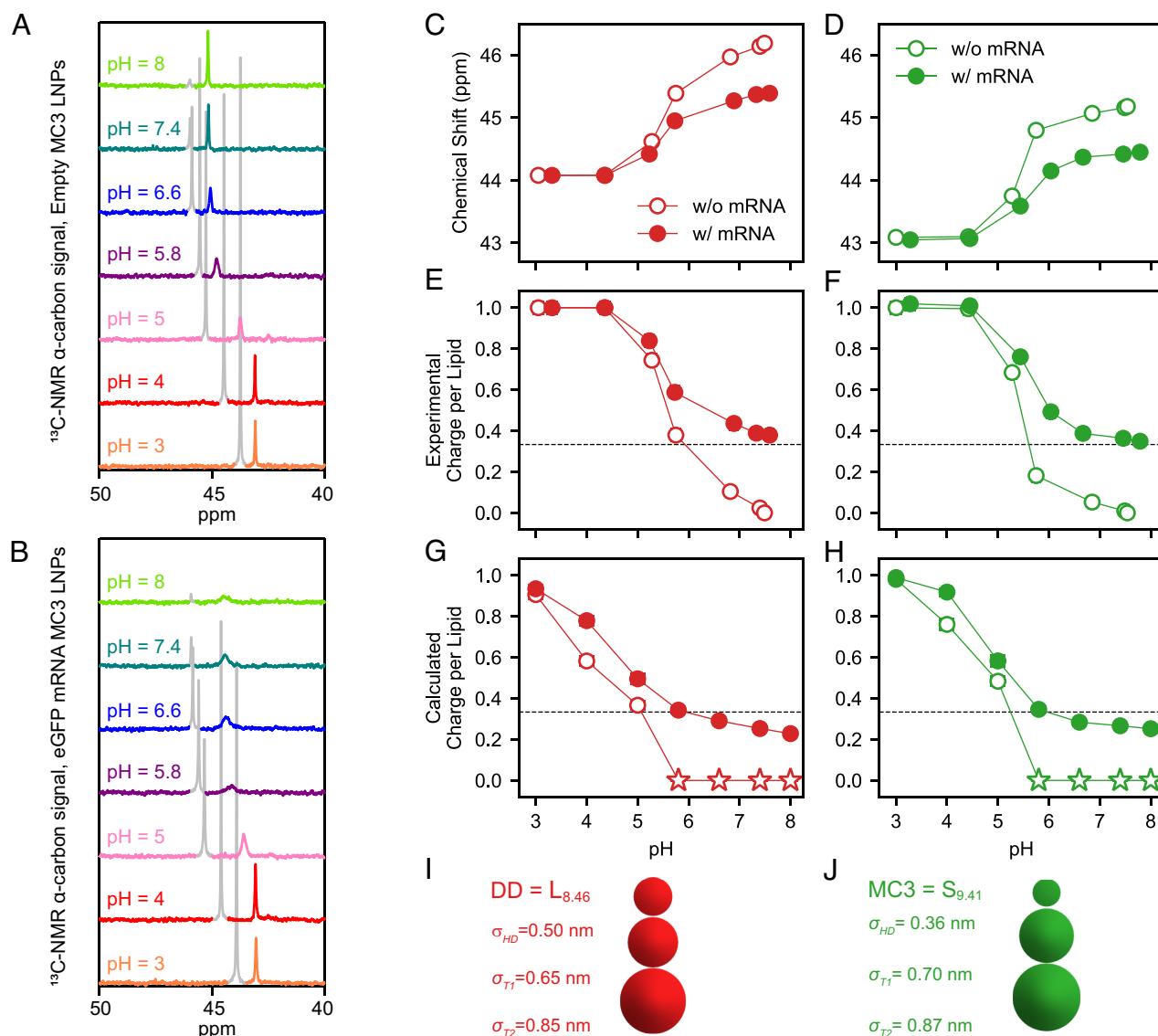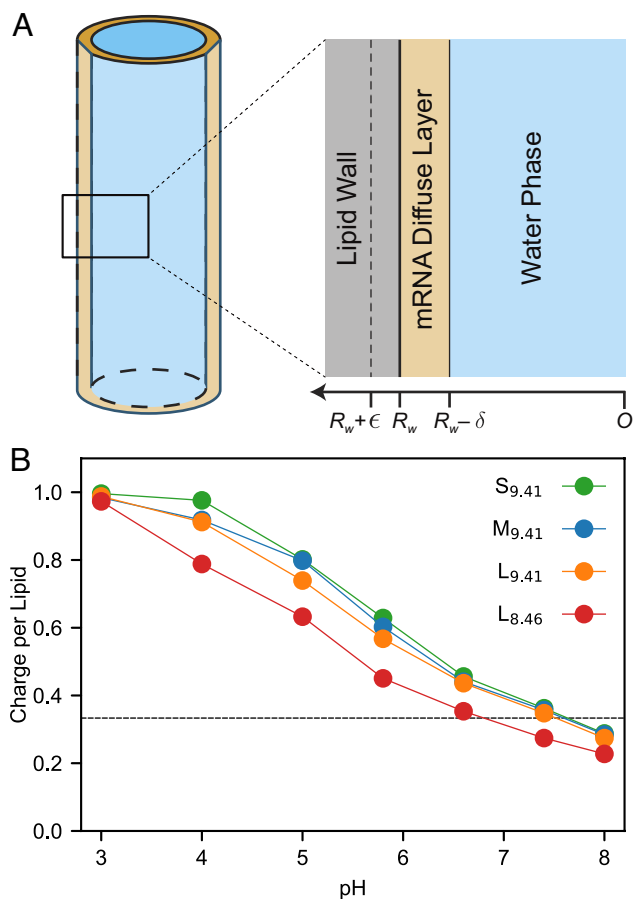
**Fig. 4.** Estimating the charge per lipid in the core of LNPs using NMR experiments and CG simulations. (*A* and *B*) $^{13}$C NMR spectra measured for empty (*A*) and mRNA-containing (*B*) LNPs of MC3 lipid. Peaks assigned to citrate molecules in the buffer served as an internal standard for pH calibration and are shown in gray. (*C* and *D*) $^{13}$C NMR chemical shifts measured for empty (open circles) and mRNA-containing (closed circles) LNPs made of **DD** (*C*) and **MC3** (*D*) lipids at various pH values. (*E–H*) Comparison between estimates from NMR experiments (*E* and *F*) and CG simulations (*G* and *H*) of the charge per lipid as a function of pH for DD (*E* and *G*) and MC3 (*D* and *H*) lipids in empty (open markers) and mRNA-containing (closed markers) LNPs. Circles (○) and stars (★) indicate systems in $H_{II}$ and disordered phase, respectively. Dashed lines indicate the nucleotide-to-lipid ratio in the experimental composition of the mRNA-containing LNPs. (*I* and *J*) Schematic illustrations of the DD (*I*) and MC3 (*J*) coarse-grained models.

headgroup. This perturbation is caused by electrostatic repulsion from i) other, like-charged headgroups and ii) desolvation due to the nonpolar environment created by the surrounding layer of lipid tails. That is, the acid–base equilibrium is strongly shifted toward the neutral, deprotonated state to minimize electrostatic repulsion. This effect is slightly less pronounced for L lipids, for which the large headgroup size sets a limit on how close the embedded charge can approach other charged moieties. It is noteworthy that the lipids do not fully deprotonate at high pH values in the $H_{II}$ phase containing mRNA. As further discussed below, this effect vanishes in the absence of mRNA and we attribute it to the negative electrostatic potential emanating from the central nucleic acid chain.

We now turn to NMR experiments and CG simulations of DD and MC3 lipids (Fig. 4). Whereas chemical shifts with and without mRNA are very similar at low pH, they differ significantly at high pH (Fig. 4 *C* and *D*). In the absence of mRNA, the system of neutral lipids at pH > 6 may adopt a disordered morphology distinct from

the $H_{II}$ phase. Because of this change in the chemical environment of the head groups, chemical shifts at pH ≥ 6 are considerably larger in the absence than in the presence of mRNA. This argument agrees very well with our simulations without mRNA (Fig. 4 *G* and *H*) where we indeed find that the $H_{II}$ phase is unstable for neutral lipids. Conversely, in the presence of mRNA, the lipids never fully deprotonate due to the ever-present negative electric potential from the central nucleotides, and the $H_{II}$ phase is maintained over the sampled pH range. Accepting this argument, the NMR chemical shifts can be reinterpreted as lipid protonation states (Fig. 4 *E* and *F*).

Both NMR experiments and CG simulations unambiguously show that the lipid degree of ionization vs pH is highly similar for DD and MC3, despite the difference of about one unit between their intrinsic $pK_a$ values. Our phenomenological lipid model offers a clear explanation for this effect: The acid–base equilibrium is governed by i) the intrinsic $pK_a$ value and ii) the local chemical environment, where the latter is highly affected by the lipid shape through packing.

**Fig. 5.** Estimating the charge per lipid in the core of LNPs using Poisson–Boltzmann calculations. (*A*) Schematic illustration of the geometry used for the Poisson–Boltzmann calculations. (*B*) Charge per lipid as a function of pH calculated from Poisson–Boltzmann calculations for $S_{9.41}$ (green), $M_{9.41}$ (blue), $L_{9.41}$ (orange), and $L_{8.46}$ (red) lipid models, i.e., the same color scheme as in Figs. 2 and 4.

obtained from CG simulations, neglect several aspects that might affect the titration behavior, namely: i) ion–ion, ion–lipid, and lipid–lipid correlations (32, 33); ii) excluded volumes; and iii) geometrical fluctuations. Moreover, in our PB calculations, all lipids are modeled with the same capacitance, whereas in principle, the capacitance might increase with decreasing headgroup size, as it is inversely proportional to the closest lipid–ion separation. However, accounting for this effect in our PB calculations would result in a larger difference between the apparent $pK_a$ value predicted for MC3 and DD, thereby increasing the discrepancy between PB and simulation predictions. Thus, PB theory captures the ionization general trend but not the dependence on the size of the headgroups.

## Conclusions

The pH-dependent ionization profile of the lipids is key for the efficacy of LNP delivery systems since mRNA release is triggered in a narrow pH window. It is therefore tempting to target the intrinsic headgroup $pK_a$ value when designing new lipids. However, lipids with significantly different intrinsic $pK_a$ values can have surprisingly similar ionization states at a given pH. Our explanation, which is based on NMR and computer simulations, is that the lipid shape has a large effect on the local chemical environment via its influence on lipid packing. Notably, we show that in the biomolecule-rich core of LNPs, subtle changes in lipid conicity strongly perturb the ionization state of the lipid headgroup and can shift the apparent $pK_a$ by several pH units. Specifically, our results indicate that the protonated state of poorly packed lipids is destabilized due to two electrostatic mechanisms: i) loss of contacts with the oppositely charged mRNA phosphate groups and ii) repulsion from the apolar tail groups. These effects are exemplified by the two differently shaped lipids investigated in this work, DD and MC3, which have almost identical apparent $pK_a$ values, despite a difference of one unit in the intrinsic $pK_a$ of their headgroups. Taken together, this study provides a molecular-level description of the interplay between the internal structure of the LNPs and the ionization state of the lipids. Our findings suggest that both lipid shape and headgroup intrinsic $pK_a$ are key determinants of the pH response and, hence, contribute to the efficient endosomal escape of mRNA-containing LNPs. This insight may facilitate the rational design of ionizable lipids for more efficacious and safer delivery systems, whereby the synthesis of polar headgroups and hydrophobic chains may target optimal combinations of intrinsic $pK_a$ and steric properties. Moreover, this work presents a computational approach to study the pH-dependent structural reorganization of biomolecule-rich phases, which is applicable to other multicomponent systems, such as biomolecular condensates, where changes in pH have important biological implications (34, 35).

## Materials and Methods

**Materials.** The ionizable cationic lipids (6Z, 9Z, 28Z, 31Z)-heptatriaconta-6, 9, 28, 31-tetraen-19-yl-4-(dimethylamino)butanoate (DLin-MC3-DMA or MC3) and 1, 2-dilinoleyloxy-n, n-dimethyl-3-aminopropane (DLin-DMA or DD) were synthesized at AstraZeneca. 1, 2-distearoyl-*sn*-glycero-3-phosphocholine (DSPC) was obtained from Avanti Polar Lipids, 1, 2-dimyristoyl-sn-glycero-3-phosph oethanolamine-N-[methoxy(polyethyleneglycol)-2000] (DMPE-PEG$_{2000}$) from NOF Corporation and Cholesterol (Chol) from Sigma-Aldrich. Enhanced Green Fluorescent Protein (eGFP) mRNA (996 nucleotides) ARCA capped modified with methoxiuridine was purchased from TriLink Biotechnologies. Hydrogeneous phosphate-buffered saline (PBS, 1 mM KH$_2$PO$_4$, 155 mM NaCl, and 3 mM Na$_2$HPO$_4$·7H$_2$O, pH 7.4) was obtained from Life Technologies. Citrate buffer was purchased from Teknova, and HyClone RNAse-free water was obtained from GE Healthcare Cell Culture. The polar headgroups of MC3 and DD were synthesized as described below.

Fig. 2*E* shows that regardless of pH, the charge per lipid decreases with decreasing headgroup size. This means that the apparent $pK_a$ value downshifts more for smaller head groups than for larger head groups. Therefore, MC3 (intrinsic $pK_a$ = 9.41, small head) and DD (intrinsic $pK_a$ = 8.46, large head) display very similar protonation profiles as observed with both NMR and simulations (Fig. 4 *E–H*). Our interpretation is that because of the larger critical packing parameter, neutral MC3 lipids may form a less ordered phase than neutral DD lipids (Fig. 2 *I* and *J*). This suggests two effects that may come into play at pH ≥ 6 and favor the downshift in apparent $pK_a$: Lipid heads are exposed i) less to the oppositely charged nucleotides and ii) more to the apolar lipid tails, which penalizes the protonated state through desolvation. Moreover, the tighter packing of MC3 lipids, compared to DD lipids (*SI Appendix*, Figs. S7 *D* and *E*), may imply stronger electrostatic repulsions between like-charged headgroups, which also contribute to the downshift of the apparent pKa (31). Note that the concept of a single apparent or stoichiometric $pK_a$ value inside a complex, pH-responsive environment has little predictive power as the shape of the proton titration curve can be highly asymmetric and ill-described by the Henderson–Hasselbalch equation. Here, we simply define the apparent $pK_a$ as the inflection point of the pH titration curve and use it merely qualitatively.

Explicitly accounting for molecular packing, desolvation, and correlations, the CG simulations well capture the local chemical environment of titratable head groups. Our PB calculations are more approximate and, although based on geometrical parameters

**Synthesis of Polar Headgroups.**

*Methyl 4-(dimethylamino)butanoate (DLin-MC3-DMA headgroup).* Thionyl chloride (1.669 mL, 22.87 mmol) was added slowly to a solution of 4-(dimethylamino)butanoic acid (1 g, 7.62 mmol) in anhydrous MeOH (20 mL) at 0 °C. The reaction mixture was warmed to 25 °C over a period of 1 h under nitrogen. The resulting mixture was stirred at 40 to 50 °C for 3 h. The reaction mixture was evaporated to dryness and redissolved in acetone (5.00 mL). The resulting solution was precipitated in ether (100 mL), and the precipitate was collected by filtration, washed with Et2O (100 mL), and dried under vacuum. The precipitate was dissolved in DCM (10.00 mL), and potassium carbonate (1.159 g, 8.39 mmol) was added in one portion. The mixture was stirred at room temperature for 24 h, filtered, and dried under reduced pressure to afford methyl 4-(dimethylamino)butanoate (0.499 g, 45.1 %) free base as a white solid. $^1$H NMR (500 MHz, DMSO-d6) 1.86–1.95 (m, 2H), 2.43 (t, 2H), 2.72 (s, 6H), 2.99–3.06 (m, 2H), 3.61 (s, 3H). UPLC, ms detection (ES+) ([M + H]+) = 146.1 Da; RT ELSD = 0.11 min.

*2, 3-dimethoxy-N, N-dimethylpropan-1-amine (DLin-DMA headgroup).* Into a 100-mL round-bottom flask, sodium hydride (2.63 g, 65.85 mmol) was added portion-wise to a stirred dissolved of 3-(benzyloxy)propane-1, 2-diol (5 g, 27.44 mmol) in DMF (50 mL) at 0 °C over a period of 5 min under nitrogen. methyl iodide (5.15 mL, 82.32 mmol) added dropwise to the stirred mixture at 0 °C over a period of 15 min under nitrogen. The resulting mixture was stirred at room temperature for 4 h. The reaction mixture was quenched by the addition of water/ice (100 mL) and washed sequentially with ethyl acetate (3 × 100 mL). The organic layers were combined and extracted with (water) (3 × 100 mL). The organic layer was dried over Na$_2$SO$_4$, filtered, and concentrated under reduced pressure to afford crude product. The resulting residue was purified by flash silica chromatography, elution gradient 0 to 100% EtOAc in hexanes. Product fractions were evaporated to dryness to afford the intermediate ((2, 3-dimethoxypropoxy)methyl)benzene (3.6864 g, 63.9 %) as a pale yellow oil. $^1$H NMR (500 MHz, CHLOROFORM-d) 3.39 (s, 3H), 3.49 (s, 3H), 3.51–3.64 (m, 5H), 4.58 (s, 2H), 7.28–7.40 (m, 5H). ((2, 3-dimethoxypropoxy)methyl)benzene (3.6864 g, 17.53 mmol) and palladium on carbon (2.61 g, 24.54 mmol) in methanol (30 mL) were stirred under an atmosphere of hydrogen at atmospheric pressure for 16 h. The reaction mixture was filtered through Celite and the reaction mixture was concentrated under reduced pressure to dryness to afford pale yellow oil. The resulting residue was purified by flash silica chromatography, elution gradient 0 to 100% EtOAc in hexanes. Product fractions were concentrated under reduced pressure to dryness to afford a second intermediate 2, 3-dimethoxypropan-1-ol (1.901 g, 90 %) as a pale, yellow oil. $^1$H NMR (500 MHz, DMSO-d6) 3.18–3.46 (m, 11H), 4.56 (br s, 1H). Finally, trifluoromethanesulfonic anhydride (0.773 mL, 4.58 mmol) was added slowly to a stirred mixture of 2, 3-dimethoxypropan-1-ol (0.5 g, 4.16 mmol) and 2, 6-dimethylpyridine (1.015 mL, 8.74 mmol) in DCM (20 mL) at −40 °C over a period of 5 min under nitrogen. The resulting mixture was stirred at −40 °C for 1 h. The reaction mixture was warmed to room temperature and stirred for an additional 1 h. The reaction mixture was evaporated to dryness and redissolved in dimethylamine in THF (31.2 mL, 62.42 mmol) at 0 °C. The reaction was stirred overnight and concentrated under reduced pressure to dryness to afford crude product. The resulting residue was purified by flash silica chromatography, elution gradient 0 to 100% (20% MeOH and 1% NH$_4$OH in DCM) in DCM. Product fractions were concentrated under reduced pressure to dryness to afford 2, 3-dimethoxy-N, N-dimethylpropan-1-amine (0.343 g, 56.0 %) as a yellow oil. $^1$H NMR (500 MHz, DMSO-d6) 2.77 (s, 6H), 3.11–3.21 (m, 2H), 3.27–3.31 (m, 3H), 3.37 (s, 3H), 3.39–3.51 (m, 2H), 3.73 (m, 1H). UPLC, ms detection (ES) ([M + H]+) = 148.9; tR ELSD = 0.23 min.

**Lipid Nanoparticles (LNPs) Preparation and Characterization.** LNPs were prepared using the microfluidic setup described in detail elsewhere (36). Briefly, stocks of lipids were dissolved in ethanol and mixed in the appropriate molar ratios to obtain a lipid concentration of 12.5 mM (1.85 mg/mL). mRNA was diluted in RNAse-free 50 mM citrate buffer pH 3.0 to obtain a mRNA:lipid weight ratio of 10:1 (CIL:nucleotide 3:1 molar ratio). Empty LNPs were also prepared using 50 mM citrate buffer as the aqueous phase. The aqueous and ethanol solutions were mixed in a 3:1 volume ratio using a microfluidic apparatus NanoAssemblr, from Precision NanoSystems Inc., at a mixing rate of 12 mL/min. LNPs were dialyzed overnight against 500 × sample volume using Slide-A-Lyzer™ G2 dialysis cassettes from Thermo Scientific with a molecular weight cutoff of 10 K. For $^{13}$C-NMR

measurements, the particles were concentrated after dialysis to approximately 0.3 mg/mL of lipid using Amicon ultracentrifugation filters.

The size of LNPs was determined by dynamic light scattering (DLS) measurements using a Zetasizer Nano ZS from Malvern Instruments Ltd. The encapsulation and concentration of mRNA were determined using the RiboGreen assay. The encapsulation in all the samples was typically 90–99%. All the samples were prepared within a week of use to ensure the chemical stability of the components.

**NMR.** $^{13}$C-NMR experiments were acquired on a 600-MHz Bruker Avance III NMR spectrometer equipped with a QNP ($^1$H, $^{31}$P, $^{13}$C, $^{15}$N) cryoprobe at 37 °C. In order to obtain an NMR lock signal for the LNP preparations, samples were prepared using 10% deuterated water. A standard Bruker zgpg30 pulse sequence with $^1$H decoupling was used with an acquisition time of 0.6 s and a delay of 0.1 s. Typically, 40,000 pulses were acquired with a sweep width of 33,333 Hz using 32 k data points. The data were processed using zero filling to 64 k and a line broadening factor of 3 Hz.

**All-Atom Molecular Dynamics Simulations.** The CHARMM-GUI (37) membrane builder module (38) was used to generate a system of dilinoleylphosphatidylcholine (DLiPC) lipids in inverse hexagonal phase. The hexagonal cell has a length of 20 nm and a radius of 3 nm. Lipids are oriented radially around a cylindrical region parallel to the *z*-axis which encloses the aqueous phase. We chose DLiPC because it has the same lipid tail as DD and MC3 (18 carbon atoms and double bonds between C9 and C10, and C12 and C13), and subsequently modified the headgroups of DLiPC to those of DD and MC3. After an initial equilibration step, the system is placed in direct contact with two 5-nm-wide bulk water regions located at the ends of the long axis of the hexagonal cell. The excess aqueous phase allows water and ions inside the cylindrical region to equilibrate with the bulk solution, which results in a fully hydrated inverse hexagonal phase. We built systems of both neutral and fully protonated lipids. Whereas neutral systems contain only lipids and water molecules, ionized systems additionally contain chloride counterions, for electroneutrality, and cholesterol in 3:1 lipid-to-cholesterol molar ratio to enhance the stability of the H$_{II}$ phase. The addition of excess water and the ionization of the system was carried out using VMD (39) scripts. CHARMM36 all-atom lipid parameters (40) and the TIP3P water model (41) were employed. All molecular dynamics simulations were carried out using NAMD 2.12 (24, 42, 43). Long-distance electrostatic interactions were evaluated using the particle-mesh-Ewald method (44). A cutoff of 1.2 nm was used for the nonbonded interactions. The nearest-neighbor list cutoff was set to 1.4 nm. After an initial energy-minimization of the starting configurations, three equilibration steps were performed with a timestep of 1 fs and for a total simulation time of 21 ns. The first equilibration step was performed in the NVT ensemble, whereas the NPT ensemble at 1 atm was used for the next two equilibration steps as well as for the production runs. In all simulations, the ratio between the lateral (*x*- and *y*-) dimensions was kept constant, and periodic boundary conditions were applied in three dimensions. The ionized systems were simulated at 298 K for 400 ns since these systems underwent a phase transition to lamellar state at 310 K, whereas the inverse hexagonal phase remained stable throughout the simulations at the lower temperature. The last 300 ns of the trajectories were used for the analyses. $R_w$ and $\phi$ were estimated by counting the number of water molecules within the cylindrical water channel, assuming a bulk water density of 1 g/mL. Since the calculation of theoretical SAXS spectra requires a hexagonal phase spanning the periodic boundaries along the *z*-axis, we constructed systems without bulk water regions based on the densities of molecules and ions calculated for systems where the H$_{II}$ phase is flanked by water regions. An additional 500-ns simulation was performed, and configurations sampled every 50 ns were used for calculating theoretical SAXS spectra.

**Coarse-Grained Monte Carlo Simulations.** We developed an implicit-solvent coarse-grained (CG) model for reactive Monte Carlo simulations of the inner core of LNPs comprising ionizable lipids, mRNA, and NaCl. Our starting point was a three-bead lipid model which i) spontaneously self-assembles into lipid bilayers, ii) reproduces experimental mechanical properties and iii) captures the dielectric mismatch between bilayer core and surrounding aqueous medium (45, 46). To favor the formation of the inverse hexagonal phase over the lamellar phase, we modified the shape of the CG lipid from cylindrical to conical. The size of the lipid beads was derived from the geometrical constraints imposed by the structure of the core of the mRNA-containing LNPs, which was characterized experimentally

by Yanez Arteta et al. (7) and in silico in the present work (center–center distance $d = 6.2$ nm, volume per lipid of ~1 nm$^3$ and ~22 lipids per a cross-section of the H$_{II}$ phase). The radius of the terminal tail bead is set to $R_{T2} = 0.43$ nm, and the radii of the middle-tail and head beads are derived as $R_{T1} = \frac{\left(d/\sqrt{3} - 2^{7/6}R_{T2}\right)}{\cot(\pi/22)+1}$

---

$\approx 0.33$ nm and $R_{HD} = \frac{d/\sqrt{3} - 2^{7/6}(R_{T2}+R_{T1})}{\cot(\pi/22)+1} \approx 0.235$ nm. These equations are obtained by considering each lipid to be inscribed in cones lying perpendicular to the water core with the vertexes at its axis. Each cone has radius $R_{T2} = 0.43$ nm and height $d/\sqrt{3} - R_{T2}$. The size of the remaining lipid beads can be derived from the following system of equations:

$$\begin{cases} \tan\frac{\theta}{2} = \frac{R_{T2}}{d/\sqrt{3} - R_{T2}} \\ R_{T1} = \left(d/\sqrt{3} - 2R_{T2} - R_{T1}\right)\tan\frac{\theta}{2}, \\ R_{HD} = \left(d/\sqrt{3} - 2R_{T2} - 2R_{T1} - R_{HD}\right)\tan\frac{\theta}{2} \end{cases}$$

where $\theta$ is the aperture of the cones. Compared to DD, our all-atom simulations show that MC3 forms a H$_{II}$ phase with larger monolayer thickness and smaller area per lipid. Therefore, for the CG model of MC3, we used a more pronounced conical shape than for DD. The conicity is quantified by the critical packing parameter CPP $= \frac{v}{\pi \times R_{HD}^2 \times h}$ (47), where $v$ is the volume of a truncated cone of height $h = R_{HD} + 2R_{T1} + R_{T2}$, and with upper and lower radii $R_{HD}$ and $R_{T2}$, respectively. To systematically investigate the effect of lipid shape on head-group protonation state, we scanned the bead sizes and generated ten lipid models with CPP values between 2.06 and 3.03, while keeping $h = 1.32 \pm 0.01$ nm and the sum of the bead volumes within $0.54 \pm 0.01$ nm$^3$ (*SI Appendix*, Table S1).

Bonded interactions within lipids are modeled via the FENE potential,

$$w_{FENE}(r_{ij}) = -\frac{1}{2}k_{FENE}r_\infty^2 \ln\left[1 - \left(\frac{r_{ij}}{r_\infty}\right)^2\right],$$

where $r_{ij}$ is the distance between bead $i$ and $j$, $k_{FENE} = 120$ kJ mol$^{-1}$ nm$^{-2}$, and $r_\infty = 3(R_{T2} + R_i)/2$ with $i = T2$ or $HD$. Lipid molecules are kept in extended conformation by a harmonic bond between the terminal beads,

$$w_{harm}(r_{ij}) = \frac{1}{2}k_{harm}(r_{ij} - r_{eq})^2,$$

where $k_{harm} = 40$ kJ mol$^{-1}$ nm$^{-2}$ and $r_{eq} = 3R_{T1} + 2.5(R_{HD} + R_{T2})$. The NaCl model consists of oppositely charged beads of radius $R_{Na} = R_{Cl} = 0.23$ nm and reproduces experimental activity coefficient derivatives for NaCl electrolyte solutions of a wide range of salt concentrations. The mRNA molecule is modeled as a chain of neutral unphosphorylated nucleotide beads of radius $R_{Nuc} = 0.35$ nm, each connected to negatively charged monophosphate beads of radius $R_P = 0.23$ nm. All mRNA beads are bound via harmonic bonds with $k_{harm} = 480$ kJ mol$^{-1}$ nm$^{-2}$ while $r_{eq} = 0.5$ nm between unphosphorylated nucleotide beads and $r_{eq} = 0.35$ nm between unphosphorylated nucleotide and monophosphate beads. To validate the mRNA model, we performed Monte Carlo simulations of a single chain of 30 nucleotide beads at different ionic strengths and found good agreement between radii of gyrations predicted from simulations and the values obtained by Plumridge et al. from small-angle X-ray scattering experiments (48) (*SI Appendix*, Fig. S9A).

All particles interact via the Coulomb and the Weeks–Chandler–Andersen potential

$$w_{WCA}(r_{ij}) = \begin{cases} 4\epsilon_{ij}\left[\left(\frac{b_{ij}}{r_{ij}}\right)^{12} - \left(\frac{b_{ij}}{r_{ij}}\right)^6 + \frac{1}{4}\right], r_{ij} \leq r_c, \\ 0, r_{ij} > r_c, \end{cases}$$

where $b_{ij}$ is the sum of the radii of beads of type $i$ and $j$, $r_c = 2^{1/6}b_{ij}$, and $\epsilon_{ij} = 8.3145 \times T$ J mol$^{-1}$ except for $\epsilon_{NaCl} = 0.01$ kJ mol$^{-1}$. In the implicit solvent, the self-assembly of the lipids into the H$_{II}$ phase is driven by an effective long-range attractive interaction between the tail beads,

$$w_{cos}(r_{ij}) = \begin{cases} -\epsilon_{ij}, r_{ij} \leq r_c \\ -\epsilon_{ij}\cos^2\left[\frac{\pi(r_{ij}-r_c)}{3.4\,b_{ij}}\right], r_c \leq r_{ij} \leq r_c + 1.7\,b_{ij} \\ 0, r_{ij} < r_c + 1.7\,b_{ij} \end{cases}$$

where $\epsilon_{ij} = 8.3145 \times T$ J mol$^{-1}$. Charged beads interact via the Coulomb potential

$$w_{CC}(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r_{ij}},$$

where $q_i$ and $q_j$ are the charges of the interacting beads and $\epsilon_0$ is the vacuum permittivity. The dielectric constant was set to $\epsilon_r = 49$ to account for the lower average dielectric constant exhibited by water nanochannels (29, 30). For comparison, a subset of simulations was also performed using $\epsilon_r = 74$, i.e., the dielectric constant of bulk water at 310 K (*SI Appendix*, Fig. S6) (49).

In the single-chain simulations used to validate the mRNA model, the presence of salt in solution is instead treated implicitly via the Debye–Hückel potential

$$w_{DH}(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r_{ij}}e^{-r_{ij}/\lambda_D},$$

where $q_i$ and $q_j$ are the charges of the phosphate beads, $\epsilon_r = 78.5$ (49), and $\lambda_D = \sqrt{0.0924 \text{ M nm}^2/c_s}$ is the Debye length for an aqueous solution of NaCl molar concentration $c_s$ at 298 K.

The repulsion of a charged bead from the core of the lipid monolayer is approximated by an ion-induced dipole interaction (46),

$$w_{CN}(r_{ij}) = -\frac{q_i^2 \alpha_j R_j^3}{8\pi\epsilon_0\epsilon_r r_{ij}^4},$$

where $R_j$ and $\alpha_j$ are, respectively, the radius and the unitless excess polarizability of the second lipid tail beads, T2, and $\epsilon_r = 74$ is the dielectric constant of water at 310 K (49). The value of $\alpha_j$ used in this study was determined by matching the free energy barrier for the translocation of a sodium ion across a DOPC bilayer at 320 K. Via Wang–Landau Monte Carlo simulations of a lipid bilayer consisting of 150 cylindrical lipids with $R_{T1} = R_{T2} = 0.45$ nm $R_{HD} = 0.4275$ nm (46), we determined that $\alpha_{T2} = -18$ yields a free energy barrier in good agreement with the reference value of $53 \pm 4$ kJ mol$^{-1}$ estimated from all-atom simulations by Pokhrel et al. (50) (*SI Appendix*, Fig. S9B).

The simulation cell is a hexagonal prism of volume 320 nm$^3$ with periodic boundary conditions. Monte Carlo isochoric moves allow for the aspect ratio of the simulation cell to fluctuate while keeping the volume constant (46). The RNA molecule is initially aligned to the $z$-axis with its termini are connected through the periodic images. Nucleotide and monophosphate beads are confined in a cylindrical region of radius 1.2 nm parallel to the $z$-axis using a harmonic potential of force constant of 10.0 kJ mol$^{-1}$ nm$^{-2}$. The system is simulated at $T = 310$ K (and also at 298 K for empty LNPs) in the semigrand canonical ensemble using the reactive canonical Monte Carlo method (21), wherein the simulation cell is at equilibrium with a bulk solution that acts as a reservoir of lipids, salt, and protons. The concentrations of neutral lipids and NaCl fluctuate throughout the simulations via Monte Carlo insertion/deletion moves, which are associated with the following energy change:

$$\frac{\Delta U}{k_B T} = -\sum_i \ln\left(\frac{N_i!}{(N_i + \nu_i)!}V^{\nu_i}\right) - \ln\prod_i a_i^{\nu_i},$$

where $k_B$ is the Boltzmann constant, $V$ is the volume of the simulation cell, and $N_i$ is the number of particles of species $i$ in the current state. $\nu_i$ is the stoichiometric coefficient of species $i$, which is equal to $+1$ or $-1$ depending on whether the species is inserted or deleted. $a_i$ is the activity of species $i$, which is pH-dependent for both neutral lipids and salt. Assuming an overall activity in the bulk solution of $10^{-3}$ (molar scale) for all lipids, the Henderson–Hasselbalch equation is used to estimate the activity of the neutral lipids via

$$a_L = \frac{10^{-3}}{1 + 10^{pK_a - pH}},$$

where $pK_a$ is the acidic dissociation constant of the polar headgroup measured experimentally in the present study for DD and MC3. The activity of sodium and chloride ions is approximated to the product of the monovalent ion concentrations in the McIlvain buffer (51) and the experimental activity coefficient of NaCl aqueous solutions of equivalent concentrations at 298.15 K (52) interpolated as described elsewhere (53, 54). Lastly, the charge state of the lipid head groups can fluctuate by means of (de)protonation Monte Carlo moves, which are coupled with insertion/deletion of NaCl, to ensure electroneutrality, and included in the energy change, $\Delta U$, given in the equation above. All Monte Carlo simulations were performed using Faunus version 2.4.2 (55).

**Modified Poisson–Boltzmann Calculations.** We use a modified Poisson–Boltzmann approach in cylindrical geometry. The radius of the cylinder was set to the radius of the water core, $R_w$, obtained from the CG simulations at each corresponding pH. The presence of the mRNA molecule is modeled by smearing out the charge of the macromolecule within an inner cylindrical layer of thickness $\delta = 0.3$ nm (Fig. 5A). To model the water core, we used a Bjerrum length, $l_B$, of 0.7 nm, corresponding to a dielectric constant of 80 and room temperature. We solved the following differential equation:

$$\frac{d^2\psi(r)}{d^2r} + \frac{1}{r}\frac{d\psi(r)}{dr} = -4\pi l_B\left[-\rho_{mRNA}(r) + \sum_i \rho_i z_i \exp\left(-z_i\psi(r)\right)\right],$$

between $r = 0$ and $r = R_w$, and

$$\frac{d^2\psi(r)}{d^2r} + \frac{1}{r}\frac{d\psi(r)}{dr} = 0,$$

between $r = R_w$ and $r = R_w + \epsilon$, where $r$ is the radial distance away from the center of the cylinder, $R_w$ is the radius of the cylinder, $\rho_i$ the ion density of species $i$, $z_i$ the valency of ion species $i$, $\rho_{mRNA}$ the charge density of the mRNA layer, $\epsilon$ the capacitance separation, and $\psi(r) = \beta e\phi(r)$ is the reduced form of the electrostatic potential, $\phi$. The capacitance separation, $\epsilon$, reflects the closest distance between a lipid headgroup and an ion, which gives rise to an extra electrostatic potential increase or decrease at the center of the lipid headgroups. We used a fixed value of $\epsilon = 0.25$ nm throughout our calculations. The density of mRNA was defined as

$$\rho_{mRNA}(r) = \frac{N}{L}|\sigma_L|\left(\frac{\delta R_w}{(R_w + \epsilon)} - \frac{\delta^2}{2(R_w + \epsilon)}\right)^{-1},$$

between $R_w - \delta < r < R_w$, and zero elsewhere, where $\sigma_L$ is the lipid surface density, and N/L is the nucleotide-to-lipid ratio. To solve the PB equation, we need two boundary conditions given by

$$\frac{\partial\psi}{\partial r}(r = 0) = 0,$$

and

$$\frac{\partial\psi}{\partial r}(r = R_w + \epsilon) = 4\pi l_B\sigma_L\alpha,$$

where $\alpha$ is the degree of ionization and given by

$$\alpha = \frac{\exp\left[\ln10\left(pK_a - pH\right) + \psi\right]}{1 + \exp\left[\ln10\left(pK_a - pH\right) + \psi\right]}.$$

The differential equation was solved by guessing an initial reduced potential, either at the surface or in the middle, and iteratively updated until the boundary conditions were fulfilled (37–44).

Author affiliations: ^aStructural Biology and NMR Laboratory & The Linderstrøm-Lang Centre for Protein Science, Department of Biology, University of Copenhagen, Copenhagen DK-2200, Denmark; ^bDepartment of Chemistry, Division of Computational Chemistry, Lund University, Lund SE-221 00, Sweden; ^cThe Hartree Centre, Science and Technology Facilities Council (STFC) Daresbury Laboratory, Warrington WA4 4AD, United Kingdom; ^dAdvanced Drug Delivery, Pharmaceutical Sciences, BioPharmaceuticals Research and Development, AstraZeneca, Gothenburg, Mölndal 431 83, Sweden; ^eMedicinal Chemistry, Early Respiratory & Immunology, BioPharmaceuticals Research and Development, AstraZeneca, Gothenburg, Mölndal 431 83, Sweden; and ^fData Science and Modelling, Pharmaceutical Sciences, BioPharmaceuticals Research and Development, AstraZeneca, Gothenburg, Mölndal 431 83, Sweden

1. Z. X. Chong, S. K. Yeap, W. Y. Ho, Transfection types, methods and strategies: A technical review. *PeerJ* **9**, e11165 (2021).
2. I. Koltover, T. Salditt, J. O. Rädler, C. R. Safinya, An inverted hexagonal phase of cationic liposome–DNA complexes related to DNA release and delivery. *Science* **281**, 78–81 (1998).
3. S. Sabnis *et al.*, A novel amino lipid series for mRNA delivery: Improved endosomal escape and sustained pharmacology and safety in non-human primates. *Mol. Therapy* **26**, 1509–1519 (2018).
4. A. Akinc *et al.*, Targeted delivery of RNAi therapeutics with endogenous and exogenous ligand-based mechanisms. *Mol. Therapy* **18**, 1357–1364 (2010).
5. G. Sahay *et al.*, Efficiency of siRNA delivery by lipid nanoparticles is limited by endocytic recycling. *Nat. Biotechnol.* **31**, 653–658 (2013).
6. A. Wittrup *et al.*, Visualizing lipid-formulated siRNA release from endosomes and target gene knockdown. *Nat. Biotechnol.* **33**, 870–876 (2015).
7. M. Yanez Arteta *et al.*, Successful reprogramming of cellular protein production through mRNA delivered by functionalized lipid nanoparticles. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E3351–E3360 (2018).
8. M. Schlich *et al.*, Cytosolic delivery of nucleic acids: The case of ionizable lipid nanoparticles. *Bioeng. Trans. Med.* **6**, e10213 (2021).
9. M. Jayaraman *et al.*, Maximizing the potency of siRNA lipid nanoparticles for hepatic gene silencing in vivo. *Angew. Chem. Int. Ed.* **51**, 8529–8533 (2012).
10. C. Chan, S. Du, Y. Dong, X. Cheng, Computational and experimental approaches to investigate lipid nanoparticles as drug and gene delivery systems. *Curr. Topics Med. Chem.* **21**, 92–114 (2021).
11. X. Lin, Current advances in computational and experimental approaches for nanoparticle–drug conjugates. *Curr. Top. Med. Chem.* **21**, 90–91 (2021).
12. M. Cornebise *et al.*, Discovery of a novel amino lipid that improves lipid nanoparticle performance through specific interactions with mRNA. *Adv. Funct. Mater.* **32**, 2106727 (2022).
13. A. N. Rissanou, A. Ouranidis, K. Karatasos, Complexation of single stranded RNA with an ionizable lipid: An all-atom molecular dynamics simulation study. *Soft Matter* **16**, 6993–7005 (2020).
14. M. Ramezanpour *et al.*, Ionizable amino lipid interactions with POPC: Implications for lipid nanoparticle function. *Nanoscale* **11**, 14141–14146 (2019).
15. J. A. Kulkarni *et al.*, On the formation and morphology of lipid nanoparticles containing ionizable cationic lipids and siRNA. *ACS Nano* **12**, 4787–4795 (2018).

16. G. Settanni, W. Brill, H. Haas, F. Schmid, pH-dependent behavior of ionizable cationic lipids in mRNA-carrying lipoplexes investigated by molecular dynamics simulations. *Macromol. Rapid Commun.* **43**, e2100683 (2021).
17. M. Paloncýová, P. Čechová, M. Šrejber, P. Kührová, 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).
18. B. M. H. Bruininks, P. C. T. Souza, S. J. Marrink, A practical view of the Martini force field. *Methods Mol. Biol.* **2022**, 105–127 (2021).
19. B. M. H. Bruininks, P. C. T. Souza, H. Ingolfsson, S. J. Marrink, A molecular view on the escape of lipoplexed DNA from the endosome. *eLife* **9**, e52012 (2020).
20. J. Corsi, R. W. Hawtin, O. Ces, G. S. Attard, S. Khalid, DNA lipoplexes: Formation of the inverse hexagonal phase observed by coarse-grained molecular dynamics simulation. *Langmuir* **26**, 12119–12125 (2010).
21. J. K. Johnson, A. Z. Panagiotopoulos, K. E. Gubbins, Reactive canonical Monte Carlo. *Mol. Phys.* **81**, 717–733 (1994).
22. S. C. Semple *et al.*, Rational design of cationic lipids for siRNA delivery. *Nat. Biotechnol.* **28**, 172–176 (2010).
23. R. Wang, Z.-G. Wang, Continuous self-energy of ions at the dielectric interface. *Phys. Rev. Lett.* **112**, 136101 (2014).
24. M. W. Tate, S. M. Gruner, Temperature dependence of the structural dimensions of the inverted hexagonal (HII) phase of phosphatidylethanolamine-containing membranes. *Biochemistry* **28**, 4245–4253 (1989).
25. F. Sebastiani *et al.*, Apolipoprotein E binding drives structural and compositional rearrangement of mRNA-containing lipid nanoparticles. *ACS Nano* **15**, 6709–6722 (2021).
26. K. Takács-Novák, K. J. Box, A. Avdeef, Potentiometric pKa determination of water-insoluble compounds: Validation study in methanol/water mixtures. *Int. J. Pharm.* **151**, 235–248 (1997).
27. I. R. Cooke, M. Deserno, Solvent-free model for self-assembling fluid bilayer membranes: Stabilization of the fluid phase based on broad attractive tail potentials. *J. Chem. Phys.* **123**, 224710 (2005).

28. I. R. Cooke, M. Deserno, Coupling between lipid shape and membrane curvature. *Biophys. J.* **91**, 487–495 (2006).

29. L. Fumagalli *et al.*, Anomalously low dielectric constant of confined water. *Science* **360**, 1339–1342 (2018).

30. S. Mondal, B. Bagchi, Comment on "investigation of dielectric constants of water in a nano-confined pore" by H. Zhu, F. Yang, Y. Zhu, A. Li, W. He, J. Huang and G. Li, RSC Adv., 2020, 10, 8628. *RSC Adv.* **11**, 5179–5181 (2021).

31. J. Philipp *et al.*, pH-dependent structural transitions in cationic ionizable lipid mesophases are critical for lipid nanoparticle function. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2310491120 (2023).

32. C. Labbez, B. Jönsson, M. Skarba, M. Borkovec, Ion−ion correlation and charge reversal at titrating solid interfaces. *Langmuir* **25**, 7209–7213 (2009).

33. M. Trulsson, Ion and Site Correlations of Charge Regulating Surfaces: A Simple and Accurate Theory. *Langmuir* **39**, 7642–7647 (2023).

34. F. G. Quiroz *et al.*, Liquid–liquid phase separation drives skin barrier formation. *Science* **367**, eaax9554 (2020).

35. S. L. Currie *et al.*, Quantitative reconstitution of yeast RNA processing bodies. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2214064120 (2023).

36. A. K. K. Leung *et al.*, Lipid nanoparticles containing siRNA synthesized by microfluidic mixing exhibit an electron-dense nanostructured core. *J. Phys. Chem. C* **116**, 18440–18450 (2012).

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

38. E. L. Wu *et al.*, CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *J. Comput. Chem.* **35**, 1997–2004 (2014).

39. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. *J. Mol. Graphics* **14**, 33–38 (1996).

40. J. Huang, A. D. MacKerell Jr., CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135–2145 (2013).

41. P. Mark, L. Nilsson, Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A* **105**, 9954–9960 (2001).

42. J. C. Phillips *et al.*, Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).

43. J. C. Phillips *et al.*, Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **153**, 044130 (2020).

44. U. Essmann *et al.*, A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).

45. N. Dharan, O. Farago, Coarse-grained molecular simulations of membrane adhesion domains. *J. Chem. Phys.* **141**, 024903 (2014).

46. G. Tesei, M. Vazdar, M. Lund, Coarse-grained model of titrating peptides interacting with lipid bilayers. *J. Chem. Phys.* **149**, 244108 (2018).

47. C. V. Kulkarni, Calculating the 'chain splay' of amphiphilic molecules: Towards quantifying the molecular shapes. *Chem. Phys. Lipids* **218**, 16–21 (2019).

48. A. Plumridge, K. Andresen, L. Pollack, Visualizing disordered single-stranded RNA: Connecting sequence, structure, and electrostatics. *J. Am. Chem. Soc.* **142**, 109–119 (2020).

49. G. C. Akerlof, H. I. Oshry, The dielectric constant of water at high temperatures and in equilibrium with its vapor. *J. Am. Chem. Soc.* **72**, 2844–2847 (1950).

50. N. Pokhrel, L. Maibaum, Free energy calculations of membrane permeation: Challenges due to strong headgroup-solute interactions. *J. Chem. Theory Comput.* **14**, 1762–1771 (2018).

51. J. Schell, E. Zars, C. Chicone, R. Glaser, Simultaneous determination of all species concentrations in multiequilibria for aqueous solutions of dihydrogen phosphate considering Debye–Hückel theory. *J. Chem. Eng. Data* **63**, 2151–2161 (2018).

52. R. A. Robinson, R. H. Stokes, *Electrolyte Solutions* (Butterworths, London, UK, ed. 2, 1959).

53. M. Gee *et al.*, A Kirkwood–Buff Derived force field for aqueous alkali halides. *J. Chem. Theory Comput.* **7**, 1369–1380 (2011).

54. G. Tesei, V. Aspelin, M. Lund, Specific cation effects on SCN−in bulk solution and at the air–water interface. *J. Phys. Chem. B* **122**, 5094–5105 (2018).

55. B. Stenqvist, A. Thuresson, A. Kurut, R. Vácha, M. Lund, Faunus–A flexible framework for Monte Carlo simulation. *Mol. Simul.* **39**, 1233–1239 (2013).

56. G. Tesei *et al.*, Electronic Notebooks: pH-sensitive coarse-grained model of mRNA lipid nanoparticles (Version v2). Zenodo. https://doi.org/10.5281/zenodo.10380674. Accessed 14 December 2023.