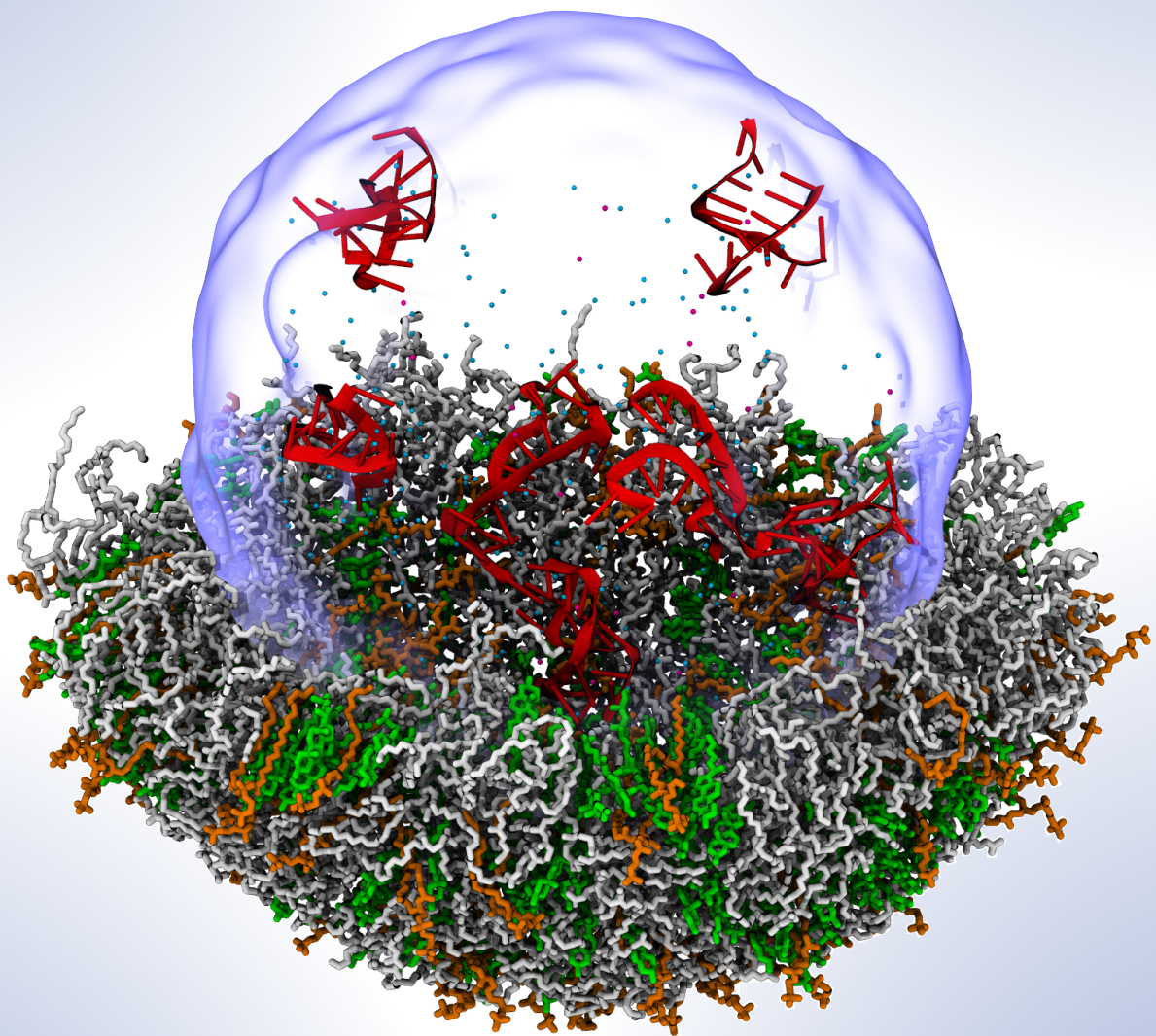# Modeling and Simulation of Ionizable Lipid Systems and Nucleic Acid-Surface Interactions

Mohd Ibrahim

Modeling and Simulation of Ionizable Lipid Systems and Nucleic Acid-Surface Interactions

Dissertation

zur Erlangung des akademischen Grades
Dr. rer. nat.

eingereicht an der
Mathematisch-Naturwissenschaftlich-Technischen Fakultät
der Universität Augsburg

von
Mohd Ibrahim
aus Achambore, Kargil, Indien

Augsburg, September 2023

MAX PLANCK GESELLSCHAFT

Universität Augsburg University

1. Erstgutachter: Prof. Dr. Nadine Schwierz

2. Zweitgutachter: Prof. Dr. Abhinav Sharma

Tag der mündlichen Prüfung: 14.11.2023

Dedicated to My Beloved Late Grandfather
Mohammad Hassan
and
To
Mukhtar, Abbas, Liyakat and Bashir

Contents

List of Abbreviations

AFM	Atom force microscopy
AMBER	Assisted model building with energy refinement
CG	Coarse-grained
CHARMM	Chemistry at Harvard macromolecular mechanics
CNT	Carbon nanotube
cryo-TEM	Cryogenic transmission electron microscopy
DNA	Deoxyribonucleic acid
DOPC	1,2-Dioleoyl-sn-glycero-3-Phosphocholine
DOTAP	1,2-Dioleoyl-3-trimethylammonium propane
dsDNA	double-stranded DNA
DSPC	1,2-Distearoyl-sn-glycero-3-phosphocholine
dsRNA	double-stranded RNA
e.g.	Latin *exempli gratia*; for example
Eqn.	Equation
ESP	Electrostatic potential
FF	force field
GAFF	General amber force field
GIXOS	Grazing incidence X-ray off-specular scattering
HF	Hartree-Fock
i.e.,	Latin *id est*; that is
IL	Ionizable lipid
KC2	DLin-KC2-DMA
LINCS	Linear constraint solver
LNP	Lipid nanoparticle
MC3	DLin-MC3-DMA (neutral)
MC3H	DLin-MC3-DMA (protonated)
MD	Molecular dynamics
mRNA	Messenger RNA
MSD	Mean squared displacement

NMR	Nuclear magnetic resonance
NSLD	Neutron scattering length density
PEG	Polyethylene glycol
POPC	1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
Ref.	Reference
RESP	Restrained electrostatic potential
RNA	Ribonucleic acid
SANS	Small-angle neutron scattering
SAXS	Small-angle X-ray scattering
SDP	Scattering density profile
SLD	Scattering length density
SMD	Steered molecular dynamics
SMFS	Single-molecule force spectroscopy
SS	Secondary structure
ssDNA	single-stranded DNA
ssRNA	single-stranded RNA
TIP3P	Transferable intermolecular potential 3-point
TEM	Transmission electron microscopy
tRNA	Transfer RNA
TRXF	Total reflection X-ray fluorescence
VMD	Visual molecular dynamics

List of Figures

List of Tables

Summary

The self-assembling property of lipid molecules enables their biological function as cellular membranes. This property has been also harnessed for various technological applications, especially in the field of drug delivery. Herein, a nucleic acid (RNA or DNA) is encapsulated in a self-assembled lipid moiety known as a lipid nanoparticle (LNP) and then transported into the cells. This encapsulation protects the nucleic acid (NA) cargo from degradation while simultaneously enhancing its cellular uptake since NA as such can not easily cross the cellular membrane. Nucleic acids are chains of nucleotides that are responsible for the storage, regulation, and transmission of genetic information. The sequence of nucleotides which is made up of five building blocks (adenine, cytosine, guanine, thymine/uracil) contains the blueprint of life. According to the *central dogma of molecular biology*, genetic information flows from DNA → RNA → proteins. It is this property that is harnessed in drug delivery, i.e., hacking the cellular machinery to execute a list of instructions that are tailored from outside to enhance or inhibit the production of certain proteins or even alter the course of certain disease-causing mutations that run across generations. This technology holds the potential to cure almost any disease and has already been used against different forms of cancer and diabetes. However, currently, it is very far from realizing its full potential. The incredibly complex nature of the protocol poses enormous challenges at each step, from optimizing the RNA to delivering it into the cells and releasing the cargo inside the cell while minimizing unintended responses.

LNPs are currently the most advanced delivery systems for nucleic acid-based drugs. These are multi-component systems composed of different types of lipids that provide structural stability to the LNP and enhance its encapsulation and fusogenic properties for efficient drug release. A crucial component of LNP that forms almost half of its composition is the *ionizable lipid* molecule. Ionizable lipids have a pH-dependent protonation state, at low pH they are positively charged and hence facilitate the encapsulation of the NA cargo which is highly negatively charged. At physiological pH, the ionizable lipid is neutral allowing efficient circulation inside the blood without degradation. Current LNP formulations still suffer from poor delivery efficiency partly owing to a lack of full understanding of its structure. Here a detailed understanding of interactions between different LNP components like nucleic acids and lipids, between ionizable lipids and other lipid species is crucial. However, LNP as a whole is an incredibly complex system, therefore model systems containing different components have to be investigated.

Lipid bilayer surfaces represent an ideal model system that has been extensively used to understand the interaction between lipids and other bio-molecules like nucleic acids or proteins as well as to understand the effect of lipid compositions on the membrane properties. A great advantage of these model systems is the possibility to directly probe the structural and dynamical properties using various highly precise surface-sensitive techniques like atomic force microscopy (AFM), fluorescence, and scattering techniques.

Mica surface represents another substrate that has been employed to investigate the

properties of various biomolecules using AFM imaging and also acts as an excellent substrate for studying supported lipid bilayers. The atomically smooth surface of mica facilitates images with a high signal-to-noise ratio. It is also a popular substrate for growing DNA nanostructures which have innumerable technological applications e.g. as bio-sensors, micro-arrays, drug delivery, and even novel computing and storage devices. In these systems, nucleic acid-surface interaction is a crucial determinant of the properties of nucleic acids being probed or of the nanostructures. Both mica and DNA are highly negatively charged, therefore, to mediate the interaction cations are employed. DNA-mica interactions are sensitive to the valency as well as the type of the cation. For example, some cations drastically disrupt the DNA conformation while others are not too effective in mediating an attraction. Here, a detailed understanding of the influence of different cations on the interaction of DNA with the mica surface is required. Such a study will be invaluable in choosing the most optimum ionic conditions for investigating DNA on mica surfaces in physiologically relevant conditions as well as aid in fine-tuning the properties of DNA nanostructures using cations.

Molecular dynamics (MD) simulation is well suited to study interactions between biomolecules and between biomolecules are other interacting species like surfaces. All-atom MD simulations in particular could provide atomistic insights into such interactions. In this thesis, we combined MD simulations and complementary experimental techniques to obtain insights into the interaction between nucleic acids and surfaces, cationic model systems, and different LNP components.

We started with the investigation of the influence of cation type and valency on the interaction between single-stranded DNA and mica surface. Using single-molecule force spectroscopy experiments and closely matched steered MD simulations we measured the detachment force required to desorb DNA from mica surface in the presence of Li^+, Na^+, K^+, Cs^+, Mg^{2+} and, Ca^{2+}. The measured force distribution was found to depend on the cation type. Our simulations reveal that the *ion-specificity* arises from a unique interplay of cation affinity towards the DNA, towards the mica surface, and its hydration properties, all of which are highly ion specific. Depending on the ion type we observe high and low force pathways arising respectively from, direct inner shell binding of cation to mica and indirect water-mediated interactions. Strongly hydrated cations like Mg^{2+} predominantly indirect water-mediated interactions while also giving rise to few strong direct interactions. On the other hand, weakly hydrated cations like Cs^+ or K^+ result in direct interactions but due to lower low affinity towards DNA exhibit lower detachment forces. However, these ions have a high affinity towards the mica surface and therefore accumulate on the surface in excess effectively reversing the surface charge, and rendering it positively charged. This gives rise to long-range attraction between DNA and surface in the presence of these ions. Based on our results we conclude that a mixture of Mg^{2+} or Ca^{2+} and K^+ would result in the best imaging conditions to study DNA in physiological conditions, the divalent ions give rise to long-lived contacts while the monovalent ions facilitate spontaneous adsorption via long-range attraction.

Next, we sought to understand the interaction of RNA with cationic and neutral bilayer

surfaces. We also explored the possibility of resolving RNA secondary structure on bilayer surfaces using neutron scattering techniques. Using the coarse-grained Martini model of RNA and lipids we looked at the effect of membrane composition and RNA base pairing on RNA-membrane interactions. The coarse-grained model allows us to simulate for a longer time scale ($\sim \mu s$) so that RNA can explore many possible configurations on the surface, however, it cannot resolve the intricate atomistic interactions like hydrogen bonding and the effect of cations. However such simulations correctly capture the main driving forces that are at play over large time and length scales. Our simulations demonstrate that *hydrophobic* and *electrostatic* interactions are the main driving forces involved in RNA-bilayer interactions. RNA adsorbs on the bilayer surface such that the hydrophobic nucleobases of the RNA interact with the hydrophobic parts of the bilayer while the negatively charged backbone is exposed to the polar head group of the lipid. This results in a strong influence of RNA base-pairing on the interactions especially with neutral lipid bilayers. Double-stranded RNA molecules where all the nucleotides are base-paired have the nucleobases facing each other and are not readily exposed to other interacting species. Therefore, in such cases, we find that RNA only weakly interacts with the bilayer surface as evident from the smaller penetration depth of RNA into the surface. Such an influence of base-pairing led us to hypothesize that RNA-bilayer interaction will be highly influenced by the RNA secondary structure, which is defined as the unique arrangement of base-paired and non base-paired loop regions for a given RNA sequence. Conversely, the interaction of RNA with the bilayer might influence its secondary structure and hence its 3D conformation. So, if such a different conformation of RNA on bilayers exists, can we resolve it using scattering techniques like neutron reflectivity or small-angle scattering profiles? We calculated reflectivity and scattering form factor profiles from different tRNA secondary structures on the bilayer interface. Our results show that the subtle differences between different RNA secondary structures are obscured by inherent Poisson noise present in such measurements. However, by selectively deuterating the RNA molecule, which is a unique and powerful feature of neutron scattering methods, small-angle scattering form factors can differentiate between different secondary structures. The biological function of RNA is intricately linked to its secondary structure and hence to its 3D conformation. Therefore, in drug delivery applications using LNPs, the secondary structure is highly optimized to minimize modification upon interaction with the lipids. However, it's still not well understood if it's fully preserved inside the complex LNP environment. Our results combined with experiments provide a method to probe the most probable RNA conformation upon interaction with lipids.

As mentioned earlier, the ionizable lipid (IL) component is one of the most crucial determinants of LNP efficacy. Extensive screening of ILs resulted in the development of some highly promising ILs of which DLin-MC3-DMA (MC3) is the very first one to be used in clinical application to treat Hereditary Amyloidosis, a rare genetic disease. However, its wide applicability is hindered by a poor understanding of its behavior with other lipid components of LNPs in different pH conditions. Since MC3 is not a standard lipid, therefore, models for it especially compatible with the standard nucleic acid force fields did not exist.

We obtained force field parameters for cationic and neutral MC3 corresponding to MC3 at two extreme pH conditions. The accuracy of the force fields was assessed by direct comparison of simulations with neutron reflectivity experiments performed on systems containing MC3/DOPC mixtures at different pH conditions, multiple MC3 fractions, and solvent contrast (i.e. different ratios of D_2O and H_2O). Our simulations demonstrated excellent agreement with experiments which implies, that the distribution of MC3 inside a neutral lipid bilayer is accurately captured. We find that pH has a drastic effect on the behavior of MC3. At low pH, cationic MC3 stays at the bilayer water interface and exhibits amphipathic behavior (i.e. hydrophobic tail and hydrophilic head group), whereas at high pH neutral MC3 exhibits entirely hydrophobic behavior. This leads to the migration of MC3 away from the water interface and its accumulation in the bilayer center increasing the bilayer thickness.

For any ionizable/titratable group, at a given pH, there are always both protonated (cationic) or deprotonated (neutral) states existing in certain fractions. Let's call this fraction the *protonation degree* (ζ). The vastly disparate behavior of MC3 in its neutral and cationic forms immediately points to the importance of the accurate assignment of the protonation degree at a given pH value. For an ionizable species in an aqueous medium in the infinite dilution limit, ζ at a given pH can be estimated from its pK_a value using the Henderson-Hasselbalch equation. However, the pK_a depends on the chemical and electrostatic environment, and inside a complex LNP or in a bilayer, predicting ζ is a highly non-trivial problem. Using a combination of all-atom MD simulation and X-ray scattering experiments we devised a methodology to estimate the protonation degree of ionizable lipids in monolayer systems. Lipid monolayers are single layers of lipid deposited on a water surface with the hydrophilic heads buried in water and tail facing towards air/vacuum. These are highly controllable model systems where the lipid properties can be accurately measured. Two simultaneous X-ray reflection measurements are performed (i) the grazing incidence X-ray off-specular scattering (GIXOS), which probes the transverse electron density of the monolayer which in turn depends on the lateral packing of lipids or the area per lipid (APL). (ii) The fluorescence signal induced by the incident X-ray beam in the surface adsorbed anions is measured. The intensity of this signal is directly proportional to the excess surface adsorbed anions which in turn is correlated to the number of cationic MC3 molecules at a given pH. Therefore, the surface density of excess anions can be computed. To obtain the fraction of cationic MC3, the APL which is not known in the experiments is required. To that end, we run multiple simulations at different APL values and for each make a direct comparison of experimental and simulation GIXOS profiles. The simulation APL which best reproduces the experimental GIXOS curves is considered for the estimation of protonation degree. Therefore, using this method we simultaneously optimize the APL and the ζ. Simulations performed using these optimum values provide structural insights into monolayers containing MC3. Consistent with previous bilayer results, we observe neutral MC3 positioned in the hydrophobic regions of the monolayer whereas charged MC3 stays at the monolayer-water interface. Moreover, the excellent agreement of simulation and experimental GIXOS curves further underscores the accuracy of our MC3

models in capturing the transverse distribution of lipids.

The determination of LNP structure has been pursued intensively recently. Studies based on Cryo-TEM and neutron/X-ray scattering showed that pH has a drastic effect on the LNP structure. At low pH, the LNP was observed to exhibit a *vesicle* like structure i.e., a spherical shape enclosed by a lipid bilayer. At high pH, multiple studies revealed an amorphous core which was attributed to the migration of ionizable lipid towards the LNP center. To obtain atomistic insights into these observations and further test our MC3 models we performed simulations on model LNPs to investigate the effect of pH and RNA cargo on its structural properties like size, hydration state (amount of water inside LNP), and distribution of different components. We carefully designed a setup where all these properties can be simultaneously studied. We started with a vesicular bilayer structure with cationic MC3 corresponding to LNP at low pH. The LNP retains its structure throughout the simulations attaining an equilibrium size and hydration level. Replacing the cationic MC3 with neutral MC3, mimicking the pH change, led to a disruption of the vesicular bilayer-like structure and the attainment of a relatively amorphous structure, consistent with previous studies. This disruption is also marked by the movement of water out of the LNP decreasing the water content at high pH as also observed in recent studies. The addition of RNA cargo does not significantly affect the structure at low pH where RNA sticks to the inner surface of LNP via electrostatic interaction with the cationic MC3. While at high pH, the RNA is loosely bound to the inner surface, and the presence of RNA sets a lower bound to the inner LNP radius due to electrostatic repulsion between the RNA molecules.

In summary, this thesis employs computational methods to understand the interactions of nucleic acids with surfaces and the pH-dependent behavior of ionizable lipids in various lipid systems. Understanding such interactions is important in realizing the technological potential of nucleic acids and ionizable lipids in drug delivery and DNA nanostructure applications. In the first part, we combined MD simulations and atomic force microscopy experiments to understand the effect of cations on DNA mica interactions. Subsequently, we studied interactions of RNA with model lipid membranes and the effect of RNA base-pairing and membrane composition on such interactions. Here, we also explored the possibility of resolving RNA secondary structures using scattering experiments. In the second part, we obtained force field parameters for a promising ionizable lipid in different protonation states. By using the model in combination with neutron and X-ray scattering experiments we obtained structural insights into the pH-dependent behavior of ionizable lipids in the bilayers, monolayers, and model lipid nanoparticle systems. These studies provide a baseline for understanding the LNP structure using computational methods.

Introduction

Lipid molecules are essential for the existence of cellular life in its current form. They form the cellular membrane which separates the internal environment of the cell from the external environment and also acts as a barrier for intracellular compartments where specialized functions take place [1, 2]. Therefore, they provide a suitable platform for different cellular processes to take place making life possible. Lipid molecules exhibit self-assembly in an aqueous medium. Depending on the exact lipid geometry, lipid composition, and external factors like temperature, pressure, solvent content, and pH, they self-assemble into a plethora of fascinating topologies like bilayers, monolayers, spherical micelles, vesicles, hexagonal and cubic phases, etc [3–9]. This self-assembly property has also been exploited for technological purposes, especially in drug delivery applications [10–14].

Nucleic acids are another class of molecules crucial for the continuity of all life forms. They are responsible for the storage, transmission, and expression of genetic information. These functions have also been harnessed for various technological purposes e.g., in drug delivery and DNA origami nanostructures. For these applications, nucleic acids have been used in combination with lipid moieties like lipid nanoparticles for drug delivery or with artificial crystal surfaces e.g., mica, to grow DNA nanostructures. To realize the full potential of such technologies and to gain insights into biological processes involving these entities, a fundamental understanding of the structure, dynamics, and interaction between various entities and factors affecting them is necessary.

In this chapter, I will introduce various aspects related to lipids, nucleic acids, and ionizable lipids and the importance of understanding their various properties for different technological applications.

1.1   Structure of lipid molecules

Lipids are water-insoluble molecules that form the cellular membrane. They have a hydrophobic part consisting of long hydrocarbon chains called the *tail* and a polar hydrophilic part known as the *head group*. The dual character makes them what is known as *amphiphatic* molecules. There are hundreds of different kinds of lipid molecules constituting the cellular and intracellular membrane and are broadly classified into three main categories namely glycerophospholipids, sphingolipids, and sterols.

Glycerophospholipids are the most abundant lipid molecules forming biological membranes. These lipids are derived from a glycerol unit HOCH2–CHOH–CH2OH by attaching various groups to the terminal hydroxyl groups (Figure 1.1 A). The head group is formed by linking a phosphate group, which carries a polar or charge substituent (**R** group), to one of the terminal hydroxyls. The **R** group can be derived from amino acid (e.g. serine)
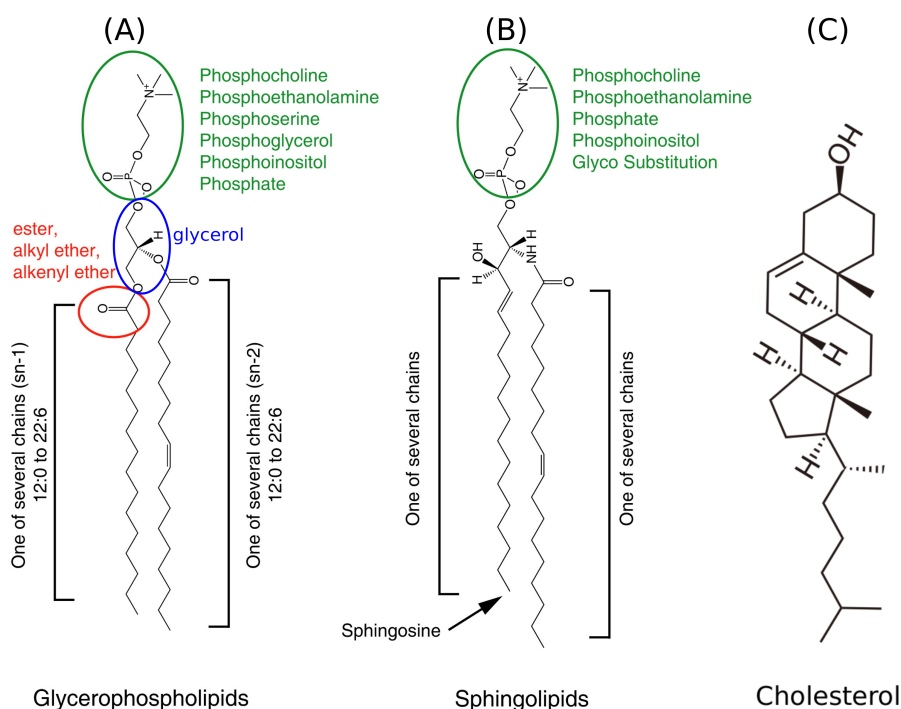
Figure 1.1: Chemical Diversity of Glycerophospholipids and Sphingolipids Chemical modification of head groups (green circle), fatty acid chain length, and number of double bonds (degree of saturation) contribute to the diversity and complexity of (A) glycerophospholipids and (B) sphingolipids. The linkage in glycerol in the sn-1 in glycerophospholipids (red circle) increases the variety further. (Adapted with permission from [15]. Copyright 2011 Elsevier Ltd). (C) Cholesterol

or other polar derivatives like choline, ethanolamine, inositol, etc. giving rise to PS, PC, PE, and PI phospholipids respectively. Most of these lipid head groups are zwitterionic with a negatively charged negatively charged phosphate group and a positively charged substituent. The lipid tail is formed by linking the remaining two terminal hydroxyl groups with a fatty acid hydrocarbon chain and can have different chain lengths and a number of double bonds (degree of saturation). The large number of possible combinations of heads, and tails as well as the linkage between them (Figure 1.1 A) gives rise to a large number of glycerophospholipids. Sphingolipids which are derived from a sphingosine replacing glycerol unit also exhibits diverse structure structure due to many possible combination of heads and tails (Figure 1.1 B). Of the sterol family cholesterol is an important constituent of biological membranes (Figure 1.1). They have rigid cyclic rings and when present in the membrane increase the membrane packing rendering it relatively rigid.

In an aqueous medium, due to their amphipathic nature, lipid molecules self-assemble such that the contact of the hydrophilic head groups with water is maximized while simultaneously minimizing the hydrophobic tail contact with water (Figure 1.2). This leads to the formation of bilayers and vesicles, which are spherical structures enclosed by a bilayer. This property enables their function as cellular membranes. When dispersed on the water surface at a suitable concentration, they tend to form monolayers with the head buried in water and tails facing towards air/vacuum.

Lipid bilayers have been extensively used in simulations and experiments to understand
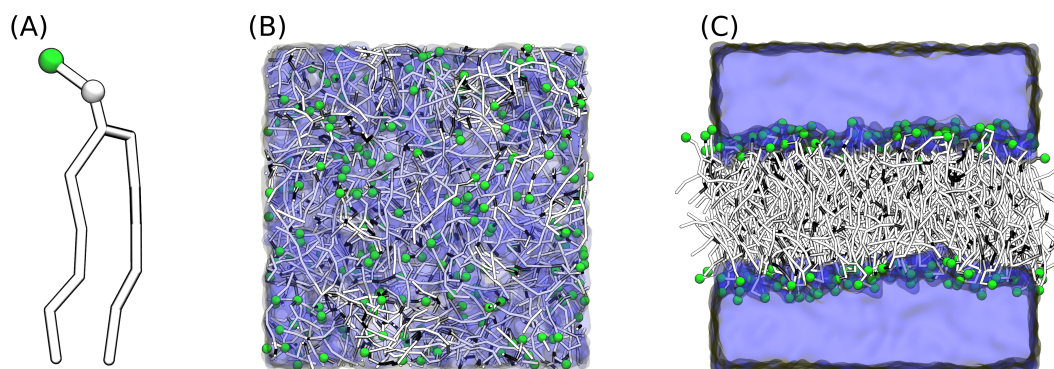
Figure 1.2: Lipid molecules self-assemble in water to form bilayers (A) A single DSPC molecule with the head group highlighted by green and white spheres. (B) DSPC molecules are randomly placed in a box of solvent which is represented by the blue transparent background (C) After a few tens of nanoseconds, the molecules self-assemble to form a bilayer. The CG Martini force field is used here.

the properties of cellular membranes. A lipid bilayer consists of two lipid layers (or leaflets) with tails facing each other and heads facing toward the solvent (Figure 1.2 C). Due to the hydrophobic tail region, a lipid bilayer acts as a barrier to polar molecules and ions. At room temperature, which is typically above the transition temperature [1] of most lipids, individual lipid molecules can diffuse freely in the leaflet plane with typical diffusivity on the order of $\mu m^2/s$. Movement of lipids across the leaflet known as a *flip-flop* happens at a very large time scale ($\sim$ minutes) since that entails the polar head group crossing the hydrophobic region which is highly energetically unfavorable. The fluidity of lipids is determined by the properties of the lipid tail. Longer tails with no or less number of double bonds exhibit high low fluidity and higher rigidity while unsaturated shorter tails display the converse behaviour. For this reason, lipids like DSPC which has no double bond in the tails are used to provide structural stability to the lipid nanoparticles for drug delivery.

1.2   Nucleic acids

Nucleic acids comprising DNA (Deoxyribonucleic acid) and RNA (Ribonucleic acid) are polymers of nucleotides which are made up of a nitrogenous base (or nucleobase), a sugar moiety, and a phosphate group (Figure1.3 A). DNA nucleotides are made up of four types of nitrogenous bases namely, adenine (A), thymine (T), guanine (G), and cytosine (C). In RNA thymine is replaced by uracil (U) and the sugar moiety is ribose instead of deoxyribose like in DNA. These five nucleotides form the building block of nucleic acids. DNA molecules are usually double-stranded helical structures with two strictly complementary strands where a specific nucleobase in one strand is paired to the corresponding complementary nucleobase via hydrogen bonds: A pairs with T and G pairs with C via hydrogen bonds, these pairs are also known as *Watson–Crick* base pairs. Such a perfectly complementary pairing ensures accurate replication and transfer of genetic information. RNA usually occurs as a single-stranded chain and due to the extra hydroxyl group on the ribose sugar

---

[1]Below the transition temperature the bilayer is in a solid phase known as the *gel phase*

is highly flexible compared to DNA and hence exhibits a diverse range of structures. The RNA strand folds onto itself forming double-stranded Watson-Crick base paired regions, loops, hairpins, and bulges. For a given RNA sequence, the unique arrangement of such structural motifs is known as its *secondary structure* (Figure 1.3 B). There are infinitely many such possible secondary structures, however, there is a unique secondary structure with minimum free energy (MFE) that can be predicted with many existing algorithms. The secondary structure folds into a 3D tertiary structure (Figure 1.3).
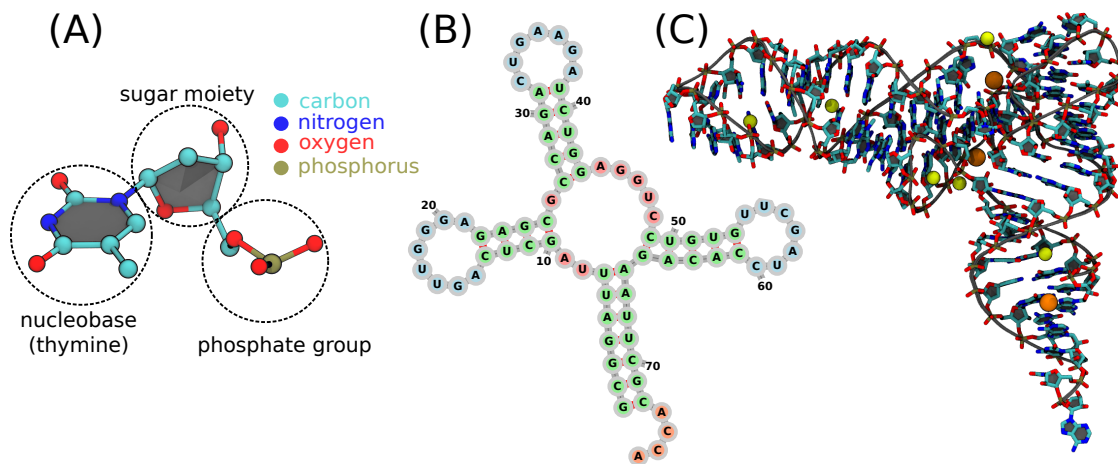
Figure 1.3: (A) A thymine nucleotide. (B The minimum free energy (MFE) secondary structure of yeast phenylalanine tRNA. (C) The corresponding tertiary structure (PDB code: `1EHZ`). The Mg^{2+} ions are shown in yellow and Mn^{2+} ions are depicted in orange.

Biologically, the DNA molecule is found inside the nucleus of the cell and stores genetic information. The information encoded in the DNA is used to create a complementary RNA molecule known as messenger RNA (mRNA) in a process known as *transcription*. The mRNA molecule serves as an intermediate messenger between DNA and the cellular machinery responsible for protein synthesis which happens in the cytoplasm outside the nucleus. Here, each set of three consecutive nucleotides, called *codons*, of the mRNA sequence, is read by the ribosomal machinery to synthesize one amino acid in a process known as *translation*. Several other types of RNA molecules also play crucial roles during translation, e.g., the transfer RNA (tRNA) molecule (Figure 1.3 B, C) that reads the mRNA sequence and brings the correct amino acids to the ribosome to ensure the protein is synthesized in the correct sequence. In these molecules, their secondary structure and hence the 3D conformation is critical to their biological role.

1.2.1 Cations are indispensable for stability and function of nucleic acids

Nucleic acids are highly negatively charged with each nucleotide carrying a single negative charge on the phosphate group. Therefore, cations are required to facilitate folding and stabilize the structure by shielding the electrostatic repulsion between the phosphate groups. Cations bind to the phosphate and the nucleobase oxygens and each cation has a specific

affinity towards these binding sites which gives rise to ion-specific effects. The affinity of cations towards the phosphate group has been shown to increase with the cation charge density whereas for affinity towards the nucleobases the the trend is reversed [16]. This results in interesting effects like twisting or unwinding of DNA in the presence of different cations [17]. Such ion specificity has also been observed for different RNA structures e.g. of several divalent cations, the stability of *Tetrahymena ribozyme* was found to be highest in present Mg^{2+} and least for Ba^{2+} [18]. Similarly, K^+ can stabilize RNA G-quadruplex structures, which are important for regulatory roles in gene expression, more effectively than many other mono and divalent ions [19, 20]. Therefore, not only the charge but the identity of the cation also plays a crucial role in the stability and function of nucleic acids. These ion-specific affinities can be harnessed to tune nucleic acid interactions with other substrates, such as mica surfaces. Mica surfaces are widely used for imaging DNA and as a substrate for growing DNA nanostructures for various technological purposes.

1.3 Nucleic acid interaction with surfaces

Nucleic acid interaction with surfaces are important for various technological application and for understanding their properties using diverse surface-sensitive experimental techniques. Mica is one such surface that has been extensively used to study the structural and dynamic properties of DNA using atomic force microscopy imaging technique. From an experimental perspective, mica can be relatively easily cleaved to obtain an atomically smooth surface which is ideal for imaging due to a very high signal-to-noise ratio. Additionally, it is the substrate of choice for growing DNA origami nanostructures. Here, the strict base-pairing property of DNA is harnessed to create arbitrary structures that are used as nano-sensors, micro-arrays, DNA-based computing and storage devices, and novel molecular machines to name a few [21, 22]. The success of these studies and applications hinges on the ability to finely tune the interaction between DNA and the mica surface.

1.3.1 Role of cations in mediating DNA-mica interactions

Mica, like DNA, is highly negatively charged with a surface charge density of around $-2.16\ e/\mathrm{nm}^2$. Therefore it's impossible to deposit or adsorb DNA on mica without cations. The earliest studies by Hansma et al. included employing divalent and transition metal cations to immobilize DNA on mica surfaces to facilitate AFM imaging [23]. The ability of cations to bind DNA on mica was found to correlate with the ionic radius owing the the specific chemistry of the mica surface. The cation type also strongly affects the conformation of DNA on the surface. High charge density transition metal ions (e.g. Ni^{2+}, Co^{2+}) and other multivalent ions strongly disrupt the DNA structure leading to a collapsed conformation that is far from its physiologically relevant structure [24]. However, the presence of monovalent ions e.g., Na^+ or K^+ in combination with divalent ions like Mg^{2+} has been demonstrated to enhance DNA-mica adsorption while simultaneously preserving the DNA conformation thus facilitating imaging in physiologically relevant conditions [25]. Furthermore, such mixtures were also found to be most suitable for well-ordered DNA origami structures on mica surfaces [26]. Here, the addition of monovalent ions increased

the mobility of the nanostructure on the surface, as opposed to keeping it tightly bound as happens for transition metals or in the presence of only divalent ions. High mobility enhances the propensity to form more well-ordered structures.
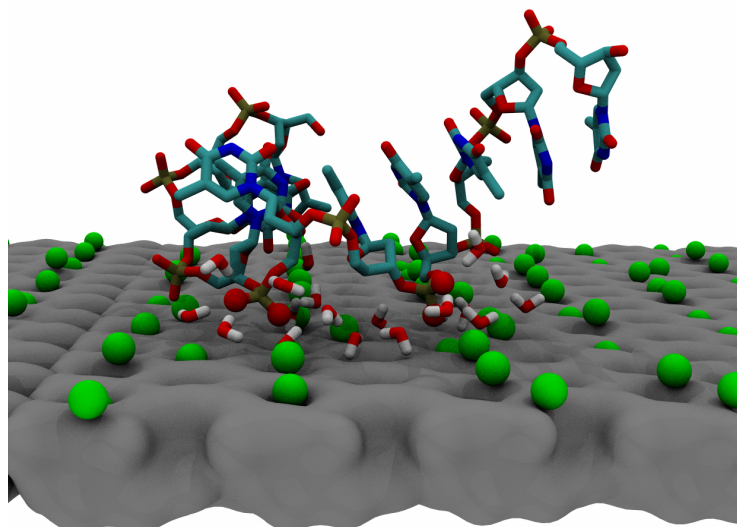
Figure 1.4: A ssDNA on a mica surface: Cations are depicted as green spheres. DNA interaction with mica involves a complex interplay of ion-DNA, ion-mica, and hydration properties of the ion.

The interaction of cations with mica exhibits ion-specific effects. A recent study combining AFM and molecule simulations on mica-electrolyte systems revealed that cations influence the hydration structure at the mica-electrolyte interface [27]. In the presence of high charge density ions like Li^+, a highly structured hydration layer was observed while low charge density ions like Cs^+ lead to a relatively diffused layer. Therefore, the ion-specificity of DNA-mica interactions has a possible origin in the ion-specific nature of multiple interactions such as ion-DNA, ion-mica, and ion hydration properties (Figure 1.4). However, specific studies to reveal the interplay of these different effects are still lacking. Such studies, while elucidating the mechanism of ion-specific effects in DNA-mica interactions, could also provide invaluable insights that could aid in choosing proper ionic conditions for fine-tuning the interactions and, consequently, the properties of DNA nanostructures.

1.3.2   Nucleic acid interaction with lipid bilayer surfaces

Lipid bilayers serve as one of the most versatile and ideal model systems for understanding the structural and dynamical properties of lipid membranes. These systems have been extensively used to study the interaction between lipid membranes and various biomolecules e.g., DNA, RNA, proteins, drug molecules, and as well as nanostructures like nanoparticles. To great advantage, these systems can be directly probed using numerous surface-sensitive techniques like atomic force microscopy (AFM), fluorescence correlation spectroscopy (FCS), ellipsometry, neutron/X-ray reflectivity, total reflection X-ray fluorescence, QCM-D (Quartz Crystal Microbalance with Dissipation), and GIXOS (grazing angle X-ray off-specular scattering), etc. These techniques provide precise information about the structure, dynamics, and interaction strength between lipid membranes and other molecular species.

The interaction of nucleic acids with lipid membranes (especially zwitterionic phospholipid) has been explored via computational as well as experimental methods like QCM-D and neutron reflectivity [28–31]. For the adsorption of double-stranded DNA (dsDNA) on a zwitterionic lipid layer (e.g. POPC), a divalent ion like Ca^{2+} was shown to be critical. In the absence of ions, no interaction between DNA and lipid layers was observed [30, 31]. The Ca^{2+} ions adsorb on the surface in excess rendering it positively charged on which the negatively charged DNA could easily adsorb via electrostatic attraction [31]. In contrast, double-stranded RNA (dsRNA) like small-interfering RNA (siRNA) was shown to weakly bind to the POPC layer even in the absence of cations while the presence of cations also enhanced the adsorption [29]. This difference was attributed primarily to the presence of overhanging unpaired nucleotides that are present at the end of the siRNA molecule. The adsorption of a tRNA molecule (Figure 1.3 C) with a more complex secondary and tertiary structure exhibited even more complicated behavior. Using QCM-D experiments, tRNA was shown to weakly and reversibly adsorb to zwitterionic phospholipid bilayers and the presence of Ca^{2+} did not have a significant effect on the interaction [28]. In the same study, the single-stranded RNA was shown to interact more effectively compared to double-stranded ones. This was attributed to the hydrophobic interaction between ssRNA nucleobases and hydrophobic parts of the bilayer which is not affected the in case of dsRNA since bases are less exposed. The interaction between RNA and lipids also affects the conformation of RNA as well as the phase-transition properties of the bilayer [28, 32, 33]. The conformation of RNA is important for its biological function. Therefore, understanding the effect of lipid bilayers on RNA conformation and devising methods to probe the RNA conformation upon interaction with lipids is highly relevant. This is especially crucial for drug delivery applications where RNA and lipid interaction has direct implications for the function and effectiveness of the drug.

1.4 Lipid nanoparticles as drug delivery vehicles: role of ionizable lipids

The biological function of RNA has been harnessed for drug delivery applications by transporting RNA from outside the cell. Using a synthetic mRNA sequence, any desired protein can be expressed or by using another class of RNA known as small interfering RNA (siRNA), the production of a desired protein can be inhibited. However, RNA can not be transported into the cells as it is highly unstable, can be quickly degraded by ribonuclease or via self-hydrolysis, and, due to its negative charge, cannot easily penetrate the cell membrane. Therefore, a delivery vehicle is usually employed for the best efficiency, where, the RNA drug cargo is encapsulated inside the delivery agent to protect it from degradation. Such a therapeutic method has the potential to cure many currently untreatable diseases like HIV, cancer, diabetes, etc. [34–43]. The ease of production of the desired RNA sequences further enables the possibility of personalized medicine where the genetic code of RNA can be tailored to the specific needs of an individual patient [44–47].

Lipid nanoparticles represent the most advanced nucleic acids-based drug delivery system and have also been used in many clinically approved applications (Figure 1.5).

Current lipid nanoparticle formulations are composed of an ionizable lipid (IL), cholesterol, a helper lipid usually, DSPC, and PEG-lipid (polyethylene glycol-lipid conjugate) roughly in the mole fraction 50:10:38.5:1.5 respectively. The helper lipid and cholesterol provide structural stability, and the PEG-lipid, which resides on the LNP surface, controls its size. These components are mixed together with the RNA cargo at low pH (~ 4) where the IL is positively charged. The lipid self-assembly property and electrostatic attraction between the positively charged IL and negatively charged RNA lead to the formation of LNP with the RNA encapsulated inside.
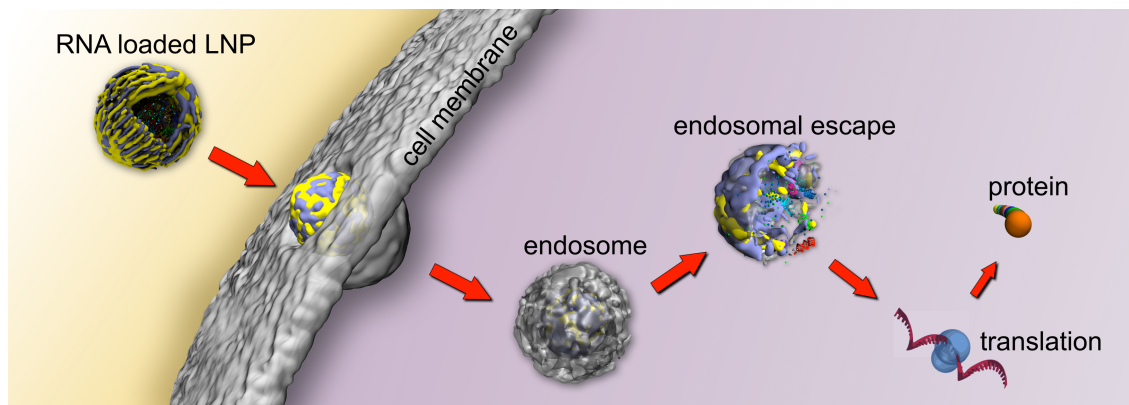
Figure 1.5: Schematic of the principle of lipid nanoparticles based delivery systems.

Since the IL forms a major component of LNPs, the delivery efficiency of LNP critically depends upon the chemical features of IL. The chemical structure of IL can be divided into three segments, heads, linkers, and tails (Figure 1.6) each of which can be tailored independently [48–50]. The head can be positively charged or neutral depending on pH and therefore is key to entrapping the RNA cargo, stabilizing RNA LNP interactions, and endosomal escape. The pK_a of the ionizable head group was found to be a key parameter for potent LNPs with an optimum range between 6.0-7.0 for efficient delivery [51]. Once the RNA cargo is released inside the cell, the rest of the LNP should be quickly degradable without creating significant toxic side effects. This is determined by the linker connecting head and tail groups [52]. The properties of the tail like the degree of saturation, and hydrocarbon chain length have a large impact on the fusogenic properties of the LNP. The overall geometry of the molecule has also been shown to affect the fusogenic properties of the LNP with a conical geometry tapered towards the head group exhibiting better transfection properties compared to a cylindrical or inverse conical geometry [48].
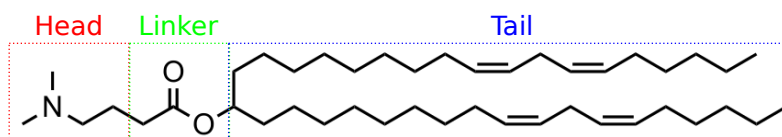
Figure 1.6: Chemical structure of ionizable lipid DLin-MC3-DMA (MC3). The nitrogen atom in the head group can be protonated (positively charged) or deprotonated (neutral) depending on the pH.

Extensive screening for the head, linker, and tail of ILs [48, 51] has led to the development of several promising ILs with DLin-MC3-DMA (MC3) (Figure 1.6) being the first IL to be approved for drug delivery for treating amyloidosis [53]. Even though IL-based LNPs represent significant and promising advancement in RNA delivery systems, the efficiency and stability of current LNPs still remain low, hampering their wide applicability [54]. One of the key reasons is a lack of a full understanding of the internal structure of LNPs and an understanding of the structure-function relationship. To improve the design of LNPs a detailed understanding of the interactions of nucleic acids with the LNP components is crucial.

Molecular dynamics (MD) simulation is a powerful method to gain atomistic insights into biological systems and interactions between different components of the system. By describing the system at an atomistic level, the role of each component and the effects of different external factors like ionic conditions, pH, temperature, and solvent content on the system can be elucidated. However, the accuracy of molecular simulation results critically relies upon the accuracy of the underlying potential energy parameters, called *force fields*, used to describe the interaction between different atoms. The accuracy is judged by comparing the results from MD simulations with experiments. Therefore, studies that combine MD simulation and experimental techniques are of particular importance as the experiments serve to validate the simulations and the simulations provide an atomistic picture of the phenomena of interest.

1.5 Aims and outline of the thesis

In this thesis, we employed computational methods in combination with complementary experimental techniques to investigate the interaction between nucleic acids and surfaces like mica and lipid bilayers. We modeled an ionizable lipid and studied its pH-dependent behavior in various lipid systems like bilayers, monolayers, and LNPs. Understanding these interactions is important to realize the full potential of their various technological application in drug delivery and DNA-based nanostructures.

- **Chapter 2** briefly outlines the methods used in the thesis.

- **Chapter 3** presents results of cation-mediated interaction between DNA and mica surface. We combined single-molecule atomic force spectroscopy and closely matched simulations to gain insights into the influence of cations $Li^+, Na^+, K^+, Cs^+, Mg^{2+}$ and Ca^{2+} on DNA-Mica interactions.

- **Chapter 4** deals with the interaction of RNA with lipid membranes using the Martini coarse-grained model. We further explored the possibility of using neutron reflectivity and small-angle scattering form factors to probe RNA conformation on the bilayer surface.

- **Chapter 5** presents the derivation of AMBER force field parameters for ionizable lipid MC3 in neutral and cationic states. The MC3 parameters are employed to simulate MC3/DOPC systems in different pH conditions and make a direct comparison

9

of simulations with neutron reflectivity experiments.

· **Chapter 6** combines MD simulation and X-ray reflection techniques to assign the protonation degree and lipid lateral packing in MC3/POPC monolayer systems. Using correctly assigned protonation degree and packing, structural insights into MC3/POPC monolayer systems are presented.

· **Chapter 7** uses the MC3 parameters to simulate a model LNP system to study the effect of pH and RNA cargo on LNP structure.

Chapter 2

Computational Methods

The classical equations of motion describe motion from atoms to galaxies. However, in only very few cases such as two body problem and a few specialized three body cases, the trajectories can be reduced to simple analytical expressions. For many bodied systems like a biomolecular system numerical methods are employed to solve the equation of motion and track the dynamics of the systems. In this chapter, I will briefly describe the basics of computational methods employed to study the dynamics of systems throughout this thesis.

2.1 Molecular dynamics simulations

Molecular dynamics simulation is a computational method to study the dynamics of a system of atoms or molecules using the classical equations of motion. Therefore, it allows us to calculate the dynamical, time-averaged structural, and thermodynamical properties of the system. With current computational resources system sizes, of the order of a few tens of nanometers could be simulated for a few microseconds. In the following, I will briefly outline the important aspects of MD simulations roughly based on the book by Frenkel and Smit [55].

2.1.1 Integrating equations of motion

In MD simulations, we begin by initializing the position of constituent atoms with a set of coordinates. The force on each atom due to all other atoms in the systems is evaluated and the positions are updated according to Newton's equations of motion. The force on a given atom due to all the other atoms is given by,

$$\mathbf{F_i} = \sum_{j \neq i}^{N} \mathbf{f_{ij}} \tag{2.1}$$

Where, $\mathbf{f_{ij}}$ is the force on i^{th} atom due to j^{th} atom and it has three components, $\mathbf{f_{ij}} = f_x\hat{\mathbf{x}} + f_y\hat{\mathbf{y}} + f_z\hat{\mathbf{z}}$. The force between two atoms is determined by the interaction potential $U(r)$, therefore, the x-component of force for example is,

$$f_x = -\frac{x}{r}\frac{\partial U(r)}{\partial r} \tag{2.2}$$

Where r is the distance between the two interacting particles. Knowing the forces on each atom at a given time instant, the positions can be updated using the following

algorithm, known as the Verlet algorithm [56],

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + \frac{f(t)}{m}\Delta t^2 + \mathcal{O}(\Delta t^4) \qquad (2.3)$$

Where m is the mass of the atom, Δt is the integration time step and is determined by the fastest motion in the system which is usually bond vibrations. By constraining hydrogen involving bonds, in a typical atomistic simulation, a time step of 2.0 femtosecond could be used. The velocity is updated using,

$$v(t) = \frac{r(t + \Delta t) - r(t - \Delta t)}{2\Delta t} + \mathcal{O}(\Delta t^2) \qquad (2.4)$$

The above equations could be obtained by simply summing up the Taylor expansion of position at time $t - \Delta t$ and $t + \Delta t$. The above algorithm preserves the time reversibility of the equations of motion i.e if the momenta are reversed at any given instant, the system will trace back same the trajectory.

2.1.2 Potential cut-offs

The evaluation of forces in Eqn. 2.1 is the most expensive part of an MD simulation as it involves evaluating distances between all possible pairs of atoms to obtain the forces. Therefore, for a system with N atoms, $N(N-1)/2 \sim N^2$ evaluations are required, i.e., the simulation time increases with the square of the system size, which is highly undesirable. It turns out that the interaction potentials in MD are short-ranged (except the Coloumb potential) and therefore can be truncated at a certain distance called "cut-off", distance r_c. For each atom, a "neighbor list" is built which is the set of atoms within a distance $r_v > r_c$ of the given atom. To evaluate the force on a given atom, only the atoms in its neighbor list are considered. The neighbor list is updated with a certain frequency. Such tricks in combination with a few more bring down the simulation time scaling with systems size to N from N^2. The truncated potential can be shifted to zero at the cut-off or switched to zero using some switching functions. The Coloumb potential due to its long-range nature is divided into short-range and long-range parts. The short-range part is truncated as described while the long-range part is obtained by solving the Poisson equation in the reciprocal space by a method known as Particle Mesh Ewald (PME) [57].

2.1.3 Periodic boundary conditions

A typical simulation e.g. water box with 100,000 atoms has around 13 % of the atoms on the box surface. In general, for a system with N atoms, the fraction of atoms near the surface varies as $N^{-1/3}$. Therefore, a typical MD simulation has a significant surface effect and therefore can not be compared with real experimental setups which are usually bulk systems ($N \sim 10^{23}$). To mimic a bulk system in MD simulations periodic boundary conditions are applied. Here, the system is surrounded by its identical periodic copies (Figure 2.1). An atom near the box edge can now have periodic images of other atoms in its neighbor list. Further, if an atom leaves the box from a side its image enters the box
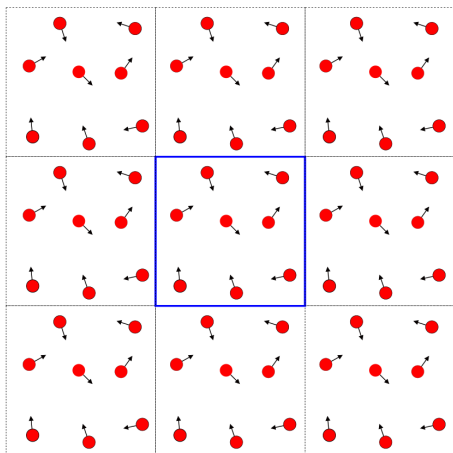
Figure 2.1: In periodic boundary conditions, the simulation box (blue) is surrounded by identical periodic copies.

from the opposite side.

2.1.4 Thermodynamic ensembles

The MD algorithm presented in section 2.1.1 will produce an ensemble with constant particle, volume, and energy (NVE) i.e. a microcanonical ensemble. However, the experimental systems which we aim to emulate in MD are constantly exchanging energy with the surroundings and are usually at constant pressure or/and temperature i.e. a canonical NPT ensemble. Two types of canonical ensemble are most popularly used in MD simulations, the NVT ensemble where the volume and temperature are fixed, and the NPT ensemble where temperature and pressure are fixed. A variety of techniques exists to perform MD at constant temperature and pressure.

The temperature can be naively controlled by exploiting the equipartition theorem, which relates the temperature to the velocity as,

$$\left\langle \frac{1}{2}mv_\alpha^2 \right\rangle = \frac{1}{2}k_B T \tag{2.5}$$

The velocities are re-scaled to keep a constant temperature. However, this method does not reproduce the expected temperature fluctuation for a canonical ensemble which becomes more significant for small systems. This scheme was modified by Bussi et al by introducing a stochastic method to re-scale the velocities [58]. This modified method called "stochastic velocity rescaling" method produces the correct properties of canonical ensemble and will be used throughout this thesis. The Berendsen thermostat maintains a constant temperature by adding an extra temperature-dependent friction term to the equation of motion. This thermostat is mostly used in the equilibration phase and is not recommended for production runs since it does not reproduce the Maxwell-Boltzmann velocity distribution.

To maintain constant pressure, commonly used barostats are the Berendsen barostat [59] and the Parrinello-Rahman barostat [60]. The Berendsen barostat scales the position of particles and box volume to keep the pressure constant. However, like the Berendsen thermostat, this barostat does not reproduce the correct NPT ensemble. It is however useful to equilibrate the system to a given pressure since it works well even when the initial

pressure is far from the target pressure. Parrinello-Rahman barostat reproduces the correct ensemble and hence is used during the production run. It, however, becomes unstable when the initial pressure deviates significantly from the target pressure. Therefore, throughout the thesis, Berendsen barostat is used during the equilibration phase while for production runs Parrinello-Rahman barostat is employed.

2.1.5 Force fields

The dynamics of an atom in MD is completely determined by the forces it is subjected to (Eqn. 2.3), which in turn is determined by the interaction potential between the atoms (Eqn. 2.2). The functional form of the potential and the parameters that define the functions is termed as "force field". The potential has contributions from bonded and non-bonded interactions Therefore,

$$U(r) = U_{bonded}(r) + U_{non-bonded}(r) \tag{2.6}$$

Bonded potential

The bonded potential is required to keep the molecular conformation intact i.e., it maintains the proper geometry of the molecule by keeping the bonds, angles, and dihedrals close to certain equilibrium values. Therefore, it has terms

$$U_{bonded} = U_{bond} + U_{angle} + U_{dihedral} + U_{improper} \tag{2.7}$$

The bond and angle potential are modeled as harmonic springs with two parameters namely (i) a bond or angle stiffness (k_b, k_a) and (ii) equilibrium values for bond length and angles (r_0, θ_0). The dihedral potential defines the angle between two planes formed by four consecutive atoms i,j,k,l, it is the angle between the plane ijk and jkl. It is modeled with a periodic function with a stiffness parameter (k_ϕ), a multiplicity (n), and a phase factor (δ). The improper potential is included to maintain the geometry of planar groups like aromatic rings, it is again defined between the planes formed by four atoms. It is modeled with a harmonic potential with a stiffness parameter (k_ψ) and an equilibrium angle value (ψ_0). The non-bonded potential therefore becomes,

$$U_{bonded} = \frac{1}{2} \sum_{\text{bonds}} k_b(r - r_0)^2 + \frac{1}{2} \sum_{\text{angles}} k_\theta(\theta - \theta_0)^2 +$$
$$\frac{1}{2} \sum_{\text{dihedral}} k_\phi(1 + \cos(n\phi - \delta) + \frac{1}{2} \sum_{\text{impropers}} k_\psi(\psi - \psi_0) \tag{2.8}$$

Non-bonded potential

The interaction between non-bonded atoms is described by the Lennard-Jones (LJ) and the Coloumb potential. The LJ potential is short-ranged and it has an attractive term that varies as r^{-6} and a repulsive term that varies as r^{-12}. The attractive term arises from dipole-dipole or induced dipole-induced dipole interaction, and the repulsive term arises from the Pauli exclusion principle when electron clouds of two atoms overlap, the

fact that no two electrons can occupy the same quantum state leads to a repulsive force. If the atoms are charged, in addition to LJ potential they also interact via the Coloumb potential. The non-bonded potential is given by,

$$U_{non-bonded} = \sum_{j \neq i} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \sum_{j \neq i} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \qquad (2.9)$$

The first term is the LJ potential and is described by two parameters, ϵ_{ij}, and σ_{ij} which denote the potential well depth and the point where the potential becomes zero respectively, r_{ij} denotes the distance between the two interacting particles. The interaction parameters ϵ_{ij} and σ_{ij} between a pair of atoms, i and j is given by the Lorentz-Berthelot rule as,

$$\sigma_{ij} = \frac{\sigma_i + \sigma_j}{2} \text{ and } \epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j} \qquad (2.10)$$

Where, ϵ_i, ϵ_j, σ_i, and σ_j are LJ parameters for interaction between the same type atoms.

The second term is the Coloumb potential between two particles with charge q_i and q_j, and ϵ_0 is the absolute permittivity of free space. All the parameters that appear in the above potentials constitute a force field. These parameters are either derived using quantum chemical calculations or by empirical methods to reproduce certain experimental observables.

2.1.6 Force fields (FFs) for lipid bilayer simulations

As described in the previous chapter, self-assembling to form bilayers is one of the most characteristic properties of lipids that enables these molecules to perform their biological function as cellular membranes [1]. The important parameters characterizing bilayers are, melting temperature, area per lipid, bilayer thickness, acyl chain order parameters (S_{CD}), and lateral diffusion coefficients. These properties depend on the specific chemistry of the constituent lipid molecules that form the bilayer. A lipid force field, ideally, should be able to simultaneously reproduce all these properties accurately which is an enormously challenging problem.

A variety of force fields at different resolutions exists to simulate lipid bilayers. The coarse-grained (CG) Martini force field maps four non-hydrogen atoms into a single bead, thus enabling longer timescale simulations with relatively larger systems [61]. Therefore, Martini FF is used to study bilayer processes like vesicle formations and fusions, phase transitions, etc., which take place over long timescales ($> \mu$s) [61–63]. To simulate a lipid bilayer with an atomistic resolution three force fields namely CHARMM36 [64], Slipids [65], and the AMBER Lipid14 [66] are widely used. All these FFs have been rigorously parametrized to reproduce the previously mentioned bilayer properties. CHARMM36 FF provides highly accurate parameters for a large number of different lipid molecules compared to AMBER and Slipids and is, therefore, the most commonly used FF for bilayer simulations [67]. However, nowadays, pure bilayer simulations are seldom performed, in most cases, bilayers are simulated with other biological molecules like proteins and nucleic acids e.g. to study ion channels, pore formations , and for applications in drug delivery. Proteins and

especially nucleic acids are best described by the AMBER family of force fields and hence are more popular for simulating such systems [68, 69]. Both CHARMM and AMBER FFs parameters are derived using different strategies and are consistent within their FF families i.e. CHARMM lipid force field is consistent with CHARMM protein forcefield and similarly for AMBER. These force fields even differ in their best simulation protocols e.g., for the CHARMM force field the truncated LJ potential is switched to zero using a particular switching function, and CHARMM uses a slightly modified water model compared to AMBER. Most importantly, in AMBER FF, the non-bonded LJ and Coulomb interactions (Eqn. 2.9) are scaled by a fudge factor of 0.5 and 0.833 respectively. It is therefore strongly discouraged to mix two force field families to describe individual components within the same simulation, such as lipids by CHARMM and simultaneously nucleic acids by AMBER [70]. Slipids FF was developed to be consistent with AMBER FF, it uses the same charge derivation protocol as AMBER, and non-bonded interactions are scaled by the same factor as in AMBER. Slipids were developed when the available AMBER FF for lipids namely General AMBER FF (GAFF) lipid [71], was not sufficiently accurate for lipid simulations. However, new AMBER lipid FFs namely Lipid14 [66], Lipid17 [72] and the most recent Lipid21 [73] has been refined to reproduce the structural and dynamic properties of lipids with the same accuracy as CHARMM and Slipids. Therefore, the AMBER lipid and nucleic acid FFs stand out to be the most appropriate choice to simulate lipid-nucleic acid systems and hence have been used in many simulation studies involving both lipids and nucleic acids.

2.1.7 General amber force field (GAFF)

As described earlier Amber force fields describe nucleic acids and proteins most accurately. Since, proteins and nucleic acids are composed of a limited number of building blocks (4 nucleotides for nucleic acids and 20 amino acids for proteins), any arbitrary sequence of nucleic acid or proteins can be simulated using a limited set of force field parameters. However, there are many biologically relevant small molecules with unique chemical structures that are often relevant to the function of nucleic acids and proteins (e.g., drug molecules). Due to the sheer number of such molecules, deriving force field parameters for each of such small molecules would be unfeasible. The General Amber Force Field addresses this issue by creating a library of transferable force field parameters for any arbitrary molecule composed of H, C, N, O, P, S, and Halogens. A total of 57 *atom types* [1] was defined based on the element type, aromaticity, hybridization state, and chemical environment. The LJ parameters were taken from existing Amber Parm94 or Parm99 FF [74, 75]. Bonds, angles, and dihedrals were obtained from experiments, empirical, and ab initio methods. The steps to obtain parameters for an arbitrary molecule are (i) obtain the 3D structure of the molecule (ii) assign GAFF atom types (iii) assign bonds, angles and dihedral parameters from the GAFF library (iv) derive partial charges using the RESP protocol as described in the following section.

The first AMBER lipid simulations were performed using the GAFF parameters with

---

[1] In simulations, an *atom type* is defined by the element type, and the two LJ parameters (σ and ϵ)

RESP charge derivation [76]. The simulations correctly reproduced experimental bilayer thicknesses and showed a qualitative agreement of electron density profiles with experiments. Such highly encouraging results led to the development of the GAFFLipid force field [71] and the successive AMBER lipid force fields (11/14/17/21) [66, 72, 73, 77]. Therefore, GAFF is a highly useful library to simulate unknown compounds and it is used in this thesis to obtain parameters for an important non-standard ionizable lipid molecule.

2.1.8 Restrained electrostatic potential (RESP) charge derivation

The RESP protocol [78] is used to derive partial charges (q_i, q_j in Eqn. 2.9) in AMBER FF. For a non-standard molecule, while we can take the non-bonded parameters (atom types, bonds, angles) from GAFF, the partial charges specific to the molecule must be derived using the RESP method. In this method, the full electron cloud of the molecule is represented at the quantum mechanical (QM) level at a certain level of theory (e.g., HF/6-31G*). Several potential energy surfaces enclosing the molecule within a distance of 1-2 times the Van der Waals radius of atoms are defined. Each surface is divided into multiple grids with a density of around 1 grid/Å^2. The QM electrostatic potential due to the electron cloud is calculated at each grid point. The set of atomic point charges that best reproduce the QM electrostatic potential at the grid points is taken as the partial charge. The RESP method utilizes restraints (and hence the "Restrained") to prevent the derived partial charges, especially of buried atoms (e.g. "C" in methyl groups), from wildly fluctuating which was a problem observed in previous ESP methods that lacked such constraints. Obtaining partial charge using the RESP method involves (i) geometry optimization by using QM methods and (ii) charge derivation as described, for the optimized geometry. To reduce any bias coming from the initial conformation, the derivation is performed for multiple conformations (~ 100) simultaneously. This method is used in this thesis for charge derivation.

2.2 X-ray/Neutron scattering techniques to probe bilayer structure

As mentioned before, the validation of molecular force fields requires accurate experimental data. X-ray and neutron scattering techniques are considered one of the most reliable methods to validate lipid force fields and to gain invaluable structural insights. These experiments can be quantitatively compared with simulations in a model-free manner. Here, I will describe some scattering techniques used in this thesis.

2.2.1 Small angle X-ray scattering (SAXS)

X-rays due to their wavelength in the range of inter-atomic distances ($\sim \text{Å}$) have been used to probe the structure of matter at the atomic level. X-rays primarily interact with the electrons of constituent atoms through the processes of elastic scattering (Thomson scattering), inelastic scattering (Compton scattering), photoelectric effect, X-ray fluorescence, and pair production [79]. The exact mode of interaction depends on the energy

of incident X-rays. The elastic and inelastic scattering processes are exploited to obtain structural information as the scattered radiation contains information about the electron distribution. The elastic scattering intensity from a volume element $d\vec{R}$ is related to the Fourier transform of the electron density $\rho(\vec{R})$,

$$I(q) \propto |F(q)|^2 \propto \left| \int_{-\infty}^{\infty} \rho(\vec{R}) e^{-i\vec{R}.\vec{q}} d\boldsymbol{R} \right|^2 \tag{2.11}$$

Where $F(q)$ which is given by the Fourier transform of the electron density is known as the form factor. The momentum transfer q is related to the scattering angle θ by $q = \frac{4\pi}{\lambda} sin(\theta/2)$. Like all scattering processes, the scattering angle (or momentum transfer) is inversely related to the probed length scale. For biological systems in solution, such as a protein solution or solution of lipid vesicles, due to the absence of any long-range ordering, the scattered waves undergo constructive interference only within a small angular range. Therefore, it's important to resolve small-angle scattering intensities which contain information about the size, shape, and average electron densities of scattering molecules. When the scattering occurs within a narrow angular range i.e. $< 10°$ or $q < 1.0$ Å^{-1}, it is referred to as Small-Angle X-ray Scattering (SAXS). Detecting such scattering requires specialized experimental setups with a considerable distance (several meters) between the sample and the detector. For crystallized samples, wide-angle X-ray scattering (WAXS) also known as the X-ray diffraction technique provides structural information with atomistic resolution since for such systems due to the long-range ordering scattering intensities are observed up to very high $q-$values.

SAXS studies to obtain lipid bilayer properties are usually performed on two types of systems (i) oriented multilamellar arrays which are stacks of bilayer supported on a substrate (ii) unilamellar vesicles (ULV) in solutions [80]. The scattering intensity measured in such experiments is modeled by [81],

$$I(q) = \frac{S(q)|F(q)|^2}{q^n} \tag{2.12}$$

Where, the denominator is a Lorentz factor, $n = 1$ for oriented samples, and $n = 2$ for ULVs [81]. $S(q)$, is the structure factor that describes the sphericity and polydispersity of the vesicles. For oriented multilamellar arrays and dilute vesicle solutions $S(q)$ can be considered to be constant, therefore,

$$I(q) \approx \frac{|F(q)|^2}{q^n} \tag{2.13}$$

Therefore, the experiments on lipid systems include obtaining the form factor using Eqn. 2.12 or Eqn. 2.13. The form factor so obtained is given by the Fourier transform of electron density along the direction normal to the bilayer plane as,

$$F(q_z) = \int_{-\infty}^{\infty} (\rho_e(z) - \rho_0) e^{-iq_z z} dz \tag{2.14}$$

For a bilayer with symmetric electron density, the *sine* part of the integral is zero so $F(q_z) = \int_{-\infty}^{\infty} (\rho_e(z) - \rho_0) \cos(q_z z) dz$.
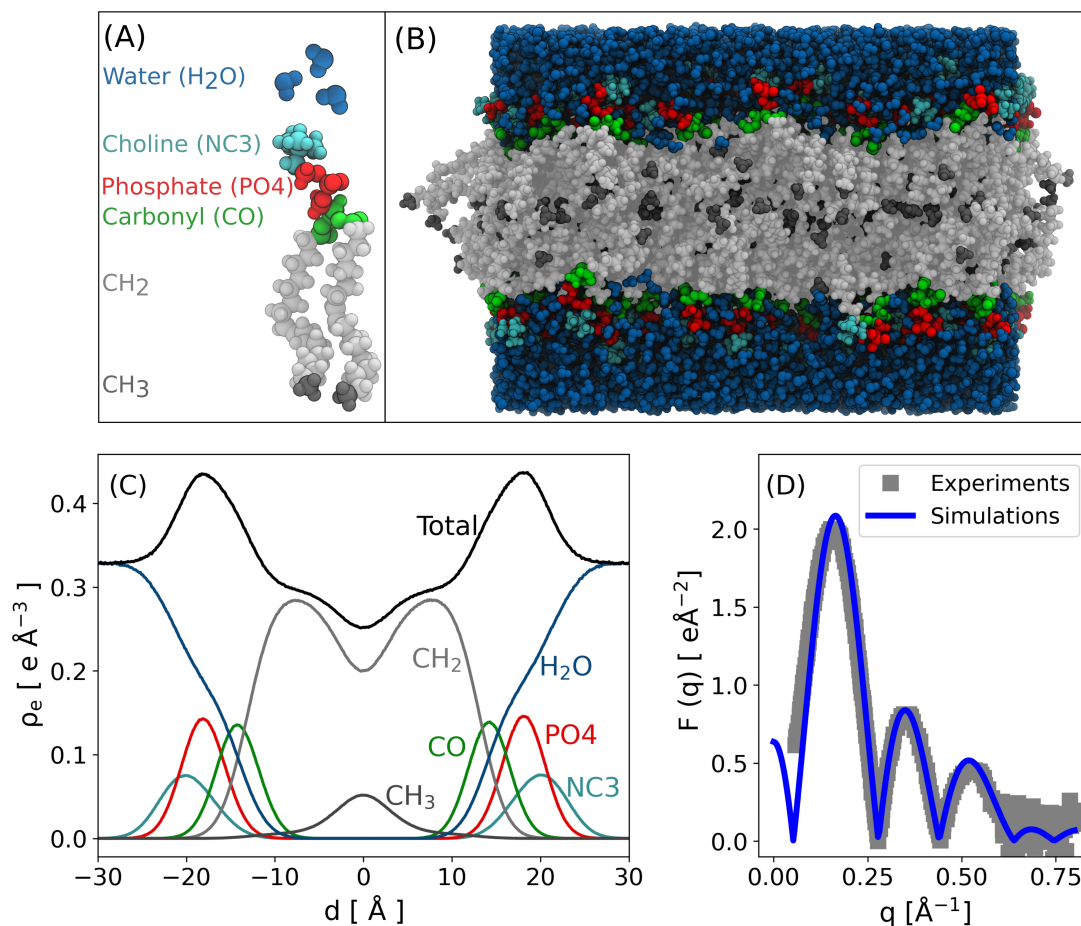
Figure 2.2: (A) DOPC molecule with different parts used in the scattering density profile (SDP) highlighted. (B) Bilayer with highlighted parts (C) The electron density of each part is modeled as a Gaussian with a mean position and standard deviation as free parameters. The total electron density is the sum of individual components. These parameters for each component are varied to reproduce the experimental form factor (D).

In practice, the electron density is heuristically modeled using slabs characterized by thickness, electron densities, and roughness. The parameters are fitted until Eqn. 2.14 is satisfied. A successful method to model the electron density is the scattering density profile (SDP) model [81]. Here, the lipid is divided into different groups like choline, phosphate, carbonyl, methylene, and methyl groups (Figure 2.2A). Each group is modeled by a normalized Gaussian probability distribution [81] with a mean position and standard deviation (Figure 2.2C). These parameters are varied until the Fourier transform of total electron density reproduces the experimental form factor (Figure 2.2D, Eqn. 2.14). From simulations, the form factor can be trivially obtained since we have access to the full electron density profile (Figure 2.2C). The above formalism is also valid for the analysis of Small angle neutron scattering (SANS) experiments with the only difference being, instead of the electron density we have the neutron scattering length densities of the nuclei. Therefore, a direct model-free comparison of experiments and simulations can be made making SAXS/SANS one of the most reliable techniques to validate lipid simulations.

2.2.2 Neutron reflectivity

Specular neutron reflectivity (NR) is another powerful scattering technique that can probe interfaces with sub-nanometer resolution. In this technique, the specular reflection of a neutron beam incident upon a planar surface e.g., a lipid bilayer, is measured. It probes the neutron scattering length density (SLD) along the direction perpendicular to the surface. It's an emerging tool to study lipid bilayers and the interaction of lipid membranes with proteins and nucleic acids in near physiological conditions. With NR, it's possible to determine protein orientations and penetration depths on a membrane surface, which are crucial information to understand protein-membrane interactions. A very powerful aspect of neutron scattering is the process of *contrast matching* where specific parts of the system can be highlighted or rendered invisible to neutrons. This is possible due to the vast difference between the scattering length of hydrogen and its heavy isotope deuterium. Hydrogen scatters neutrons out of phase and hence has a negative scattering length density whereas deuterium has a high positive scattering length (Table 2.1). Therefore, replacing hydrogen with deuterium in different proportions makes *contrast matching* possible. Further, neutrons are sensitive to light elements as apparent from high scattering lengths for lighter elements (Table 2.1), whereas X-rays are more sensitive to heavy elements (more electrons). Therefore, neutron and x-ray scattering provide complementary information and hence have been used extensively to elucidate the structural properties of biological systems at physiologically relevant conditions [80–84].

Table 2.1: Neutron scattering lengths of common elements. Data reproduced from NIST website[2].

Nuclei	^{1}H	^{2}H	C	N	O	Na	Si	P	Cl
b [fm]	-3.7390	6.671	6.6460	9.36	5.803	3.63	4.1491	5.13	9.5770

The exact nature of the interaction of neutrons with matter is described by the Schrödinger's wave equation. As a first-order approximation, the reflectivity, R (q) (ratio of reflected and incident intensity), in the weakly interacting regime is given by the Born approximation as [85],

$$R(q) = \frac{16\pi^2}{q^4} \left| \int \frac{d\rho(z)}{dz} e^{-izq_z} dz \right|^2 \qquad (2.15)$$

where, $\rho(z)$ is the SLD profile along z i.e normal to the plane of the surface, q is the momentum transfer, $q = \frac{4\pi}{\lambda} \sin(\theta/2)$ with θ being the scattering angle and λ the wavelength of incident neutrons. The above approximation works well only in the high$-q$ region and diverges in the low-q regime. This is demonstrated in Figure 2.3, where the reflectivity from a silica-D_2O interface is calculated. For such systems at low incidence angles below a certain critical value (q_c), there is total external reflection i.e. R ($q < q_c$)= 1. The Born approximation fails to capture such characteristic properties.

It turns out that the reflection of neutrons (as well as X-rays) can be treated using

---

[2]https://www.nist.gov/ncnr/neutron-scattering-lengths-list
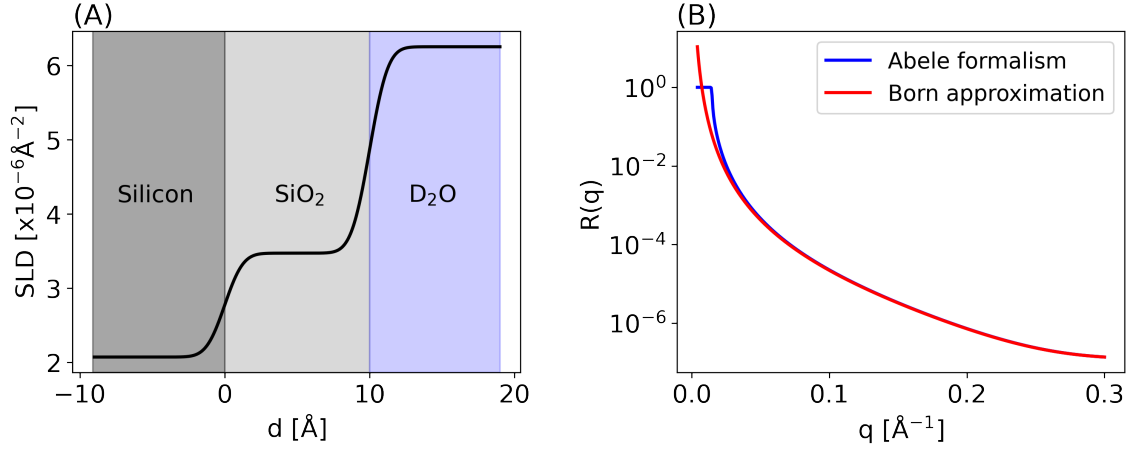
Figure 2.3: Neutron reflectivity from Silica-D_2O system. (A) The system is modeled with three slabs corresponding to silicon, SiO_2, and D_2O layers. (B) Neutron reflectivity from profile in (A) using Abele's matrix formalism and the Born approximation. The Born approximation diverges for the low$-q$ regime.

the *transfer-matrix* formalism as in optics. Such formalism is exact and is the primary method to model reflectivity profiles. One such *transfer matrix* method is the Abeles matrix formalism [86]. In this formalism, the system is divided into layers of constant scattering length densities (SLD) along the direction normal to the surface. Each layer is characterized by three parameters, the scattering length density (ρ_n), thickness (d_n), and roughness (σ_n). Due to differences in these parameters in each layer, the incident neutron beam is partly reflected and partly transmitted. Therefore for each layer, a characteristic transfer matrix having the amplitudes of transmitted and reflected beam can be defined as,

$$c_n = \begin{bmatrix} e^{k_n d_n} & r_{n,n+1}e^{k_n d_n} \\ r_{n,n+1}e^{-k_n d_n} & e^{-k_n d_n} \end{bmatrix} \tag{2.16}$$

where, $r_{n,n+1}$ is the Fresnel reflection coefficient, between layers n and $n+1$, given by,

$$r_{n,n+1} = \frac{k_n - k_{n+1}}{k_n + k_{n+1}}e^{(-2k_n k_{n+1}\sigma_{n,n+1}^2)} \tag{2.17}$$

where, $k_n = \sqrt{q^2/4 - 4\pi(\rho_n - \rho_0)}$ is the neutron wave vector in layer n. The resultant matrix for n layers is ,

$$M = \prod_{0}^{n} c_n \tag{2.18}$$

The reflected intensity at a given q is then given by,

$$R(q) = \left|\frac{M_{21}}{M_{11}}\right|^2 \tag{2.19}$$

The Abele's matrix formalism is demonstrated in Figure 2.3, where the silica-D_2O systems are divided into three slabs corresponding to silicon, the silicon-dioxide (SiO_2) and

the D_2O layer with a certain roughness, thickness and SLD for each layer. Note that the thickness of the incoming medium (Silicon here) and outgoing medium (D_2O) does not affect the calculation.

In the conventional analysis of reflectivity data from lipid bilayers, bilayer-protein, or bilayer-nucleic acid setups, a slab model of the system based on heuristic knowledge about the system is prepared first. For a pure lipid bilayer, the slabs would correspond to the following layers: silicon, SiO_2, lipid head group, lipid tail group, and water slab. Each slab is defined by roughness, thickness, and SLD value. These values are fitted under certain physical constraints until a good agreement between the calculated and experimentally measured reflectivity profile is obtained. More slabs are required for complex systems involving other molecules like nucleic acids or multiple lipid types. Due to the high dimensionality of the modeling procedure, there is no unique solution i.e., multiple SLD profiles may generate the same experimental reflectivity curve. In neutron scattering, such ambiguity is mitigated by performing experiments with multiple solvent contrasts (*contrast matching*) i.e., by varying the relative proportion of H_2O and D_2O in the solvent. The modeling is then carried out to reproduce reflectivity profiles at all the measured contrasts simultaneously.

To obtain reflectivity from simulations, the SLD profile is calculated as,

$$SLD_i(z) = \frac{\sum b_i n_i}{V_i} \tag{2.20}$$

Where n_i is the number of atoms with scattering length b_i in a slice with volume V_i. The continuous SLD profile obtained above is sliced up into many Angstrom-wide slices and Abele's formalism as outlined above is used to obtain the reflectivity profile. Since the width of slabs is of the order an Angstrom, the roughness is set to zero for all slabs. In this thesis, the Refnx python package is used to implement Abele's matrix formalism. Exact implementation with various functionalities is available at the links in section A.1.

2.2.3 Grazing incidence X-ray off-specular scattering (GIXOS)

Like the reflectivity technique described in the previous section, GIXOS can also be used to obtain interfacial SLD profiles (electron density in GIXOS). The major difference is that GIXOS is performed at a fixed angle whereas in reflectivity the radiation source and detector are allowed to move to preserve the specular condition and hence probes only one component (q_z) of the scattered wave vector. Due to the in-plane structural roughness of the lipid bilayer surface, a significant amount of off-specular ("diffused") scattering takes place at a given angle. Off-specular scattering techniques are used to obtain in-plane structural information. However, at very low $q_{x,y}$, the q_z dependence of scattered intensity can be approximated as [87, 88],

$$I(q_z) = \frac{V(q_z) R_q(q_z)}{R_F(q_z)} \tag{2.21}$$

Where, $V(q_z)$, is the Vineyard function, it is defined as follows

$$V(Q) = \begin{cases} 0 & \text{, for Q<0} \\ |2Q|^2 & \text{, for 0<Q<1} \\ \left| \dfrac{2Q}{Q+\sqrt{Q^2-1}} \right|^2 & \text{, for Q>1} \end{cases} \tag{2.22}$$

Where, $Q = \frac{q_z}{q_v}$, $q_v = \sqrt{16\pi\rho_b}$. ρ_b is the electron density of bulk water times the classical electron radius ($2.8179403227 \times 10^{-15}$ m), i.e, $q_v = 0.0217 Å^{-1}$. R_F is Fresnel reflectivity from an infinitely sharp (zero roughness) air-water interface (Figure 2.4 (A, B)) and R_q is the specular reflectivity from the interface. A typical GIXOS profile from a lipid monolayer (Figure 2.4 C) deposited on the water is shown in Figure 2.4 D. Conventional analysis of GIXOS curves is identical to reflectivity as described before, i.e., by assuming slabs for various components and fitting the thickness, roughness, and slab density to minimize the χ^2 deviation between experimental and theoretical profile (Eqn. 2.21). The Python implementation, using Refnx, to obtain such profiles from MD simulations is available at the links in section A.3.
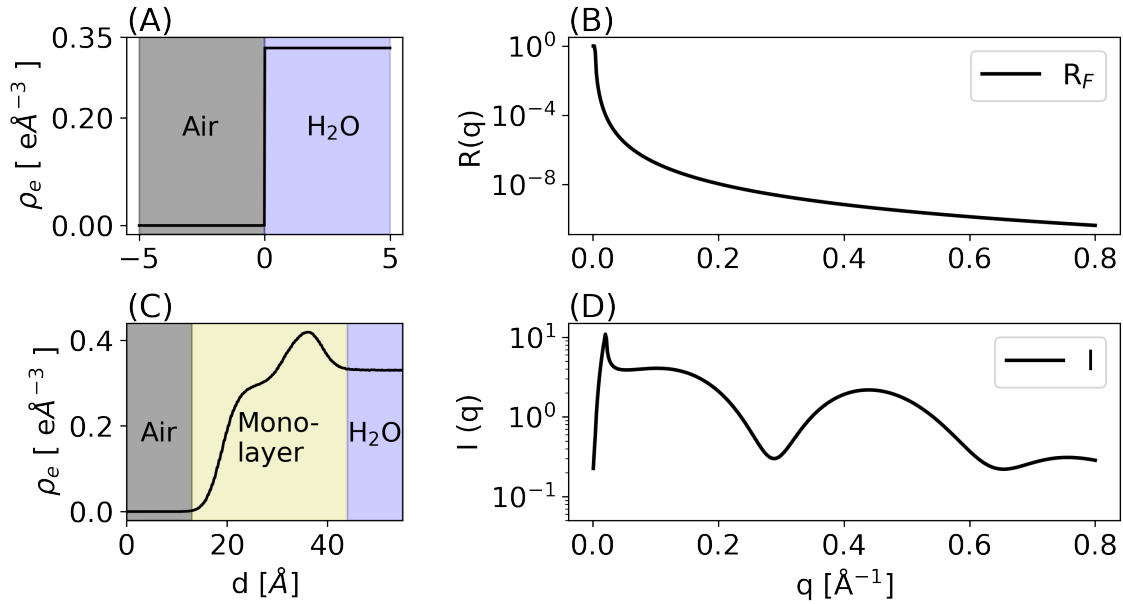
Figure 2.4: Modeling of GIXOS intensity from a lipid monolayer on water. (A) Electron density of an infinitely sharp air-water interface. (B) Specular Fresnel reflection from the interface in (A). (C) Total electron density of a lipid monolayer deposited on a water surface. (D) GIXOS intensity from the interface in (C), obtained using Eqn. 2.21.

23

DNA Interaction with Mica Mediated by Monovalent and Divalent Cations

The work presented in this chapter had been conducted under the supervision of Prof. Dr. Nadine Schwierz at the Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Frankfurt am Main, Germany, and at the Institute of Physics, University of Augsburg, Augsburg, Germany. The simulations were performed on the GOETHE HLR of the state of Hessen. The Atomic Force Microscopy experiments were performed by Christiane Wenzel, Dr. Max Lallemang, and Dr. Bizan N. Balzer at the Institute of Physical Chemistry, University of Freiburg, Freiburg, Germany. Major aspects of this chapter are published in [89]:

- **Mohd Ibrahim**,Christiane Wenzel, Max Lallemang Bizan N. Balzer, and Nadine Schweirz. "Adsorbing DNA to mica by cations: influence of valency and ion type". *Langmuir 2023, 39, 44, 15553–15562*

3.1   Introduction

The adsorption of DNA to solid substrates is crucial for many applications involving DNA nanostructures and bio-sensors [25, 90–93]. The atomically smooth surface of mica and the relative ease to cleave it makes mica an ideal substrate to deposit biomolecules for imaging purposes. The smooth surface ensures a high signal-to-noise ratio. Therefore, it has been used extensively as a substrate to image DNA and various other biomolecule using atomic force microscopy (AFM) to study their physicochemical properties at physiologically relevant conditions and to grow DNA nanostructures [26, 94–101]. Since both DNA and mica are negatively charged, cations are indispensable to bind DNA to the mica surface to facilitate imaging. Divalent cations have been widely used in this regard and the binding affinity of DNA to mica surface has been shown to be dependent on the type, radii, and concentration of cations used [23, 102, 103]. Transition metals cations (Ni^{2+}, Zn^{2+}, Co^{2+}) due to their high charge density have been especially effective in tightly binding DNA to mica surface. However, due to strong binding the DNA is kinetically trapped in a collapsed conformation which most likely differs from its conformation in physiological conditions [104]. A recent study by Heenan and Perkins demonstrated the successful adsorption of DNA to mica surfaces in physiologically relevant ionic conditions containing K^+ and Mg^{2+} [25]. However, the exact picture of the molecular interactions leading to the adsorption remains unclear. Consequently, a protocol for the selection of optimal ionic conditions for better imaging or for growing nanostructures with desired properties remains unresolved.

Several theoretical models have been put forth to explain the cation-mediated attraction between DNA and mica [105–110]. Most of these models predict an attractive force between

DNA and mica only in the presence of divalent cations. The attraction is explained by the ion-ion correlation effect which arises from different contributions like ion-density fluctuations around the mean ion distribution, non-electrostatic ion-ion interactions, and direct and water-mediated interactions between the ions and the surfaces [108]. Models based on counter-ion condensation also predict attraction in the presence of both mono and divalent cations [110]. In most of these models, both mica and DNA are assumed as rigid, uniform line charge density with fixed, structureless cation binding sites with no ion specificity. Therefore, these models disregard many important effects like ion-specific binding toward DNA and mica, hydration effects, and conformational flexibility of DNA all of which depend on the ion type [16, 17, 27]. Hence, it is crucial to understand the role of each of these effects to obtain a full understanding of DNA interaction with the mica surface.

All-atom molecular dynamics simulation provides an ideal method to resolve all the above contributions. In particular, steered molecular dynamics simulations closely mimic AFM-based Single Molecule Force Spectroscopy (SMFS) experiments and have been used to gain atomistic insights into the interaction of biomolecules and interfaces as well as intramolecular interactions involved in the unbinding mechanism of proteins[17, 111–115]. AFM-based Single Molecule Force Spectroscopy (SMFS) is a versatile method for studying the interactions between single molecules and surfaces or interfaces. Using this method it is possible to measure the molecular interaction forces in the range of piconewtons and has been successfully applied to characterize the interaction of different biomolecules with interfaces like lipid bilayers, gold, and graphite surfaces [113, 114, 116–120].

Here, we combined steered MD simulations and single-molecule AFM experiments to measure the detachment force between single-stranded DNA and muscovite mica surface in the presence of mono- and divalent cations $Li^+, Na^+, K^+, Cs^+, Mg^{2+}$ and Ca^{2+}. We find that DNA-mica interactions involve a complex interplay of the cation affinity towards DNA, mica surface, and its hydration properties leading to a non-monotonic dependence of force distribution on ion type. In both experiments and simulations, Cs^+ leads to the lowest force whereas K^+ and Na^+ result in similar forces. For the ions with high charge density i.e., Na^+, Mg^{2+} forces can be higher or lower depending on the mode of interaction i.e., whether direct or water-mediated. Our combined approach provides an atomistic picture of the origin of ion-specific forces and provides an informed way to choose proper ionic conditions for the adsorption of DNA on mica surfaces in physiologically relevant conditions.

3.2 Methods

Here, only the simulation methods are presented. For details of experimental methods refer to Ibrahim et al. [89].

3.2.1 System setup

A snapshot of the full simulation setup is shown in Figure 3.1 A. It consists of a muscovite mica $(KAl_3Si_3O_{10}(OH)_2)$ crystal with dimension $8.26 \times 7.172 \times 4.00$ nm^3 and the 001 plane exposed to the solvent. The naturally contained K^+ ions on the mica surface were

(A)   (B)   (C)

No.	Name	Charge [e]
1	at	+1.5750
2	obts	-1.1688
3	st	+2.100
4	ob	-1.0500
5	oh	-0.9500
6	ho	+0.4251
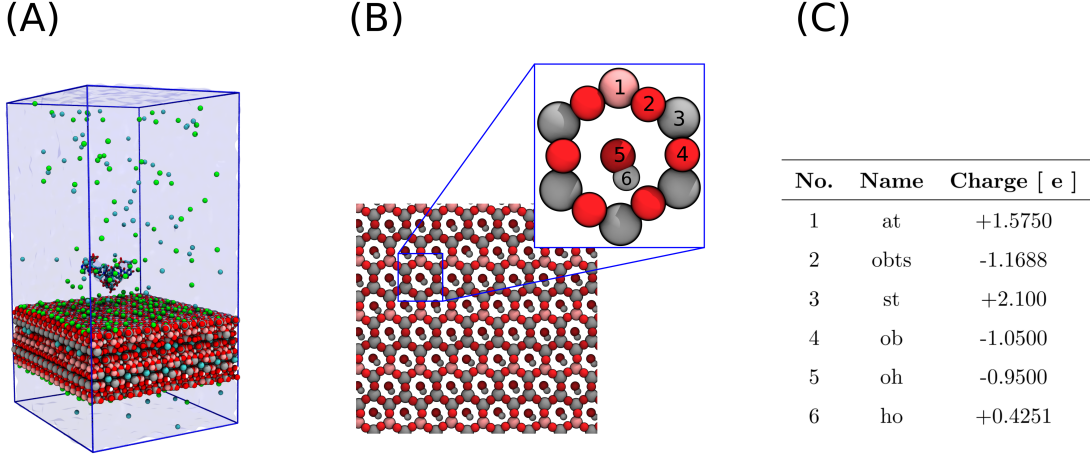
Figure 3.1: (A) Snapshot of the full simulation system with mica crystal, DNA, ions, and water (transparent background). (B) A close view of the mica surface exposed to the solvent and ions. The surface has silicon (st), aluminum (at), and two types of oxygens (i) bridging oxygen, ob, which bridges two silicon atoms (ii) obts, which bridges a silicon and an aluminum atom. These atoms form ditrigonal cavities. The hydroxyl (oh and ho) group is recessed below the surface (by around ~ 1.7Å). (C) Charges and names of surface atoms corresponding to the atoms numbered in (B).

removed to render the surface negatively charged. Each of the exposed surfaces has a charge of $-128e$ making the surface charge density, σ_{bare} = -2.16 enm^{-2}. A single-stranded DNA (ssDNA) molecule composed of 10 Thymine (poly-T) nucleotides was placed at a perpendicular distance of 2 nm from the surface. The ssDNA molecule was generated by removing one of the strands of a double-stranded helix composed of poly-T and its complementary sequence. The double-stranded helix was obtained using the `fd_helix.c` program. After solvating the system, cations and chloride ions were added to neutralize the surface negative charge and to obtain a bulk concentration of 150 mM. The final systems contained around ~ 97000 atoms and the box dimension was $8.26 \times 7.172 \times 16.7$ nm³. Six different setups corresponding to cation Li^+, Na^+, K^+, Cs^+, Mg^{2+}, and Ca^{2+} were simulated.

3.2.2 MD Simulations

The simulations were performed using the Gromacs package (v-2018, v-2021.5) [121]. The Vander Waals interactions were cut-off at 1.0 nm and switched to zero at 1.2 nm using a switching function. Short-ranged Coulomb interaction was cut off at 1.2 nm and long-range part was evaluated with the Particle Mesh Ewald method with a Fourier grid spacing of 0.12 nm [57]. All bonds involving hydrogen were constrained using the LINCS algorithm [122] which allowed an integration time step of 2.0 fs. TIP3P water model [123] was used to describe water, the Amber parmBSC1 forcefield [124] was used to describe DNA, and the Clayy.ff [125] was used for mica. The monovalent ions were described using the Mamatkulov-Schwierz forcefield [126], whereas, for divalent cations Mg^{2+} and Ca^{2+}, recently developed parameters were used [127, 128]. Analysis was done using an inbuilt Gromacs routine or by self-written Python scripts based on the MDAnalysis python package [129]. All visual analysis were done using the VMD package [130].

The simulation protocol closely mimics the experimental procedure. In the experiments, the DNA is covalently attached to the AFM tip through a linker molecule, the tip (and hence the DNA) is pushed towards the surface and after certain *dwell time*, the DNA is pulled away from the surface and the corresponding force of desorption is measured. Similarly, the simulation consists of the following protocol (i) The system was first energy minimized with a gradient descent algorithm. (ii) Subsequently, a 50 ns equilibration run was carried out with position restraints on the DNA. This allows the cations to adsorb at the negatively charged mica surface and form an equilibrated counter-ion layer known as electric *double-layer*. Temperature and pressure were maintained using the Berendsen thermostat and barostat [59] with a time constant of 1 ps. (iii) The restraint on the DNA was removed and allowed to relax for 1 ns (iv) The DNA was pushed towards the surface by applying a constant force over a period of 50 ns (v) The DNA was allowed to relax on the surface for 250-300 ns. During this time the number of contact between the DNA and the surface is equilibrated as indicated in Figure A3.1. And finally (vi) Frames from the last 150 ns of the previous run were selected for pulling simulations using the steered MD method as described in the following section. From protocol (iii) onward, the temperature and pressure were fixed using the stochastic velocity rescaling thermostat with a time constant of 1.0 ps [58] and the Parrinello-Rahman barostat [60] with a time constant of 5.0 ps respectively.

3.2.3   Steered molecular dynamic simulations

Steered molecular dynamic simulations (SMD) mimic single molecule techniques like SMFS and magnetic tweezers experiments where inter-molecular or intra-molecular interaction strength is measured by applying an external force.

Constant velocity SMD

Here, a harmonic potential is applied between two groups, in our case between DNA and surface.

$$U(\mathbf{x}, t) = \frac{k(\mathbf{x} - (\mathbf{x_0} + \mathbf{v}t))^2}{2} \tag{3.1}$$

which leads to a time-dependent harmonic force as,

$$\mathbf{F}_{\text{ext.}}(\mathbf{x}, t) = k(\mathbf{x_0} + \mathbf{v}t - \mathbf{x}) \tag{3.2}$$

Where k is the spring constant of the harmonic potential, $\mathbf{v}$ is the constant pulling velocity. $\mathbf{x}$ is the reaction coordinate vector along which pulling is performed. In, our case $\mathbf{x}$ is the $z-$distance between the center of mass of mica surface atoms and the center of mass of $5'$-end Thymine residue with initial value $\mathbf{x}(t=0) = \mathbf{x_0}$. Both $\mathbf{x}$ and $\mathbf{v}$ are along z-direction. For each system with a given cation, two pulling velocities were used $v = 0.1 \ m/s$ and $v = 10 \ m/s$. For the Na$^+$ setup, three additional pulling velocities ($v = 0.01, 0.05, 1.0 \ m/s$), to study the force scaling with velocity, were used. For $v = 10 \ m/s$, 100-130 pulling curves were generated whereas for $v = 0.1 \ m/s$ 6-10 force-extension curves were measured
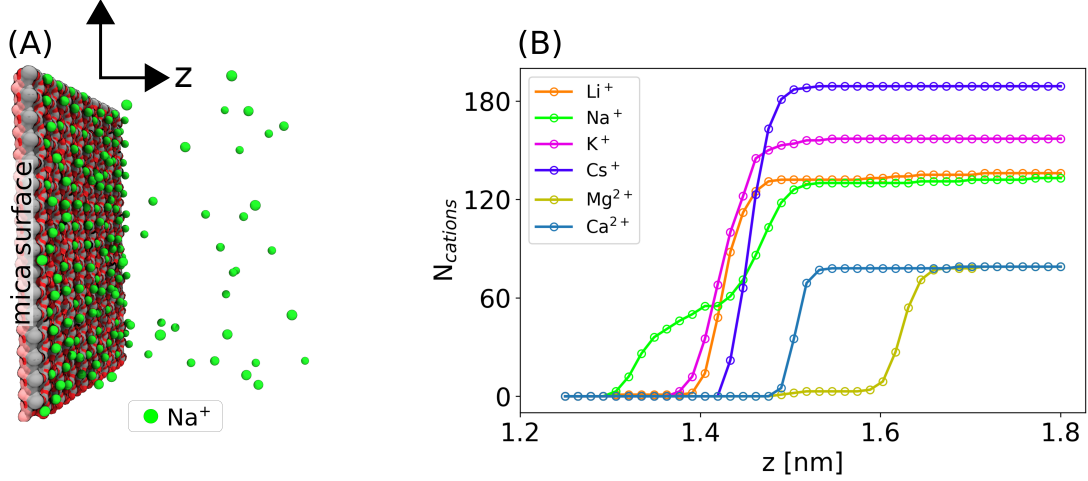
Figure 3.2: (A) Cations accumulate at the mica surface to compensate the surface negative charge. The surface has a net charge of $-128e$. (B) The absolute number of cations of different types accumulated at the surface as a function of distance from the center of mica crystal. We obtain 136 Li$^+$, 133 Na$^+$, 156 K$^+$, 189 Cs$^+$, 79 Ca^{2+} and 79 Mg^{2+} surface adsorbed ions.

for each ion.

Constant force SMD

In constant force SMD a linear potential, $U(x) = -\alpha|\mathbf{x}|$, is applied between two groups. We used this technique to push the DNA toward the mica surface. Force was applied on all the non-hydrogen atoms of DNA, and the value of α for each atom was chosen such that each atom is accelerated by the same amount $\mathbf{a}$, i.e, the total force on the DNA is,

$$\mathbf{F_{ext.}} = \sum_i \alpha_i = \sum m_i \mathbf{a} \tag{3.3}$$

where, m_i is mass of ith non-hydrogen atom. a is set such that $F_{ext} = 470$ pN, which is closer to the value used in experiments (500 pN) to push the DNA towards the surface.

3.2.4 Obtaining the correct bulk salt concentration

For a system with N_w water molecules, number of ions, N_{ions}, required to obtain a bulk salt concentration c (in Molar) is given by,

$$N_{ion} = \frac{N_w c}{55.55} \tag{3.4}$$

In the case of highly charged systems like mica, ions are also required to neutralize the charge of the system. As mica is highly negatively charged, cations adsorb to it to compensate for the negative charge (Figure 3.2 A). However, we observed that cations accumulate on the mica surface in excess, i.e. in more quantity than what is required to neutralize the surface negative charge. The amount of excess-adsorbed cation depends on the type of cation as clear from Figure 3.2 B, where the absolute number of surface
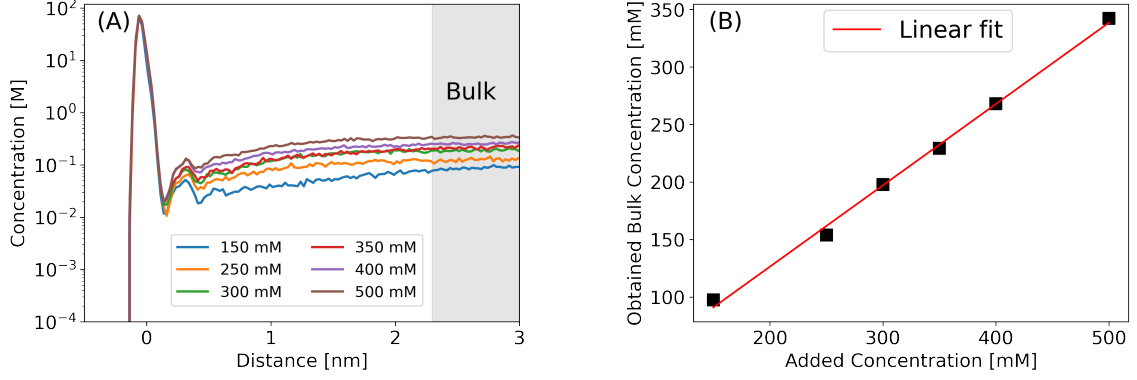
Figure 3.3: (A) Concentration profile for K $^+$ ion for different added concentrations according to Eqn. 3.4. The added concentrations are shown in the legend. These profiles were obtained by considering the last 50 ns of a 100 ns equilibration run. (B) We obtain a decreased bulk concentration due to the over-accumulation of cations on the negatively charged (-128 e) mica surface. We used a linear fit of this curve to obtain a bulk concentration of 150mM for a given ion.

adsorbed ion ($N_{cations}$) is plotted,

$$N_{cations}(z) = \int_0^z \rho_n(z) l_x l_y dz \qquad (3.5)$$

Where, the ρ_n is the cation number density along z, l_x and l_y are box dimensions along x and y respectively. Once d is outside the adsorbed cation layer, $N_{cations}$ saturates (for small d) as the concentration of ions in the bulk is orders of magnitude lower than at the surface. Depending on cation type each $N_{cations}$ has different saturation values (Figure 3.2 B). Due to the finite number of ions in our system, such excess accumulation of cations on the surface leads to a decrease in bulk salt concentration. Figure 3.3 B shows the concentration added according to Eqn. 3.4 and the obtained bulk concentration (Figure 3.3). However, the added vs obtained concentration shows a linear dependence in the concentration range of interest. Therefore, for a given ion, to obtain a bulk concentration of 150 mM (i) two 50 ns long simulations with 150 mM and 300 mM added concentrations were run (ii) from a plot like Figure 3.3 B for the two concentrations, the concentration required to obtain a 150 mM bulk value was interpolated. In summary, 327 Li$^+$, 325 Na$^+$, 360 K$^+$, 409 Cs$^+$, 214 Mg^{2+}, and 212 Ca^{2+} ions were added to obtain a bulk concentration of 150 mM for systems with corresponding cations.

3.2.5 Rupture force definition

In pulling experiments, the usual quantity of interest is the rupture force, i.e. the force it takes to break certain interactions. Here, in both experiments and simulations, we defined it as the peak force just before the DNA desorbs from the surface. Rupture force is assigned as the peak value in the force-extension curve within a 1.5 nm region around the desorption point (Figure 3.4). In most cases, this also corresponds to the maximum force measured (Figure 3.4 B).
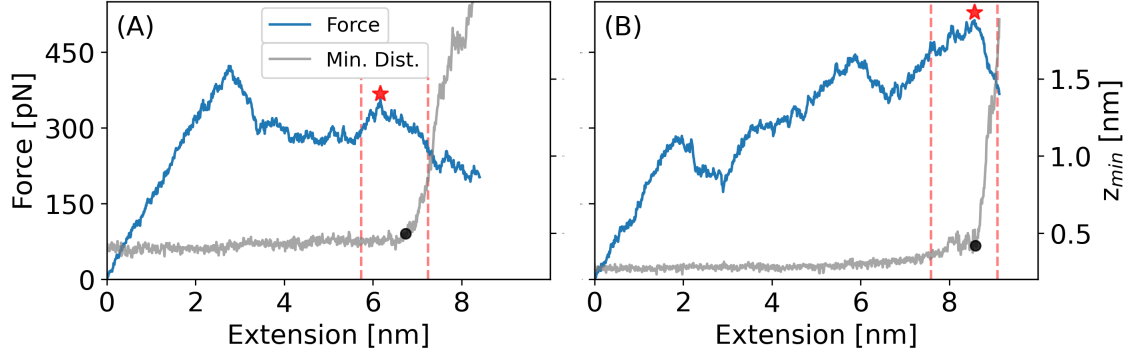
Figure 3.4: The desorption point is determined by the point at which the minimum distance of DNA from the mica surface (Min. Dist.) starts increasing sharply (black dot markers) (A) In this case, force at detachment is smaller than the maximum force over the course of desorption. (B) Here, the detachment force is the same as the maximum force over the whole force-extension curve.

3.3   Results

3.3.1   Ion-specific interfacial structure of DNA, mica, and cations

The negative surface charge of mica leads to the formation of a pronounced double layer. In addition to the diffusive ions, which are loosely associated with the surface, the first layer of cations adsorbs specifically at the mica surface. Depending on the cation type, these ions occupy different binding sites: The ditrigonal cavities or the negatively charged O atoms (obts) bridging Al and Si atoms (Figure 3.1 B, Figure 3.5 A). For monovalent ions, small-sized Li^+ exclusively binds to obts atoms, Na^+ occupies both obts atoms and the ditrigonal cavities, whereas K^+ and Cs^+ exclusively occupy the ditrigonal cavities (Figure 3.5 A). For divalent ions, Mg^{2+} binds to both the bridging oxygens (obts and ob) whereas Ca^{2+}, like Li^+, binds only to obts (Figure 3.5 A). The sharing of cation between DNA-phosphate oxygen and mica surface stabilizes the DNA on the surface (Figure 3.5 B). The different binding modes are also reflected in the cation density profile along the direction normal to the surface (Figure 3.5 (C-H)), e.g, Na^+ binding in both the ditrigonal cavities and to obts appears as two peaks in the density profile. The area under the highest peaks (Eqn. 3.5) gives the total cations adsorbed on the surface (Figure 3.2 B). From the amount of adsorbed cations ($N_{cations}$), the effective surface charge density for a given cation is,

$$\sigma_{eff} = \frac{z N_{cations} + q_{mica}}{l_x l_y} \tag{3.6}$$

Where, z is cation valency, $q_{mica} = -128.0e$ is the total charge of bare mica surface. σ_{eff} values for different cations show that Cs^+ ion reverses the surface charge and renders it effectively positively charged, Na^+ and Li^+ just neutralize the negative surface charge and the other ions lie in between (Table 3.1).

The density profile of interfacial water exhibits pronounced differences due to differences in cation distribution on the surface and in cation hydration properties (Figure 3.2 C-H), which is also reflected in the magnitude of the dominant peak in water density profiles, e.g, the water layer for Cs^+ is highly diffused and least dense, while Mg^{2+} exhibits a highly

31

structured and very dense water layer (Figure 3.5 F, G). For, Mg^{2+}, both bulk and surface-adsorbed cations are always hydrated, resulting in the water peak appearing before the cation peak (Figure 3.5G). This is also evident from the snapshot in Figure 3.5B, where a gap can be seen between Mg^{2+} and the surface, indicating the presence of a water layer within the gap. In contrast, for all other ions, the cation peak appears first, indicating direct binding of the cation to the mica surface.
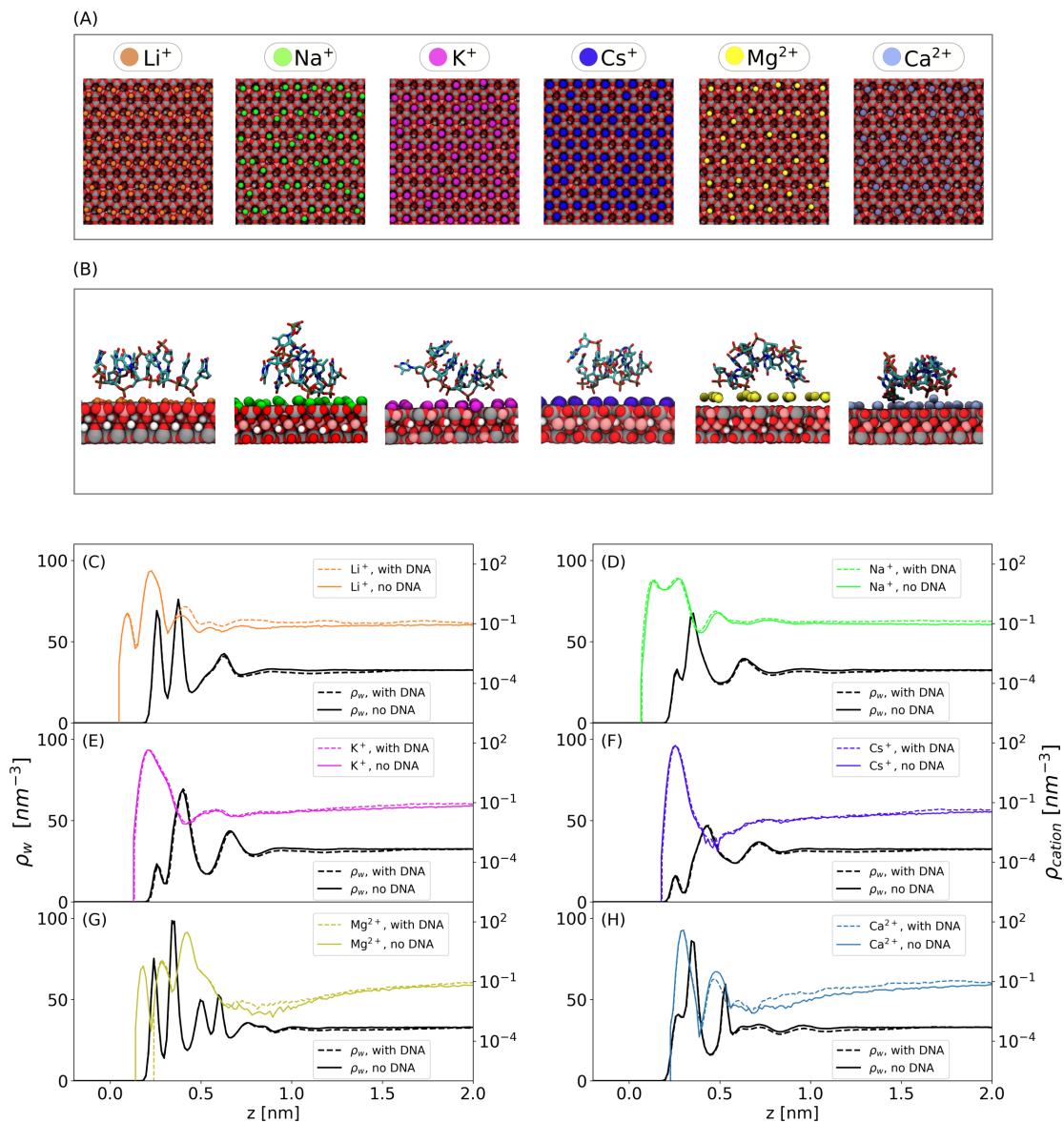
Figure 3.5: Distribution of cations, water, and DNA at the mica/water interface. (A) Simulation snapshot of the specifically adsorbed cations at the mica surface. The ion binding sites shift with increasing charge density from the oxygen triad to the center of the ditrigonal cavity. (B) Simulation snapshot of the ssDNA at the mica/water interface for different cations. (C-H) Ion distribution profiles and density profiles of the water oxygens. The profiles are shown for both sides of the mica crystal: Without DNA (solid lines) and with DNA (dashed lines)..
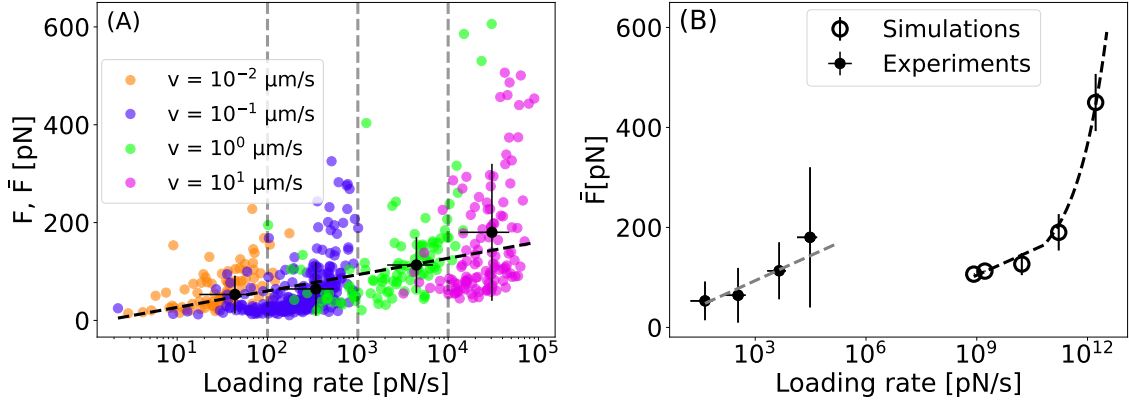
Figure 3.6: Dynamics of DNA desorption from mica in 150 mM NaCl. (A) Rupture force, F, and average rupture force $\overline{F}$ as a function of the loading rate. Different colors indicate the four different pulling velocities. The individual data points from each measurement are sorted into four blocks according to the pulling velocity applied to the AFM system, and the average rupture force is calculated (large black points). Errors indicate the standard deviation. The dashed line indicates a fit based on the Bell-Evan model. (B) Mean force at rupture as a function of the loading rate from experiments and simulations. The dashed line corresponds to fit from using the Bell-Evens model for experiments and $\sqrt{r}$ scaling for simulations with r being the loading rate.

3.3.2 Rupture forces scaling in experiments and simulations

Single-molecule force spectroscopy can be used to extract kinetic information about the interacting molecules/surfaces. There are many theoretical models [131–134] which relate the average rupture force to the force loading rate making it possible to extract characteristic kinetic parameters like the barrier height, transition point, and transition rate (k_0) from bound to the unbound state. To employ such models, the average rupture force is measured at multiple force loading rate values (r). Figure 3.6 shows the resulting rupture force vs loading rate profiles for a system with Na^+ cation. The experiments were performed at four different pulling velocities (0.01, 0.1, 1.0, 10.0 $\mu m s^{-1}$), indicated by different colors in Figure 3.6A. The data points are divided into four different bins according to the loading rate value and the average rupture force is evaluated for each bin. The average rupture force vs loading rate is fitted with a Bell-Evans model [135],

$$\bar{F} = \frac{1}{\beta x^{\ddagger}} ln \frac{\beta r x^{\ddagger}}{k_0} \tag{3.7}$$

where, $\beta = \frac{1}{k_B T}$, $T = 300\ K$, $x^{\ddagger}$ is the distance from the bound state to the transition point, beyond which the DNA is completely desorbed from the surface, k_0 is the desorption rate in absence of any external perturbation and r is the loading rate. From the fit we obtain, $x^{\ddagger} = (0.33 \pm 0.08)\ nm$, $k_0 = (0.366 \pm 0.3)\ s^{-1}$.

In order to compare the scaling regime of experiments and simulations, we performed simulations at different pulling velocities ranging from $(0.005-10\ m/s)$. Since the simulation pulling velocities are several orders of magnitude higher than in experiments, the force vs loading rate from simulations follows a different scaling behavior of form, $F \propto r^{\frac{1}{2}}$, which is in accordance with analytical theories [136].

Table 3.1: The effective surface charge density (σ_{eff}) of mica surface, the diffusion coefficient (D) of DNA on the mica surface and average rupture forces from simulations and experiments, in presence of different cations. The forces correspond to the distribution shown in Figure.

Ion Type	σ_{eff} [e nm^{-2}]	D [10^{-7} cm^{2}s^{-1}]	$\overline{F}_{\text{sim}}$ [pN]	$\overline{F}_{\text{exp}}$ [pN]
Li$^+$	0.13	0.65 ± 0.51	575 ± 173	72 ± 53
Na$^+$	0.08	0.64 ± 0.20	421 ± 59	79 ± 61
K$^+$	0.47	2.17 ± 0.61	404 ± 69	124 ± 149
Cs$^+$	1.03	2.92 ± 1.15	318 ± 55	75 ± 138
Mg^{2+}	0.51	5.32 ± 3.09	337 ± 131	189 ± 127
Ca^{2+}	0.51	0.81 ± 0.16	627 ± 112	159 ± 176
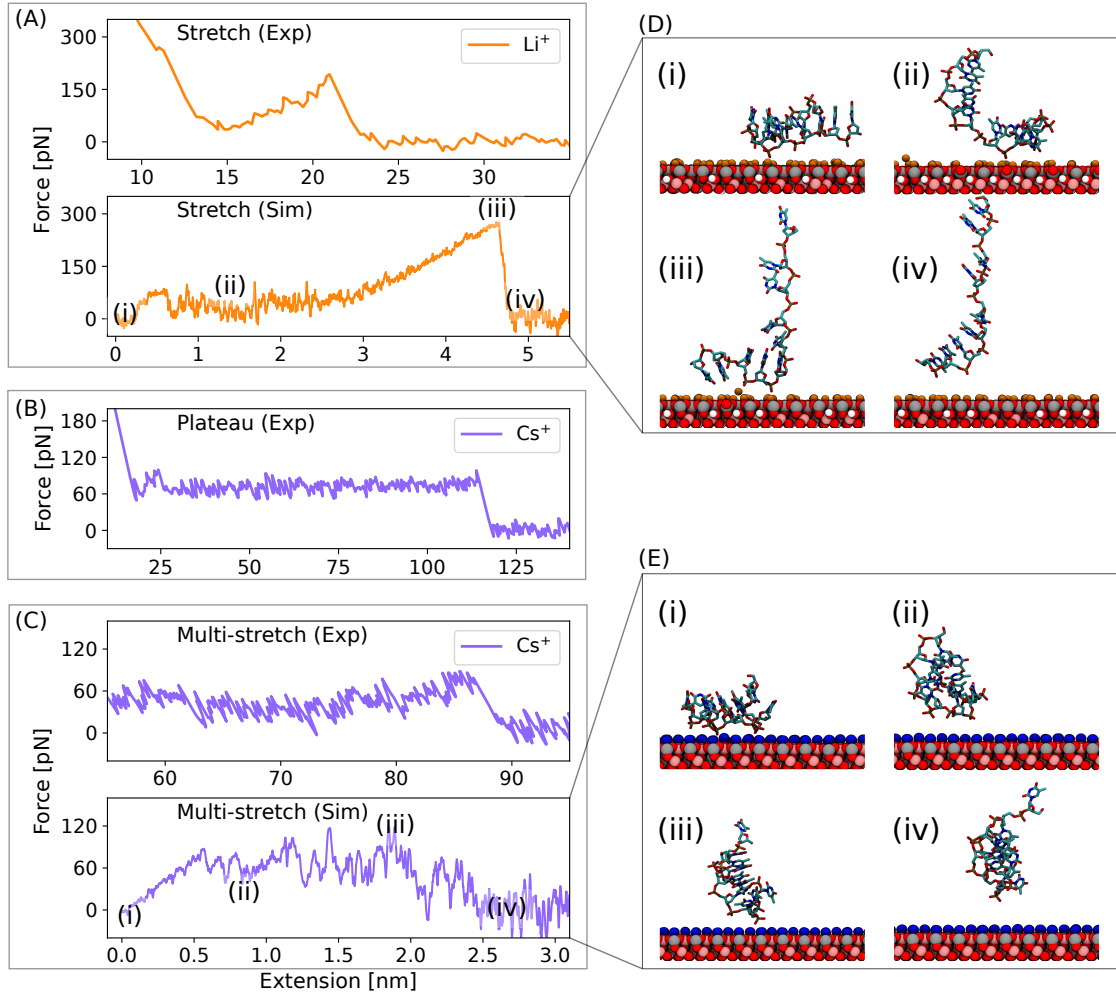
Figure 3.7: Comparison of the curve shapes. (A-C) Representative force-extension curves obtained from experiments and simulations for Li+ and Cs+. Plateau-type curves are not observed in the simulations. (D-E) Snapshots of the DNA at different points during the desorption process are indicated in the corresponding simulated force-extension curves. Here, the force-extension curves correspond to pulling velocities of 0.1 m/s for simulations and 1.0 $\mu m/s$ experiments. Further force-extension curves from simulations can be found in Figure A3.2-A3.4.

3.3.3 Force extension curves from experiments and simulations

The force-extension curves from experiments and simulations show similar shapes indicating a similar desorption mechanism (Figure 3.7 A-C). The curve shapes from experiments are broadly divided into three categories namely plateaus, stretch, and multi-stretch. The stretch curve shapes are most frequently observed in the experiments compared to plateaus and multi-stretches. Whereas in the simulations, the plateau forces are never observed which is consistent with previous studies involving steered MD simulations of peptides from lipid membranes and other surfaces [113–115]. The plateau force (Figure 3.7B) arises from the fast relaxation time of DNA between the mica surface and the AFM tip compared to the pulling speed [114]. Such a stationary equilibrium is not achieved in the simulation due to the very high pulling velocity. The stretch curves arise due to stretching of the DNA between the mica surface and the AFM tip (Figure 3.7D). Here, the DNA is tightly anchored to the surface by an ion that has a strong affinity for both the phosphate group of the DNA and the mica surface (e.g. Li^+, Ca^{2+}) leading to the stretching of DNA. The multi-stretch arises from multiple binding and unbinding events during the course of desorption (Figure 3.7E), this happens for ions with weaker hydration shells which facilitates multiple rebinding leading to a characteristic stick-slip motion of DNA on the surface. Here the conformation of DNA stays intact. Thus, in addition to ion-mediated DNA-surface interactions, the conformational changes of the DNA (Figure 3.7D-E) also have an impact on the shape of the force-extension curves.

3.3.4 Origin of ion-specific rupture forces

Figure 3.8 show the distribution of rupture forces from experiments and simulations normalized by the corresponding average rupture force for Na^+ ion. The average forces from experiments and simulations do not show a monotonic trend with the charge density of ions (Table 3.1). The ion-specific contributions are obscured by the broad distribution of forces. In general, for both experiments and simulations display similar behavior: Cs^+ results in the lowest forces, while Na^+ and K^+ exhibit similar force distribution. For Mg^{2+}, Ca^{2+}, the individual forces can be lower or higher. In the simulations, Li^+ displays a broad distribution of forces with very high and low forces in contrast to the experimental distribution which is relatively sharply peaked around a lower force value.

To explain the observed trend of rupture forces we looked at the type of interaction that breaks at rupture. We classify such interactions into four categories: (i) Water-mediated (WM), in this case, there is a hydration layer between the DNA and the surface-bound ions (Figure 3.9A). Rupture force here arises from the breaking of H-bonds between the hydration layer and DNA phosphate oxygens. (ii) Ion-phosphate contact (IP), in this case, the shared cation between the DNA phosphate group and mica surface stays bound to the surface leading to the breakage of an ion-phosphate contact (Figure 3.9B). (iii) Ion-surface (IS), here, the shared cation leaves with the DNA causing the breakage of an ion-surface contact. This category also includes the special case for Mg^{2+} (IS_{Mg}), where, contact between hydrated Mg^{2+} and surface breaks (Figure 3.9C). (iv) Ion-phosphate and Ion-surface (IP&IS), in this case, the non-bridging phosphate oxygen is coordinated to two
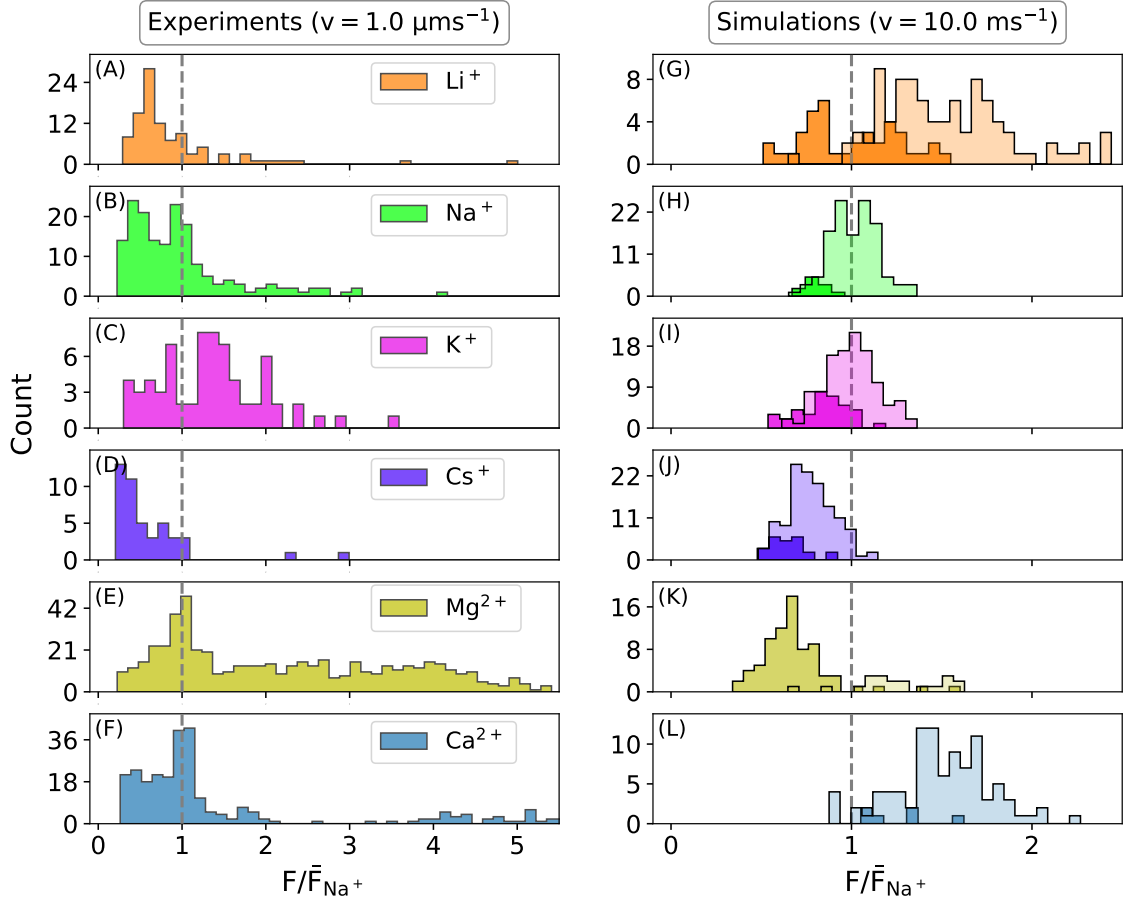
Figure 3.8: Ion-specific distribution of rupture forces of DNA from mica from simulations and experiments for different cations. Left (A-F): Distribution from AFM experiments with v = 1.0 μm/s. The average rupture force is given in Table 3.1. For comparison, the forces are shown relative to the average rupture force in NaCl. Right (G-L): Distribution from pulling simulations with $v = 10$ m/s. The simulation distribution in each panel is shown in two shades. The lighter shades correspond to the WM desorption mechanism and the darker shades correspond to the IP, IS, or IP&IS mechanism. See the main text for an explanation.

surface-bound cations, at rupture, one cation stays bound to the surface while the other leaves with the DNA leading to a simultaneous breakage of ion-surface and ion-phosphate contact (Figure 3.9D). This is observed for small-sized cations i.e., Li^+ and Na^+. In general, the strength of these interactions varies as WM < IP < IS_{Mg} $\leq$ IS < IP&IS. Figure 3.9 E-J shows that depending on the ion type the rupture force has varying contributions from all the outlined contact types. For Li^+, rupture force has an almost equal contribution from all the contact types (Figure 3.9 E), whereas, for other monovalent ions, the rupture forces arise mostly from ion-phosphate interactions (Figure 3.9 F-H). For Mg^{2+} the interactions are mostly water-mediated and for Ca^{2+} ion-phosphate contact are predominant (Figure 3.9 I-J). Therefore, ion-specific rupture force originates from a complex interplay of ion-DNA, ion-surface, and ion-water interactions.
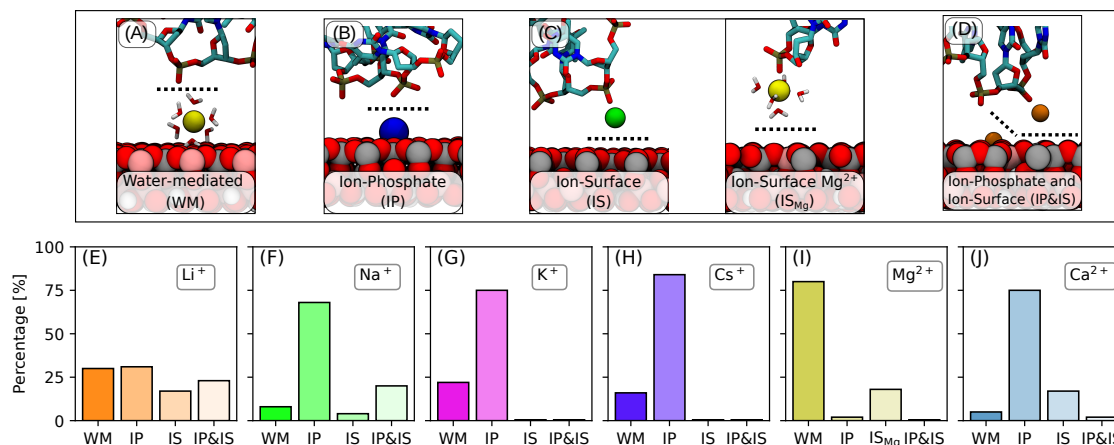
Figure 3.9: Depending on ion type rupture force arises from the breaking of different types of contacts. (A-D) Snapshots at rupture depicting different types of contacts that break, the dashed line shows the point where contact breaks. (E-J) Percentage of different contact types that give rise to the measured rupture forces. For a given panel the lighter shades correspond to contacts that in general give rise to higher rupture forces.

3.4 Discussions

Our results reveal that DNA interaction with the mica surface primarily occurs via the sharing of cations between the oxygens of DNA-phosphate backbone and the mica surface. The interaction is influenced by the cation's affinity towards the mica surface, the DNA phosphate oxygen, and as well as the cation hydration properties. Due to their unique size and charge density, the cations exhibit unique affinity towards the mica surface, occupy different binding sites at the surface, and accumulate at the surface in different amounts. Small-sized cations like Li^+ and Na^+ adsorbs to just neutralize the surface negative charge, whereas cations like K^+ and Cs^+ adsorb to the surface in excess effectively reversing the surface charge and rendering it positively charged (Table 3.1). This gives rise to a relatively long-ranged electrostatic attraction between DNA and mica as evident from the weak spontaneous adsorption of DNA to the mica surface without any external force in the case of Cs^+ and K^+. Such spontaneous adsorption is also a result of a weak hydration layer which acts as a barrier for the DNA to directly access the surface adsorbed cation layer. Strongly hydrated ions prevent spontaneous adsorption and in such cases, a small force is necessary to overcome this hydration barrier. The hydration properties also influence the observed shape of the force-extension curves. At low pulling speeds, for cations with weak hydration layers, the DNA can bind and unbind multiple times before fully detaching from the surface, giving rise to multi-stretching stick-slip curves. Note that, in addition to the weak hydration properties, for observing stick-slip-like motion, the cation must also have a weak affinity towards the DNA-phosphate group, which leads to a short lifetime of ion-DNA contact (∼100 ps) [16]. This weak affinity facilitates rapid binding and rebinding. For strongly hydrated ions with a high affinity toward the DNA phosphate group, such rebinding does not occur due to a large hydration barrier and due to relatively stable ion-DNA contact pairs resulting in a stretching force-extensions curve.

A direct quantitative comparison of rupture forces between experiments and simulations

is not possible since they lie in different scaling regimes (Figure 3.6). At a low loading rate as in experiments, the Bell-Evans model yields a value of $x^{\ddagger} = (0.33 \pm 0.08)$ nm for the distance between the bound state and the transition state and an intrinsic desorption rate $k_0 = (0.37 \pm 0.30)$ s^{-1}. These values indicate that the desorption process involves the breaking of strong local interactions instead of the gradual unbinding of the single-stranded DNA molecule as it is pulled away from the surface. From our simulations, these local contacts would correspond to any of the contacts, outlined in the previous section, that break during rupture. In literature, the measurement of polytryptophan and polytyrosin on a lipid bilayer showed a $x^{\ddagger}$ of 0.65 nm and 0.63 nm, respectively [114]. The unbinding of a FITC-E2 wt mutant revealed a $x^{\ddagger}$ of about 0.4 nm [137].

The rupture forces originate from different interactions which depend on the affinity of ions towards DNA, mica, and as well as water. Due to convoluted contributions from these interactions, a clear ion-specific trend of rupture forces is not observed. The magnitude of rupture forces from ion-phosphate and water-mediated interactions are dictated by the affinity of cation towards DNA phosphate oxygen and it follows the order, $Cs^+ < K^+ < Na^+ < Li^+ < Ca^{2+} < Mg^{2+}$ [16, 138]. Whereas, ion-surface interactions are dictated by the cation-mica affinity which follows $Mg^{2+} \leq Ca^{2+}$ and $Na^+ \leq Li^+ < K^+ < Cs^+$. For monovalent ions Cs^+, K^+, and Na^+ rupture force predominantly arises from ion-phosphate interaction (Figure 3.9 F-H). Therefore, rupture force is determined by the ion-phosphate affinity which is the least for Cs^+. This explains the smallest force observed for Cs^+ in both experiments and simulations. For Mg^{2+}, both experiments and simulation exhibit a bimodal distribution but the high force part is underrepresented in the simulations. The simulations reveal that low forces arise from indirect water-mediated interactions and high forces arise from ion-surface interactions (Figure 3.9I). Due to the strong hydration of Mg^{2+} and slow water exchange rate, the water-mediated contact is dominant. However, by using a high temperature pre-equilibration procedure that allows inner-sphere binding of Mg^{2+} to DNA phosphate oxygens, we could reproduce the high force part of the experimental distributions (Figure 3.10). In this procedure, simulations were first conducted at a high temperature (400 K) allowing the strongly hydrated Mg^{2+} to shed off their hydration shells and form direct contact. The temperature was then reduced to normal (300 K) and pulling simulations were performed as usual. For both Li^+ and Ca^{2+}, the higher forces are over-represented in the simulations compared to experiments. The discrepancy might be due to the force field overestimating Li^+ affinity towards both mica surface and DNA leading to very high rupture forces. To improve the agreement between experiments and simulations, affinity-optimized force fields that correctly and simultaneously reproduce the binding affinities of the cations to DNA and mica. However, to proceed along those lines accurate experimental data on the binding affinities between mica and DNA is required [127, 128].

The mobility (or the diffusion coefficient, D) of DNA on the mica surface is an important quantity of interest for creating well-ordered DNA nanostructures [26]. For such purposes, a mixture of divalent ions typically Mg^{2+}, and a monovalent ion is used. The monovalent ion (e.g., Na^+) is employed to enhance the mobility while the divalent ion is supposed to
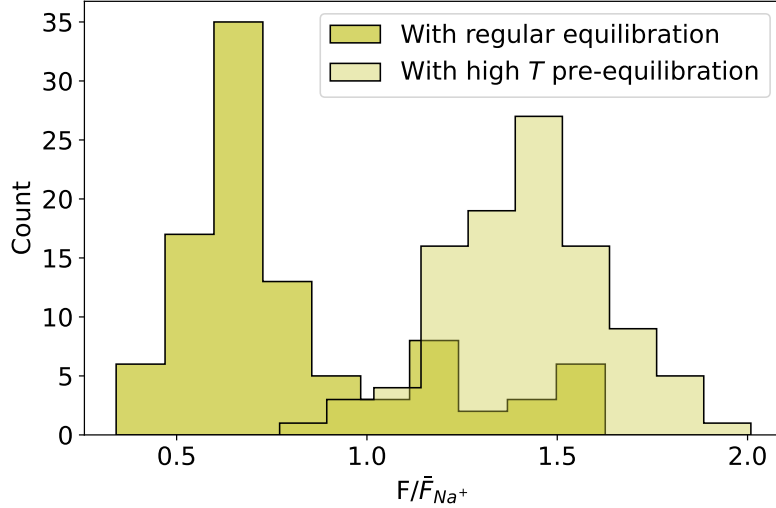
Figure 3.10: Comparison for force distribution for Mg^{2+} from straightforward and high temperature pre-equilibration procedure. The straightforward sampling method is the same as shown in Figure 3.8K.

keep the DNA tightly bound to the surface. However, our results show that Mg^{2+} gives rise to higher mobility than monovalent ions like Na^{+} (Table 3.1). The reason is again due to the strong hydration layer for Mg^{2+}, due to which the DNA is not in direct contact with the ionic layer allowing for higher surface mobility. Such an indirect contact is also a result of the much smaller timescale of our simulations compared to the water exchange rate from the first hydration shell of Mg^{2+}. Even though for Mg^{2+} we used a recent force field that accelerates the water exchange process, for our setup, it still remained challenging to observe exchanges on a few 100 ns timescale. On the other hand, in experiments, due to larger timescale DNA can make direct contact with adsorbed Mg^{2+} layer resulting in a tightly bound DNA and possibly the high rupture force that we observed. For all other ions, the mobility is simply governed by the affinity of the cation to the DNA phosphate oxygen, with Cs^{+} having the highest mobility due to its lowest affinity to the DNA phosphate oxygen.

Finally, the interaction between DNA and mica is commonly explained using models based on ion correlation theory. Such models predict repulsion between DNA and mica in the presence of only monovalent ions. Our results show that the DNA-mica interaction involves an interplay of ion affinity towards DNA, mica, and its hydration properties. The existing theoretical models do not account for all these effects and hence are unable to explain our experiment and simulation results.

Based on our results, we could speculate on the optimal ionic conditions for adsorbing DNA on mica surfaces in physiologically relevant conditions. Among the monovalent ions, K^{+} gives a fine balance between better surface coverage of the mica and affinity towards DNA. Other ions either exhibit low affinity towards DNA e.g. Cs^{+} or have low coverage on the mica surface (Li^{+}orNa^{+}) which could hinder spontaneous adsorption of DNA. Therefore, a mixture of K^{+} and divalent ions like Mg^{2+}or Ca^{2+} would be the best choice as was also observed previously [25]. The monovalent ions facilitate spontaneous adsorption

and the divalent ions form stable contacts of DNA and mica thereby enhancing overall adsorption.

3.5   Conclusions

In this chapter, we combine molecular dynamics simulations and single-molecule atomic force microscopy experiments to characterize the detachment forces of single-stranded DNA from mica surface mediated by the metal cations $Li^+, Na^+, K^+, Cs^+, Mg^{2+}$ and Ca^{2+}. Cations adsorb to the mica surface to compensate or overcompensate the bare negative surface charge. Each cation exhibits a unique affinity towards the mica surface and modifies the interfacial structure in distinct ways. DNA adsorption at the mica surface is determined by cation affinity towards mica, DNA phosphate oxygen, and hydration properties. Therefore, DNA adsorption at mica results from a complex interplay of ion-DNA, ion-surface, and hydration effects. Consequently, the ion-specificity of average detachment forces is not discernible due to the convoluted contribution of these various effects. Weakly hydrated ions and water-mediated contacts lead to low detachment forces and high mobility of the DNA on the surface. Direct ion-DNA or ion-surface contacts lead to significantly higher forces. Our results could be used to choose optimal ionic conditions for imaging purposes or for fine-tuning the properties of DNA nanostructures on mica surfaces.

3.6   Appendix

3.6.1   Data archive

The simulation input file for all the simulations presented in this chapter is archived in the following tree. Go to any of the directories e.g. `CIO` and run the command

```
gmx grompp -f ../mdp/md-pr.mdp -c npt.gro -p dna-mica.top -n index.ndx \
-o npt.tpr
```

```
$HOME/Data-Archive-Mohd_Ibrahim/ssDNA-Mica/
├── CIO
│   └── dna-mica.top   index.ndx   npt.gro
├── Ca
│   └── dna-mica.top   index.ndx   npt.gro
├── KIO
│   └── dna-mica.top   index.ndx   npt.gro
├── LIO
│   └── dna-mica.top   index.ndx   npt.gro
├── Mg
│   └── dna-mica.top   index.ndx   npt.gro
├── NIO
│   └── dna-mica.top   index.ndx   npt.gro
├── create-mica-crystal
├── forcefield
│   ├── amber99bsc1.ff
│   ├── clay.ff
│   ├── cruz_ca.ff
│   ├── grotz_mg.ff
│   ├── nschwierz.ff
│   └── itp
│
└── mdp
```

3.6.2 Distance of each DNA residue from the mica surface during equilibration
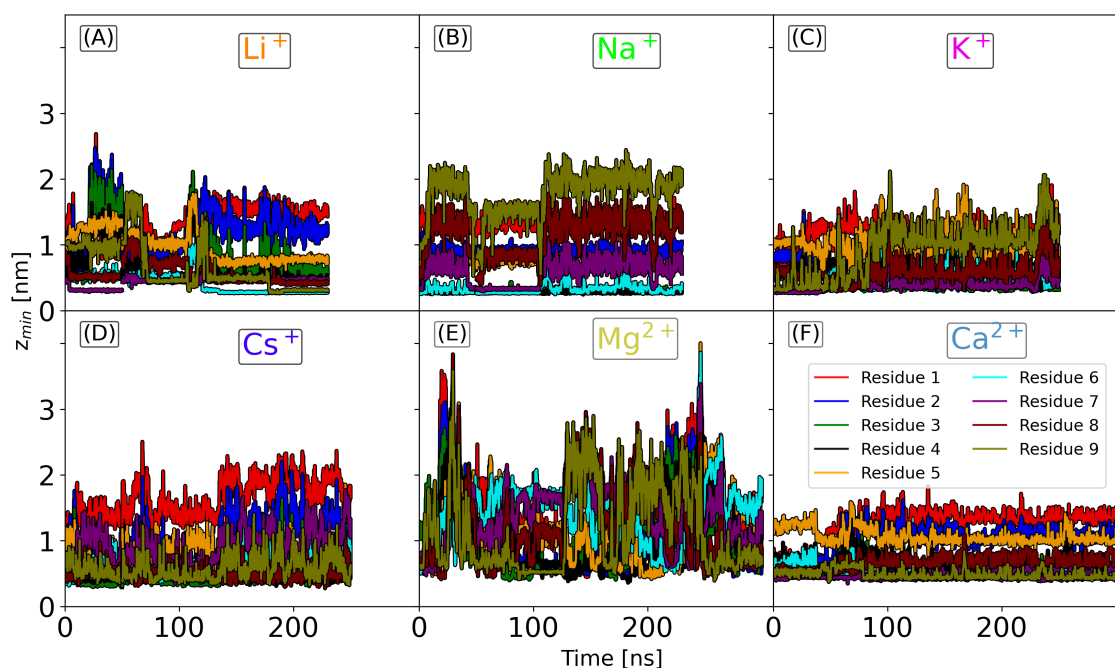
Figure A3.1: Equilibration after forced adsorption with $F_{ext} = 470$ pN. The plots show the evolution of the minimum distance of the DNA phosphate group of each residue from the mica surface. Configurations equally spaced in time after 100 ns or 150 ns are selected for pulling.

3.6.3 Force-extension profiles from simulations at low pulling speed (0.1 m/s)
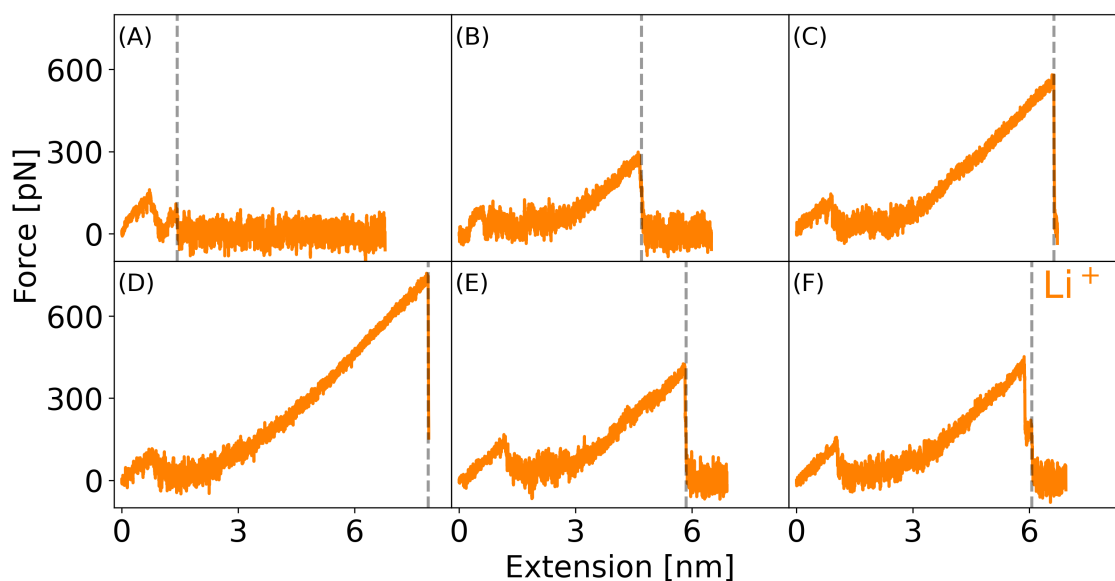
Figure A3.2: Force-extension profiles from simulations for Li^+ at a low simulation pulling speed of 0.1 m/s. The dashed vertical lines show the point where the DNA fully detaches from the mica surface.
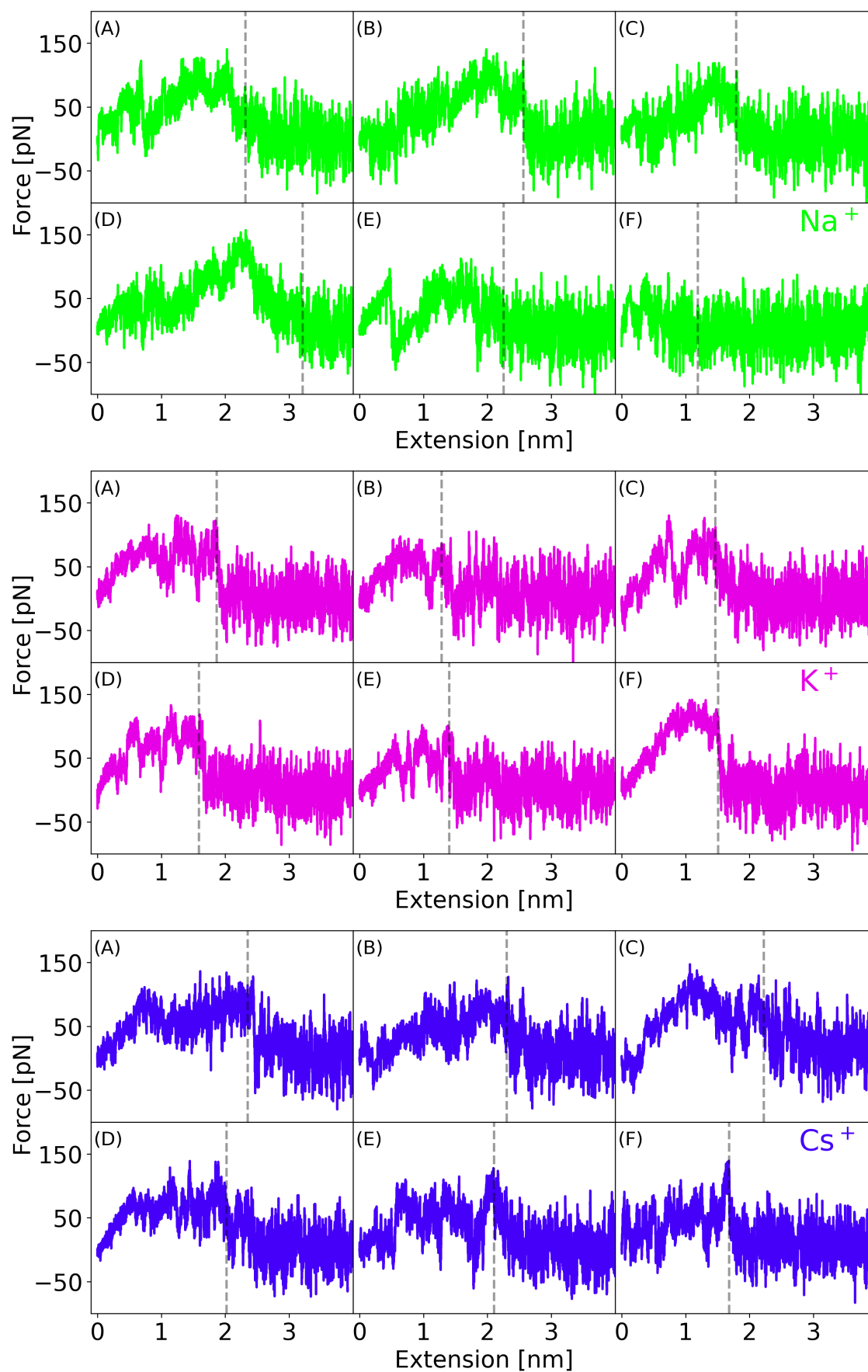
Figure A3.3: Force-extension profiles from simulations for Na$^+$ (green), K$^+$ (magenta), and Cs$^+$ (violet) at a low simulation pulling speed of 0.1 m/s. The dashed vertical lines show the point where the DNA fully detaches from the mica surface.
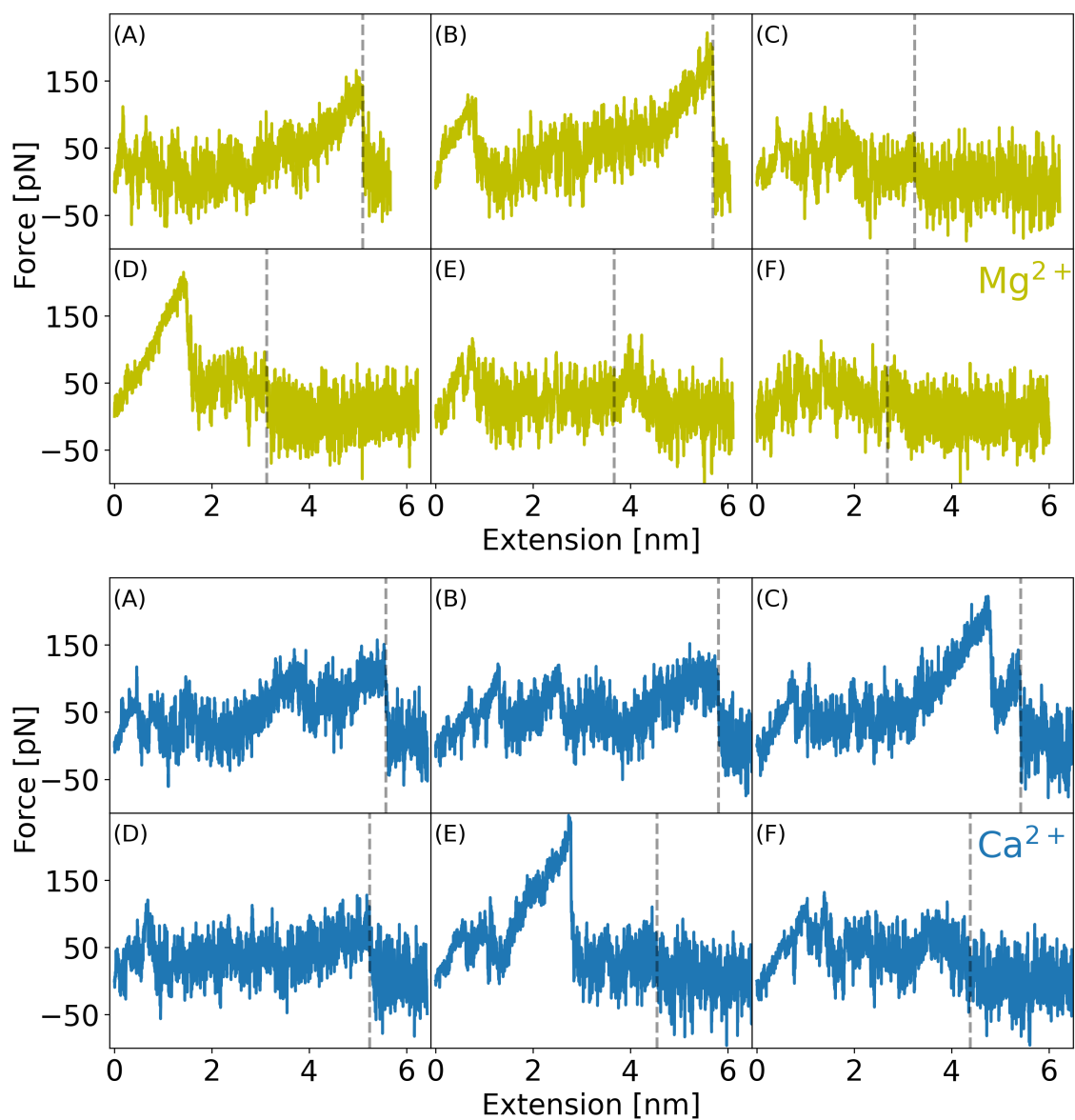
Figure A3.4: Force-extension profiles from simulations for Mg^{2+} (yellow) and Ca^{2+} (blue) at a low simulation pulling speed of 0.1 m/s. The dashed vertical lines show the point where the DNA fully detaches from the mica surface.

3.6.4 Analysis of the type of contacts that break at rupture

The analysis of interactions that break at rupture (main text Figure 3.9) is done using a custom Python script based on MDAnalysis [129]. The script does the following:

1. Identify the detachment point in the force-extension profile. The detachment point is detected from the minimum distance of DNA from mica shown by the gray curve in Figure 3.4. At detachment, the minimum distance increases sharply as shown by the black circular marker in Figure 3.4 A-B.

2. Obtain the frames of interest (FOI) in the simulation trajectory that is within $\sim$1 nm on either side of the detachment point in the force-extension curve.

3. For the FOI, three cation lists are created: (i) L1: those within a cut-off distance (r_1) of the non-bridging oxygen of the phosphate backbone, i.e., O1P and O2P, (ii) L2: Those within a cut-off distance (r_2) of the DNA nucleobase oxygens, i.e., O2 and O4, (iii) L3: Those within a cut-off distance (r_3) of mica surface oxygen atoms. The pair lists are created using the MDAnalysis `capped_distance` function. The cut-off distances (r_1, r_2, r_3) are determined from the radial distribution function (RDF) of the cation and the water oxygen and are shown as vertical dashed lines in Figure A3.5. For all ions except Mg^{2+} $r_1 = r_2 = r_3$. For Mg^{2+}, since all the surface-bound Mg^{2+} is in hydrated state, r_3 is set by the second RDF maximum (Figure A3.5 E), while $r_1 = r_2$ is the same as for other ions.

4. The number of shared cations between the phosphate oxygens and the mica surface (N_p) is given by the intersubsection of L1 and L3, while the number of shared cations between the DNA nucleobase oxygen and the mica surface (N_n) is given by the intersubsection of L2 and L3.

5. If $N_p = N_n = 0$: There are no shared cations at any point during the detachment process. Therefore, the rupture force is assumed to arise from indirect water-mediated contacts.

6. If $N_p \neq 0$ or $N_n \neq 0$: We look at the last FOI, where the DNA is fully detached from the mica surface. We count the number of initially shared cations that are still bound to the mica surface. For n_p out of N_p and n_n out of N_n cations bound to the surface, it follows:
   (1) If $n_p = N_p$: All shared cations are bound to the mica surface and the rupture force is attributed to the breakage of the ion-phosphate (IP) type of contact.
   (2) If $n_p = 0$: All shared cations are desorbed from the mica surface. Therefore, the rupture force is attributed to breakage of the ion-surface (IS) type of contact.
   (3) If $n_p \neq 0$ and $n_p < N_p$: Some shared cations are bound to the mica surface and some are bound to the DNA phosphate. Therefore, the rupture force arises from the simultaneous breakage of both ion-phosphate (IP) and ion-surface (IS) types of contacts. In almost all cases for all types of ions, we find that N_n is close to zero, i.e., the nucleobase is not involved in the interaction. Therefore, the classification for ion-nucleobase is omitted.

Visual inspection (in VMD) of random trajectories (> 10) accurately matched the results from the script given above.
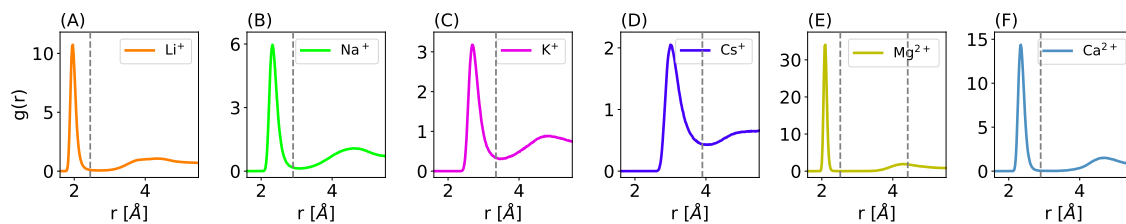
Figure A3.5: The radial distribution function (RDF), $g(r)$, between water oxygen and the cation is used to obtain the cut-off for interactions. The vertical dashed lines represent the cut-off for each cation type. The RDF provides insight into the stability of the first hydration shell. The height of the first peak reflects the ion-water interaction strength and follows the order $Cs^+ < K^+ < Na^+ < Li^+ < Ca^{2+} < Mg^{2+}$ in agreement with the experimentally measured hydration free energies [138]. The position of the first peak corresponds to the radius of the hydration shell.

Probing RNA Conformation on Bilayer Surfaces from Scattering Techniques

The work presented in this chapter had been conducted under the supervision of Prof. Dr. Nadine Schwierz at the Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Frankfurt am Main, Germany, and at the Institute of Physics, University of Augsburg, Augsburg, Germany. The simulations were performed on the GOETHE HLR of the state of Hessen.

4.1   Introduction

The secondary structure of mRNA cargo is an important factor for the efficient design of lipid nanoparticles for drug delivery as it's an important determinant of the protein translation properties of the mRNA [139–142]. Factors such as ionic conditions, pH levels, and interacting species have a substantial impact on RNA conformation, particularly those with charged properties [33, 143]. These environmental factors can either stabilize or disrupt secondary structures thus affecting the translation efficiency and also the interaction of RNA with other interacting species [144]. Due to the negative charge of the RNA, the secondary structure and interaction are particularly more sensitive to electrostatic environments or the charge of the interacting species [33]. For LNPs used for drug delivery, almost half the lipid composition is positively charged at low pH. Therefore, for such systems understanding the influence of membrane charge on the RNA secondary structure (or conformation) is important for efficient drug design. Previous studies on tRNA in POPC/DOTAP liposomes have shown that the secondary structure of tRNA molecules is influenced in the presence of membranes due to electrostatic and hydrophobic interaction with the membrane [32, 33]. Scattering techniques, especially reflectivity and small angle scattering present an invaluable tool to gain insights into the interaction as well as structure of RNA/bilayer systems and the possibility to resolve RNA conformation on bilayers using such techniques has not been explored. Reflectivity techniques, in particular, can probe structures at bilayer interfaces and for RNA-lipid complexes, such techniques could provide information about the adsorption behavior as well as the RNA conformation on the bilayer surface.

In this chapter, we employed the coarse-grained Martini model to study the interaction of RNA with neutral and charged model bilayer systems to explore the impact of membrane composition and RNA base-pairing on the adsorption and dynamics of RNA on lipid bilayers. We find that on neutral bilayers, single-stranded RNA adsorbs more strongly than double-stranded RNA, and the interaction is driven by hydrophobic attraction between exposed RNA bases and hydrophobic parts of the bilayer. On charged bilayer, the

interaction is driven by electrostatic interactions. Subsequently, we explored the possibility of resolving RNA conformation on a bilayer using small-angle neutron scattering form factors and neutron reflectivity techniques. We conclude that due to inherent noise in such measurements neutron reflectivity and small-angle neutron scattering (SANS) can not differentiate between different RNA conformations, however, selective deuteration of RNA in the case of SANS shows clear differences between different RNA conformations.

4.2 Methods

System setup

We simulated DOPC/DOTAP lipid bilayer and their mixtures. The initial setup was prepared using the `insane.py` script [145]. The lipids were described with the coarse-grained (CG) martini model (version-2.1) and RNA was described using the corresponding RNA force field [146, 147]. The systems were solvated with CG water, and to avoid freezing (a common artifact of CG water), 10% anti-freezing water was added. The systems with single-stranded or double-stranded RNA contained 100 lipids per monolayer, while the system with tRNA contained 324 lipids per monolayer.

To create a coarse-grained RNA structure, a double-stranded atomistic structure was first created using the `fd_helix.c` script[1], which was then coarse-grained using the `martinize-nucleotide.py` script. Single-stranded RNA was created by removing the complementary sequence from a double-stranded structure. The RNA so created was placed at a distance (>2nm) from the bilayer surface.

The simulations were done using the Gromacs (v-2018) package [121]. The Van der Waals and short-ranged electrostatic interactions were truncated at 1.1 nm. The long-range electrostatic interaction was evaluated using a reaction-field method with the relative dielectric constant of water set to 15. A time step of 10 fs was used to integrate the equations of motion.

tRNA secondary structure and 3D-structure

We used the tRNA molecule with sequence `GCGGAUUUAGCUCAGUUGGGAGAGCGCCAGACUGAAGA UCUGGAGGUCCUGUGUUCGAUCCACAGAAUUCGCACCA`. The secondary structure of the sequence was predicted using the RNAfold package [148] and has a minimum free energy structure (MFE) of -22.40 kCal/mol. Ten sub-optimal secondary structures were predicted that lie within 18 kCal/mol of the MFE structure and are separated by at least 1.5 kCal/mol in energy. The RNAComposer webserver [2] was used to obtain an atomistic 3D structure for each secondary structure [149], which was then coarse-grained using the `martinize_nucleotide.py` script.

---

[1]https://casegroup.rutgers.edu/fd_helix.c
[2]http://rnacomposer.ibch.poznan.pl/;jsessionid=7DC7B05B04585826145EE5E924726

Noise model for reflectivity [3] and form factors

The total intensity measured in the experiments, $I_{meas}(q)$, contains contribution due to background scattering. Subtracting the background contribution, the reflected intensity from the sample is given by,

$$I_{refl}(q) = I_{meas.}(q) - I_{bkg}(q) \tag{4.1}$$

Assuming we've perfectly subtracted the background signal, the incident and reflected neutron beams are subject to Poisson noise. Therefore the variance in reflected and incident intensity is given by,

$$var(I_0) = I_0, \quad var(I_{refl}(q)) = I_{refl}(q) \tag{4.2}$$

Consequently, the error model of the reflectivity is given by the ratio distribution of two Poisson random variables. For large enough mean values, these Poisson distributions can be approximated by Gaussians. Under certain conditions, the ratio distribution $Z = \frac{X}{Y}$ of two Gaussian distributed random variables X and Y, is again a Gaussian distribution with mean $\mu_z = \frac{\mu_x}{\mu_y}$ and variance, σ_z given by [150],

$$\sigma_z^2 = \frac{\mu_x^2}{\mu_y^2}\left(\frac{\sigma_x^2}{\mu_x^2} + \frac{\sigma_y^2}{\mu_y^2}\right) \tag{4.3}$$

Using the variances for the incident and reflected beam from eqn. 4.2 in eqn. 4.3, we get, the variance in $R(q)$, as

$$var(R(q)) = \frac{R(q)}{I_0} + \frac{R(q)^2}{I_0} \tag{4.4}$$

In order to check the validity of eqn. 4.4 we compare it with the numerical results obtained using explicit Poisson distributed intensities, for the reflected and incident neutron beams, with mean and variance given by eqn.4.2. Figure 4.1 shows that our noise model matches well with the numerical results for $I_0 \geq 10^2$, which is well within the experimental noise range. Therefore, the noise model is valid for values typical of experiments.

To simulate noise for scattering form factor profiles, we first calculate the scattering intensities using the relation, $I(q) = F(q)^2/q^2$. Noise is then added to the scattering intensities by random drawing from a Gaussian distribution with variance given by, [151],

$$var(I(q)) = \frac{I(q) + const.}{kq} \tag{4.5}$$

Where, we set $k = 0.1$, $const. = 0$, and q is the momentum transfer in the z-direction. The

---

[3] The noise model for reflectivity was derived by Dr. Jürgen Köfinger

variance in form factor is then given by,

$$\sigma_F^2(q) = \frac{q^2 \sigma_I(q)^2}{4I(q)} \tag{4.6}$$
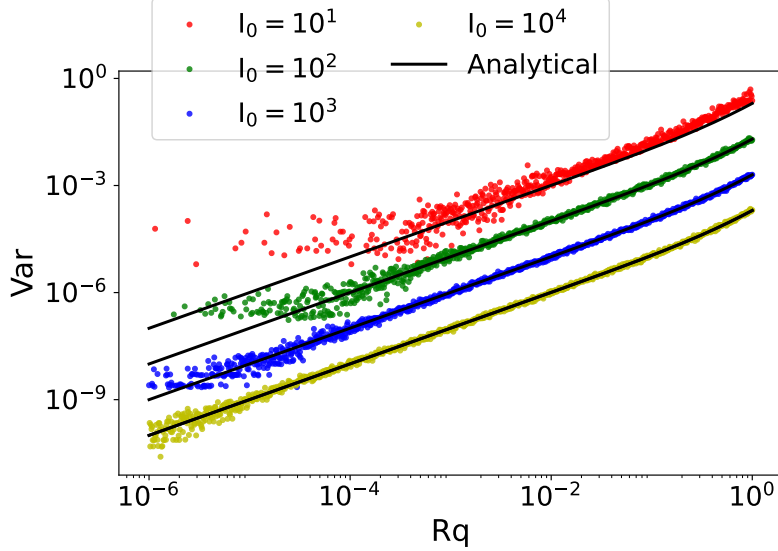
Figure 4.1: Noise model for reflectivity Comparison of analytical variances (Eqn. 4.4) with variances generated assuming Poisson distribution.

Scattering length density (SLD) profile from coarse-grained Martini simulations

As mentioned in the previous chapter (section 2.2.2), the SLD along the bilayer normal from simulations is obtained by dividing the box into small slices as,

$$SLD(z) = \frac{\sum_i n_i(z) b_i}{dV} \tag{4.7}$$

Where n_i is the number of atoms with scattering factor b_i in a slice with volume dV. The b_i for atoms are well defined (Table 2.1). However, for the CG-martini bead it was not reported, we used two equivalent methods to assign obtained SLD from CG simulations:

(i) Backmapping: In this method few thousand frames were back-mapped to all-atom using the `backward.py` script [152]. The back-mapped frames were then read by a custom script to calculate SLD according to Eqn. 4.7 using reported b_i for atoms. This method is tedious, takes up a lot of storage space and SLD calculation is slower.

(ii) Direct calculation from CG simulation: Here we assign b_i for CG-beads and use Eqn. 4.7 for SLD calculations. The b_i for beads is assigned using the mapping file provided with the force field for each molecule. An example of part of a mapping for DOPC is shown in Table 4.1, where the contribution of each atom to a CG-bead is specified. Based on the above mapping file (Table 4.1), the b_i for CG-bead $NC3$ is assigned as,

$$b_i(NC3) = b_i(N) + \frac{3b_i(C)}{4} + 3b_i(C) + \frac{3b_i(H)}{4} + \frac{3b_i(H)}{4} + ... \tag{4.8}$$

Index	Atoms	CG-Beads
1	N	NC3
2	C12	NC3 NC3 NC3 PO4
3	C13	NC3
4	C14	NC3
5	C15	NC3
6	H12A	NC3 NC3 NC3 PO4
7	H12B	NC3 NC3 NC3 PO4
..	...	...

Table 4.1: An example martini mapping file: here a part of DOPC mapping file is shown.
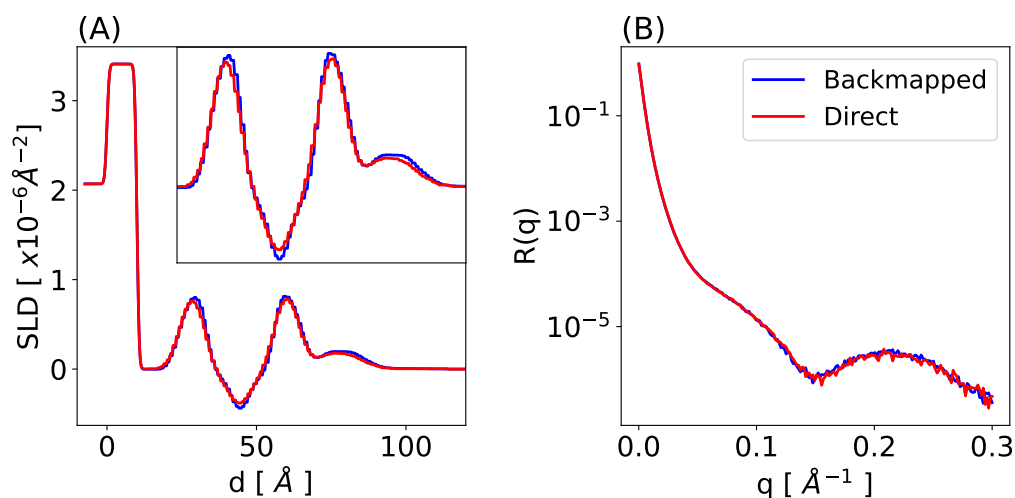
Figure 4.2: (A) SLD from back mapped and direct calculation methods. The zoomed view in the inset shows differences between the two methods (B) The corresponding reflectivity profiles are practically identical. Here 8% solvent deuteration is assumed and the system consists of DOPC, tRNA, Water, and ions. At higher deuteration, the differences become even smaller.

Where $b_i(X)$ is the scattering factor of atom X. This method is highly efficient and the absence of back mapping saves a lot of disk space. However, the scattering factor of atoms constituting a bead is lumped at the center of the bead, hence the detailed distribution of nuclei that maps in a given CG-bead is lost. A typical SLD profile obtained at 8% solvent deuteration level from the two methods is depicted in Figure 4.2 A. The subtle difference between the SLD is lost in the scattering curve (Figure 4.2 B) and both methods give the same result.

4.3 Results

4.3.1 Single and double-stranded RNA on neutral and charged bilayers

Interaction of single and double-stranded RNA with charged and neutral lipid bilayers

We looked at the influence of membrane composition and RNA base-pairing on RNA membrane interaction using the coarse-grained Martini model. The bilayer membrane is
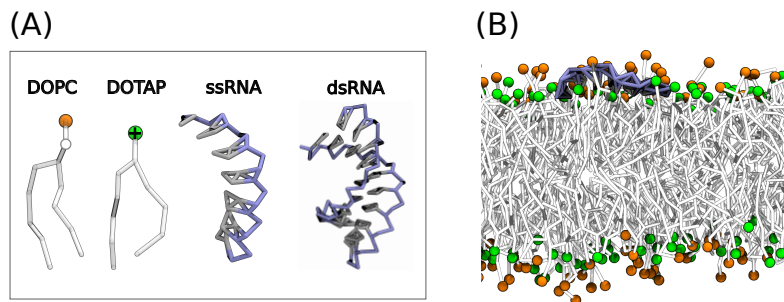
Figure 4.3: (A) Lipid and RNA molecules simulated. DOTAP is positively charged. For RNA gray color corresponds to the nucleobases and blue corresponds to the phosphate backbone. (B) A simulation setup with a mixture of DOPC and DOTAP and as single-stranded RNA.

composed of a zwitterionic lipid DOPC and a cationic lipid DOTAP (Figure 4.3 A). We simulated membranes composed of only DOPC lipids, only DOTAP lipids, and mixtures of DOPC and DOTAP. To understand the effect of RNA base pairing on the RNA membrane interaction both single-stranded (ssRNA) and double-stranded RNA (dsRNA) RNA were used (Figure 4.3). The ssRNA was composed of eight Uracil nucleotides (poly-U) and the dsRNA was composed of the sequence `UCUUCUACUU` and its complementary sequence. The CG 3D structure of each of these components is depicted in Figure 4.3A. Each simulation was 10 μs long.
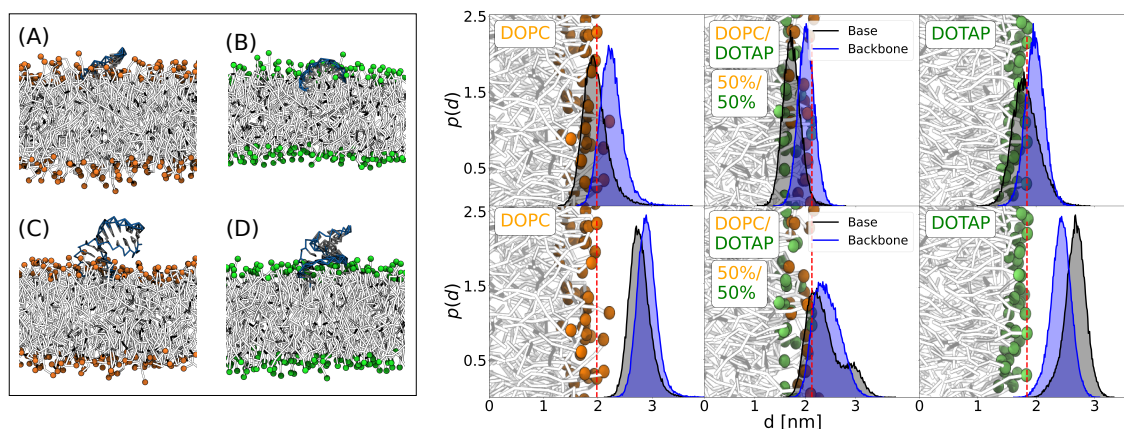
Figure 4.4: (A) Single-stranded RNA on pure DOPC surface (B) and on pure DOTAP surface (C) Double-stranded RNA on pure DOPC surface (D) on pure DOTAP surface. These snapshots are after 10 μs of simulation. Right: (Top row) probability distribution of the distance p(d) of ssRNA bases (black) or ssRNA backbone (blue) from the lipid membrane center for different membrane compositions. The red dashed vertical line shows the average position of the lipid membrane head group. (Bottom row): Probability distribution for dsRNA for different membrane compositions.

The final snapshots demonstrate that both ssRNA and dsRNA adsorb to the bilayer surface regardless of the charge of the bilayer (Figure 4.4 (A-D)). However, the compositions of the bilayer and RNA base pairing influence the extent of RNA penetration into the bilayer surface, with single-stranded RNA exhibiting more effective interaction with the bilayer. To gain further insights into the adsorption mechanism, we calculated the probability distribution ($p(d)$) of base and backbone positions relative to the bilayer center. For ssRNA,
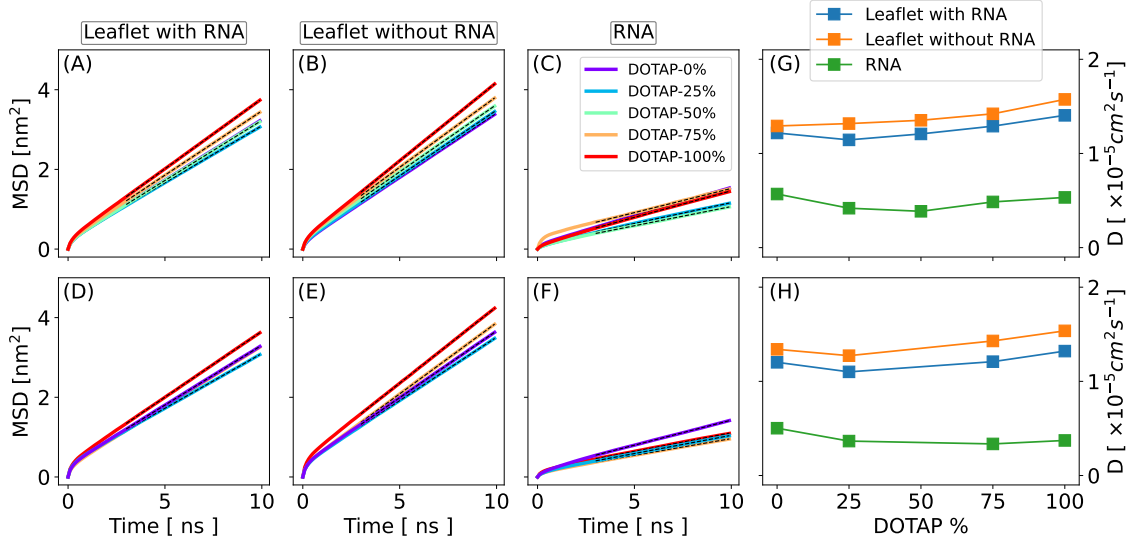
Figure 4.5: (A-F) Mean square displacement of different components. The black dashed line corresponds to the linear fit using eqn. 4.9. (G-H) Corresponding diffusion coefficients

the peak in $p(d)$ for bases consistently lies closer to the bilayer center compared to the backbone, indicating a preferential orientation of RNA on the bilayer surface. Specifically, ssRNA assumes an orientation where bases interact with the hydrophobic regions of the bilayer while the backbone interacts with the polar head groups. As the overall bilayer charge increases (i.e., with an increasing DOTAP fraction), this preferential orientation becomes less prominent, as evident from the lower peak of $p(d)$ for bases compared to the backbone in the pure DOTAP setup (Figure 4.4). On the other hand, dsRNA exhibits weak interaction with the neutral, pure DOPC bilayer, with the peak position of $p(d)$ appearing at relatively larger distances from the bilayer center, suggesting that dsRNA adsorbs to the bilayer surface in a more upright position (Figure 4.4 C). With an increasing DOTAP fraction, the interaction is driven by electrostatic attraction between the negatively charged RNA backbone and the bilayer charge, as seen from the appearance of backbone $p(d)$ peaks closer to the bilayer center than the bases.

Dynamics of RNA on bilayer surface

The dynamics of lipids and RNA can be characterized by the diffusion coefficient (D) which can also be measured in experiments (e.g., Fluorescence microscopy). The diffusion coefficient in two dimensions (in the bilayer plane) is estimated from the mean square displacement (MSD) as,

$$MSD(t) = 4Dt \tag{4.9}$$

Therefore, D can be estimated by fitting the MSD profile in the diffusive regime (Figure 4.5 (A-F)).

We calculated the diffusion coefficients for the center of mass of RNA and the lipids in each leaflet (Figure 4.5 G, H). In general, the lipids in the leaflet with RNA have slightly slower dynamics compared to the leaflet without the RNA and this becomes more
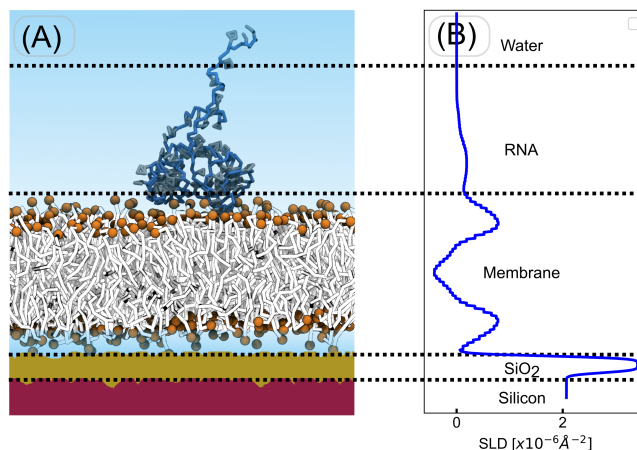
Figure 4.6: System snapshot for neutron reflectivity, showing an RNA molecule on the membrane surface and the contributions to the SLD from different components including the silicon substrate. SLD profile is calculated by considering 8% deuteration level.

differentiable with increasing DOTAP fraction indicating the strong electrostatic attraction between RNA and DOTAP. Such interactions lead to slow dynamics of lipids in contact with RNA giving rise to an overall smaller diffusion coefficient. The overall lipid dynamics also becomes faster with increasing DOTAP fraction which again arises due to repulsion among positively charged DOTAP. Finally, the dynamics of RNA is significantly slower than the lipids and it becomes more prominent with increasing DOTAP fraction. The similar diffusion coefficient of ssRNA and dsRNA indicates that the heavier mass of RNA is not responsible for slower RNA dynamics compared to lipids. A plausible cause could be some dynamical constraints imposed on the RNA due to the specific nature of the surface, arising from e.g., a relatively large time scale of unbinding for RNA-lipid interactions.

4.3.2 Probing RNA conformation on bilayer surface

We use the CG-martini model of RNA and membrane to explore the possibility of resolving the RNA conformation on the bilayer surface using neutron reflectivity and small angle x-ray scattering (SAXS). The snapshot of the typical system with a tRNA on the bilayer surface is shown in Figure 4.6. The silicon substrate is theoretically modeled into the SLD and is not part of the simulations. Since the reflectivity or SAXS form factor is fully determined from the SLD or electron density profiles, it is important to first validate that the CG model reproduces the correct SLD/electron density profiles.

Comparison of all-atom and CG-martini form factors

To validate the structure produced by CG simulations we compared the electron density of a pure DOPC bilayer with all-atom simulations. The CG model correctly reproduced the approximate shape of the electron density and also the peak-to-peak distance which represents the bilayer thickness (Figure 4.7 A). Compared to all-atom, CG electron density achieves bulk value at a larger distance. This arises from an increase in water density near the bilayer surface which is a consequence of low water penetration in the CG-bilayer due

Figure 4.7: (A) Electron density from all-atom and CG-martini simulations for pure DOPC bilayer (B) The x-ray form factors obtained from Fourier transformation of corresponding electron densities. The experimental data is the same as used to validate Lipid14 FF [66].

to the larger size of CG beads, this is a known issue for the CG-Martini model [153]. The form factor obtained from these electron densities compares well with the experiments (Figure 4.7 B)

Different tRNA conformations on neutral and charged membrane

The timescale of RNA folding using straightforward molecular dynamics simulation is beyond the reach of current computational power. Therefore, to simulate different RNA secondary structures we created the 3D conformation, corresponding to a given secondary structure, which stays intact throughout the simulation. We simulated 20 different setups corresponding to 10 tRNA conformations and two bilayer compositions composed of pure DOPC and pure DOTAP. Each simulation was run for 3 μs. The tRNA secondary and tertiary structures are shown in Figure 4.8.

The interaction between the tRNA molecule and the bilayer varies depending on its conformation and the bilayer composition. This is clearly visible in the average SLD profiles (Figure 4.9 A-B). The contribution of tRNA to the SLD profile appears as a bump relative to the bilayer with no RNA. The tRNA10 conformation on DOPC displays a relatively sharp bump between 25 Å to 50 Å whereas for tRNA7, the corresponding bump is broadly spread out between 25 Å to 75 Å (Figure 4.9 A). This indicates that tRNA10 interacts more strongly with the DOPC bilayer as compared to tRNA7. On DOTAP, the whole SLD profile with tRNA appears distorted compared to the pure DOTAP profile (Figure 4.9B). The peaks in the SLD, which correspond to the DOTAP head groups, become asymmetric. Additionally, there is a slight increase in SLD at the center of the bilayer (d=0 Å), and the SLD of the bilayer leaflet facing away from the tRNA becomes broader and smaller. These observations can be attributed to the strong electrostatic interaction between the positively charged bilayer and the negatively charged tRNA. Such strong interactions lead to the distortion of the bilayer upon tRNA binding, as depicted in the snapshots in Figure 4.9 (tRNA7/DOTAP).
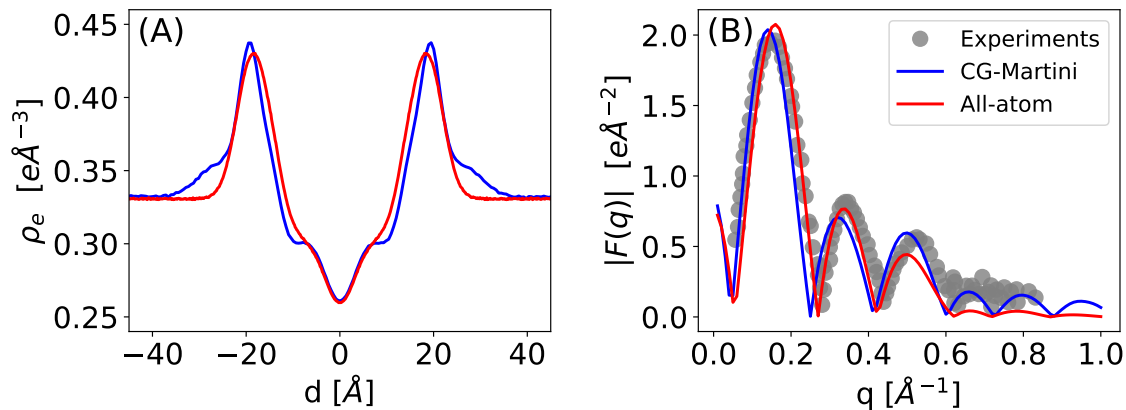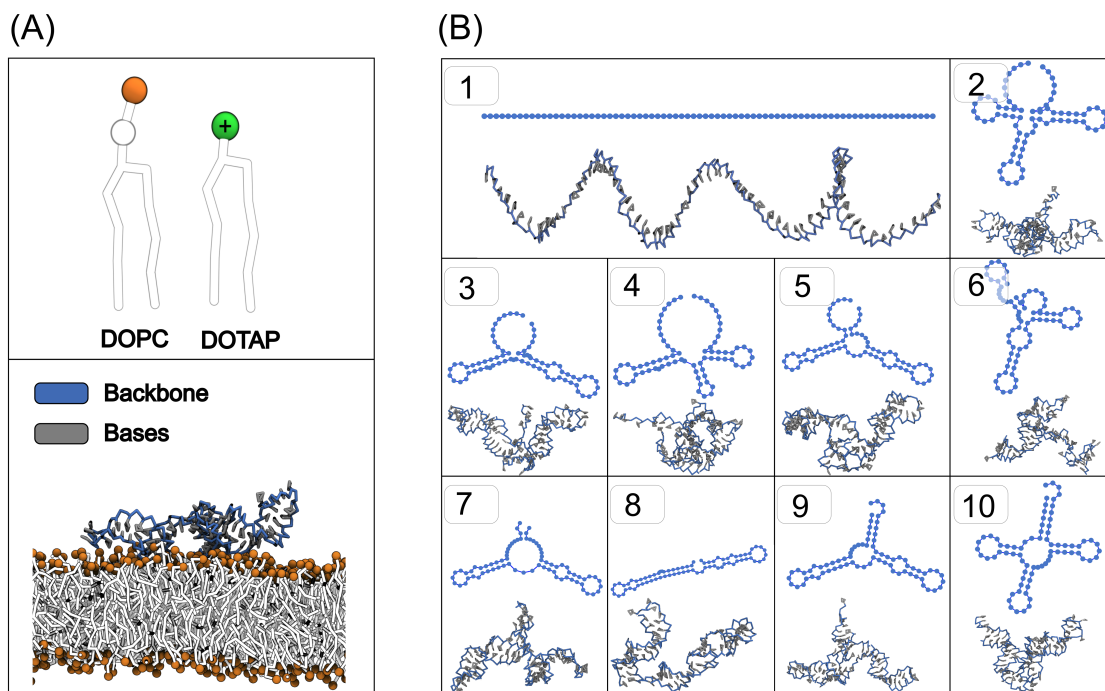
Figure 4.8: (A) Electron density from all-atom and CG-martini simulations for pure DOPC bilayer (B) The x-ray form factors obtained from Fourier transformation of corresponding electron densities.

Reflectivity from tRNA bilayer setups

We now turn to the direct calculation of neutron reflectivity profiles to see if it can resolve the differences, between different RNA conformations for different membrane compositions, observed above. The SLD profiles and corresponding reflectivity profiles at 8.0 % deuteration level are shown in Figure 4.10 (A-D). At this deuteration level, the solvent contribution to the SLD becomes zero as indicated by zero SLD in the bulk. The reflectivity profiles show that it's challenging to even distinguish between systems with and without RNA, therefore differences between secondary structures are even more subtle and can not be discriminated by neutron reflectivity. The neutron beam passes through the silicon substrate before interacting with the membrane, resulting in a significant amount of neutron beam being reflected at the silicon-SiO_2 and SiO_2-bilayer interfaces before reaching the sample. The reflectivity profiles without considering the substrate are able to detect the presence of RNA (Figure 4.10 F, H). However, the subtle differences between secondary structures are still obscured by the inherent Poisson noise that accompanies such experiments. Further, the differences between reflectivity signals from dissimilar secondary structures were not discernible at any deuteration level between 0% to 100%. A powerful feature of neutron reflectivity is the possibility to enhance the signal from part of the system by selectively deuterating it. Deuterating tRNA increases its SLD (Figure 4.11 A), however, it does not result in a significant difference in the measured reflectivity profile (Figure 4.11 B). Therefore, we conclude that at this tRNA concentration (1 tRNA per 324 lipids) neutron reflectivity can not differentiate the subtle differences between different RNA secondary structures.
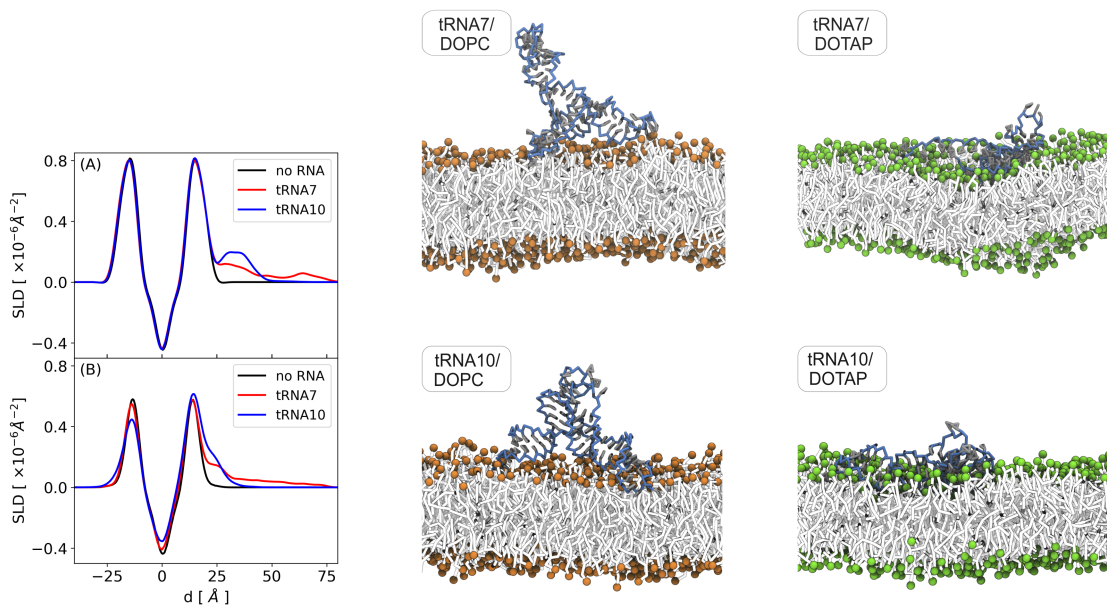
Figure 4.9: (A) SLD profile for pure DOPC and DOPC with two different tRNA conformations. (B) SLD profile for pure DOTAP and DOTAP with two different tRNA conformations. The solvent is ignored in these calculations for clarity. The corresponding snapshots after 3 μs.
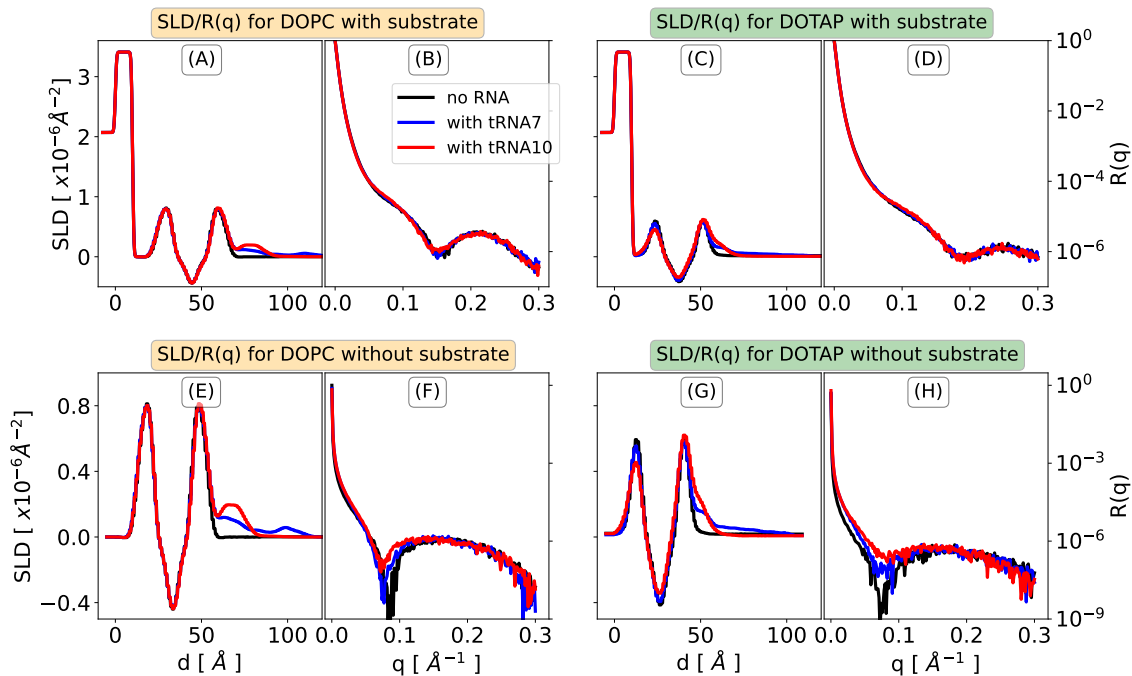
Figure 4.10: (A-B) SLD and reflectivity profile from pure bilayer and bilayer with two different tRNA conformations. Here bilayer is composed of pure DOPC (C-D) Same profiles with DOTAP as the bilayer composition. (E-F) SLD and reflectivity profile in the absence of silicon support for DOPC/tRNA setup (G-H) same for DOTAP-composed setups.
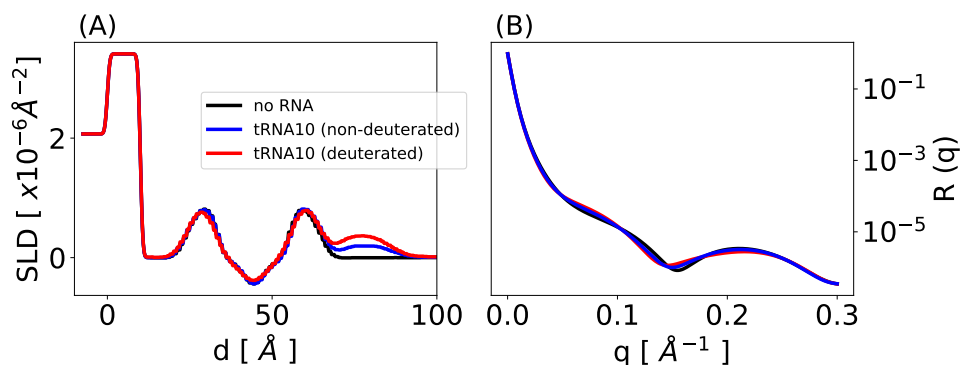
Figure 4.11: (A) SLD profile for pure DOPC and DOPC with two different tRNA conformations. (B) SLD profile for pure DOTAP and DOTAP with two different tRNA conformations. The solvent is ignored in these calculations for clarity. The corresponding snapshots after 3 μs.
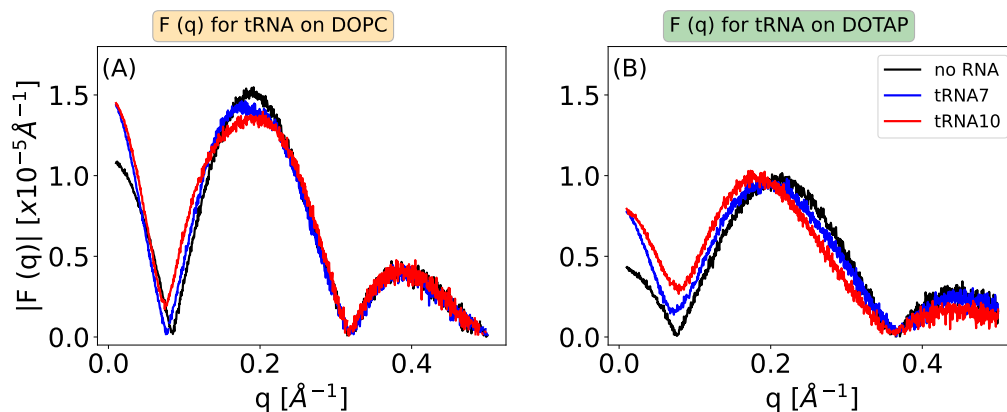
Figure 4.12: SANS form factor for pure bilayer (no RNA) and bilayer with different tRNA conformations. (A) for tRNA/DOPC systems (B) for tRNA/DOTAP systems.

Small-angle neutron scattering (SANS) form factors for tRNA bilayer setups

Figure 4.13 (A-B) shows SANS form factors for the previous setup at the same solvent deuteration level (8.0%). Compared to reflectivity, SANS profiles show discernible differences between pure bilayer and bilayer-tRNA setups especially at low q. For the system with DOTAP, differences between different secondary structures are also visible since in this case, the tRNA has more drastic effect on the bilayer structure. However, for the DOPC setup, such differences are not very prominent. For such cases, the selective deuteration of the tRNA enables clear differentiation of the secondary structures from the SANS form factor (Figure 4.13B). This differentiation is evident through the broadening of the first Bragg peak, and distinct differences between various secondary structures are clearly noticeable in the low q regime.

4.4 Discussion

Being able to predict the most probable RNA conformation in different environments has great significance in RNA-based drug delivery applications using lipid nanoparticles since
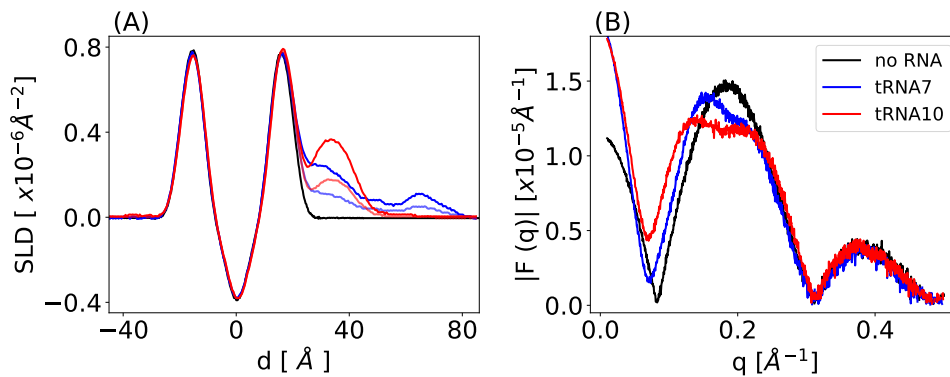
Figure 4.13: Effect of selective deuteration of tRNA on SANS form factor. (A) SLD profile for pure bilayer system (black) and with deuterated tRNA. The SLD profile for the non-deuterated case is also depicted using the corresponding slightly transparent colors. Deuteration of tRNA increases its contribution to the total SLD as indicated by an increase in SLD for $d > 35$ Å (B) SANS form factor obtained from deuterated SLDs in (A). Here, the bilayer corresponds to a pure DOPC membrane.

RNA conformation affects its translation properties. Earlier studies [28] have reported that RNA interacts with lipid bilayers primarily via hydrophobic and electrostatic interactions resulting in a change in its conformation. In this chapter, using the coarse-grained Martini model for RNA and lipids, we explored the possibility of resolving RNA conformation on bilayer surfaces by employing the widely used scattering techniques, namely neutron reflectivity and SANS. We first looked at the interaction of ssRNA and dsRNA with neutral and charged bilayer systems. ssRNA interacts with the bilayer such that the hydrophobic nucleobases are closer to the hydrophobic tail parts of the bilayer whereas the charged phosphate backbone interacts with the polar parts of the bilayer. This results in a stronger interaction of ssRNA with neutral bilayers compared to double-stranded RNA for which hydrophobic nucleobases are not exposed in agreement with previous experimental studies. Therefore, even though the CG-Martini model can not resolve the intricacies of interaction, correctly captures the major driving forces that influence the RNA conformation at the bilayer surface i.e., hydrophobic and electrostatic interactions. The model also correctly captures the bilayer structure as indicated by a good agreement of CG and all-atom electron density profiles. Furthermore, our reflectivity and form factor calculation relies on well-converged SLD profiles and the fast dynamics of RNA with the CG model ensures that the RNA explores all possible configurations on the surface. These results justify our use of CG-model to study the RNA conformation at surfaces. Each secondary structure has a unique number of double-stranded and single-stranded regions hence each conformation attains a unique configuration on the surface (snapshots in Figure 4.9) shaping the RNA contribution to the total SLD profile in unique ways. Neutron reflectivity can not differentiate between such subtle signatures from different RNA conformations even after enhancing the signal from RNA by selective deuteration. The substrate with silicon and silicon dioxide (SiO_2) layers has a major influence in obscuring the signal from the sample. Similarly, SANS form factors, especially for neutral bilayer, where the RNA relatively weakly interacts with the bilayer, do not show a significant difference between different RNA conformations. However, in this case, selective deuteration of RNA leads to a clear distinction of different

secondary structures. Therefore, a combination of coarse-grained simulations with SANS experiments on simple liposomes with and without RNA can be used to obtain RNA conformation LNP-like environment. These simulation results could further help interpret the experimental results of such a system.

4.5 Conclusions

In this chapter, we first established that coarse-grained Martini simulations closely reproduce bilayer structures by comparing X-ray form factors from CG-Martini FF with all-atom simulations and experiments. The FF also captures the driving force involved in RNA interaction with neutral and charged bilayer systems. The main driving forces involved, which influence RNA conformations on bilayers, are hydrophobic and electrostatic interactions. To resolve RNA conformation on the bilayer, we calculated the neutron reflectivity profiles and SANS form factors from bilayer-tRNA systems with different tRNA conformations corresponding to different secondary structures. Neutron reflectivity and SANS form factor with undeuterated RNA can not resolve differences in secondary structures but selective deuteration of RNA in the case of SANS leads to a clear distinction between different RNA conformations. These results suggest that such coarse-grained simulation can be used to elucidate RNA secondary structure on bilayer systems from SANS experiments.

4.6 Appendix

4.6.1 Data archive

The simulation input file for all the simulations presented in this chapter is archived in the following tree. Go to any of the directories e.g. `./dsRNA-membrane/DOTAP0/` and run the command

```
gmx grompp -f ../../mdp/em.mdp -n index.ndx -p topol.top -c npt.gro  \
-r npt.gro -o npt.tpr
```

```
$HOME/Data-Archive-Mohd_Ibrahim/tRNA-DOPC-DOTAP-Martini
├── dsRNA-membrane
│   ├── DOTAP0
│   │   └── index.ndx npt.gro topol.top
│   ├── DOTAP100
│   ├── DOTAP25
│   ├── DOTAP50
│   ├── DOTAP75
│   └── martini-itp
├── ssRNA-membrane
│   ├── DOTAP0
│   ├── DOTAP100
│   ├── DOTAP25
│   ├── DOTAP50
│   ├── DOTAP75
│   └── itp
└── tRNA-membrane
    ├── tRNA-1
    │   ├── DOPC
    │   ├── DOTAP
    │   └── rna-itp
    ├── tRNA-2
    │   ├── DOPC
    │   ├── DOTAP
    │   └── rna-itp
    ├── tRNA-3
    │   ├── DOPC
    │   ├── DOTAP
    │   └── rna-itp
    ├── tRNA-4
    │   ├── DOPC
    │   ├── DOTAP
    │   └── rna-itp
    ├── tRNA-5
    │   ├── DOPC
    │   ├── DOTAP
    │   └── rna-itp
    ├── tRNA-6
    │   ├── DOPC
    │   ├── DOTAP
    │   └── rna-itp
    ├── tRNA-7
    │   ├── DOPC
    │   ├── DOTAP
    │   └── rna-itp
    ├── tRNA-8
    │   ├── DOPC
    │   ├── DOTAP
    │   └── rna-itp
    ├── tRNA-9
```

```
│       │       ├── DOPC
│       │       ├── DOTAP
│       │       └── rna-itp
│       ├── tRNA-10
│       │       ├── DOPC
│       │       ├── DOTAP
│       │       └── rna-itp
│       └── membrane-itp
├── pure-bilayer
│       ├── dopc
│       ├── dotap
│       └── membrane-itp
├── scripts
│       ├── sld_cg_rna.py
│       └── sld_cg_rna_deutrated.py
└── mdp
```
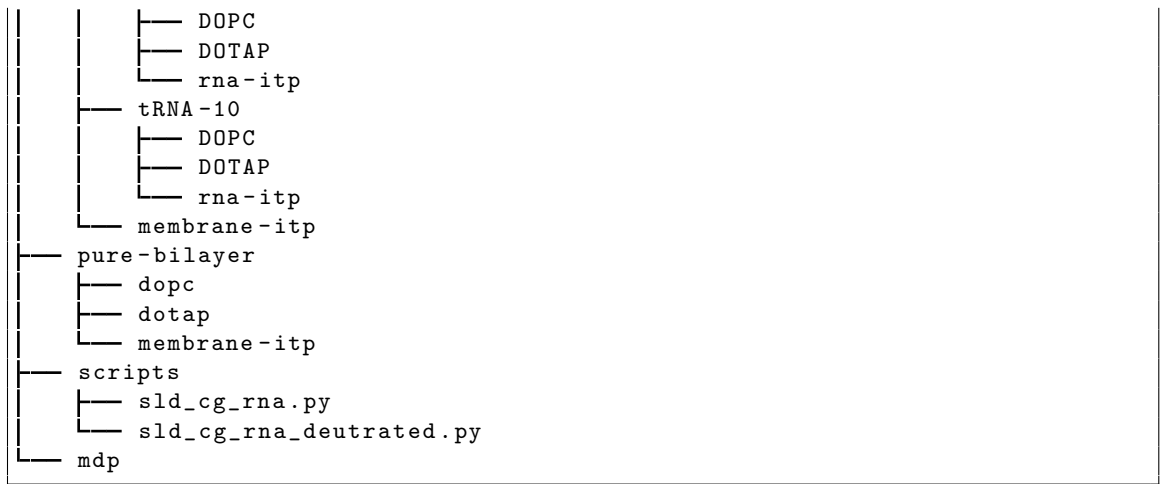
Parameterization of DLIN-MC3-DMA and Comparison with Neutron Reflectivity

The work presented in this chapter had been conducted under the supervision of Prof. Dr. Nadine Schwierz at the Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Frankfurt am Main, Germany, and at the Institute of Physics, University of Augsburg, Augsburg, Germany. The simulations were performed on the GOETHE HLR of the state of Hessen. The experiments were performed by Jennifer Gilbert and Prof. Dr. Tommy Nylander at the Department of Chemistry, Lund University, Lund, Sweden. Major aspects of this work is published in:

- **Ibrahim, M.**, Gilbert, J., Heinz, M., Nylander, T., and Schwierz, N. 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. doi: doi.org/10.1039/D3NR00987D [154]

5.1 Introduction

Lipid nanoparticles (LNPs) represent one of the most advanced delivery systems to transport RNA-based drugs [46, 155–157]. LNPs contain an ionizable lipid to aid RNA encapsulation and release, helper lipid and cholesterol to provide structure stability, and PEG-lipids to control the size and the mRNA drug cargo. Ionizable lipids form almost 50% mole fraction of the total LNP lipid compositions and hence are a crucial determinant of its delivery efficiency. DLin-MC3-DMA (MC3) has emerged as one of the most promising ionizable lipids and has been employed in one of the first clinically approved RNA-based drugs [14, 51, 156, 158]. It has a pKa value of 6.44 inside the LNP [51], which renders it positively charged at low pH, ensuring efficient RNA encapsulation. Additionally, it remains neutral at physiological pH inside the bloodstream, preventing degradation and promoting efficient circulation. Despite their importance, the structure of current LNP formulations remains poorly understood thus hindering a rational design of efficient LNPs [54]. A proper understanding of the LNP structure would involve resolving the distribution of different components inside the LNP, the interaction between these components, and the influence of pH and RNA cargo on these interactions.

Molecular dynamics (MD) simulations are well suited for resolving the interactions between different components of LNP and provide atomistic structural insight, however, the accuracy of MD results critically depends on the underlying force field (FF) which describes the interaction between different atoms. For MC3 FF parameters based on CHARMM36 [70] and Slipids [65] FF have been reported. Ermilova and Swenson derived parameters

for MC3 in combination with the Slipids FF [159]. Unfortunately, the authors did not parametrize the protonated state of the MC3 molecule which is important to understand MC3 behavior at low pH conditions. More recently, Park et al. derived parameters for neutral and protonated MC3 in combination with the CHARMM36 lipid FF [160]. Due to a lack of experimental data on simple systems with MC3, the predictions of these FFs were not validated with any experiment. Lastly, no MC3 FF compatible with the AMBER FF family has been reported. Since AMBER FFs are frequently used for proteins, nucleic acids, and lipids, therefore, parametrizing an MC3 FF compatible with the AMBER family of FFs is of utmost relevance.

Scattering experiments are one of the most reliable experiment techniques to validate molecular dynamics simulations, especially for lipid bilayers where characteristic structural quantities like electron and neutron scattering length density profiles can be probed with sub-nanometer resolution [81, 161, 162]. The results from such experiments can be compared with MD simulations in a model-free manner. Therefore, a combined approach of MD simulations and scattering techniques could provide invaluable insights into the understanding of the structures of lipid bilayers and other more complex structures involving lipids like lipid nanoparticles.

In this chapter, we obtain FF parameters for cationic and neutral MC3 based on the AMBER Lipid17 protocol. We then assess the accuracy of the currently developed and existing FFs by making a direct comparison with Neutron reflectivity measurements, performed on bilayer setups containing MC3 and DOPC mixtures, at different pH conditions, multiple deuteration contrasts, and different MC3 percentages. For the currently developed FF parameters, we obtain excellent agreement with experiments for almost all cases. The existing CHARMM-based MC3 FF also shows a similar level of agreement whereas the Slipids-based MC3 underestimates the bilayer thickness.

Throughout this chapter and next, cationic MC3 is denoted by MC3H and neutral MC3 simply as MC3.

5.2 Methods

5.2.1 Partial charge derivation of protonated and neutral MC3

The partial charge derivation for two protonation states of DLin-MC3-DMA was carried out following AMBER lipid14/11 charge derivation protocol [66, 77]. In this protocol, the lipid molecule is split into head and tail groups. The split groups are capped such that the chemical environment at the split position is preserved during the charge derivation. The charges for the capping groups are constrained to pre-derived values during the charge derivation, the capping molecule and its partial charges are shown in Figure 5.1. 20 random conformations were used to obtain the partial charges for the capping group. 100 random capped tail group conformations and 50 random capped head group conformations were used for charge derivation for each protonation state. To obtain random conformations, 100 ns pure MC3 and pure MC3H simulations were set up using the recently released CHARMM36-based FF [160]. Random conformations were selected from the last 50 ns trajectory of a 100 ns simulation run. The geometry of these conformations was further

optimized using the Gaussian09 software at the HF/6-31G* level in the gas phase [163].
The RESP charge derivation procedure [78] was then used to derive the partial charges
for all the conformations simultaneously using the RED-vIII.52.pl software [164]. The
electrostatic potential was evaluated at HF/6-31G*level of theory.
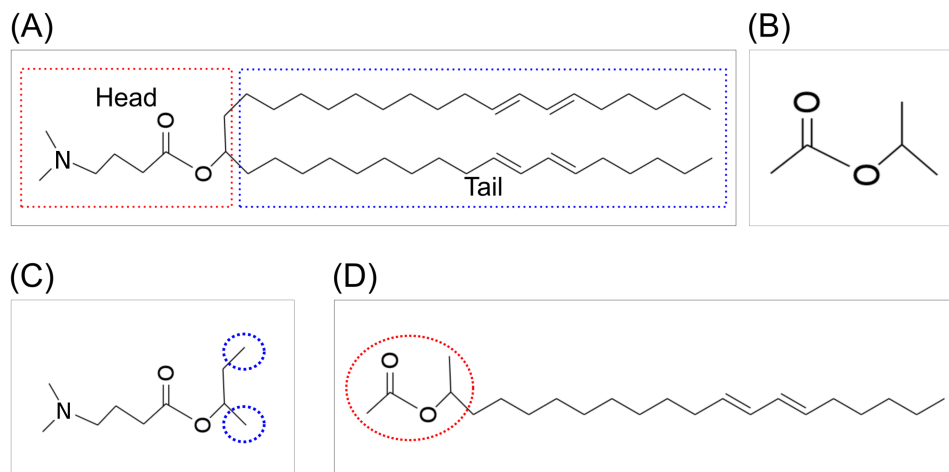
Figure 5.1: (A) In AMBER protocol lipid molecule is divided into head and tails groups. The head
and tail for MC3 are depicted by the red and blue boxed regions respectively. The head and tails are
capped during the charge derivation with a capping molecule. (B) The capping group (C) capped
head group, the cap is highlighted with the blue dashed circle and (D) capped tail group, the cap
is highlighted with the red dashed circle.

5.2.2 Assigning atom types, angles, dihedrals

The next step is to assign other non-bonded parameters namely, atom types, bonds, angles
and dihedrals. To that end, a dummy `.itp` file was created for the whole molecule using
the `acpype.py` script [165] assigning atom types according to the GAFF FF [166]. The
dummy parameter file created was then modified to have the correct partial charges derived
above and the atom types were modified in accordance with the AMBER Lipid17 force
field. The modified parameter file was then used as input for a custom Python script based
on MDAnalysis [129] and ParmEd Python packages [167]. The script searches for a given,
bond, angle, or dihedral in the AMBER Lipid17 parameter file `Lipid17.dat`. In case a
parameter is not found in `Lipid17.dat` it looks for the exact parameters in the GAFF2
parameter file `gaff2.dat` after properly replacing the lipid atom types with corresponding
GAFF types.

5.2.3 MD simulations

We simulated MC3-DOPC system with MC3 mole fraction varying from 5 % - 15% (Table
5.1). The CHARMM-GUI webserver [168] was used to obtain the initial bilayer setup for
systems simulated with CHARMM36 FF whereas for simulations with AMBER Lipid17
or Slipids, the MemGen webserver [169] was used. The CHARMM-based systems were
solvated with mTIP3P water model [64] while for others TIP3P water model [123] was
used. Sodium and chloride ions were added to neutralize the systems and to obtain a bulk
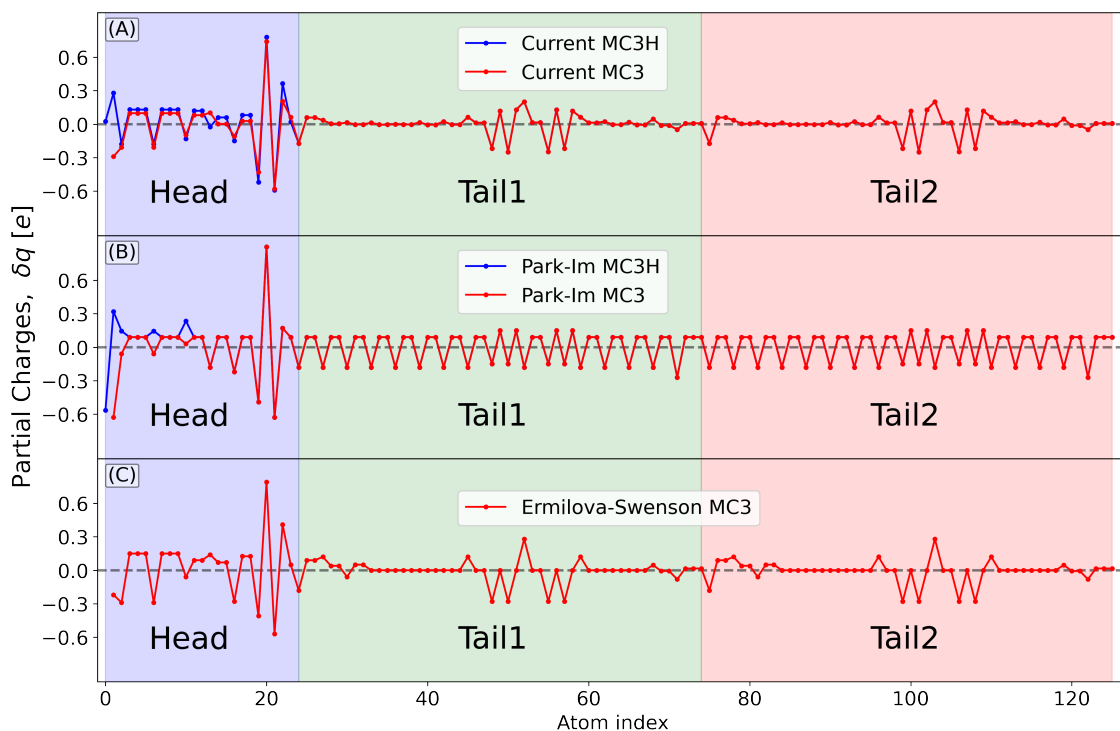
Figure 5.2: Partial charges for protonated and neutral MC3 from AMBER and CHARMM FF. The 0th index corresponds to the Nitrogen atom in the head group. For CHARMM36 FF the nitrogen atom of both neutral and protonated MC3 are negatively charged to the same extent ($\sim$ -0.6e).

salt concentration of 150 mM.

To describe MC3 and DOPC we used the following combination of FFs (i) CHARMM36 DOPC FF and Park-Im MC3 FF (ii) Slipids DOPC FF and Ermilova-Swenson MC3 FF (iii) Lipid17 DOPC FF and currently developed MC3 FF. For ions, the Mamatklov-Schwierz FF [126] parameters optimized in combination with TIP3P water were used.

The structures generated with the Mem-Gen web server were minimized using a gradient descent algorithm followed by 500 ps in the NVT ensemble with an additional flat bottom potential acting on the water molecules to prevent them from penetrating the membrane. The system was further equilibrated for 20.0 ns with the semi-isotropic Berendsen barostat with a time constant of 5.0 ps and with the Berendsen thermostat [59] with a time constant of 1.0 ps. The production run was carried out for 600.0 ns. The pressure was fixed at 1.0 bar using the semi-isotropic Parinello-Rahman barostat with a coupling constant of 1.0 ps [60]. The temperature was fixed at 300.0 K using the velocity rescaling method with stochastic term [58] and with coupling constant 1.0 ps.

Van der Waals interactions were cut-off at 1.2 nm and a force-switch between 1.0 nm and 1.2 nm was used. Short-range electrostatic interactions were cut off at 1.2 nm and the long-range part was evaluated using the Particle Mesh Ewald [57] method with a Fourier grid spacing of 0.12 nm. All bonds involving hydrogens were constrained using the LINCS algorithm with LINCS order 4 [122]. A time step of 2.0 fs was used to integrate the equations of motion. Simulation trajectories were visualized using VMD [130]. All the simulations were performed using GROMACS (version-2018) [121].
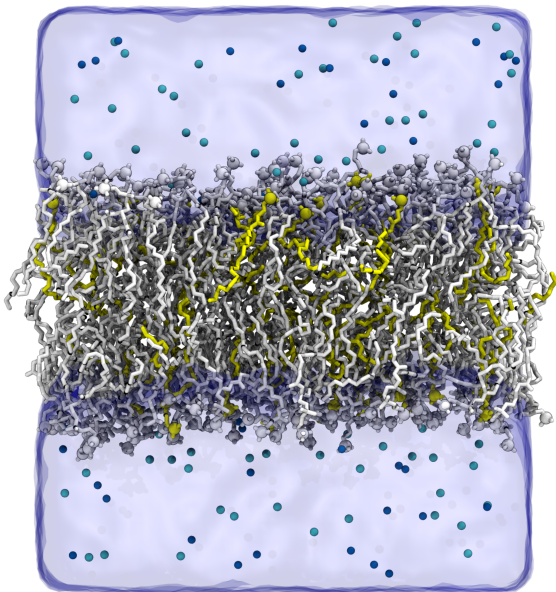
Figure 5.3: Snapshot of full simulation system with MC3 (yellow), DOPC (white) bilayer. Water is shown by the transparent background and ions as blue and cyan spheres.

Table 5.1: Simulation setup for the different systems. † denotes the setups for which all three force fields were used. Otherwise, the only the current parameters for MC3 or MC3H and the Amber Lipid17 parameters for DOPC were used. The box size of † setups were around 8.3 nm × 8.3 nm × 10 nm. For the rest the box size were approx. 11.8 nm × 11.8 nm × 8.7 nm. N_{xx} represents number of molecules of xx and N_{wl} is number of water per lipid molecule.

System	MC3 [mol %]	N_{DOPC}	N_{MC3}	N_{MC3H}	N_{wl}	Na^+	Cl^-
Pure DOPC†	0.0%	200	0	0	∼ 50	27	27
MC3/DOPC†	15.0%	170	30	0	∼ 72	40	40
MC3H/DOPC†	15.0%	170	0	30	∼ 71	38	68
MC3H/DOPC	10.0%	360	0	40	∼ 60	105	65
MC3H/DOPC	5.0%	380	0	20	∼ 60	85	65
MC3/DOPC	5.0%	380	20	0	∼ 60	65	65

5.2.4 Modeling of silicon substrate and obtaining reflectivity profile

A key difference between the simulation and the experimental setup is that, in experiments, the bilayer is supported on a silicon substrate whereas in simulations the bilayer is surrounded by water on both sides (Figure 5.3). The substrate has a significant impact on the measured reflectivity profile. Therefore, to make a direct comparison of experiments and simulations, the effect of the substrate must be taken into account. As described in the previous section, the reflectivity profile is fully determined by the scattering length density profile (SLD). Therefore, modeling the substrate would entail obtaining the substrate SLD and incorporating it into the bilayer SLD to get the full SLD profile. This is achieved in a three-step procedure: The silicon substrate SLD is obtained by fitting the experimental data of the substrate in solution to a three-slab model (Figure 5.4). Subsequently, the position of the substrate/bilayer interface is determined from the simulations (Figure 5.5). Finally, we follow the procedure of previous work [170, 171] and introduce two physically motivated fit parameters corresponding to the fraction of water in the bilayer leaflet closer to the substrate (α) and to the substrate area covered by water patches (γ).

Obtaining substrate SLD

The silicon substrate SLD is obtained by fitting the experimental data of the substrate in solution to a three-slab model. For each experimental sample cell, reflectivity experiments were performed for substrate-D_2O system. The reflectivity profiles were fitted with a three-slab model: Si-layer, SiO_2 and D_2O. The Refnx python package [172] was used to obtain the substrate SLD profiles. The resulting fits and SLD profiles are shown in Figure 5.4. Finally, the SLD profiles of the three-slab model were added to the SLD of the simulated free bilayer.
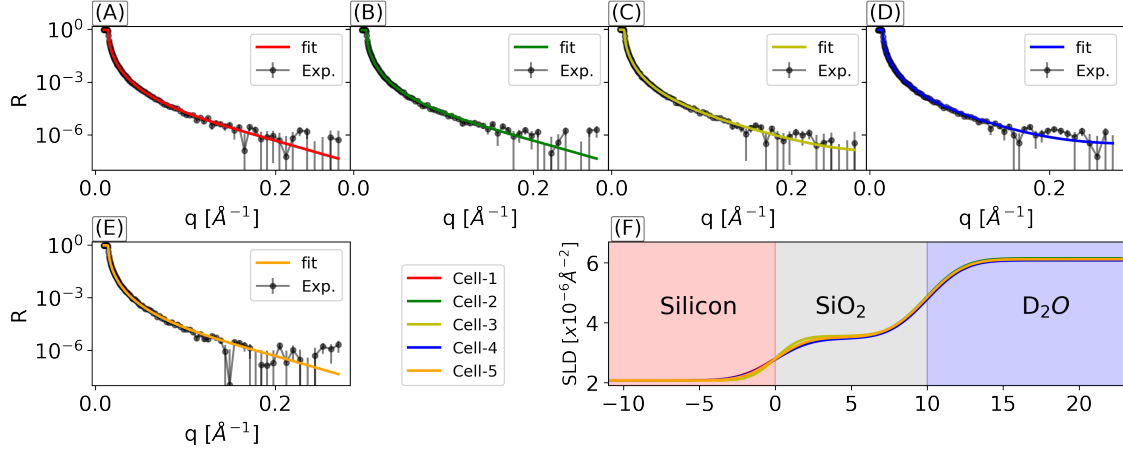
Figure 5.4: Modeling of the substrate: (A-E) Reflectivity profile R for the silicon/silicon dioxide/D_2O system for each experimental sample cell. (F) SLD corresponds to the best fit of the experimental reflectivity profiles. The SiO_2 layer thickness in all cases is around 10 Å

Position of the substrate/bilayer interface

The position of the substrate/bilayer interface is determined using the SLD from the free bilayer simulations. Here, the point at which the SLD assumes zero is taken to be the position of the substrate. These points are depicted by vertical dashed lines in Figure 5.5.

Fitting of water fraction (α) and water patches (γ)

Following previous work, we introduce two fit parameters [170, 171]. The first parameter, α corresponds to the amount of water in the bilayer leaflet closer to the substrate. It scales only the SLD of solvent in that leaflet and was chosen to vary between 0 and 1. $\alpha = 1$ corresponds to the unscaled and fully hydrated profile. $\alpha = 0$ corresponds to the fully dehydrated profile (Figure 5.6B). The second parameter, γ is the fraction of water patches on the substrate. $\gamma = 0$ corresponds to a situation where the substrate is perfectly covered by the bilayer and hence coverage is 100% (Figure 5.6A). Note that the phospholipid bilayer surface coverage on a silicon substrate can be as low as 70% [173–175]. The total SLD as a function of the distance d from the center of the bilayer can be written as

$$SLD(d) = \begin{cases} (1-\gamma)\rho_b(d) + (1-\gamma)\alpha\rho_s(d) + \gamma\rho_{s0} & \text{for} \quad d < d_{\min} \\ (1-\gamma)\rho_b(d) + (1-\gamma)\rho_s(d) + \gamma\rho_{s0}, & \text{otherwise} \end{cases} \tag{5.1}$$
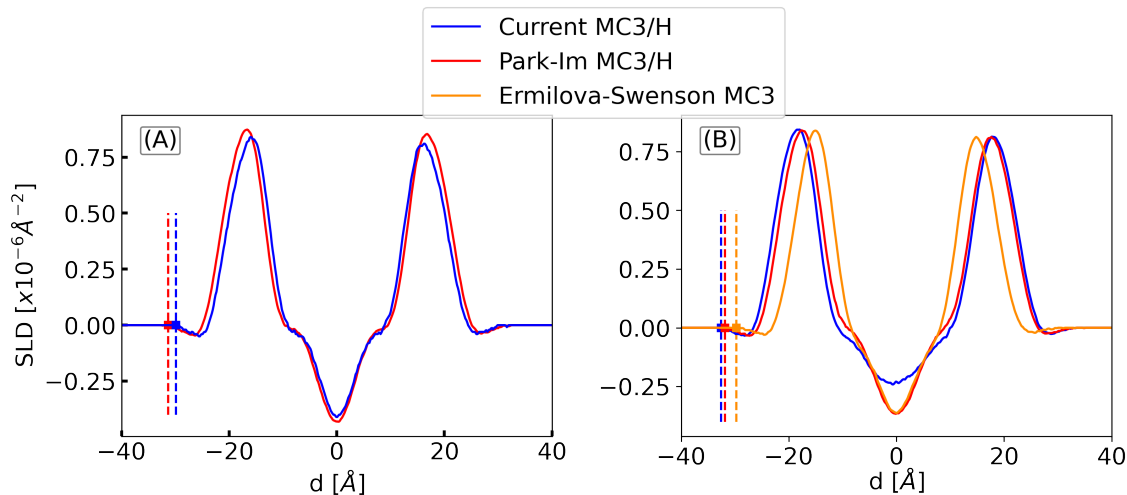
Figure 5.5: Position of the substrate/bilayer interface. SLD for cationic MC3H/DOPC (A) and neutral MC3/DOPC systems (B) for the different force fields. The vertical dashed lines indicate the position of the substrate.

ρ_s and ρ_b are the SLD of solvent and bilayer, respectively. ρ_{s0} is the SLD of bulk solvent obtained from the fits in Figure 5.4: $\rho_{s0} = -0.56 \times 10^{-6}$ Å^{-2} for H$_2$O, $\rho_{s0} \sim 6.1 \times 10^{-6}$ Å^{-2} for D$_2$O, and $\rho_{s0} = 2.07 \times 10^{-6}$ Å^{-2} for contrast matched conditions. $d_{\min}$ denotes the position of the minimum in the SLD of the free bilayer (Figure 5.5). Note that in the center of the bilayer water content is zero over a distance of about 2 nm. Therefore the choice of $d_{\min}$ does not affect the calculation as long as it is in the water-free region of the bilayer.
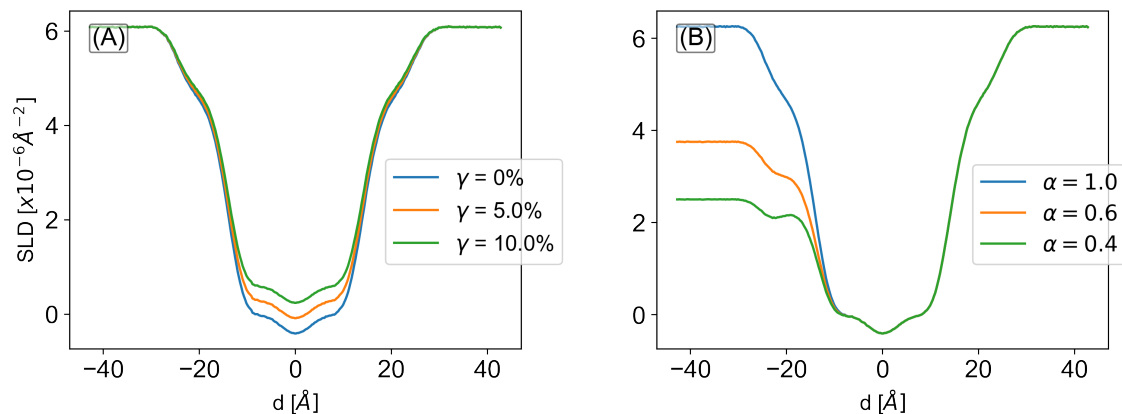
Figure 5.6: (A) Effect of fit parameter γ on the SLD profile. $\gamma = 0\%$ corresponds to the profile obtained from the simulations. (B) Effect of fit parameter α on the SLD profile. $\alpha = 1$ corresponds to the unperturbed SLD profile from the simulations. In both these cases, the solvent is 100% D$_2$O and the system is the MC3H-DOPC setup with the current force field parameters.

Optimization procedure

To optimize the values of α and γ for a given setup, a grid search was used in the parameter space $\alpha \in [0, 1]$ and $\gamma \in [0, 0.5]$. 10,000 grid points in the (α, γ) plane corresponding to 100 points along each axis were used. At each grid point, we calculated the χ^2 deviation

between experiments and simulations at three contrasts as,

$$\chi^2(\alpha, \gamma) = \frac{1}{(N-1)} \sum_q \frac{(R_q^{\mathrm{exp}} - R_q^{\mathrm{sim}}(\alpha, \gamma))^2}{\sigma_q^2} \tag{5.2}$$

R_q^{exp} denotes the reflectivity at point q in the reflectivity profile obtained from the experiments. $R_q^{\mathrm{sim}}(\alpha, \gamma)$ is the reflectivity from the simulations for fixed values of α and γ. N is the number of data points in the reflectivity profiles and σ_q is the experimental error at each q point. The total χ^2 value for fixed values of α and γ is the sum over the three different contrasts:

$$\chi^2_{\mathrm{global}}(\alpha, \gamma) = \chi^2_{D_2O}(\alpha, \gamma) + \chi^2_{H_2O}(\alpha, \gamma) + \chi^2_{CMSi}(\alpha, \gamma) \tag{5.3}$$

5.2.5 Experimental methods

Neutron reflectivity experiments were performed over q range $0.009 - 0.27$ Å^{-1}. The experimental setup consisted of MC3-DOPC bilayers supported on a silicon substrate and was carried out at two different pH values, pH=6.0 and pH=7.0, using 5%, 10%, and 15% MC3 mole fractions. Each experiment was performed at three different solvent contrasts, 100 % H_2O, 100 % D_2O, and the contrast matched silicon (CMSi) which has 38 % D_2O. To characterize the silicon substrate, experiments on D_2O-substrate systems were performed. More details on experiments can be found in Ibrahim et al. [154].

5.3 Results

5.3.1 MC3 distribution depends on pH and force field

We simulated MC3-DOPC systems using different FFs in high and low pH conditions. At low pH (i.e, pH $<$ pK$_a$) all MC3 lipids are assumed to be positively charged whereas at high pH (i.e, pH $>$ pK$_a$), all MC3 lipids are assumed to be neutral. Note that the pK$_a$ of MC3 inside an LNP is 6.44. Figure 5.7 shows the final snapshots of MC3-DOPC setup, after 600 ns of simulation time, from different force fields. The neutral MC3 from the current force field segregates into the bilayer center increasing bilayer thickness compared to a pure DOPC bilayer (Table 5.2). The bilayer thickness increases systematically from 36.8 Å to 41.5 Å with increasing MC3 fraction from 0 to 15 % (Table 5.3). The neutral MC3 from Park and Im [160] in combination with CHARMM36 FF also shows a tendency to accumulate in the bilayer center, however, it's not as pronounced as the current FF. Similarly to the current FF, the accumulation leads to an increase in bilayer thickness compared to the pure DOPC bilayer. The neutral MC3 from Ermilova and Swenson [159] in combination with Slipids FF show the contrary behavior, here, both the tail and head group of the MC3 molecule tend to stay away from the bilayer center and accumulate at the bilayer water interface. The bilayer thickness in this case is reduced compared to a pure DOPC bilayer (Table 5.2).

The cationic MC3H from the current and the Park-Im FF show similar behavior, both tend to stay at the bilayer water interface. In both cases, the bilayer thickness does not change significantly compared to the pure DOPC bilayer (Table 5.2). Overall, the Park-Im

MC3H has a slight tendency toward the bilayer center as apparent from the increased mass density profiles (Figure 5.7 B). The differences between the observed behavior of neutral MC3 from Park-Im and our current model can be attributed to the partial charges of the head group nitrogen atom. Neutral MC3 from the Park-Im model has almost twice the negative partial charge, at the head group nitrogen, compared to the current MC3 (Figure 5.2 A, B). Consequently, the electrostatic interaction between the MC3 nitrogen and the highly positively charged carbonyl carbons (situated at the junction of the head and tail group) of DOPC is much stronger, thus preventing the MC3 from accumulating strongly in the bilayer center. This is evident in the radial distribution functions between these two atoms, which show a significantly prominent first peak for the Park-Im case compared to other FFs (Figure 5.8 B). The different behavior of Ermilova-Swenson MC3 compared to the current model most likely stems from differences between other non-bonded parameters such as Lennard-Jones, dihedrals, or angles. Both models use the AMBER RESP charge derivation protocol, therefore, partial charges do not differ significantly (Figure 5.2 A, C).
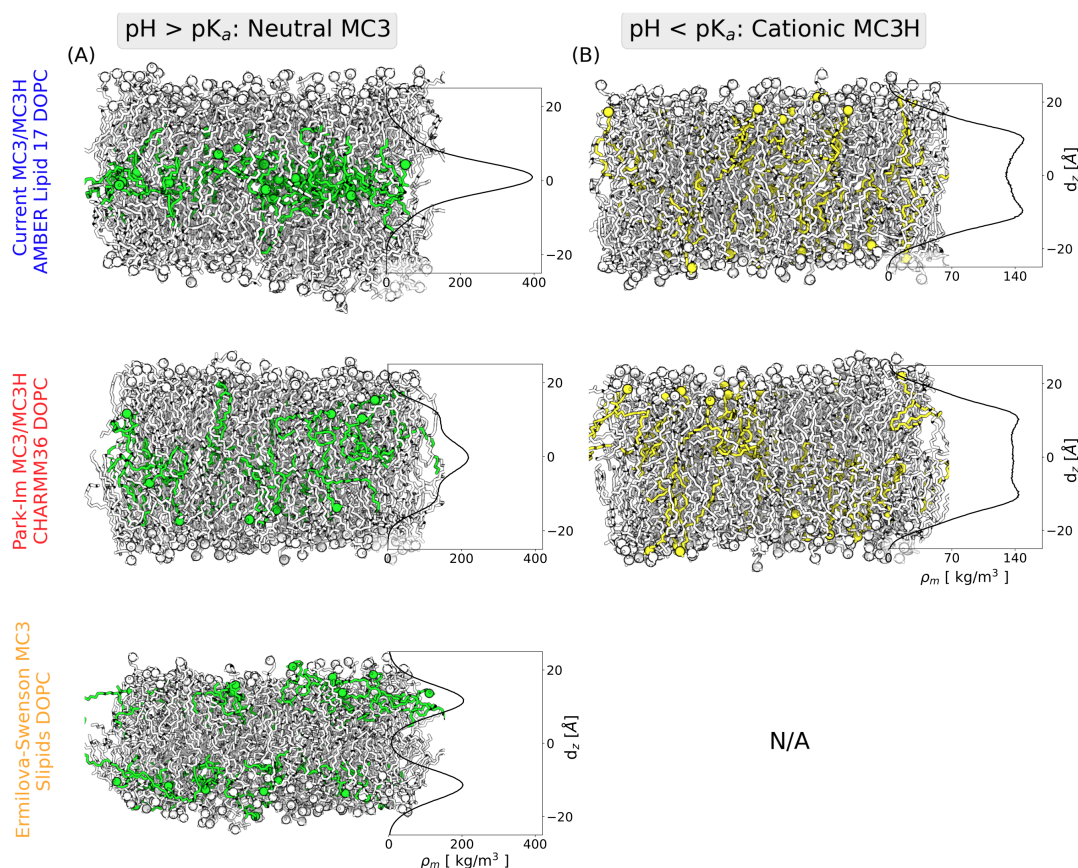
Figure 5.7: Simulation snapshots and mass density profiles of 15% MC3 in a DOPC bilayer for the different force fields at high pH where MC3 is neutral (A) and at low pH where MC3 is cationic (B). The spheres represent the nitrogen atoms in the head groups. Neutral MC3 is shown in green, cationic MC3 in yellow, and DOPC in white.
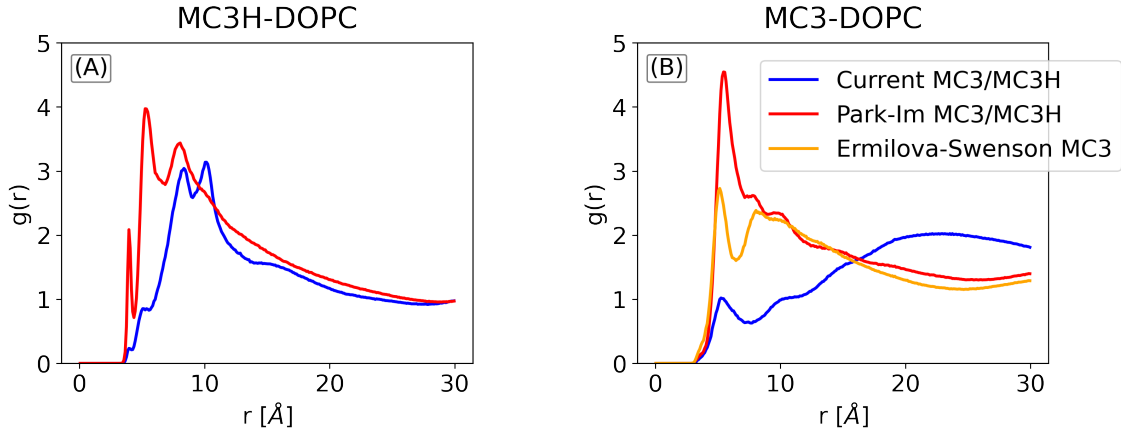
Figure 5.8: Radial distribution function between the positively charged carbonyl carbon of DOPC and the nitrogen atom of cationic MC3H (A) and neutral MC3 (B) for different force fields.

System	Current MC3/H and Lipid17 DOPC	Park-Im MC3/H and CHARMM36 DOPC	Ermilova-Swenson MC3 and Slipids DOPC
Pure DOPC	36.8 ±0.6 Å	38.2±0.6 Å	36.7±0.3 Å
MC3-DOPC	41.5 ±0.2 Å	39.7±0.8 Å	33.2±0.4 Å
MC3H-DOPC	37.4 ±0.6 Å	38.8±0.8 Å	N/A

Table 5.2: Bilayer thickness (D_{HH}) for the different systems using different force fields. The bilayer thickness corresponds to the distance between the two highest peaks in the electron density profile. Errors correspond to standard deviations obtained from block averaging (4 blocks for the last 400 ns of the simulations). The reported literature values for pure DOPC from X-ray scattering experiments are 35.3 Å [176], 36.7 Å[177], 36.9 Å [178], 37.1 Å [179] at $T = 30$ °C.

System	0% MC3/H	5% MC3/H	10% MC3/H	15% MC3H/H
MC3H-DOPC	36.8 ±0.6 Å	36.7±0.4 Å	36.5±0.4 Å	37.4 ±0.6 Å
MC3-DOPC	36.8 ±0.6 Å	37.8±0.3 Å	N/A	41.5 ±0.2 Å

Table 5.3: Thickness of DOPC bilayers containing different mole fractions of cationic or neutral MC3. The bilayer thickness corresponds to the distance between the two highest peaks in the electron density profile. Errors correspond to standard deviations obtained from block averaging (4 blocks for the last 400 ns of the simulations). All simulations were done with the current force fields for cationic and neutral MC3 and the Amber Lipid 17 force field for DOPC.

The minor difference between the behavior of cationic MC3H from the current and Park-Im model could also be attributed to the high partial charges of head group nitrogen. This again results in a relatively stronger interaction with DOPC carbonyl carbon as evident from the RDF (Figure 5.8 A). Since these carbons are positioned closer to the center compared to the head group, a relatively favorable interaction with these requires the MC3 to slightly move toward the bilayer center.
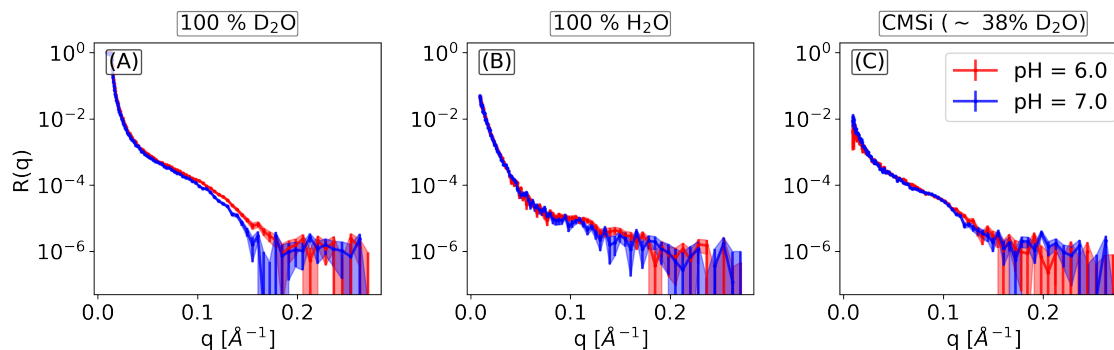
Figure 5.9: Experimental reflectivity profiles for systems with 15%MC3:85%DOPC at different pH
and at different solvent contrasts.

5.3.2 Direct comparison with neutron reflectivity experiments

To validate the above observations we make a direct comparison with neutron reflectivity
experiments performed on an MC3-DOPC system. The high pH experimental profile was
compared with the neutral MC3-DOPC setup while the low pH profiles were compared with
the cationic MC3H-DOPC simulation setup. The silicon substrate on which the bilayer
in experiments is supported but absent in the simulations was modeled as described in
Section 5.2.4. To compare different force fields, we used the experimental reflectivity profiles
measured on the system with the highest MC3 fraction. The experimental reflectivity
profiles with the highest MC3 mole fraction at different solvent contrast and pH values are
presented in Figure 5.9. Only the 100% D_2O contrast exhibits significant pH-dependent
differences, where the position of the characteristic minima ($q = 0.16 - 0.22$ Å^{-1}) is shifted
towards low $q-$value with increasing pH. Figure 5.10 shows the experimental profile at
different fractions of MC3 for the two pH conditions at 100% D_2O contrast. At low pH,
the reflectivity profiles are insensitive to change in MC3 fraction whereas at high pH, the
minima shift to lower $q-$values with increasing MC3 fraction.
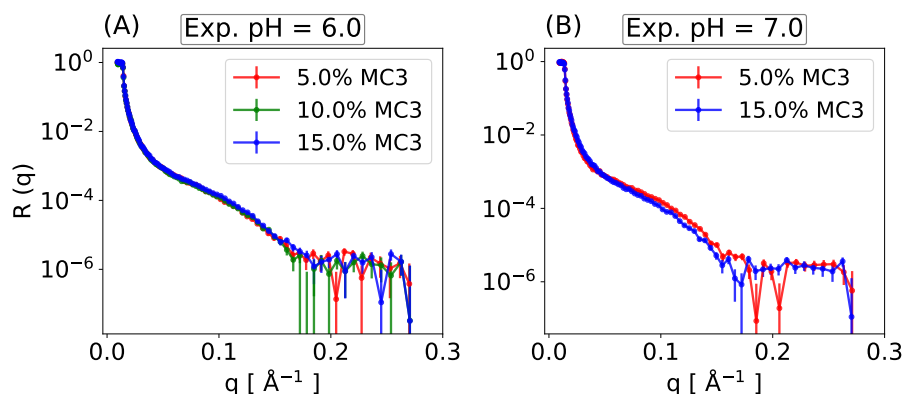
Figure 5.10: Comparison of the experimental reflectivity profile at different MC3 fractions for pH
6 and pH 7. The profiles correspond to 100% D2O solvent contrast. For pH 7, the shift of the
characteristic minimum to lower q upon increasing the MC3 fraction is clearly visible.

Cationic MC3H-DOPC comparison with experiments

The SLD profiles and corresponding neutron reflectivity profile at three different solvent contrasts for MC3H-DOPC systems from current and CHARMM36 FF are shown in Figure 5.11. Both FFs show excellent agreement with experiments at all contrasts as evident from visual inspection and from the low global χ^2 values (5.12). For both cases, we obtain a single global minimum in the specified $\alpha - \chi$ parameter space. For the current FF parameters with Lipid17, the optimum values are $\alpha = 0.53$, $\gamma = 0.05$, for Park-Im with CHARMM36, $\alpha = 0.62$, $\gamma = 0.08$. Even though there are slight differences between the SLD profiles from current and Park-Im models (Figure 5.11 A-C), these differences do not translate into significant variations in the reflectivity profiles possibly due to convoluted contribution from different components, i.e., water, DOPC, and MC3H.

For the current MC3H FF, the comparison with low MC3 fractions with MC3:DOPC mole fraction of 5%:95% and 10%:90% also exhibits excellent agreement with experiments (Figure 5.13). In these cases $(\alpha, \gamma) = (0.51, 0.116)$ and $(\alpha, \gamma) = (0.52, 0.101)$ for 5% and 10% MC3H systems respectively. These values indicate that the surface coverage of the substrate at low pH is in the range 89%-95% which is well within the reported values [173–175].

Figure 5.11: Direct comparison of simulations and neutron reflectivity experiments for 15% cationic MC3 in a DOPC bilayer. Neutron scattering length density profiles from the simulations with two different force fields at three different deuteration levels (top, A-C) and reflectivity profiles from simulations and experiments (bottom, D-F). The contrast corresponds to (A,D) 100% D_2O , (B, E) 100 % H_2O and (C,F) $\sim$ 38% D_2O also known as contrast matched silicon (CMSi). The dashed lines in (A) indicate the regions of substrate, bilayer, and water. Simulation snapshots for the two force fields are shown in Figure 5.7B. The experiments were performed at pH 6 and the error bars are calculated from error propagation of the square root of the number of counts per bin on the detector during the data reduction.
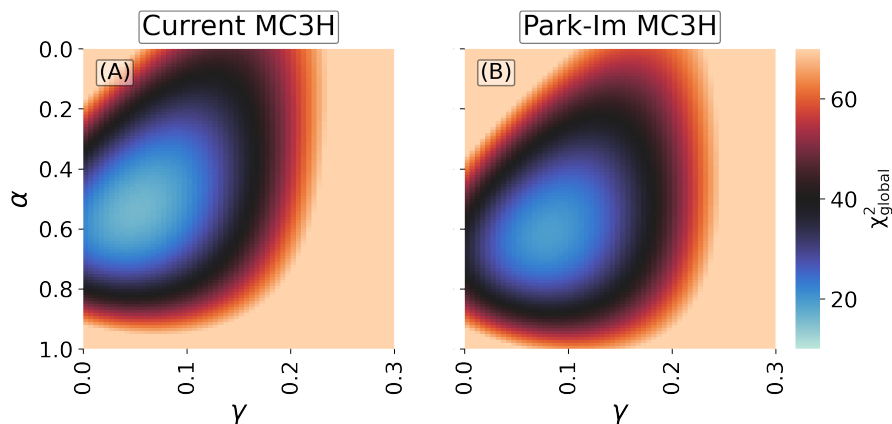
Figure 5.12: χ^2_{global} (Eqn. 5.3) deviation between experiments and current FF and between experiments and Park-Im MC3H FF. We obtain global minima for both FFs within the specified parameter space. α denotes the water fraction at the substrate-bilayer interface and γ depicts the area fraction of substrate not covered by the bilayer but by possible water patches.

Figure 5.13: Neutron reflectivity profile from experiments at pH 6.0 and simulations for systems with MC3H:DOPC mole fraction 5%:95% and 10%:90%. The current MC3H force field with Lipid17 DOPC FF is used here. The reflectivity profiles at different MC3H fractions are shifted in y for clarity.

Neutral MC3-DOPC comparison with experiments

The neutral MC3-DOPC systems are compared with experimental reflectivity profiles performed at pH=7.0. The SLD and corresponding reflectivity profiles from simulations and experiments for all FFs are shown in Figure 5.14. Both the current and Park-Im model show equally good agreement with experiments. Like in the previous case the differences in the distribution of MC3 between the two models do not significantly affect the reflectivity profiles. The Ermilova-Swenson model predicts a reflectivity profile with the characteristic minima for pure D_2O contrast shifted to higher q-values compared to the experiments. Note that, the 100% D_2O system provides the highest contrast between the solvent and the hydrocarbon tails, therefore, the position of characteristic minima (first minima) in this case is highly sensitive to bilayer thickness. The shift to high q-values for the Ermilova-Swenson model is a clear indication that it predicts a thinner bilayer compared to experiments. For the three force fields, we obtain similar values: $\alpha = 0.52$, $\gamma = 0.156$ (current MC3 and

AMBER DOPC), $\alpha = 0.51$, $\gamma = 0.151$ (Park- Im MC3H and CHARMM36 DOPC) and $\alpha = 0.40$, $\gamma = 0.121$ (Ermilova-Swenson MC3 and Slipids). The visual inspection and χ^2 values indicate that MC3-DOPC systems slightly deviate from experiments compared to the MC3H-DOPC setups (Figure 5.15). For the current FF, we also obtain good agreement with experiments for setup with 5% MC3 and 95 %DOPC (Figure 5.16). The optimum values of γ suggest a surface coverage between 85%-90% at pH=7.0. These values are well within the range reported from experiments [173–175] and simulations [170, 171] for bilayers on silica surfaces. In general, the surface coverage at high pH is slightly less than at low pH.
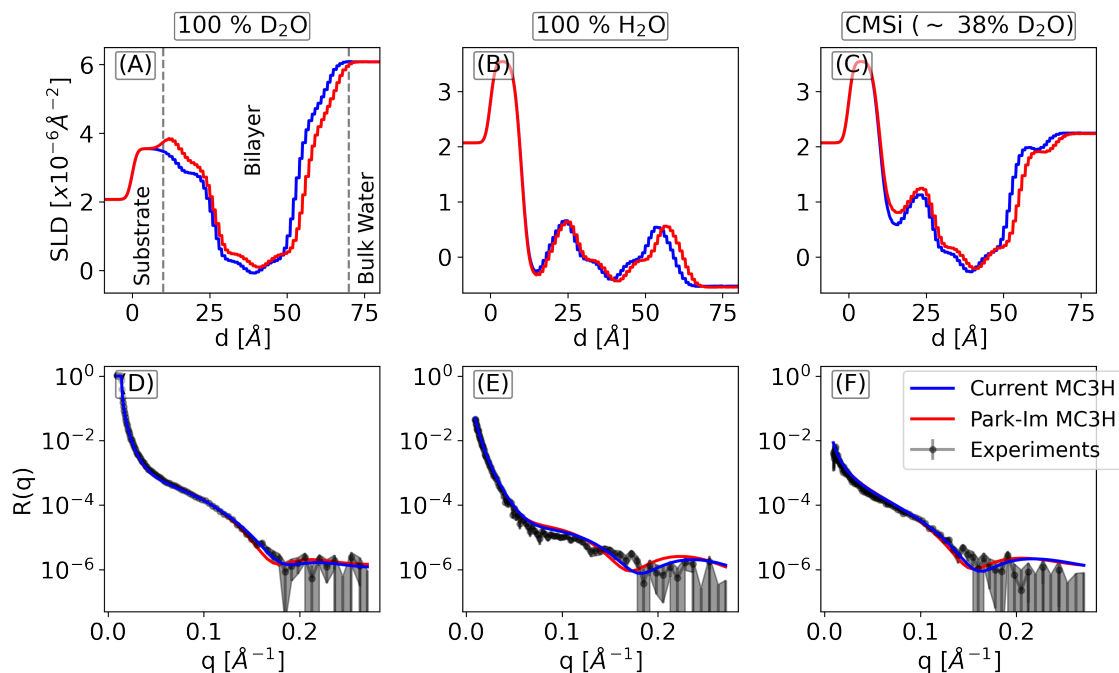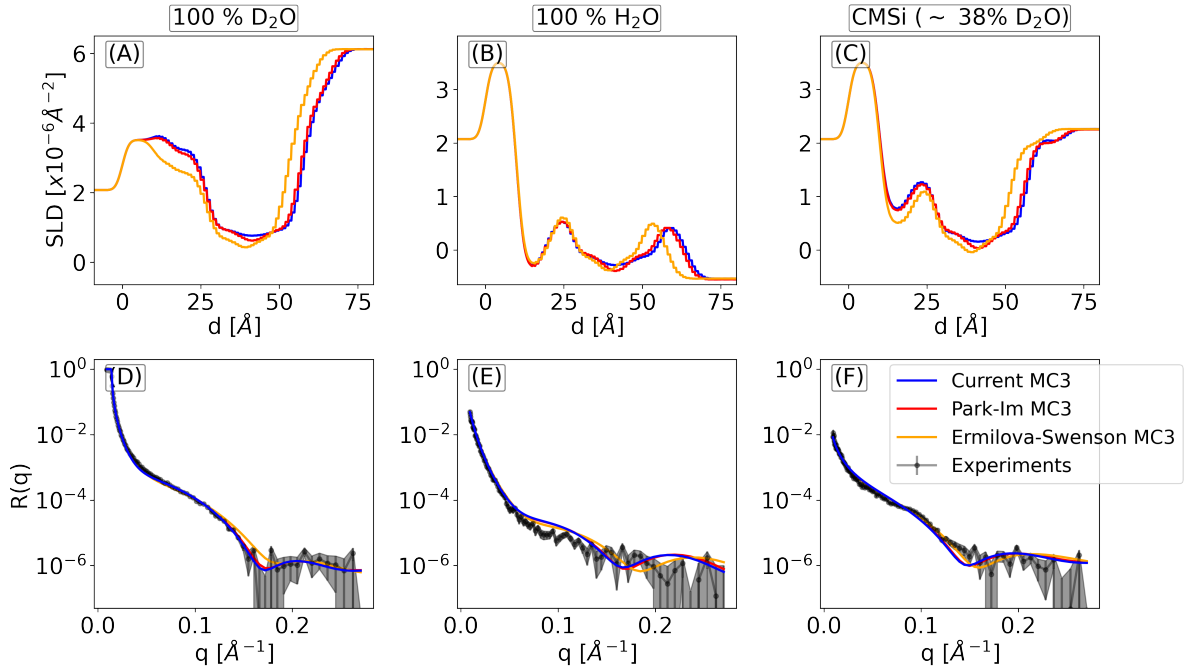

Figure 5.14: Direct comparison of simulations and neutron reflectivity experiments for 15% neutral MC3 in a DOPC bilayer. Neutron scattering length density profiles from the simulations with three different force fields at three different deuteration levels (top) and reflectivity profiles from simulations and experiments (bottom). The contrast corresponds to (A,D) 100% D_2O , (B,E) 100 % H_2O and (C,F) $\sim$ 38% D_2O. The experiments were performed at pH 7.0. The corresponding simulation snapshots are shown in Figure 5.7 A.

Effect of degree of protonation

In the comparison of experiments and simulations, we assumed all the MC3 to be protonated at pH=6.0 and all to be neutral at pH=7.0. However, in the real system that is a very unlikely case given that MC3 has a reported pK_a=6.44. The degree of deprotonation of an ionizable group at a given pH can be estimated by the Henderson–Hasselbalch equation as,

$$\zeta = \frac{1}{1 + 10^{pK_a - pH}} \tag{5.4}$$

This equation has also been used to obtain fractions of charged and uncharged MC3 in the blood and inside the endosome. According to the above equation, for $pK_a = 6.44$, at
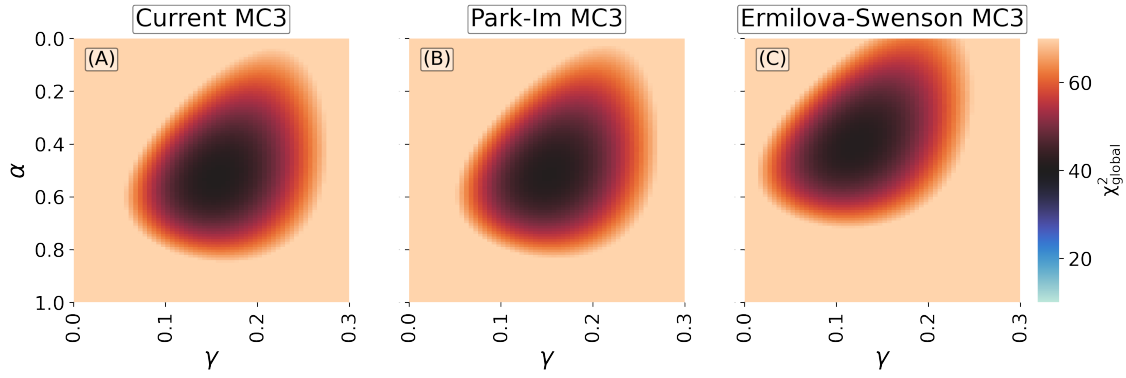
Figure 5.15: χ^2_{global} (Eqn. 5.3) deviation between experiments and current (A), Park-Im (B) and Ermilova-Swenson (C) FFs. We obtain global minima for all the FFs within the specified parameter space. α denotes the water fraction at the substrate-bilayer interface and γ depicts the area fraction of substrate not covered by the bilayer but by possible water patches.
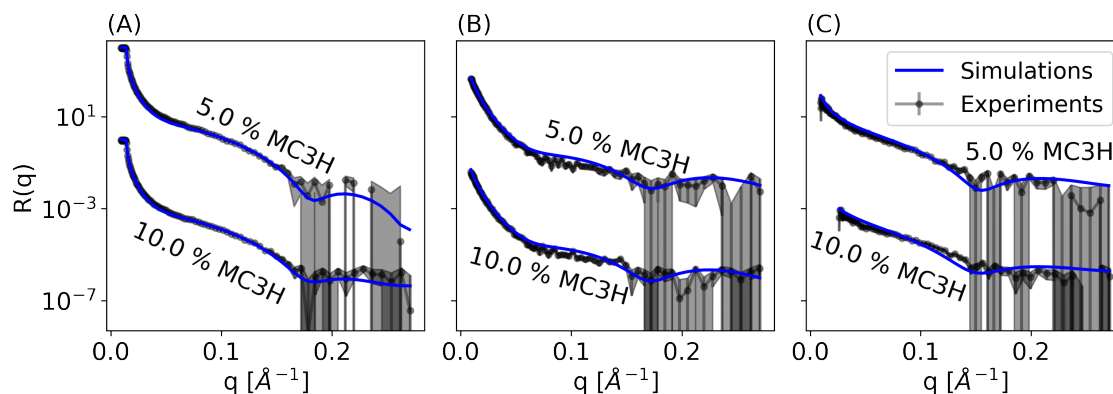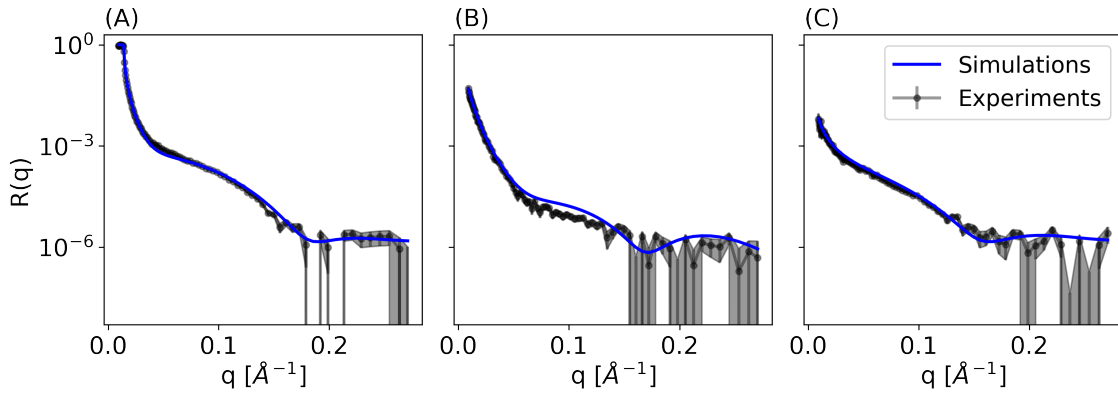

Figure 5.16: Neutron reflectivity profile from experiments (pH = 7.0) and simulations for systems with 5.0 % MC3 and 95 % DOPC at different solvent contrast (A) 100 % D_2O (B) 100 % H_2O and (C) $\sim$ 38.0% D_2O. The current MC3 force field with Lipid17 DOPC FF is used here. Here the optimum values of α and γ are 0.77 and 0.106 respectively.

pH=6.0: 73.0% of MC3 is charged, and at pH=7.0: 78% of MC3 is neutral. The simulation results using these numbers do not have a significant impact on the agreement between experiments and simulations as evident from Figure 5.17
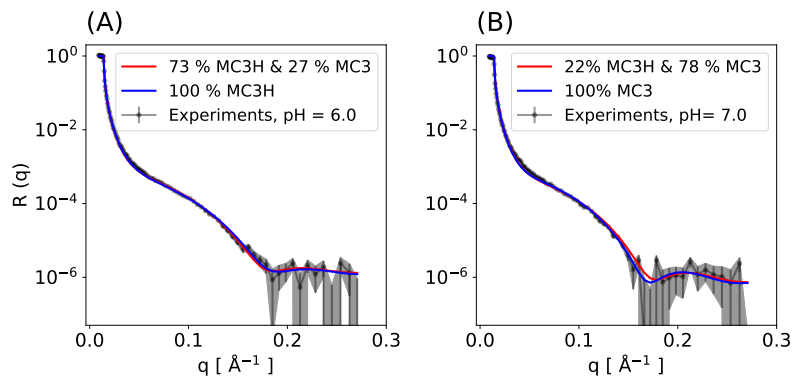
Figure 5.17: Effect of degree of protonation on the agreement of experiments and simulations. The degree of deprotonation (fraction of neutral MC3) is given by the Henderson-Hasselbalch equation. Here total MC3 (charged + neutral) mole fraction is 15%.

5.4   Discussions

We parametrized neutral and cationic MC3 corresponding to MC3 in two extreme pH conditions. To validate our parameters and to obtain structural insights, we performed experiments and simulations on MC3/DOPC setup in different pH conditions and different MC3 mole fractions. We also performed simulations using existing MC3 force fields and compared them with experiments enabling us to assess their predictions. Note that, due to a lack of experimental data, the existing force fields were not originally validated against any experimental data.

The simulations with cationic and neutral MC3 shows very different distribution of MC3. For neutral MC3 current FF predicts an increase in the bilayer thickness, relative to pure DOPC bilayer, due to the accumulation of MC3 in the bilayer center. The Park-Im model shows the same behavior but the tendency of MC3 to accumulate in the bilayer center is less prominent than the current FF. The Ermilova-Swenson model shows the contrary behavior, the MC3 has a tendency to accumulate at the bilayer-water interface, and the bilayer thickness is reduced compared to the pure DOPC system (Table 5.3). The cationic MC3 from the current and the Park-Im FF stays at the bilayer water interface and does not significantly affect the bilayer thickness compared to pure DOPC. A qualitative inspection of experimental reflectivity profiles at different pH values displays a shift of characteristic minima towards low-q value with increasing pH (Figure 5.9 A). Such a shift is a clear indication of an increase in bilayer thickness with pH which goes in line with the prediction from the current and Park-Im FFs. Moreover, at high pH, the characteristic minima shift towards low-q value upon increasing MC3 fraction whereas at low pH no such shift is observed (Figure 5.10). This again goes in line with the observation from the current FF (Table 5.3), where bilayer thickness increases with an increasing fraction of MC3 but varying the fraction of MC3H has no significant effect on the thickness. Previous cryo-TEM imaging studies of LNPs containing MC3 or an analogous amino lipid DLin-KC2-DMA (KC2) revealed a high electron-dense core at high pH and a vesicular bilayer structure at low pH [180, 181]. The high electron density core was attributed to the accumulation of neutral KC2 in the LNP core, whereas at low pH the cationic KC3 stays at the bilayer water interface giving rise to a vesicular bilayer LNP structure. Recent simulations on KC2/POPC bilayer showed that neutral KC2 has a tendency to segregate in the bilayer center whereas protonated KC2 stayed at the bilayer water interface [182]. The observed behavior of MC3 from current and Park-Im FFs are in line with these studies whereas the Slipid-based Ermilova-Swenson FF does not explain the observations.

The direct quantitative comparison of simulations with neutron reflectivity profiles further validates the observed MC3 distribution from different FFs. The current and the Park-Im MC3/H FFs exhibit very good agreement with experiments suggesting that the simulations correctly capture the transverse distribution of different components. The current FF also displays excellent agreement with experiments at low MC3 fractions (5 % and 10%) further substantiating its accuracy. The Ermilova-Swenson MC3 does not reproduce the characteristic minima for 100 % D_2O contrast and a shift of minima towards high q-values indicates a thinner bilayer compared to experiments. For all FFs at both pH values, at

100 % H_2O contrast, the simulations show slight deviations from experiments between q = 0.07-0.12$Å^{-1}$. Due to the low scattering length of H_2O, the pure H_2O signals are relatively noisy and are insensitive to the fit parameters α and γ but are more sensitive to substrate properties and effects on the bilayer. Therefore, the still good agreement for pure H_2O contrast substantiates the robustness of our substrate modeling protocol but also highlights the intricate effects of explicit substrate presence that are not captured in our modeling process. The presence of explicit substrate affects the interfacial water density and the bilayer SLD profiles[171].

Visual inspection of reflectivity profiles and χ^2 values indicate that the agreement with experiments is better in the case of cationic MC3 at low pH than for neutral MC3 at high pH (Figure 5.12, Figure 5.15). It is surprising that even though all FFs predict very different distributions of neutral MC3 (Figure 5.7 A), none of them perfectly captures the measured reflectivity profiles. Possible sources of deviations could be (i) our assumption of fully neutral MC3 at ph=7.0, (ii) the absence of explicit substrate, or (iii) imperfect force fields. Assuming the MC3 $pK_a = 6.44$ inside LNP, the Henderson–Hasselbalch equation predicts 78 % of MC3 to be neutral at pH=7.0 as opposed to our 100 % assumption. However, simulation involving 78 % of MC3 neutral and the remaining 22 % protonated did not result in significant improvement in agreement (Figure 5.17 B) indicating that the current technique is not sensitive to such detailed changes. As mentioned earlier, the substrate effect is most prominent in pure H_2O case and the good agreement between experiments and simulations at low pH might indicate that such effects are not very significant. However, the consistently lower surface coverage for neutral MC3 from our modeling protocol could be an indication of a pH-dependent substrate effect, since silica substrate is known to exhibit a rich pH-dependent surface chemistry [183]. Finally, improvement in the force field could lead to a better agreement. A starting point would be, the dihedral angles which were not found in the Lipid17 force field and were taken from the GAFF force field, these dihedrals are indicated in the `itp` files (available at: https://github.com/bio-phys/ForceFieldsMC3). Accurate dihedral angles are crucial for accurately reproducing lipid properties, in fact, the improvement of the dihedral angle was one of the important factors for the accuracy of the Lipid14 force field compared to earlier versions. Furthermore, MC3 has a non-standard head group therefore, accurate Lennard-Jones parameters for the head group based on thermodynamic quantities like heat of vaporization and NMR order parameter could further improve the agreement with experiments. Finally, to validate structures in detail SAXS experiments would be a better choice as the SAXS form factors can resolve intricate differences between different force fields much better compared to reflectivity as they can be measured up to much higher q-values (Figure 5.18). Also, neutron reflectivity with selective deuteration of MC3 could better elucidate the extent of MC3 accumulation in the bilayer center since as shown in the previous chapter signal from the deuterated part, MC3 in this case would be highly enhanced.
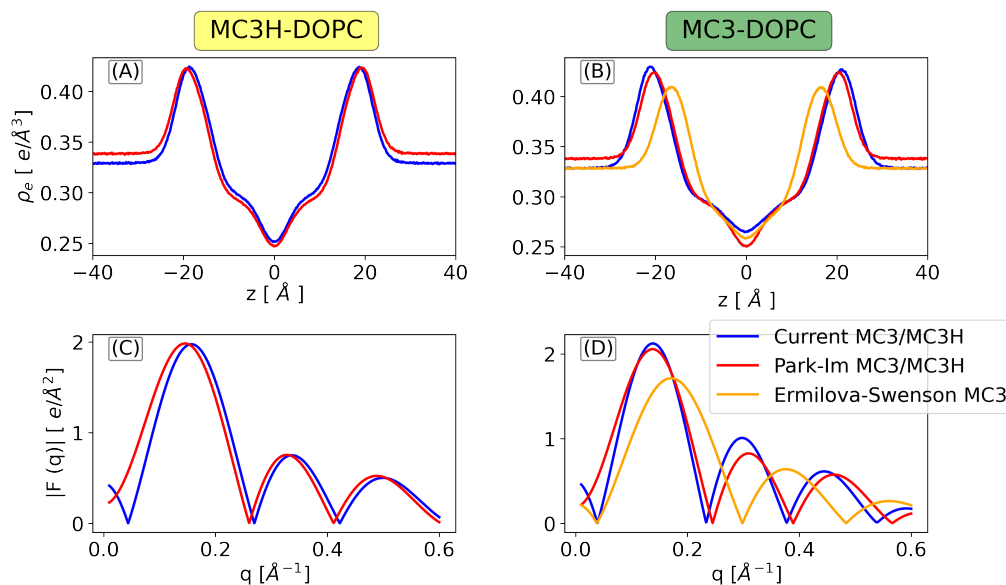
Figure 5.18: X-ray form factors from the simulations with the different force fields. (A, B) Electron densities for MC3H-DOPC and MC3-DOPC systems. (C, D) Calculated X-ray scattering form factors F(q)

5.5 Conclusions

In this chapter, we derived force field parameters for cationic and neutral MC3. We performed a direct comparison of current and existing models of MC3 with neutron reflectivity experiments performed on MC3/DOPC bilayers. The current and CHARMM-based FF displayed equally good agreement with experiments whereas the Slipid-based Ermilova-Swenson model underestimated the thickness of the bilayer at high pH. The combination of simulation and experiments gives invaluable structural insights into the pH-dependent behavior of MC3 in DOPC. At low pH, the cationic MC3 stays at the bilayer water interface whereas at high pH the neutral MC3 segregates into the bilayer center increasing its thickness. Such results on these model systems have significant implications for the behavior of these lipids in the complex LNP environment. Furthermore, the compatibility of the force fields with the AMBER family of force fields, which is widely used for simulating nucleic acids, enables accurate simulations of MC3-RNA interactions which is crucial for rational drug design.

5.6 Appendix

5.6.1 Data archive

The simulation input file for all the simulations presented in this chapter is archived in
the following tree. Go to any of the directories e.g. `./MC3-DOPC/MC3-15%` and run the
command

```
gmx grompp -f ../../mdp/md-pr.mdp -c npt.gro -p mc3.top -n index.ndx \
-o npt.tpr
```

```
$HOME/Data-Archive-Mohd_Ibrahim/MC3-DOPC-bilayers/
├── Amber
│   ├── MC3-DOPC
│   │   ├── MC3-15%
│   │   │   └── index.ndx   mc3.top   npt.gro
│   │   └── MC3-5%
│   │       └── index.ndx   mc3.top   npt.gro
│   ├── MC3H-DOPC
│   │   ├── MC3-15%
│   │   │   └── index.ndx   mc3.top   npt.gro
│   │   ├── MC3-10%
│   │   │   └── index.ndx   mc3.top   npt.gro
│   │   └── MC3-5%
│   │       └── index.ndx   mc3.top   npt.gro
│   └── forcefield
│       ├── Lipid17-dopc.ff
│       ├── amber_na.ff
│       └── nschwierz.ff
├── Charmm
│   ├── MC3-DOPC
│   │   │   └── index.ndx   mc3.top   npt.gro
│   │   └── toppar
│   ├── MC3H-DOPC
│   │   │   └── index.ndx   mc3.top   npt.gro
│   │   └── toppar
│   │
│   └── nschwierz.ff
├── Slipids
│   ├── MC3-DOPC
│   │   └── index.ndx   mc3.top   npt.gro
│   ├── slipids.ff
│   └── nschwierz.ff
├── mdp
└── Experiments
```

5.6.2 Simulation convergence

To obtain average values of observables from correlated data and also to check the conver-
gence of simulation block averaging method is used [184]. In this method, the time series
for an observable is divided into blocks of size n, and the average value of the observable
in each block is calculated. Let σ_n be the standard deviation among the averages of M
blocks each with length n. The overall standard error in the observable from M blocks is
given by, [185],

$$\sigma^2 = \frac{\sigma_n^2}{\sqrt{M-1}} \tag{A5.1}$$

The block size n is increased i.e. the number of blocks M is decreased and the above equation is evaluated for each n. Beyond a certain block size, the data becomes uncorrelated, and σ^2 saturates which is taken as the standard error for the observable. The variance in σ^2 is given by [55]

$$\sqrt{\frac{2\sigma^4}{(M-1)^3}} \tag{A5.2}$$

For the area per lipid (A), σ^2 with its variance is plotted against the block size and it saturates after around 75 ns (Figure A5.1C).
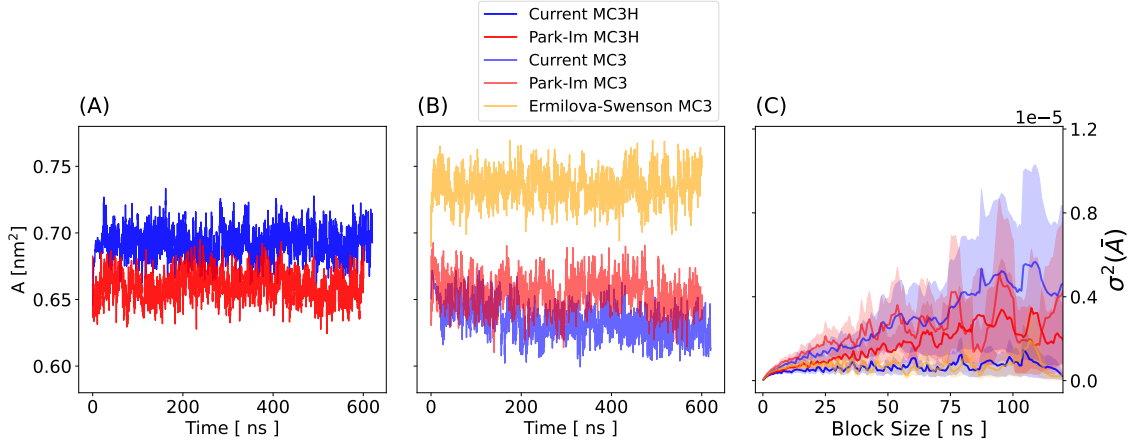
Figure A5.1: Time trace of the area per lipid, A, for the cationic MC3H/DOPC systems (A) and the neutral MC3/DOPC system (B). (C) Result from block averaging for the standard deviation of the ensemble average of the area per lipid (σ^2) as a function of the block size. The shaded region indicates the statistical error of σ^2.

5.6.3 Characteristic lipid layer thicknesses D_{HH}, $2DC$ and D_B from different force fields

Here we report characteristic bilayer thicknesses, D_{HH}, $2DC$ and D_B (Figure A5.2). The bilayer thickness, D_{HH} is defined as the distance between the two highest peaks of the total electron density (Figure A5.2 A). The hydrophobic thickness $2DC$ is defined as the distance between the two points where the DOPC hydrocarbon tails electron density is half of the maximum value [162] (Figure A5.2 A). The Luzzati thickness (D_B), is defined as [65],

$$D_B = d_z - \int_{-d_z/2}^{d_z/2} p(\mathrm{d})d\mathrm{d} \tag{A5.3}$$

Where $p(d)$ is the water probability distribution along z and is defined as $p(d) = n_w(d)V_w/dV$, where, n_W is the number of water molecules in the slice with volume dV, V_w volume of a single water molecule ($V_w \sim 30.5$ Å^3). d_z is the z-dimension of the simulation box. Further differences between MC3 added and pure DOPC setups are shown in Figure A5.3.
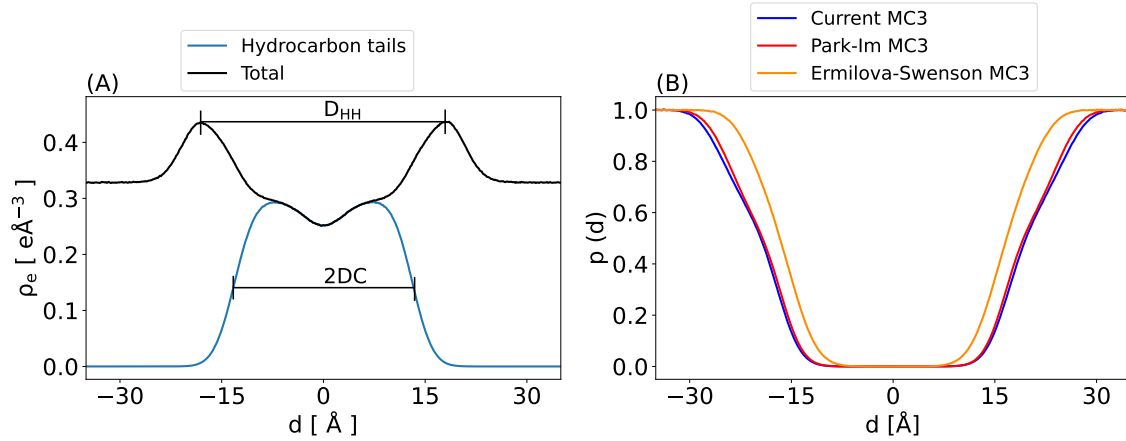
Figure A5.2: (A) Definition of D_{HH} and $2DC$. (B) Normalized probability distribution of water for the different force fields.

State	System	Current MC3/H (15%) and Lipid 17 DOPC	Park-Im MC3/H (15%) and CHARMM36 DOPC	Ermilova-Swenson MC3 (15%) and Slipids DOPC
Reference	Pure DOPC	34.9 ±0.4 Å	37.9±0.5 Å	37.3±0.7 Å
Neutral	MC3/DOPC	39.7 ±0.3 Å	38.7±0.7 Å	34.1±0.5 Å
Charged	MC3H/DOPC	35.7 ±0.5 Å	38.6±0.5 Å	N/A

Table A5.1: Luzzati thickness (D_B) for the different systems using different force fields. Errors correspond to standard deviations obtained from block averaging (4 blocks for the last 400 ns of the simulations).

State	System	Current MC3/H (15%) and Lipid 17 DOPC	Park-Im MC3/H (15%) and CHARMM36 DOPC	Ermilova-Swenson MC3 (15%) and Slipids DOPC
Reference	Pure DOPC	27.0 ±0.3 Å	28.9±0.5 Å	27.9±0.1 Å
Neutral	MC3/DOPC	32.5 ±0.3 Å	30.6±0.3 Å	24.3±0.3 Å
Charged	MC3H/DOPC	28.6 ±0.4 Å	29.9±0.3 Å	N/A

Table A5.2: DOPC hydrocarbon thickness (2DC) for the different systems using different force fields. Errors correspond to standard deviations obtained from block averaging (4 blocks for the last 400 ns of the simulations).
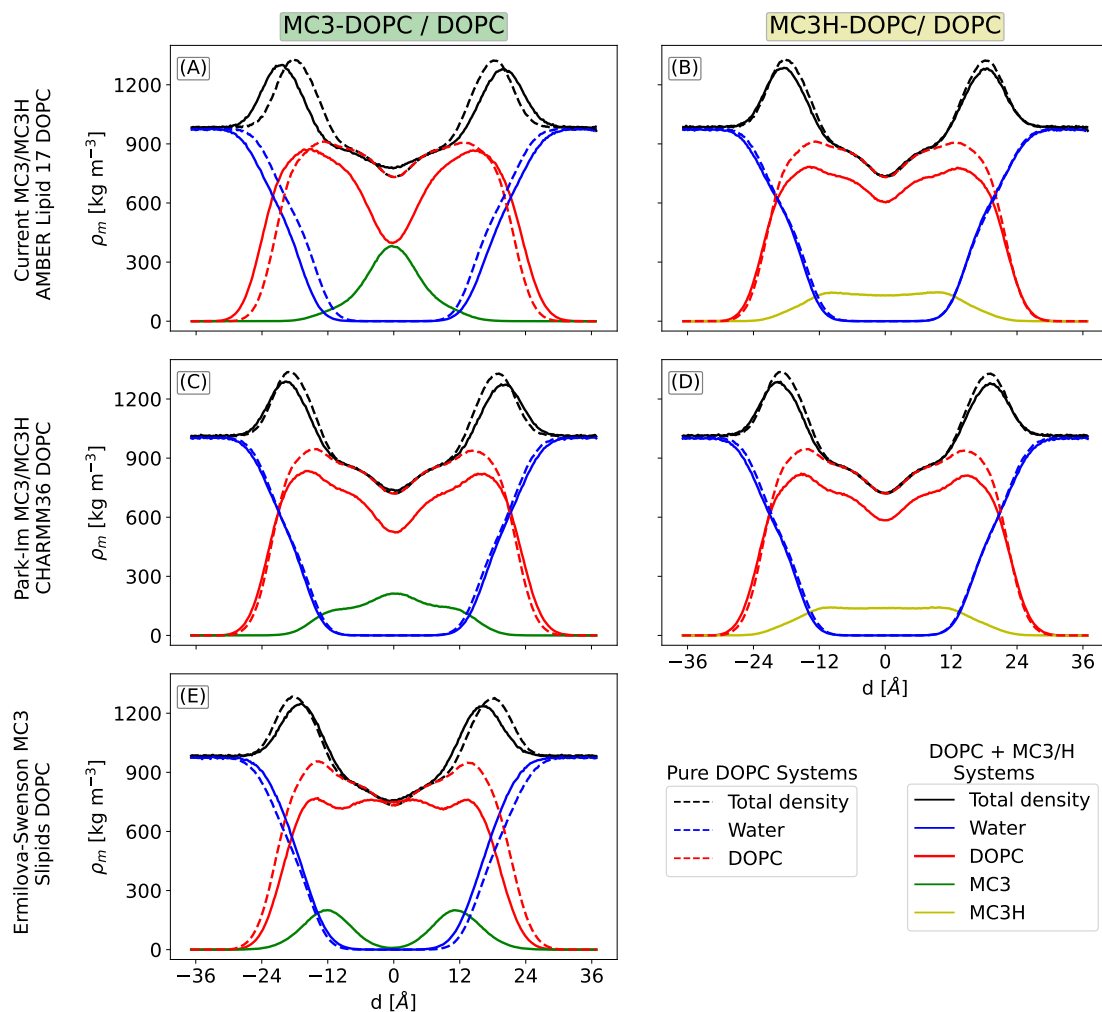
Figure A5.3: Mass density profiles for pure DOPC systems and systems with mixtures of DOPC and cationic or neutral MC3 for the different force fields. The dashed lines in all cases correspond to a pure DOPC setup and the solid lines correspond to the DOPC-MC3/H setups with 15% MC3/H. The density of neutral MC3 is shown in green, cationic MC3H is shown in yellow, water is shown in blue, and DOPC is shown in red. In each panel, for a given component, the density in both MC3/H-DOPC and pure DOPC systems is plotted.

A Methodology for Accurate Treatment of Protonation Degree and Packing of Monolayers with Ionizable Lipids

The work presented in this chapter was conducted under the supervision of Prof. Dr. Nadine Schwierz at the Institute of Physics, University of Augsburg, Augsburg, Germany. The simulations were performed on the GOETHE HLR of the state of Hessen. The experiments were mainly performed by Miriam Grava and Prof. Dr. Emanuel Schneck at the Institute for Condensed Matter Physics, TU Darmstadt, Darmstadt, Germany. Major aspects of this work are presented in the manuscript [186]:

· Miriam Grava, **Mohd Ibrahim**, Akhil Sudarsan, Julio Pusterla, Julian Philipp, Joachim Rädler, Nadine Schweirz, Emanuel Schneck . "Combining molecular dynamics simulations and x-ray scattering techniques for the accurate treatment of protonation degree and packing of ionizable lipids in monolayers". *J. Chem. Phys.,* *159, 154706 (2023)* (Featured Article)

6.1   Introduction

The structure of ionizable lipid-based LNPs is highly dependent on the pH of the system [180, 187]. In fact, the structural changes due to the sudden decrease of pH inside the endosome are believed to be the key factor responsible for the release of RNA cargo [188]. At pH above the apparent pK_a the ionizable lipids are mostly neutral whereas at pH below pK_a most of the ionizable lipids are expected to be protonated i.e. positively charged. The pK_a of the lipid however is dependent on its chemical and electrostatic environment, therefore the fraction of protonated lipids in a given setup at a given is not known a priori. For instance, pK_a of ionizable lipid Dlin-MC3-DMA in infinite dilution is reported to be around 9.4 [189] whereas inside LNP it is shifted to 6.4 [51]. In the previous chapter, we assumed all MC3 to be protonated at low pH ($<pK_a$) and all neutral at high pH ($>pK_a$) assuming the pK_a inside an LNP. For accurate simulation at a given pH, the degree of protonation has to be assigned.

Many approaches have been proposed to perform constant pH MD simulations for both coarse-grained and all-atom simulations. Many of these methods are based on the so-called $\lambda-$dynamics, where an additional degree of freedom λ, corresponding to each ionizable group, is introduced in the Hamiltonian to take into account the protonation or deprotonation of the ionizable groups[190–192]. Other methods introduce explicit particles to represent the protons which depending on the electrostatic environment and pH can bind or unbind from the ionizable groups [193]. Most of these methods do not scale linearly with the number of ionizable groups [190, 191], and also require evaluation of expensive
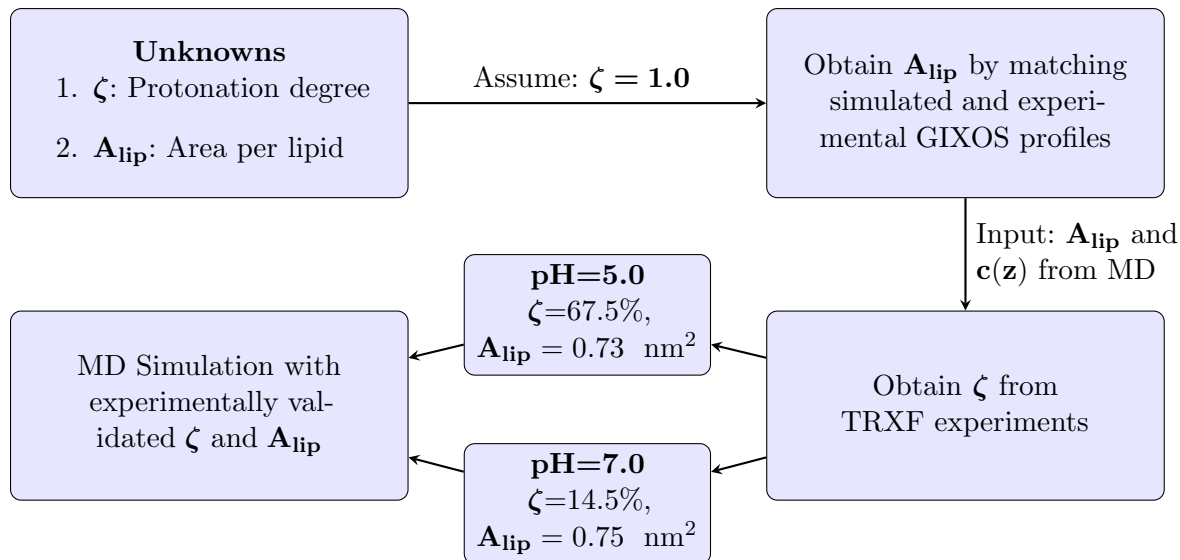
Figure 6.1: Protocol to assign protonation degree

free energy difference between the protonated or deprotonated state and a reference state beforehand. For proteins, such methods have been applied successfully [191, 192], although computationally expensive and exhibit slow convergence. For systems with ionizable lipids, such approaches might be even more challenging due to a large number of ionizable sites (equal to the number of lipids) and complex pH-dependent phase transitions [188] which may lead to very slow convergence.

Here, we introduce a methodology to assign the degree of protonation and packing of monolayers containing ionizable lipids. First, we optimize the area per lipid which quantifies the lateral packing of lipids in the monolayer by combining molecular dynamics simulations and grazing incidence x-ray off-specular reflection (GIXOS). Subsequently, by combining simulations at the optimum area per lipid and total reflection X-ray fluorescence (TXRF) we measure the excess ions bound to the monolayer from which the protonation degree is extracted. Finally, we perform simulations with the determined correct area per lipid and degree of protonation to gain structural insights into MC3-POPC monolayers.

6.2   Methods

The general protocol that combines MD simulations and scattering experiments to obtain the protonation degree (ζ) is outlined in Figure 6.1. In the first step, we assume $\zeta = 100.0\%$ i.e. all MC3 is charged and obtain the area per lipid (A_{lip}) from simulations. A_{lip} is obtained by matching simulation and experimental GIXOS profiles. The optimized A_{lip} and the concentration profile from the simulation are used as input for the TRXF experiments from which the correct value of ζ is estimated at each pH. Finally, simulations are performed using the optimum A_{lip} and correct ζ value to obtain structural insight into POPC/MC3 monolayers. The simulation and experimental details are described below.

6.2.1 Experimental methods

The experimental setup consisted of MC3/POPC monolayer on a Langmuir trough. This
setup allows accurate control of surface pressure (π) which is related to the surface tension
of pure water (γ_0) and with monolayer (γ) as,

$$\pi = \gamma_0 - \gamma \tag{6.1}$$

A monochromatized synchrotron X-ray beam with energy 15 keV (wavelength, $\lambda =$
0.827 Å) was incident on the surface at an angle, $\theta_i = 0.07°$ which is just below the
critical angle ($\theta_c = 0.082°$) for total reflection. Two complementary measurements namely
GIXOS (grazing incidence X-ray off-specular scattering) and TRFX (total reflection X-ray
fluorescence) were performed simultaneously.

GIXOS

As mentioned before, GIXOS measures the q_z dependence of the diffuse scattering intensity,
$I(q_{xy} \neq 0, q_z)$. Like reflectometry in the previous chapter, GIXOS has all the information
about the transverse scattering length density (electron density in this case). The intensity
measured in GIXOS is given by [88],

$$I(q_{xy} \neq 0, q_z) = I(q_z) = R(q_z)\frac{V(q_z)}{R_{\mathrm{F}}(q_z)} \tag{6.2}$$

Where, $R(q_z)$ conventional reflectivity profile, $V(q_z)$ and $R_{\mathrm{F}}(q_z)$ is the vineyard function
[87] and Fresnel reflection respectively. In the current experiment, $q_{xy} = 0.04$ Å^{-1}.

TRXF

The incident X-ray induces a fluorescence signal from the Br$^-$ ions present on the surface
due to photoelectric ionization and therefore, is highly sensitive to the interfacial accumula-
tion of ions. The intensity of the fluorescence signal recorded by a detector placed parallel
to the water surface depends on the interfacial ion distribution profile $c(z)$ as [194],

$$I_{\mathrm{F}} = A \int_{-\infty}^{\infty} c(z)\phi(z)dz \tag{6.3}$$

Where $\phi(z)$ is the standing wave profile and is obtained using a two-slab representation
of electron density profiles [195]. The slabs correspond to the heads and tails of the lipids.
A is a prefactor that is calibrated using a reference bare KBr solution for which $c(z)$ has
a known constant profile [194, 195]. I_F is modeled by a Gaussian function with the mean
constrained to the characteristic K_α ($\approx$ 12 keV) or L_β($\approx$ 1.5 keV) emission lines of Br$^-$
[196].
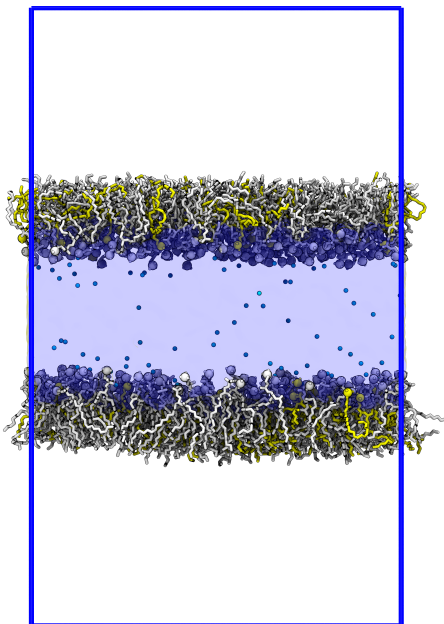
Figure 6.2: MC3/POPC monolayer simulation setup. A water slab (blue background) is placed between two monolayers. Here colors are: white (POPC), yellow (MC3H), and cyan spheres (Cl^- ions). The white and yellow spheres represent the Nitrogen atoms in the head group of respective lipids.

6.2.2 Computational methods

System setup

We simulated MC3/POPC monolayers with a mole fraction of 20% MC3 and 80% POPC. The initial monolayer setup was prepared by translating the leaflets of an equilibrated bilayer which in turn was created using the MemGen web server [169]. A snapshot of the simulation system is shown in Figure 6.2. Each monolayer contains 160 POPC, 40 MC3 molecules, 60 water molecules per lipid, and chloride ions to neutralize the positive charge of cationic MC3. POPC lipids were described using the AMBER Lipid17 force field [72], for MC3, the current force field developed in this thesis was used [154]. Ions were described with the Mamatkulov-Schwierz parameters [126] and the TIP3P water model was used [123].

Simulation details

The systems were energy minimized using a gradient descent algorithm. The temperature of the systems was maintained at 293.15 K using the velocity rescaling thermostat with stochastic term [58] and a time constant of 1.0 ps. For the NPT simulations, semi-isotropic Parrinello-Rahman was used [60]. The Van der Waals interaction potential was truncated and shifted to zero at 1.2 nm. Electrostatic interactions were truncated at 1.2 nm and the Particle Mesh Ewald (PME) method [57] with a Fourier grid spacing of 0.12 nm was used to evaluate the long-range electrostatics part. All bonds involving hydrogens were constrained using the LINCS algorithm [122]. A time step of 2.0 fs was used to integrate the equations of motion. Analysis was performed using Gromacs inbuilt routines and with MDAnalysis [129] and trajectories were visualized with VMD [130]. The GIXOS curves were calculated from the electron density profile using the Refnx Python package [172].

Two different sets of simulations were carried out. In the first part, we obtained the optimum area per lipid (A_{lip}) by matching experimental and simulations GIXOS profiles.
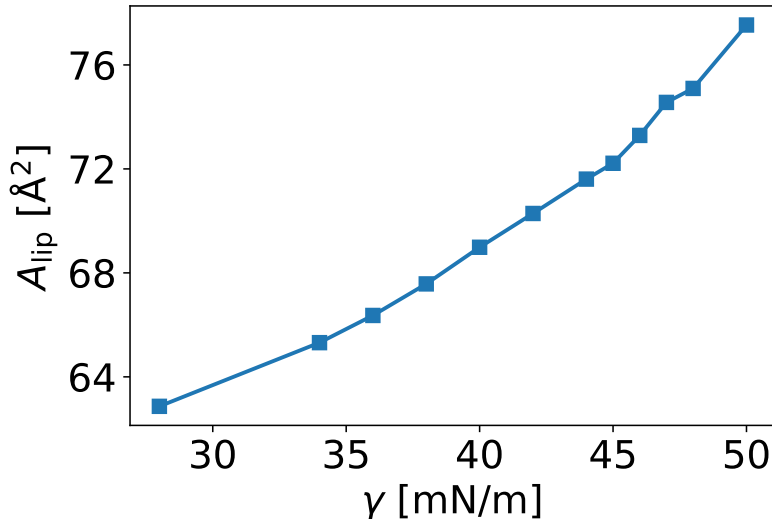
Figure 6.3: At each value surface tension (or lateral pressure), the monolayer attains an equilibrium value of A_{lip}.

The optimum A_{lip} is used as input for the experiments to obtain the correct protonation degrees (ζ) at each pH value. In the second part, simulations with correct A_{lip} and correct ζ were performed to obtain structural insights into the monolayers at each pH.

NPT simulations at different lateral pressure

To obtain the optimum area per lipid (A_{lip}) multiple simulations in semi-isotropic NPT ensemble were performed by varying the lateral components P_{xx} and P_{yy} of the pressure tensor. The semi-isotropic barostat ensures $P_{xx} = P_{yy}$. These pressure components are varied from -50 bar to -34 bar which corresponds to varying the surface tension from 34.0 mN/m to 50.0 mN/m. A total of 12 simulations each 100 ns long were performed. The surface tension can be obtained from the pressure tensor components as [197, 198],

$$\gamma = \frac{L_z}{n} \left[P_{zz} - \frac{P_{xx} + P_{yy}}{2} \right] \tag{6.4}$$

where $n = 2$ is the number of interfaces in the simulation box. The P_{zz} component was kept constant (~ 0) by setting the compressibility in $z-$direction to zero. At each value of surface tension, the system attains an equilibrium value of A_{lip} while the box height in $z-$direction remained fixed at $L_z = 20$ nm. The variation of A_{lip} with surface tension (or lateral pressure components) is depicted in Figure 6.3. The optimum simulations parameters obtained for the two different pH values are shown in Table 6.1.

NVT simulations with correct A_{lip} and ζ

After obtaining the optimum value for A_{lip} and ζ simulations were performed in NVT ensemble. The size of the simulations box in x and y direction was fixed such that the A_{lip} value obtained above was maintained. The simulations were run for 500 ns. The simulation parameters are summarized in Table 6.2.

Table 6.1: Simulation parameters that best reproduce the experimental GIXOS curves for $\alpha = 1.0$.

pH	Fixed Parameters			Calculated Parameters				
	ζ	L_z	P_{xx}/P_{yy}	P_{zz}	L_x/L_y	A_{lip}	γ	π
		[nm]	[bar]	[bar]	[nm]	[nm^2]	[mN/m]	[mN/m]
5.0	1.0	20.0	46.0	-0.36 ± 0.72	12.10 ± 0.02	0.73 ± 0.002	45.9 ± 0.9	12.1 ± 0.9
7.5	1.0	20.0	48.0	1.26 ± 1.0	12.25 ± 0.02	0.75 ± 0.003	49.8 ± 1.2	8.2 ± 1.2

Table 6.2: Parameters used for the simulations with experimentally validated A_{lip} and ζ.

pH	Fixed Parameters				Calculated Parameters				
	ζ	L_x/L_y	L_z	A_{lip}	P_{xx}	P_{yy}	P_{zz}	γ	π
		[nm]	[nm]	[nm^2]	[bar]	[bar]	[bar]	[mN/m]	[mN/m]
5.0	0.675	12.10	20.0	0.73	-47.8 ± 0.35	-48.2 ± 0.4	-0.36 ± 0.46	47.6 ± 0.6	10.4 ± 0.6
7.5	0.150	12.25	20.0	0.75	-48.8 ± 0.6	-49.2 ± 0.8	-0.42 ± 0.56	49.1 ± 0.8	8.9 ± 0.8

6.3   Results

6.3.1   Matching area per lipid from experiments and simulations

In the first part, we optimized the lateral packing of lipids in the monolayer which is quantified by the area per lipid (A_{lip}). For this optimization, the degree of protonation was assumed to be 100% at both pH 5.0 and pH 7.5 The GIXOS curves for simulations at each value of A_{lip} were calculated and compared with the experimental profile. The electron density profiles at low A_{lip} values are more extended in the $z-$direction compared to high values of A_{lip} (Figure 6.4 A) indicating relatively ordered, tightly packed and a thicker monolayer at low A_{lip} values. The GIXOS profiles are highly sensitive to such thickness changes as indicated by a clear systematic shift of characteristic peaks between q_z=0.3 Å^{-1} and 0.5 Å^{-1} to low q_z values with increasing monolayer thickness i.e. lower A_{lip} (Figure 6.4 B).

The value of A_{lip} which minimizes the χ^2 deviation between the experimental GIXOS profile (I_{exp}) and the simulation profile (I_{sim}) is taken as the optimum value. χ^2 is defined as follows:

$$\chi^2 = \sum_{q_z} \frac{(I_{\text{sim}}(q_z) - I_{\text{exp}}(q_z))^2}{N} \tag{6.5}$$

Where N is the number of data points.

For both pH values, we obtain global minima for χ^2 within the specified range of A_{lip} (Figure 6.5A). The optimum values of A_{lip} are 73 Å^2 and 75 Å^2 for pH 5.0 and pH 7.5 respectively which is in the typical range for lipid monolayers at comparable lateral pressure [199]. The GIXOS profiles corresponding to the optimum A_{lip} for the two pH values show excellent agreement with experiments (Figure 6.5B).

The experiments are performed at a fixed value of surface pressure, $\pi = 30$ mN/m which according to Eqn. 6.1 corresponds to a surface tension of 28 mM/m with $A_{\text{lip}} = 63$ Å^2. This surface tension is calculated assuming $\gamma_0 = 58$ mN/m for TIP3P water model at 293.15 K as obtained by extrapolating data from Vega et. al. The GIXOS profile corresponding to $A_{\text{lip}} = 63$ Å^2 depicted by dashed black line deviates significantly from the experimental
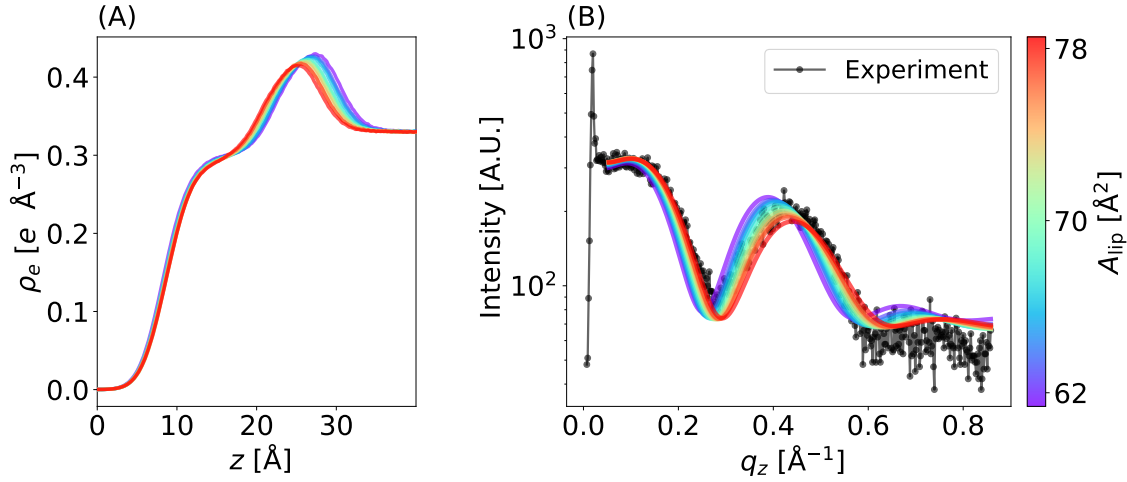
Figure 6.4: (A) Electron density from simulations and (B) corresponding GIXOS profiles for different values of area per lipid, A_{lip}. Here, the experimental GIXOS profile corresponds to pH 5.0.

profile.
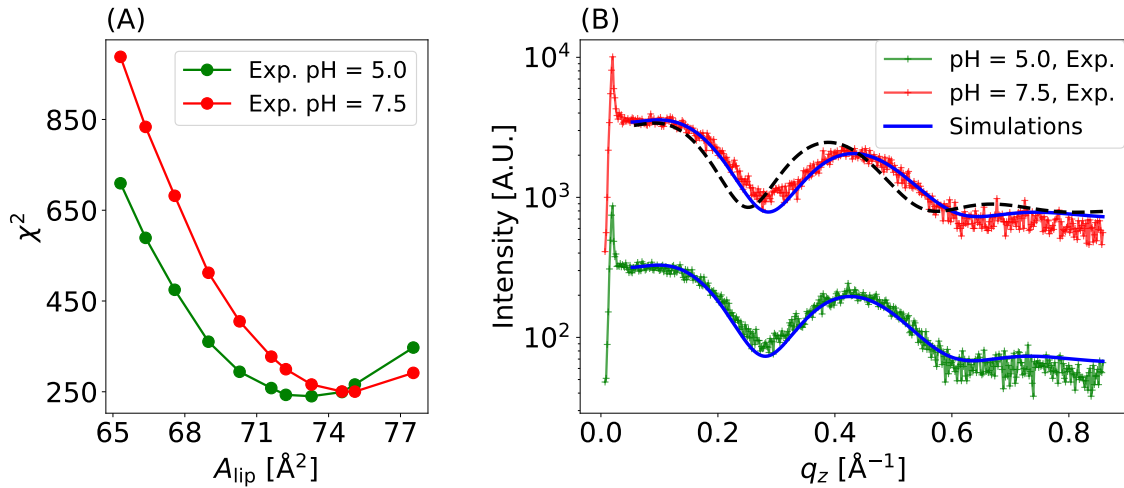
Figure 6.5: (A) χ^2 deviation between experiments and simulations for different values of A_{lip}. The optimum values of A_{lip} are 73 Å^2 and 75 Å^2 for pH 5.0 and pH 7.5 respectively. (B) Comparison of experimental and simulation GIXOS curves at optimum A_{lip}. The dashed black line corresponds to a setup at the same surface pressure ($\pi = 30$ mN/m) as in experiments with $A_{\text{lip}} = 63$ Å^2.

6.3.2 Assigning protonation degree [1]

The degree of protonation (ζ) of a titratable group, with a given pK$_a$ value, at given pH in the infinite dilution limit is given by the Henderson–Hasselbalch (HH) equation as,

$$\zeta = \frac{1}{1 + 10^{\text{pH}-\text{pK}_a}} \tag{6.6}$$

However, the pK$_a$ of a molecule is highly dependent upon its electrostatic environment. For a lipid monolayer/bilayer system with a surface potential (Φ_0), the degree of protonation

---

[1]The analysis of TRXF spectra to assign protonation degree was performed by Miriam Grava and Prof. Dr. Emanuel Schneck

of ionizable groups at the surface can be more accurately predicted as [200],

$$\zeta = \frac{1}{1 + 10^{\mathrm{pH}-\mathrm{pK_a}+\frac{e\Phi_0}{k_{\mathrm{B}}T\ln 10}}} \qquad (6.7)$$

Where, T is the temperature, e is the elementary charge, and k_{B} is the Boltzmann constant. The pK_a of MC3 in the infinite dilution limit is reported to be around 9.4. Using this value, the HH equation predicts ζ=100 % at both pH 5.0 and pH 7.5. Using Eqn.6.7 gives $\zeta = 99\%$ for pH 5.0 and 68% for pH 7.5. Here we used $\Phi_0 = 92.5$ mV for pH 5.0 and $\Phi_0 = 33.5$ mV for pH 7.5 [2].
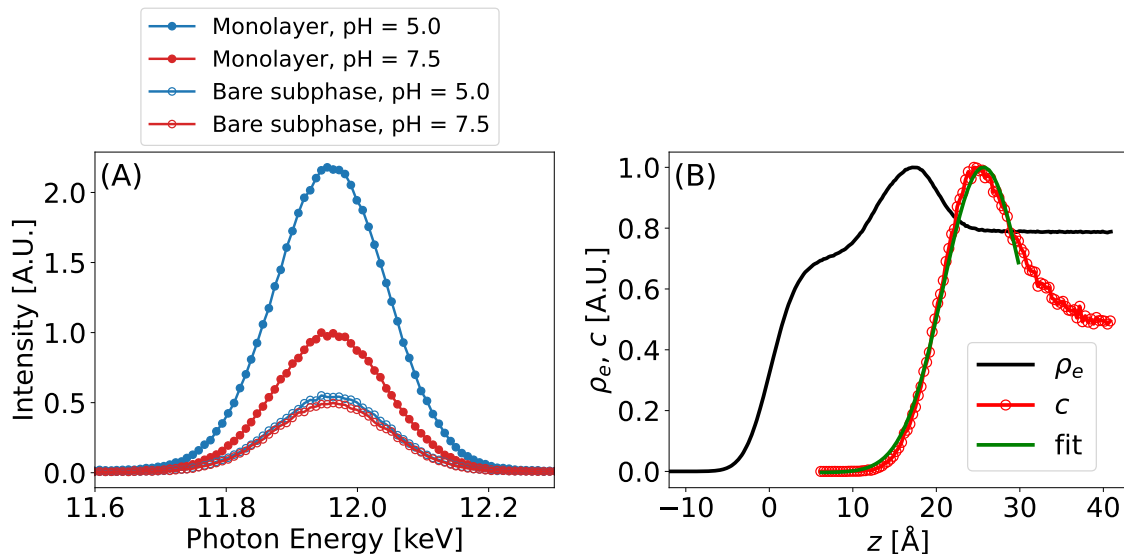
Figure 6.6: (A) Total reflection X-ray fluorescence spectra corresponding to K_α emission lines of Br^- for monolayer setup at different pH values. The spectra for subphase i.e. without monolayer is shown by open circles. (B) Cl^- ion density from simulation and the corresponding Gaussian fit according to Eqn. 6.8. The corresponding electron density is also depicted.

We used TRXF to assign the protonation degree for the ionizable MC3 at a given pH. The intensity of the measured fluorescence signal around the characteristic emission lines (K_α and L_β) of Br^- is directly proportional to the excess surface bound Br^- ions which in turn depends on the degree of protonation of the ionizable group. The fluorescence spectra for the monolayer exhibit higher intensity at lower pH values, indicating a significant proportion of protonated MC3 molecules while the relatively lower intensity at high pH suggests a reduced fraction of protonated MC3 molecules (Figure 6.6A). For reference, the spectra for the bare subphase i.e. without the monolayer are also depicted (Figure 6.6A). To obtain the protonation degree, the measured fluorescence spectra were modeled using Eqn. 6.3. The two required functions are $\phi(z)$, the standing wave profile, and $c(z)$, the Br^- ion concentration profile: $\phi(z)$ was obtained from two slab model representations of the electron density profile. $c(z)$ was modeled as a truncated Gaussian as,

$$c(z) = c_0 + (c_{\max} - c_0)e^{-(z-z_{\max})^2/(2\sigma^2)}. \qquad (6.8)$$

Where c_0 is a constant experimental bulk concentration. The parameters for the above

---

[2]The calculation of Φ_0 was performed by Akhil Sudarsan

function were obtained by fitting the simulation Cl^- ion density profile (Figure 6.6B).
We obtained $z_{\max} = z_0 + 0.56$ nm; $\sigma = 0.48$ nm for pH 5.0 and $z_{\max} = z_0 + 0.53$ nm;
$\sigma = 0.47$ nm for pH 7.5. Here, z_0 defines the interface between the tail and head group
slabs. The only unknown now in $c(z)$ is c_{max}, which is obtained by fitting the experimental
spectra (Figure 6.6A) with Eqn. 6.3. The excess counter ions accumulated on the surface
compared to the bulk is then given by,

$$\Gamma = \int_{z_0}^{\infty} [c(z) - c_0]\, dz. \tag{6.9}$$

We obtain $\Gamma = 0.185$ nm^{-2} for pH 5.0 and $\Gamma = 0.039$ nm^{-2} for pH 7.5. Assuming a one-to-
one correlation between excess surface Br^- ions and the number of protonated MC3, the
protonation degree ζ is given by,

$$\zeta = \frac{\Gamma A_{\mathrm{lip}}}{f_{\mathrm{MC3}}} \tag{6.10}$$

where $f_{\mathrm{MC3}} = 0.2$ is the MC3 mole fraction. We obtain $\zeta = 67.5$ % for pH 5.0 and $\zeta = 14.5$ %
for pH 7.5.

6.3.3 Simulation at correct A_{lip} and ζ

To gain insights into the structure of MC3/POPC monolayers at different pH values, we
performed simulations using the correct values of A_{lip} and ζ obtained above. The box size
is constrained to keep the A_{lip} as obtained above fixed. The degree of protonation was
imposed by employing varying numbers of neutral MC3 and protonated MC3. For instance,
at $\zeta = 67.5\%$, 27 out of 40 MC3 molecules per monolayer were charged, while the remaining
molecules were neutral. Similarly, for $\zeta = 14.5\%$, 6 out of 40 MC3 molecules were charged,
while the rest remained neutral.

The snapshot of the setup corresponding to $\zeta = 67.5\%$ after 500 ns and the corre-
sponding number density profiles for different components are shown in Figure 6.7. The
cationic MC3H exhibits the same behavior as POPC with the hydrophilic head group
facing toward water and the hydrophobic tail group pointing away from it. The neutral
MC3 stays away from the interfacial water region. This is also evident from the number
density profiles where MC3 density for both head and tail parts lies in the hydrophobic
parts of the monolayer (Figure 6.7B).

The different lipid species are uniformly distributed within the monolayer plane instead
of forming clusters. This distribution is evident from the in-plane radial distribution
function (RDF) of the nitrogen atoms in the lipid head groups. The position of the first
peak in the RDF represents the nearest neighbor distance between the head group nitrogen
while its magnitude gives the probability (relative to an ideal gas) of finding them at that
distance. The RDF between POPC head groups shows a first peak at approximately 0.9
nm and exhibits regular fluid behavior. However, the RDF between MC3H head groups has
a first peak magnitude below one, suggesting that the positively charged MC3H molecules
are separated due to electrostatic repulsion. In the case of MC3, the RDF peaks are also

below one, and the RDF does not approach zero at very low distances. This is because MC3 head groups can occupy different positions in the z-axis and can in principle occupy the same positions in the $xy-$plane. The lower than one nearest neighbor peak in the MC3 case may be attributed to steric clashes caused by the various possible conformations they can assume, unlike MC3H and POPC, where the head groups remain in the interfacial water region. The cross-RDFs between MC3H and POPC nitrogen exhibit behavior similar to that of POPC/POPC nitrogen, as both species occupy the interfacial water region. However, between MC3 and POPC or MC3 and MC3H, the RDFs remain closer to 1 even at smaller distances due to their different positions in the z-axis.
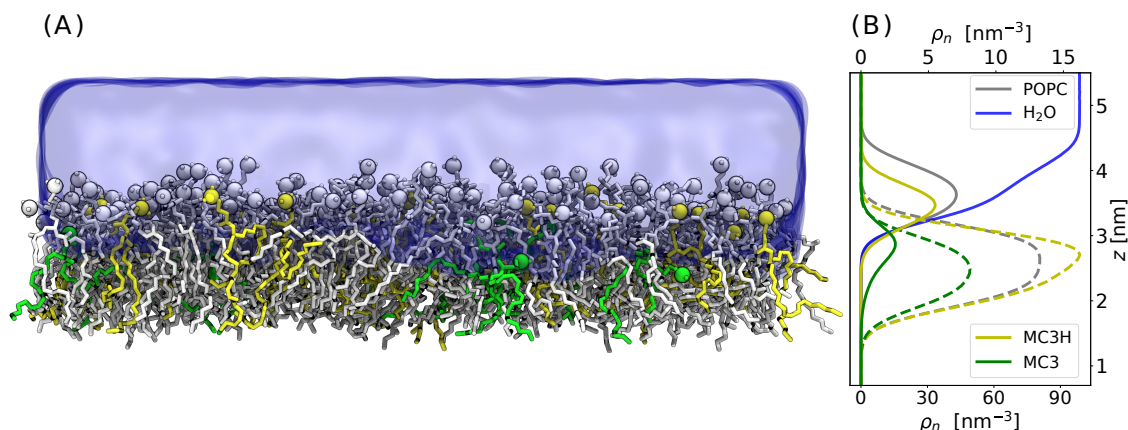
Figure 6.7: (A) Side view snapshot of monolayer systems with $\zeta = 67.5\%$. Colors are: green (MC3), yellow (MC3H), white (POPC), blue background (water). The head group nitrogen is shown as spheres. (B) Number density profile for MC3, MC3H, POPC, and water. For the lipids, The dashed lines correspond to the tail number density and the full lines correspond to the head number density. The lower $x-$scale corresponds to water and POPC head and tail groups whereas the upper $x-$scale corresponds to MC3 and MC3H head and tail groups.

Figure 6.8: (A) Top view snapshot of the monolayer at $\zeta = 67.5\%$. colors are: green (MC3), yellow (MC3H), and white (POPC). The head group nitrogen in each lipid is shown with spheres. (B) RDF between head group nitrogen atoms of lipids of the same species (C) RDF between head group nitrogen of lipids of different species.

6.4   Discussion

The pH-dependent protonation state is the key property that makes lipids like MC3 useful for applications in drug delivery [51]. The pK_a of a molecule and hence its protonation

degree is dependent on its chemical and electrostatic environment. In this chapter, we devised a method to assign protonation degrees by combining MD simulations and scattering experiments.

The first part of the method includes obtaining the area per lipid from simulations assuming all MC3 to be protonated ($\zeta = 100\%$). To that end, simulations were run at different surface tension values and the one that best reproduced the GIXOS curves was considered as the optimum value of area per lipid. From the simulations, the optimum surface tension (γ) was found to be around 47.6 mN/m and 49.1 mN/m for pH 5.0 and pH 7.5 respectively (Table 6.1). In the experiments $\gamma = 42$ mN/m as obtained using $\pi = 30$ mN/m, $\gamma_0 = 72$ mN/m in Eqn.6.1. Therefore, the surface tension in simulations and experiments is quite close with just a difference of around ~ 7 mN/m. However, when considering surface pressure, in experiments $\pi = 30$ mN/m whereas in simulations, assuming γ_0 for TIP3P water to be 58 mN/m at 293.15 K (obtained by extrapolating data by Vega et. al [198]), we obtain $\pi \approx 8$ mN/m and $\pi \approx 12$ mN/m, which is 2-3 times lower than in experiments. This disagreement most likely arises from the inaccuracy of the TIP3P water model [201] which underestimates water surface tension by around ~ 20 mN/m [198]. Therefore, water models that better reproduce air-water surface tension e.g., the OPC model might lead to a closer agreement with experiments [202].

The chloride ion density and optimum area per lipid from the simulations were used to assign the protonation degree. We obtained 67.5 % and 14.5% charged MC3 at pH 5.0 and pH 7.5 respectively. These values are significantly different from the predictions of the Henderson–Hasselbalch equation (Eqn. 6.6) as well as its modified form which takes into account electrostatic environments (Eqn. 6.7). Both these methods predict ~ 100 % charged MC3 at pH 5.0, which suggests that pK_a shift of a molecule arises from many intricate and convoluted effects like electrostatic environment, molecular conformation, and partial dehydration [203, 204] of lipids which are not accounted for by simple theories.

The GIXOS curves from simulations at the correct area per lipid as well as the correct protonation degree do not significantly differ from the GIXOS curves obtained assuming a protonation degree of 100 % (Figure 6.9A). This shows that area per lipid is the most important parameter to consider for validation of the MD simulation of monolayers. Both GIXOS curves and the normalized Cl$^-$ ion densities with the updated protonation degree do not exhibit significant differences from the initial guess of 100 % protonation degree (Figure 6.9). In case of deviation, additional iterations of the protocol would have been necessary until convergence is achieved.

Finally, the simulation at the correct area per lipid and protonation degree MC3H behaves similarly to POPC with the head group facing towards water and tail group away from it. MC3 tends to reside in the hydrophobic water-free region. This behavior is consistent with the observation in the previous chapter for bilayers where neutral MC3 accumulated in the middle of the bilayer and protonated one stayed at the bilayer water interface.
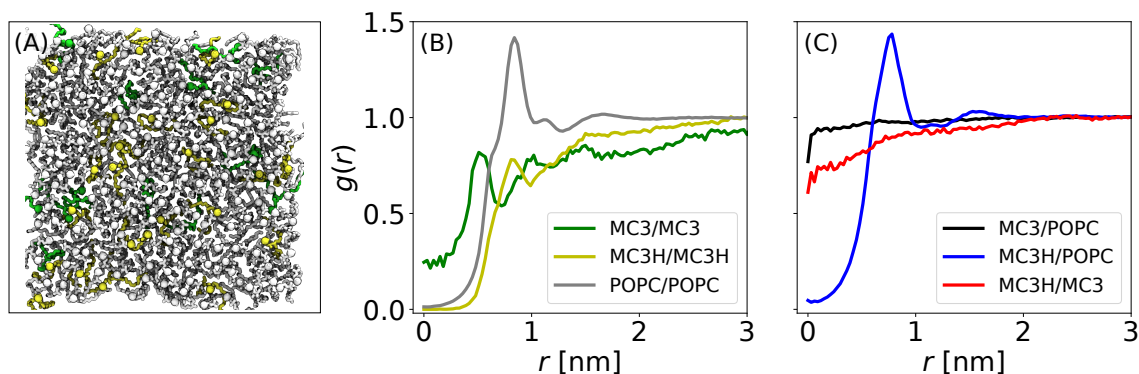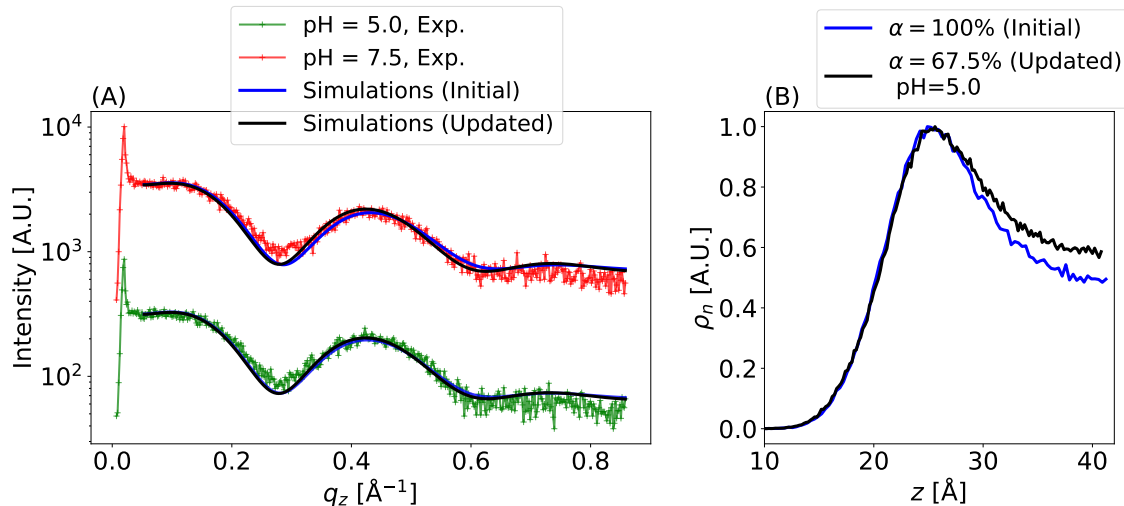
Figure 6.9: (A) Top view snapshot of the monolayer at $\zeta = 67.5\%$. colors are: green (MC3), yellow (MC3H), and white (POPC). The head group nitrogen in each lipid is shown with spheres. (B) RDF between head group nitrogen atoms of lipids of the same species (C) RDF between head group nitrogen of lipids of different species.

6.5 Conclusion

The protonation degree is necessary information required for the accurate simulation of systems containing ionizable components. In this chapter, we combined all-atom molecular dynamics simulations with total reflection X-ray scattering techniques like GIXOS and TRXF to devise a protocol to obtain the degree of protonation of ionizable lipids at a given pH. We obtained 67.5% of charged MC3 at pH 5.0 and 14.5% charged at pH 7.0. These values differ significantly from the predictions of simple theories like the Henderson–Hasselbalch equation and its modified form. The knowledge of the protonation degree of ionizable lipids at a given pH is critical to understanding the pH-dependent phase behavior of these lipids. Our methodology provides a consistent method to obtain the protonation degree. The simulations at the correct protonation degree and area per lipid show that neutral MC3 tends to stay in the hydrophobic parts of the monolayer away from the monolayer-water interface whereas protonated MC3H stays at the monolayer-water interface with the head group towards the water like the POPC lipids. From the radial distribution function, we infer that the lipids are uniformly distributed in the plane of the monolayer, and no segregation is observed. These results are consistent with the observation from the previous chapter with MC3-DOPC bilayers. Further, the excellent agreement of experimental and simulation GIXOS intensities at both pH values further indicates that our MC3 force field captures the transverse lipid distribution accurately.

6.6 Appendix

6.6.1 Data archive

The simulation input file for all the simulations presented in this chapter is archived in the
following tree. Go to any of the directories e.g. `./ph-5` and run the command

```
gmx grompp -f ../mdp/nvt.mdp -c nvt.gro -p mc3.top  -n index.ndx \
-o nvt.tpr
```

```
$HOME/Data-Archive-Mohd_Ibrahim/MC3-POPC-monolayers/
├── ph-5
│   └── index.ndx   mc3.top   nvt.gro
├── ph-7
│   └── index.ndx   mc3.top   nvt.gro
├── popc-mc3h
│   ├── index.ndx   mc3.top
│   ├── npt-28 (npt.gro)
│   ├── npt-34
│   ├── npt-36
│   ├── npt-38
│   ├── npt-40
│   ├── npt-42
│   ├── npt-44
│   ├── npt-46
│   ├── npt-47
│   ├── npt-48
│   ├── npt-50
│   └── density_scattering_profiles
├── forcefields
│   ├── amber_na.ff
│   ├── lipid.ff
│   └── nschwierz.ff
└── mdp
```

Chapter 7

On the Structure of a Model Lipid Nanoparticle

The work presented in this chapter had been conducted under the supervision of Prof. Dr. Nadine Schwierz at the Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Frankfurt am Main, Germany, and at the Institute of Physics, University of Augsburg, Augsburg, Germany. The simulations were performed on the GOETHE HLR of the state of Hessen.

7.1 Introduction

Ionizable lipid-based Lipid nanoparticles (LNPs) have emerged as the most advanced delivery systems for RNA-based drugs e.g., recent vaccines for COVID-19. LNPs are typically composed of a cationic ionizable lipid that facilitates efficient encapsulation of RNA at low pH due to its positive charge, phospholipid (typically DSPC) and cholesterol that provides structural stability, polyethylene glycol lipid (PEG) lipid that controls the LNP size during the formulation process and the mRNA cargo. Current LNP formulations suffer from limitations in delivery efficiency, primarily due to an incomplete understanding of their structure and the structure-function relationship.

Many studies based on Cryo-TEM and scattering methods have been reported to elucidate the structure of LNPs and most of them have been carried out on LNPs composed of ionizable Dlin-MC3-DMA (MC3) or an analogous Dlin-KC2-DMA (KC2) lipids. Based on Cryo-TEM and coarse-grained simulations Leung et al. [205] observed that KC2/DSPC/cholesterol lipid nanoparticles have a dense high electron density core. This was attributed to the formation of a local nanostructure or inverted micellar phases inside the LNP core i.e., little water pockets or RNA surrounded by cationic MC3 head groups. Using small-angle neutron scattering techniques Arteta et al. [206] proposed a core-shell model for LNP with the DSPC and cholesterol monolayer shell constituting the surface of LNP and the ionizable MC3 and RNA residing in the core. Consistent observations were made by Sebastiani et al. [207] on similar systems. Kulkarni et al. [180] studied the effect of pH and RNA cargo on the structure of LNPs composed of KC2 using Cryo-TEM imaging. The LNPs were observed to exhibit a vesicular bilayer structure at low pH and an electron-dense core at high pH. This was attributed to the accumulation of neutral KC2 in the LNP core at high pH. Similar observations were made for MC3-based LNPs. Molecular dynamics simulations with an accurate description of ionizable component for such a system is still lacking and could provide invaluable atomistic insights into the behavior of these lipids at different pHs.

In this chapter, I will use the cationic and neutral MC3 force field developed in this thesis to study the effect of pH and RNA cargo on the LNP structure. I will start from a

low pH vesicular bilayer model of LNP from Kulkarni et al and increase the pH, adding RNA to see their effect on the LNP structure. The increase in pH is mimicked by replacing cationic MC3H with neutral MC3. To facilitate possible structural rearrangement of LNP due to the addition of RNA or increase in pH, a carefully designed setup using carbon nanotubes (CNTs) was employed that allows the exchange of water in and out of the vesicle. At low pH i.e. with cationic MC3H, the LNP (or vesicle) stays in a vesicular bilayer structure. With an increase in pH i.e., by replacing cationic MC3H with the neutral MC3, the vesicular structure is disrupted due to the migration of MC3 away from the lipid-water interface. The RNA molecules interact strongly with the inner surface of the LNP at low pH and do not have a significant effect on the vesicular bilayer structure however at high pH, RNA exhibits loose binding to the inner surface and it sets a lower bound for the inner radius which shrinks to unexpectedly small values in absence of RNA.

7.2 Methods

7.2.1 System setup

The initial setup is modeled as a lipid vesicle composed of MC3H, cholesterol, and DSPC with their mole fraction in both inner and outer shells at 50%, 38%, and 12% respectively in accordance with previous studies on such systems [180]. The initial setup was prepared with the `Bumpy.py` python script [208], which folds a regular lipid bilayer with a given composition into a vesicle. The initial radius of the vesicle, which is defined as the average position of the peak-to-peak radial electron density with respect to the bilayer center, is around 75 Å. The vesicle consisted of 1375 MC3H, 1027 cholesterol, and 332 DSPC lipid molecules of which 868, 647, and 208 respectively are in the outer monolayer shell. The lipid ratio in the inner and outer monolayer is chosen such that the volume per lipid or the lipid packing density in each shell is as close to each other as possible keeping in view earlier coarse-grained simulation results on vesicles [209]. For simulations with RNA, 10 double-stranded RNA each with 10 Uracil nucleotides and the corresponding complementary sequences was used. The amount of water inside the vesicle is unknown therefore, water was allowed to move in and out of the vesicle by introducing two carbon nanotubes (CNT) each with a diameter of around 2 nm and a length of $\sim 6~nm$. The CNT has an armchair chirality with the terminal carbons functionalized with hydroxyl (OH) groups. Such a choice has been shown to impart high water permeability to the CNT and ensures the least tilt with respect to the radial vector [210]. The system was solvated with 404898 TIP3P water molecules and the initial box size was $26 \times 26 \times 26~nm^3$. Overall the initial system contained around 1.5 million atoms, a cross-section view of the initial setup is shown in Figure 7.1A. The Lipid17 forcefield was used for DSPC and cholesterol, and MC3/H was described by the currently developed force field in this thesis [154]. Ions were described with the Mamatkulov-Schwierz parameters [126] while RNA was described by the Amber99sb-ildn* [211] with parmbsc0 [212] and χ_{OL3} corrections [213]. An isotropic Parinello-Rahman or Berendsen barostat was used with the same settings as described in previous chapters. Cut-offs related to non-bonded potentials, LINCS constraints, integration time, thermostat, and simulation package are the same as described in previous chapters.

7.2.2 Simulation details

The initial vesicular setup prepared above, by folding a bilayer, has a highly energetically unfavorable configuration due to the presence of many steric clashes. Therefore, multiple steps of equilibration were necessary. In the first step energy minimization was performed using a gradient descent algorithm using a very low ($\sim 10^{-10}$) value for the Lennard-Jones interaction terms ϵ for the atoms constituting the vesicle. ϵ was increased to the correct value in three minimization setups. Subsequently, the system was simulated for 25 ns with varying amounts of position restraints on the vesicle components and the CNTs while water and ions were allowed to move freely. During this time, water and ions quickly flow through the CNT and the system proceeds towards an equilibrium state as depicted by the gray shaded area in Figure 7.1B-E where the radius and number of water, Na$^+$ and Cl$^-$ inside the vesicle is plotted. Afterward, the system was allowed to evolve freely without any restraints on any components.

The electron densities were evaluated using a custom Python script based on MDAnalysis [129]. The last 75-100 ns for each system were considered for the analysis. The radial electron density is calculated as,

$$\rho_e(r) = \sum_i \frac{z_i N_i}{4\pi r^2 dr} \tag{7.1}$$

Where N_i is the number of atoms with z_i electrons in a shell (with width dr) at a radial distance r.
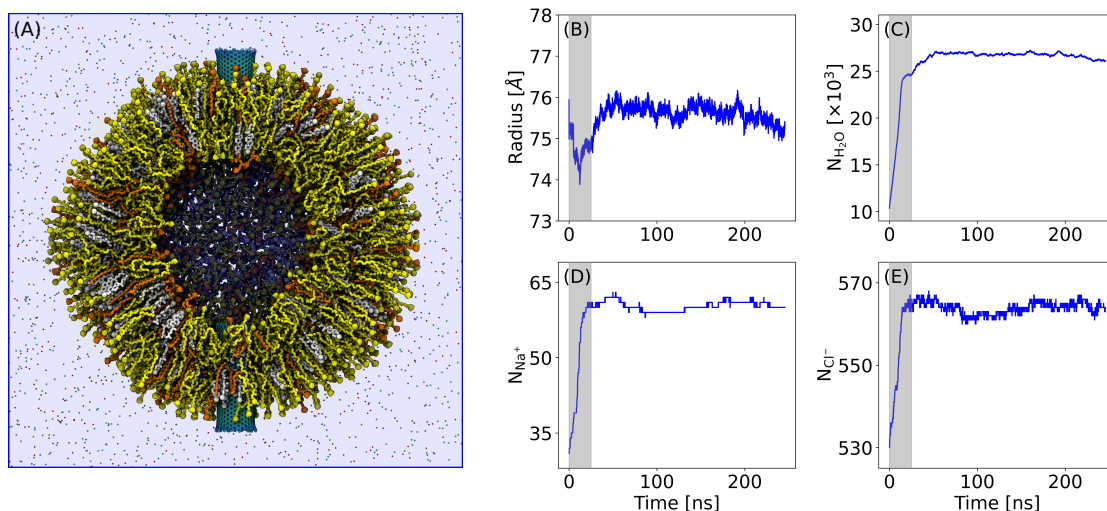
Figure 7.1: (A) A cross-sectional view of the initial setup of the full system. The CNTs are shown by cyan tubes at the poles. MC3H is represented in yellow, DSPC (orange), and cholesterol (white). Water is shown by the blue background and ions by the little spheres. (B) Radius of the vesicle with time (C) Number of water molecules (in 1000s), (D)Na$^+$ and (E)Cl$^-$ ions inside the vesicle. Inside is defined as distances less than the vesicle radius. The gray shaded area corresponds to the trajectory where some level of position restrain is applied on the lipids and CNT.

Table 7.1: Radius of LNP for different cases depicted in Figure 7.2. The systems MC3H, MC3H/RNA, MC3, and MC3/RNA correspond to Figure 7.1A, C, B, and D respectively. The inner radius is defined as the distance from the center of LNP to the position of the first peak in the full electron density profile shown in Figure 7.3. Similarly, the outer radius corresponds to the distance from the LNP center to the position of the second peak in the full electron density profile.

Systems	MC3H	MC3H/RNA	MC3	MC3/RNA
Inner radius [Å]	62.4	62.4	36.3	61.5
Outer radius [Å]	85.8	86.5	81.6	86.7

7.3 Results

7.4 Effect of RNA cargo and pH on the LNP structure

To study the effect of RNA cargo on the LNP structure, (i) we started with a vesicle with cationic MC3H. After around 50 ns of restraint-free simulations, the vesicle attains an equilibrium radius, and ions and water inside the vesicle equilibrate (Figure 7.1 B-E). (ii) After 150 ns, to mimic an increase in pH all the cationic MC3H was replaced by neutral MC3, and the corresponding number of chloride ions was removed to keep the setup electrostatically neutral. (iii) To look at the effect of RNA cargo after 150 ns 10 double-stranded RNA molecules each with 10 Uracil residues were introduced inside the vesicle. (iv) After this setup was equilibrated, to mimic the increase in pH for RNA-loaded LNP, all the MC3H was replaced by MC3. For all these systems, after adding the RNA or after replacing MC3H with MC3, the setups were carefully equilibrated with different levels of restraints on different parts of the system. I will call these different setups from here on as, MC3H, MC3, MC3H/RNA, and MC3/RNA. The snapshot after 200 ns for each case is shown in Figure 7.2. The MC3H setup stays in a bilayer vesicular configuration throughout the simulation (Figure 7.2A). Replacing MC3 with MC3H leads to disruption of the vesicular structure and water moves out of the core leading to a shrinkage of its volume as evident from the value of inner radius (Table 7.1). Adding, RNA to the MC3H setup does not seem to have any effect on the overall structure of the vesicle (Figure 7.1C), however, replacing MC3H with MC3 in the RNA-loaded vesicle again disrupts the vesicular structure. In all cases, the outer radii which represent the overall measurable size of the LNP are very similar. For the inner radii, the MC3 setup which exhibits a significantly small radii is an exception (Table 7.1).

Comparing the electron density of different setups (Figure 7.3) provides additional structural insights. The initial flat region in the electron density profile represents the water core, while the two peaks correspond to the inner and outer lipid head group positions. The water content in the core, represented by the flat region, has the largest radial extension in MC3H setups, followed by MC3H/RNA and MC3/RNA, with MC3 showing a significantly lower value. Electron density profiles for cationic and neutral MC3 systems also differ in peak positions, magnitude, and broadness. In particular, neutral MC3 setups (MC3 and MC3/RNA) at high pH exhibit a higher, broader first peak shifted towards the center, suggesting migration of components towards the LNP center. This effect is more pronounced
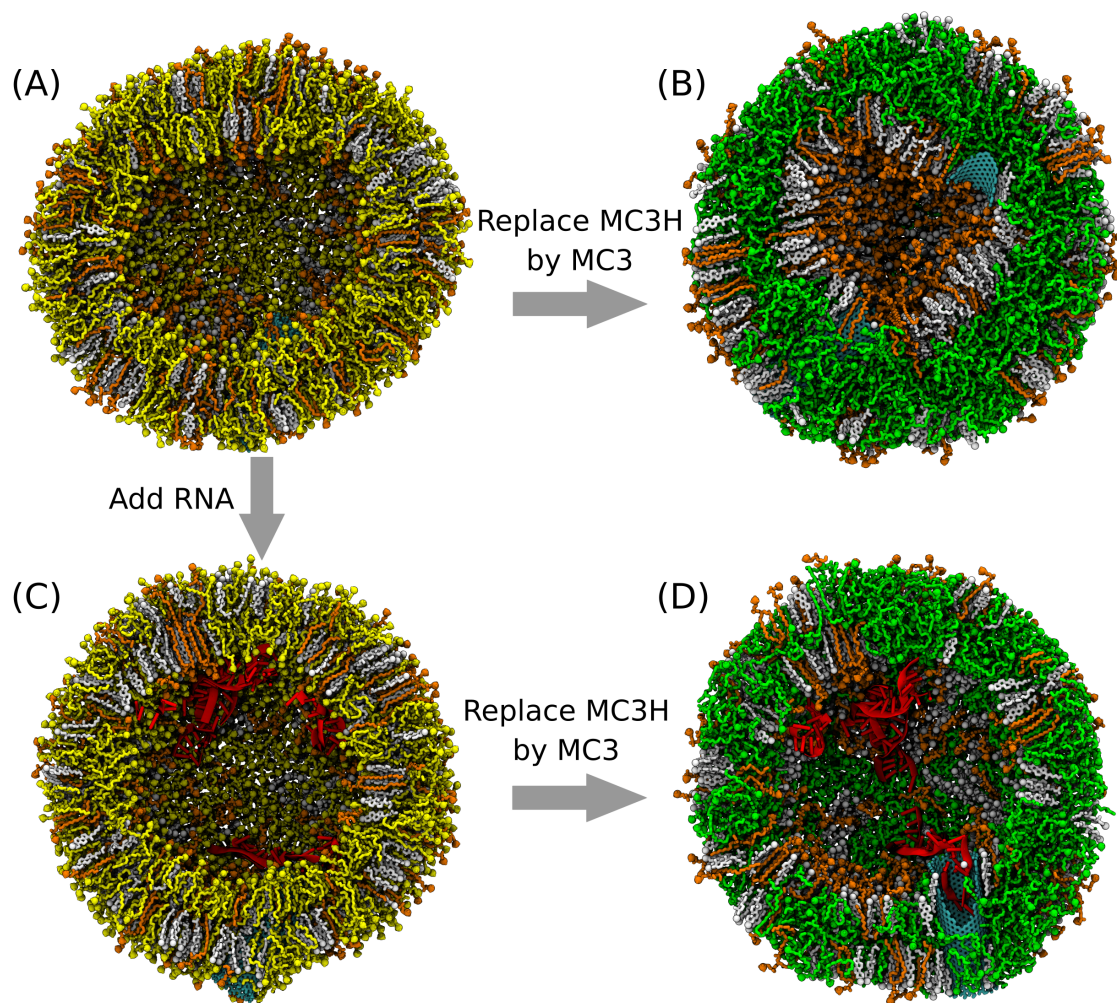
Figure 7.2: Cross-section view of different systems after 200 ns of simulations. Water and ions are omitted for clarity. (A) Vesicle with cationic MC3H (B) With neutral MC3H (C) with cationic MC3 and RNA (D) neutral MC3 and RNA. The color codes are: yellow (MC3H), orange (DSPC), white (cholesterol), red (RNA), green (MC3), and cyan (CNT). The setups are named as (A) MC3H (B) MC3 (C) MC3H/RNA and (D) MC3/RNA.

in setups without RNA, indicating that the presence of RNA sets a lower bound for the core size to avoid electrostatic repulsion between RNA molecules. The MC3H/RNA (green curve) setup shows that the addition of RNA to the MC3H setup just broadens the first peak and slightly increases the peak height. This indicates that RNA is strongly adsorbed to the inner surface of the LNP in contrast to the MC3/RNA setup (black curve)where, the contribution of RNA to the electron density appears a separate bump between 30 Åto 50 Å, which suggests a weaker interaction of RNA with the inner surface of LNP.

To gain further structural insights, the electron density of individual components for each setup is shown in Figure 7.4. I have plotted the electron density for the tail and head group of MC3/H and DSPC, for cholesterol, water, and the RNA molecule. For the MC3H setup, the vesicular structure is reflected in two distinct nearly symmetrical head peaks for MC3H and DSPC. The MC3H tails display two prominent peaks corresponding to the lipids in the inner and outer shells whereas the DSPC tails show only a single peak indicating
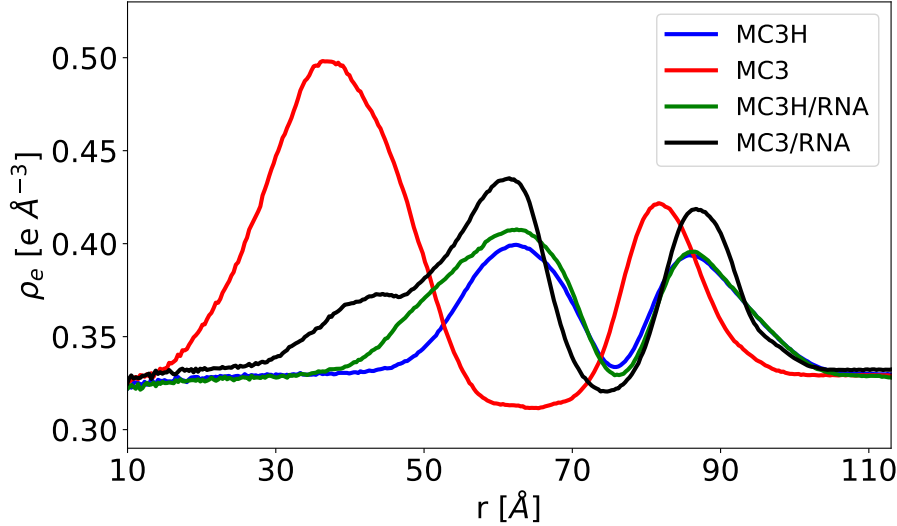
Figure 7.3: Full radial electron density of different systems obtained using Eqn. 7.1.

high interdigitation. Cholesterol also shows a certain degree of interdigitation between inner and outer shells as apparent from the low contrast between peaks and the minima between the peaks (Figure 7.4A). For the MC3 setup, the cholesterol and DSPC tails exhibit two prominent peaks which are relatively far separated from each other indicating these components in the inner and outer shells are pushed in radially opposite directions. This is a consequence of MC3 molecules moving away from the water-lipid interface into the center of the shell as evident from the drastic increase of MC3 head density in the center of the shell i.e. between 40 Å-80 Å (Figure 7.4B). The MC3H/RNA setup (Figure 7.4C) displays similar behavior as MC3H with the additional peak corresponding to the RNA appearing near the MC3H head group peaks. The MC3/RNA setup with neutral MC3 (Figure 7.4D) shows similar behavior as the corresponding MC3 setup without RNA (Figure 7.4 B). The RNA peak appears further from the DSPC head peaks of the MC3 head group peaks compared to the MC3H/RNA setup indicating loosely bound RNA to the inner surface of LNP.

7.5 Discussion

A complete understanding of the structure of ionizable lipid-based LNPs still remains lacking. Here, I used the MC3 model developed in this thesis to study the pH-dependent properties of MC3-based LNPs. I started with a vesicular bilayer structure corresponding to LNP at low pH as reported by Kulkarni et al. The initial diameter of the LNP was around 15 nm which is in the same range ($\sim 11\pm4$ nm) as reported for such LNPs at low pH (~ 4.0). The results from these simulations are consistent with various studies based on Cryo-TEM and small-angle neutron scattering experiments. In almost all the studies, an electron-dense core was observed at high pH while the morphology of the outer shell was either reported to be a monolayer shell composed mainly of the helper lipid and cholesterol or by concentric multi-lamellar shells with varying water content. Whereas at low pH a vesicular bilayer structure or a multilamellar structure has been hypothesized.

For empty LNPs without RNA, LNP stays in a vesicular bilayer structure at low pH
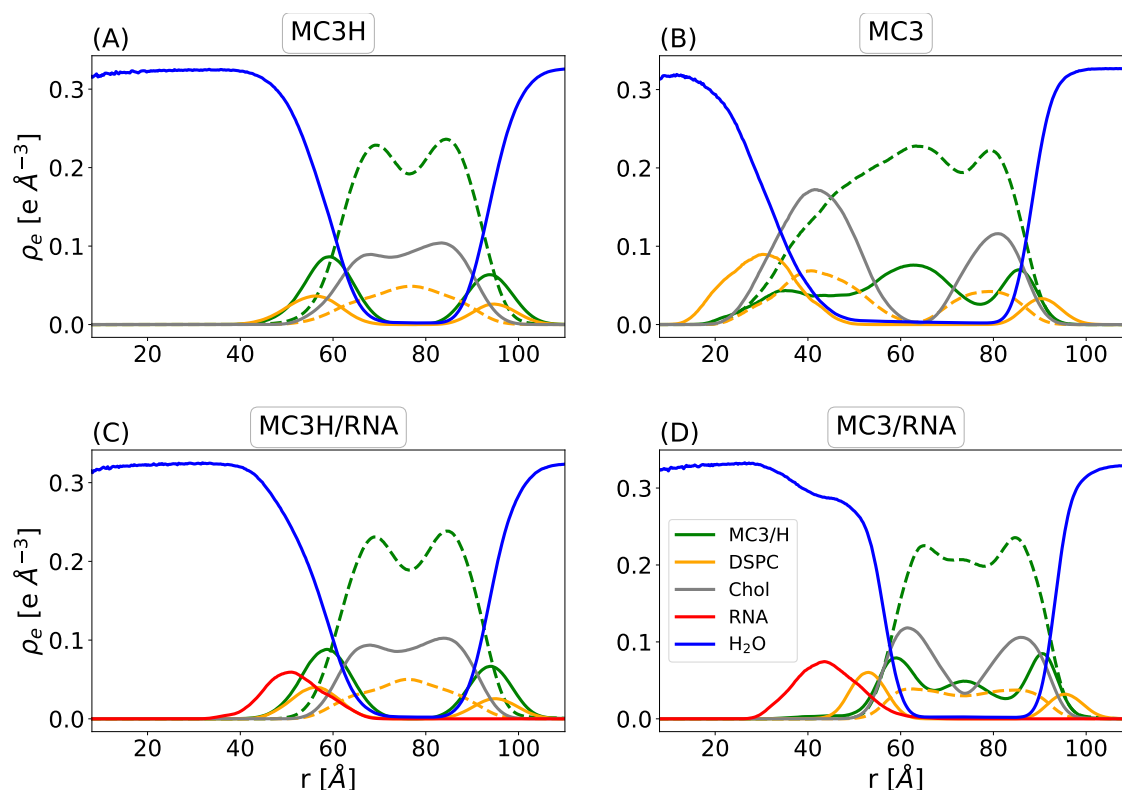
Figure 7.4: Electron density of each component for each setup. For MC3/H and DSPC, the electron density for the head and tail groups are presented separately. The tail groups are shown by the dashed profiles and the head groups are in full lines of the same color.

consistent with the reported experiments. At high pH, the bilayer structure is disrupted (Figure 7.2B) and a tendency to form an amorphous core is observed again in line with multiple reports. However, unlike previous reports, the core is not fully occupied by neutral MC3 molecules. In our case, MC3 does not migrate toward the center but due to its hydrophobic nature moves toward the lipid tail region away from the bilayer-water interface as observed in the previous chapter in the case of MC3-DOPC bilayers. This pushes the DSPC and cholesterol towards the center until the core attains a very small curvature with a high surface density of DSPC and cholesterol through which the MC3 can not penetrate into the center. For much larger LNP as in experiments at high pH, we could expect the MC3 to accumulate the LNP center with DSPC and cholesterol forming the outer shell. The hydration of the LNP core is another important structural aspect that our simulations predict to be highly pH dependent. At, low pH the water content in both empty and RNA-loaded LNP is significantly higher than at high pH. Recently, Li et al. [187] studied the pH-dependent structural changes in MC3-based LNPs where at low pH, RNA-loaded LNP was shown to exhibit a significantly larger diameter than at high pH. This was attributed to the excess amount of water inside the LNP at low pH. Our simulations are fully consistent with such studies, however, the size of LNP with RNA at low and high pH was not significantly different. For the MC3/RNA setup (Figure 7.2D), during the latter part of the simulation few RNA molecules blocked the CNT and hence the flow of water out of the LNP. Therefore, we could expect a further decrease in water content and size

105

inside the RNA-LNP at high pH with longer simulations but due to repulsion between RNA, the decrease in size can not be as small as in the absence of RNA. Finally, the addition of RNA did not lead to any significant changes in the LNP structure at low pH but at higher pH, it only seemed to set a lower limit to the LNP core size. It should also be noted that the amount of RNA which is usually measured with the ratio of ionizable lipids and the number of phosphate groups (N/P ratio) in the current study is $N/P = 7.6$ which is considered to be low RNA concentration. The effect of RNA at different pH was observed to be at $N/P <= 3$, where it was shown to induce multilamellar structure to the LNP at both high and low pHs. Such observations can not be tested with our current setup.

7.6 Conclusion

In this chapter, we used the MC3 developed in this thesis to study a model LNP system. The predictions from simulations were consistent with previously reported results. The vesicular bilayer structure is stable with fully protonated MC3 but is disrupted with neutral MC3 where it tends to form a more amorphous structure. Further, we observe a higher water content inside the LNP at low pH than at high pH. The addition of RNA ($N/P \sim 7.6$) at low pH did not affect the structure of LNP significantly compared to empty LNPs whereas at high pH it seemed to prevent the LNP core from becoming two small due to electrostatic repulsion between the negatively charged RNA molecules. These observations further validate the accuracy and the applicability of our MC3 force field to complex therapeutically relevant setups.

7.7   Appendix

7.7.1   Data archive

The simulation input file for all the simulations presented in this chapter is archived in the following tree. Go to any of the directories e.g. `./cationic-lnp-with-rna` and run the command

```
gmx grompp -f ../mdp/md-pr.mdp -c npt.gro -p topol.top  -n index.ndx \
-o npt.tpr
```

```
$HOME/Data-Archive-Mohd_Ibrahim/LNP/
├── cationic-lnp-with-rna
│   ├── index.ndx   npt.gro   topol.top
│   └── em
│
├── cationic-lnp-without-rna
│   ├── index.ndx   npt.gro   topol.top
│   └── em
│
├── neutral-lnp-with-rna
│   ├── index.ndx   npt.gro   topol.top
│   └── em
│
├── neutral-lnp-without-rna
│   ├── index.ndx   npt.gro   topol.top
│   └── em
│
├── forcefields
│   ├── amber_na.ff
│   ├── cnt.ff
│   ├── lipid.ff
│   ├── nschwierz.ff
│   └── rna-itp
│
└── mdp
```

Conclusions and Outlook

In this thesis, we used molecular dynamics simulation and complementary experimental techniques to gain atomistic insights into the interaction of nucleic acids with surfaces, cationic and neutral bilayers. We also modeled and investigated the behavior of ionizable lipids in neutral lipid bilayers.

We started off by investigating the effect of monovalent cations Li^+, Na^+, K^+, Cs^+ and divalent cations Mg^{2+}, Ca^{2+} on the interaction of single-stranded DNA with mica surface. Both DNA and mica are negatively charged, therefore cations are required to mediate the interactions. We used single-molecule force spectroscopy experiments and steered MD simulations to characterize forces required to desorb the ssDNA from the mica surface. We found that DNA-mica interaction is a complex interplay of cation-DNA interactions, cation-mica interactions, and hydration effects. These interactions are responsible for the high and low regions in the rupture force distributions observed in both experiments and simulations. Indirect water-mediated interactions give rise to lower force pathways whereas direct cation-DNA or cation-surface interactions result in high rupture forces. From both experiments and simulations, Cs^+ results in the lowest rupture forces since here force is dictated by cation-DNA affinity which is weakest for Cs^+. K^+ and Na^+ give rise to similar forces whereas for Li^+ the high force part of the distribution is over-represented in the simulations. For Ca^{2+}, the simulation under-represents the low force part. Here, accurate force field parameters with a correct affinity towards the mica surface as well as DNA are required. For Mg^{2+}, using straightforward and enhanced sampling methods, we could reproduce both the high and low force part of the distribution corresponding to direct and water-mediated interactions respectively. The results from our experiments and simulations could aid in choosing proper ionic conditions for DNA deposition on mica surfaces in physiological conditions.

Using the coarse-grained Martini model for RNA and lipids we characterized the adsorption behavior of single- and double-stranded RNA on neutral and cationic bilayers composed of DOPC and DOTAP and their mixtures. RNA adsorbs to the membrane such that the nucleobase of RNA interacts with hydrophobic tail parts of the bilayer and the backbone interacts with the polar or charged head group of the bilayer. Therefore, the base pairing of RNA plays a crucial role in facilitating adsorption as RNA with exposed nucleobase i.e. those in single-stranded regions interacts more effectively with the membrane. Consequently, RNA interaction with the bilayer also depends on the secondary structure (SS) of RNA since each SS is composed of a unique arrangement of single and double-stranded regions. This leads to each SS adopting a unique configuration on the membrane surface as was clear from the simulation of 10 different SS of tRNA molecule. Such a unique mode of adsorption of a given SS was clearly reflected in the neutron scattering length

density profiles. Based on this information we explored the possibility of resolving different SS from neutron reflectivity or small-angle neutron scattering (SANS) form factor profiles. We found that due to the inherent Poisson noise of the measurement technique, the subtle differences between different SS could not be differentiated. However, selective deuteration of tRNA results in clear differences between the SANS form factor profile of different SS. The SS of mRNA affects its stability and translation properties, therefore, probing the SS of mRNA is an important aspect of LNP structure and function. By combining coarse-grained simulation and scattering experiments on identical systems it's possible to probe the most probable RNA conformation on the membrane. Here, a method called *ensemble refinement* could be very useful where the scattering profile from simulations is represented by the weighted average of intensities from all the SS in consideration. The weights of different SS are varied until the best agreement between experiments and simulation is achieved [214]. The SS with the highest weight corresponds to the most likely RNA conformation.

To investigate the pH-dependent behavior of therapeutically relevant ionizable lipid-containing systems and their interaction with other components of LNP, such as neutral lipids and cholesterol, we developed accurate parameters for one of the most promising ionizable lipids, namely Dlin-MC3-DMA (MC3), in two different protonation states corresponding to MC3 in two extreme pH conditions. To assess the accuracy of the newly developed parameters and existing force fields, we made direct comparisons of simulations with neutron reflectivity measurements. These simulations were performed on bilayer setups containing MC3 and DOPC mixtures in different pH conditions, multiple solvent contrasts, and MC3 mole fractions. For the currently developed parameters, we obtained excellent agreement with experiments for almost all cases. Similarly, the existing CHARMM-based MC3 forcefield also demonstrated a comparable level of agreement whereas the Slipids-based MC3 underestimated the bilayer thickness and predicted MC3 distribution contrary to the current and CHARMM-based force fields. Based on these observations, we could ascertain the pH-dependent distribution of MC3 in MC3-DOPC bilayers. At low pH i.e. pH < pKa, the cationic MC3 stays at the bilayer water interface whereas at high pH (pH> pKa), the neutral MC3 segregated towards the bilayer center, leading to an increase in the overall bilayer thickness. This indicates that pH has a drastic effect on the behavior of MC3 with the protonated and neutral MC3 exhibiting vastly different behavior. Therefore, for accurate simulation at a given pH, the correct assignment of a fraction of charged and neutral ionizable lipid i.e. the protonation degree is crucial. This fraction is not known a priori and depends upon the chemical and electrostatic environment of the lipid.

We devised an iterative procedure to assign the protonation degree at a given pH for lipid monolayer systems (POPC, MC3 mixture) by combining MD simulations, grazing incidence X-ray off-specular scattering (GIXOS), and X-ray fluorescence experiments. In this procedure, we first optimize the lateral packing of lipids (or area per lipid) using the experimental GIXOS profiles and assuming a 100% protonation degree. The protonation degree is assigned based on the concentration of excess surface adsorbed anions which are clearly reflected in the X-ray fluorescence signal intensity. Here, the area per lipid from the simulation and the transverse anion (Cl^- ions) distribution are used as input for

the experiments. Simulation performed with correct protonation degree and area per lipid showed that the neutral MC3 accumulated in the hydrophobic tail parts whereas protonated MC3 stayed at the bilayer-water interface consistent with the observation for DOPC-MC3 bilayers. Further, the simulations excellently reproduced the experimental GIXOS curves indicating that the currently developed MC3 force fields capture the transverse distribution of MC3 accurately.

To study the effect of pH and RNA cargo on the LNP structure we used our MC3 model to simulate model LNPs. The starting structure corresponded to LNP at low pH i.e all cationic MC3 and was based on the reports by Kulkarni et al [180], where a vesicular bilayer structure was observed at low pH and an amorphous solid core was observed at high pH. The vesicular bilayer structure with cationic MC3 achieved an equilibrium size and remained stable throughout the simulation time. Replacing cationic MC3 with neutral MC3, mimicking an increase in pH, resulted in the disruption of the vesicular bilayer structure and the final structure resembled the reported amorphous structure at high pH from Kulkarni et al, but without the solid core. The simulations provide atomistic insights into the formation of such amorphous structures at high pH. Due to the hydrophobic nature of neutral MC3, it moves away from the water-LNP interface pushing the monolayer shells in radially opposite directions. The inward movement of the inner shell towards the LNP center expels the water out of the LNP core leading to a relatively high electron density region since it gets mostly occupied by lipids. Unlike the experiments, we do not observe a fully solid core devoid of water due to the small size of the LNP which induces high curvature in the inner monolayer shell. Such a high curvature makes the inner monolayer shell highly dense making it difficult for the neutral MC3 to penetrate through and form a solid core. The addition of RNA did not significantly affect the LNP structure at low pH whereas at high pH, the presence of RNA sets a lower bound to the inner LNP radius due to electrostatic repulsion between the RNA molecules. Further, recent scattering experiments reported low water content inside the LNP at high pH compared to at low pH which is exactly in line with the prediction of our simulations. To summarize, our LNP model aligns with and explains the previously observed pH-dependent behavior of ionizable lipids-based LNPs, further underscoring the wide applicability and accuracy of our MC3 force field. However, the size of LNP (diameter $\approx$ 16 nm) in our current simulations falls at the low end of the experimental sizes and the simulation time scale is very short. To enable quantitative comparison with scattering experiments and to observe structural transition which is now hindered due to the small size of the LNP, simulations of much bigger LNPs are required.

All-atom simulation of LNPs with sizes comparable to experiments (diameter $\sim$ 50 nm) is not feasible with current computational resources. Here coarse-grained simulations could be more suitable as they enable simulation at the level of experimental length scale and over much longer timescales. Our accurate all-atom force field provides an excellent reference for coarse-graining the MC3 lipid molecule in different protonation states. Using recent tools [215] it's possible to coarse-grain the MC3 lipid such that it reproduces the all-atom pH-dependent behavior.

Finally, our MC3 force field was also applied to understand the phase behavior of

MC3/cholesterol systems [188]. From small-angle X-ray scattering (SAXS) experiments our collaborators established that such systems transition from inverse micellar phase at high pH to inverse hexagonal phase at low pH. Such phase transition facilitates the endosomal release of mRNA cargo and therefore is a determinant of LNP efficacy [188]. Our inverse micellar and inverse hexagonal simulations correctly captured the electron density profiles thus reproducing the SAXS intensities. In these simulations, the micellar or hexagonal phase was imposed to reproduce the experimentally measured lattice spacings. However, to study the formation of these phases and their dependence on hydration state and pH, typically much larger simulation setups are required. Here, coarse-grained simulations will again be very useful.

In summary, the accurate MC3 force field developed in this thesis excellently captures the pH-dependent behavior of MC3 across many systems and provides valuable structural insights. In all these systems, the pH-dependent behavior arises from the vastly different behavior of neutral and cationic MC3. The neutral MC3 behaves as a purely hydrophobic molecule whereas cationic MC3 exhibits amphipathic characteristics. These observations along with our simulation insights into RNA/DNA-surface interactions form the baseline for the interpretation of LNP structure and aid in the development of more potent delivery systems.

List of Publications

1. **Mohd Ibrahim**, Jennifer Gilbert, Marcel Heinz, Tommy Nylander, and Nadine 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, 2023, 15, 11647-11656*

2. **Mohd Ibrahim**,Christiane Wenzel, Max Lallemang Bizan N. Balzer, and Nadine Schweirz. "Adsorbing DNA to mica by cations: influence of valency and ion type". *Langmuir, 2023, 39, 44, 15553–15562*

3. Miriam Grava, **Mohd Ibrahim**, Akhil Sudarsan, Julio Pusterla, Julian Philipp, Joachim Rädler, Nadine Schweirz, Emanuel Schneck . "Combining molecular dynamics simulations and x-ray scattering techniques for the accurate treatment of protonation degree and packing of ionizable lipids in monolayers". *J. Chem. Phys., 159, 154706 (2023)* (Featured Article)

4. Julian Philipp, Aleksandra Dabkowska1, Anita Reiser, Kilian Frank, Rafał Krzysztoń, Christiane Brummer, Bert Nickel, Clement E. Blanchet, Akhil Sudarsan, **Mohd Ibrahim**, Svante Johansson1, Pia Skantze, Urban Skantze, Sofia Östman, Marie Johansson, Neil Henderson, Kjetil Elvevold, Bård Smedsrød, Nadine Schwierz, Lennart Lindfors and Joachim O. Rädler. "pH-dependent structural transitions in cationic ionizable lipid mesophases are critical for lipid nanoparticle function"". *Proc. Natl. Acad. Sci. U.S.A.*, 120(50):e2310491120, 2023

   · *Contribution*: provided crucial contributions in the setup, analysis, interpretation, and writing of the simulation part.

Other publications are listed on `https://scholar.google.com/citations?user=8h9-8HwAAAAJ&hl=en`.

Bibliography

[1] Singer, S. J. and Nicolson, G. L. "The fluid mosaic model of the structure of cell membranes: Cell membranes are viewed as two-dimensional solutions of oriented globular proteins and lipids.". *Science*, 175(4023):720–731, 1972.

[2] Nicolson, G. L. "The fluid—mosaic model of membrane structure: Still relevant to understanding the structure, function and dynamics of biological membranes after more than 40 years". *Biochim. Biophys. Acta - Biomembr.*, 1838(6):1451–1466, 2014.

[3] Luzzati, V., Tardieu, A., Gulik-Krzywicki, T., Rivas, E., and Reiss-Husson, F. "Structure of the cubic phases of lipid–water systems". *Nature*, 220(5166):485–488, 1968.

[4] Tanford, C. *The hydrophobic effect: formation of micelles and biological membranes 2d ed.* J. Wiley., 1980.

[5] Tilcock, C. P. "Lipid polymorphism". *Chem. Phys. Lipids*, 40(2-4):109–125, 1986.

[6] Lindblom, G. and Rilfors, L. "Cubic phases and isotropic structures formed by membrane lipids—possible biological relevance". *Biochim. Biophys. Acta - Biomembr.*, 988(2):221–256, 1989.

[7] Shearman, G., Ces, O., Templer, R., and Seddon, J. "Inverse lyotropic phases of lipids and membrane curvature". *J. Phys. Condens. Matter*, 18(28):S1105, 2006.

[8] Venturoli, M., Sperotto, M. M., Kranenburg, M., and Smit, B. "Mesoscopic models of biological membranes". *Phys. Rep.*, 437(1-2):1–54, 2006.

[9] Conn, C. E. and Seddon, J. M. "Structures of lipid membranes: Cubic and inverse hexagonal phases". In *Handbook of Lipid Membranes*, pages 49–63. CRC Press, 2021.

[10] Malone, R. W., Felgner, P. L., and Verma, I. M. "Cationic liposome-mediated rna transfection.". *Proc. Natl. Acad. Sci. U.S.A.*, 86(16):6077–6081, 1989.

[11] RH, M. "Solid lipid nanoparticles (sln)-an alternative colloidal carrier system for controlled drug delivery". *Eur J Pharm Biopharm.*, 41:62–69, 1995.

[12] Wissing, S., Kayser, O., and Müller, R. "Solid lipid nanoparticles for parenteral drug delivery". *Adv. Drug Deliv. Rev.*, 56(9):1257–1272, 2004.

[13] Tenchov, R., Bird, R., Curtze, A. E., and Zhou, Q. "Lipid nanoparticles– from liposomes to mrna vaccine delivery, a landscape of research diversity and advancement". *ACS nano*, 15(11):16982–17015, 2021.

[14] Hou, X., Zaks, T., Langer, R., and Dong, Y. "Lipid nanoparticles for mrna delivery". *Nat. Rev. Mater.*, 6(12):1078–1094, 2021.

[15] Coskun, Ü. and Simons, K. "Cell membranes: the lipid perspective". *Structure*, 19 (11):1543–1548, 2011.

[16] Cruz-León, S. and Schwierz, N. "Hofmeister series for metal-cation-rna interactions: The interplay of binding affinity and exchange kinetics". *Langmuir*, 36:5979–5989, 2020.

[17] Cruz-León, S., Vanderlinden, W., Müller, P., Forster, T., Staudt, G., Lin, Y.-Y., Lipfert, J., and Schwierz, N. "Twisting dna by salt". *Nucleic Acids Res.*, 50(10):5726–5738, 2022.

[18] Koculi, E., Hyeon, C., Thirumalai, D., and Woodson, S. A. "Charge density of divalent metal cations determines rna stability". *J. Am. Chem. Soc*, 129(9):2676–2682, 2007.

[19] Bhattacharyya, D., Mirihana Arachchilage, G., and Basu, S. "Metal cations in g-quadruplex folding and stability". *Front. Chem.*, 4:38, 2016.

[20] Hardin, C. C., Watson, T., Corregan, M., and Bailey, C. "Cation-dependent transition between the quadruplex and watson-crick hairpin forms of d (cgcg3gcg)". *Biochemistry*, 31(3):833–841, 1992.

[21] Castro, C. E., Kilchherr, F., Kim, D.-N., Shiao, E. L., Wauer, T., Wortmann, P., Bathe, M., and Dietz, H. "A primer to scaffolded dna origami". *Nat. Methods*, 8(3): 221–229, 2011.

[22] He, Z., Shi, K., Li, J., and Chao, J. "Self-assembly of dna origami for nanofabrication, biosensing, drug delivery and computational storage". *Iscience*, 2023.

[23] "Dna binding to mica correlates with cationic radius: Assay by atomic force microscopy". *Biophys. J.*, 70:1933–1939, 1996.

[24] Rivetti, C., Guthold, M., and Bustamante, C. "Scanning force microscopy of dna deposited onto mica: equilibration versus kinetic trapping studied by statistical polymer chain analysis". *J. Mol. Biol.*, 264(5):919–932, 1996.

[25] Heenan, P. R. and Perkins, T. T. "Imaging dna equilibrated onto mica in liquid using biochemically relevant deposition conditions". *ACS nano*, 13(4):4220–4229, 2019.

[26] Xin, Y., Rivadeneira, S. M., Grundmeier, G., Castro, M., and Keller, A. "Self-assembly of highly ordered dna origami lattices at solid-liquid interfaces by controlling cation binding and exchange". *Nano Res.*, 13:3142–3150, 2020.

[27] Lin, S. R. V., Grotz, K. K., Siretanu, I., Schwierz, N., and Mugele, F. "Ion-specific and ph-dependent hydration of mica-electrolyte interfaces". *Langmuir*, 35:5737–5745, 2019.

[28] Michanek, A., Kristen, N., Höök, F., Nylander, T., and Sparr, E. "Rna and dna interactions with zwitterionic and charged lipid membranes—a dsc and qcm-d study". *Biochim Biophys Acta Biomembr*, 1798(4):829–838, 2010.

[29] Antipina, A. Y. and Gurtovenko, A. A. "Toward understanding liposome-based sirna delivery vectors: Atomic-scale insight into sirna–lipid interactions". *Langmuir*, 34 (29):8685–8693, 2018.

[30] Gromelski, S. and Brezesinski, G. "Dna condensation and interaction with zwitterionic phospholipids mediated by divalent cations". *Langmuir*, 22(14):6293–6301, 2006.

[31] Antipina, A. Y. and Gurtovenko, A. A. "Molecular mechanism of calcium-induced adsorption of dna on zwitterionic phospholipid membranes". *J. Phys. Chem. B*, 119 (22):6638–6645, 2015.

[32] Marty, R., N'soukpoe-Kossi, C. N., Charbonneau, D. M., Kreplak, L., and Tajmir-Riahi, H.-A. "Structural characterization of cationic lipid–trna complexes". *Nucleic Acids Res.*, 37(15):5197–5207, 2009.

[33] Suga, K., Tanabe, T., Tomita, H., Shimanouchi, T., and Umakoshi, H. "Conformational change of single-stranded rnas induced by liposome binding". *Nucleic Acids Res.*, 39(20):8891–8900, 2011.

[34] Routy, J.-P., Boulassel, M.-R., Yassine-Diab, B., Nicolette, C., Healey, D., Jain, R., Landry, C., Yegorov, O., Tcherepanova, I., Monesmith, T., et al. "Immunologic activity and safety of autologous hiv rna-electroporated dendritic cells in hiv-1 infected patients receiving antiretroviral therapy". *Clin. Immunol.*, 134(2):140–147, 2010.

[35] Van Gulck, E., Vlieghe, E., Vekemans, M., Van Tendeloo, V. F., Van De Velde, A., Smits, E., Anguille, S., Cools, N., Goossens, H., Mertens, L., et al. "mrna-based dendritic cell vaccination induces potent antiviral t-cell responses in hiv-1-infected patients". *AIDS*, 26(4):F1–F12, 2012.

[36] Allard, S. D., De Keersmaecker, B., de Goede, A. L., Verschuren, E. J., Koetsveld, J., Reedijk, M. L., Wylock, C., De Bel, A. V., Vandeloo, J., Pistoor, F., et al. "A phase i/iia immunotherapy trial of hiv-1-infected patients with tat, rev and nef expressing dendritic cells followed by treatment interruption". *Clin. Immunol.*, 142(3):252–268, 2012.

[37] Van Tendeloo, V. F., Van de Velde, A., Van Driessche, A., Cools, N., Anguille, S., Ladell, K., Gostick, E., Vermeulen, K., Pieters, K., Nijs, G., et al. "Induction of complete and molecular remissions in acute myeloid leukemia by wilms' tumor 1 antigen-targeted dendritic cell vaccination". *Proc. Natl. Acad. Sci. U.S.A.*, 107(31): 13824–13829, 2010.

[38] Heiser, A., Coleman, D., Dannull, J., Yancey, D., Maurice, M. A., Lallas, C. D., Dahm, P., Niedzwiecki, D., Gilboa, E., Vieweg, J., et al. "Autologous dendritic cells transfected with prostate-specific antigen rna stimulate ctl responses against metastatic prostate tumors". *J. Clin. Invest.*, 109(3):409–417, 2002.

[39] Weide, B., Pascolo, S., Scheel, B., Derhovanessian, E., Pflugfelder, A., Eigentler, T. K., Pawelec, G., Hoerr, I., Rammensee, H.-G., and Garbe, C. "Direct injection of protamine-protected mrna: results of a phase 1/2 vaccination trial in metastatic melanoma patients". *J. Immunother.*, 32(5):498–507, 2009.

[40] Yoon, S., Lee, J., Cho, H., Kim, E., Kim, H., Park, M., and Kim, T. "Adoptive immunotherapy using human peripheral blood lymphocytes transferred with rna encoding her-2/neu-specific chimeric immune receptor in ovarian cancer xenograft model". *Cancer Gene Ther.*, 16(6):489–497, 2009.

[41] Fotin-Mleczek, M., Duchardt, K. M., Lorenz, C., Pfeiffer, R., Ojkic-Zrna, S., Probst, J., and Kallen, K.-J. "Messenger rna-based vaccines with dual activity induce balanced tlr-7 dependent adaptive immune responses and provide antitumor activity". *J. Immunother.*, 34(1):1–15, 2011.

[42] Jirikowski, G. F., Sanna, P. P., Maciejewski-Lenoir, D., and Bloom, F. E. "Reversal of diabetes insipidus in brattleboro rats: intrahypothalamic injection of vasopressin mrna". *Science*, 255(5047):996–998, 1992.

[43] Creusot, R. J., Chang, P., Healey, D. G., Tcherepanova, I. Y., Nicolette, C. A., and Fathman, C. G. "A short pulse of il-4 delivered by dcs electroporated with modified mrna can both prevent and treat autoimmune diabetes in nod mice". *Mol. Ther.*, 18 (12):2112–2120, 2010.

[44] Shemesh, C. S., Hsu, J. C., Hosseini, I., Shen, B.-Q., Rotte, A., Twomey, P., Girish, S., and Wu, B. "Personalized cancer vaccines: clinical landscape, challenges, and opportunities". *Mol. Ther.*, 29(2):555–570, 2021.

[45] Sahin, U., Derhovanessian, E., Miller, M., Kloke, B.-P., Simon, P., Löwer, M., Bukur, V., Tadmor, A. D., Luxemburger, U., Schrörs, B., et al. "Personalized rna mutanome vaccines mobilize poly-specific therapeutic immunity against cancer". *Nature*, 547 (7662):222–226, 2017.

[46] Sahin, U., Karikó, K., and Türeci, Ö. "mrna-based therapeutics—developing a new class of drugs". *Nat. Rev. Drug Discov.*, 13(10):759–780, 2014.

[47] Sahin, U. and Türeci, Ö. "Personalized vaccines for cancer immunotherapy". *Science*, 359(6382):1355–1360, 2018.

[48] Semple, S. C., Akinc, A., Chen, J., Sandhu, A. P., Mui, B. L., Cho, C. K., Sah, D. W., Stebbing, D., Crosley, E. J., Yaworski, E., et al. "Rational design of cationic lipids for sirna delivery". *Nat. Biotechnol.*, 28(2):172–176, 2010.

[49] Han, X., Zhang, H., Butowska, K., Swingle, K. L., Alameh, M.-G., Weissman, D., and Mitchell, M. J. "An ionizable lipid toolbox for rna delivery". *Nat. Commun.*, 12 (1):7233, 2021.

[50] Tilstra, G., Couture-Senécal, J., Lau, Y. M. A., Manning, A. M., Wong, D. S., Janaeska, W. W., Wuraola, T. A., Pang, J., and Khan, O. F. "Iterative design of ionizable lipids for intramuscular mrna delivery". *J. Am. Chem. Soc.*, 145(4):2294–2304, 2023.

[51] Jayaraman, M., Ansell, S. M., Mui, B. L., Tam, Y. K., Chen, J., Du, X., Butler, D., Eltepu, L., Matsuda, S., Narayanannair, J. K., Rajeev, K. G., Hafez, I. M., Akinc, A., Maier, M. A., Tracy, M. A., Cullis, P. R., Madden, T. D., Manoharan, M., and Hope, M. J. "Maximizing the potency of siRNA lipid nanoparticles for hepatic gene silencing in vivo". *Angew Chem Int Ed Engl*, 51(34):8529–8533, 2012.

[52] Eygeris, Y., Gupta, M., Kim, J., and Sahay, G. "Chemistry of lipid nanoparticles for rna delivery". *Acc. Chem. Res.*, 55(1):2–12, 2021.

[53] Adams, D., Gonzalez-Duarte, A., O'Riordan, W. D., Yang, C.-C., Ueda, M., Kristen, A. V., Tournev, I., Schmidt, H. H., Coelho, T., Berk, J. L., et al. "Patisiran, an rnai therapeutic, for hereditary transthyretin amyloidosis". *N. Engl. J. Med.*, 379 (1):11–21, 2018.

[54] Schoenmaker, L., Witzigmann, D., Kulkarni, J. A., Verbeke, R., Kersten, G., Jiskoot, W., and Crommelin, D. J. "mrna-lipid nanoparticle covid-19 vaccines: Structure and stability". *Int. J. Pharm.*, 601:120586, 2021.

[55] Frenkel, D., Smit, B., and Ratner, M. A. *Understanding molecular simulation: from algorithms to applications*, volume 2. Academic press San Diego, 1996.

[56] Verlet, L. "Computer" experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules". *Phys. Rev.*, 159(1):98, 1967.

[57] Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. "A smooth particle mesh ewald method". *J. Chem. Phys.*, 103(19):8577–8593, 1995.

[58] Bussi, G., Donadio, D., and Parrinello, M. "Canonical sampling through velocity rescaling". *J. Chem. Phys.*, 126:014101, 2007.

[59] Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. "Molecular dynamics with coupling to an external bath". *J. Chem. Phys.*, 81 (8):3684–3690, 1984.

[60] Parrinello, M. and Rahman, A. "Polymorphic transitions in single crystals: A new molecular dynamics method". *J. Appl. Phys.*, 52:7182–7190, 1981.

[61] Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P., and De Vries, A. H. "The martini force field: coarse grained model for biomolecular simulations". *J. Phys. Chem. B*, 111(27):7812–7824, 2007.

[62] Marrink, S. J. and Mark, A. E. "The mechanism of vesicle fusion as revealed by molecular dynamics simulations". *J. Am. Chem. Soc.*, 125(37):11144–11145, 2003.

[63] Marrink, S.-J. and Mark, A. E. "Molecular view of hexagonal phase formation in phospholipid membranes". *Biophys. J.*, 87(6):3894–3900, 2004.

[64] MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D., and Karplus, M. "All-atom empirical potential for molecular modeling and dynamics studies of proteins". *J. Phys. Chem. B*, 102 (18):3586–3616, 1998.

[65] Jämbeck, J. P. M. and Lyubartsev, A. P. "Derivation and systematic validation of a refined all-atom force field for phosphatidylcholine lipids". *J. Phys. Chem. B*, 116 (10):3164–3179, 2012.

[66] Dickson, C. J., Madej, B. D., Skjevik, Å. A., Betz, R. M., Teigen, K., Gould, I. R., and Walker, R. C. "Lipid14: The amber lipid force field". *J. Chem. Theory Comput.*, 10(2):865–879, 2014.

[67] Leonard, A. N., Wang, E., Monje-Galvan, V., and Klauda, J. B. "Developing and testing of lipid force fields with applications to modeling cellular membranes". *Chem. Rev.*, 119(9):6227–6269, 2019.

[68] Šponer, J., Banáš, P., Jurečka, P., Zgarbová, M., Kührová, P., Havrila, M., Krepl, M., Stadlbauer, P., and Otyepka, M. "Molecular dynamics simulations of nucleic acids. from tetranucleotides to the ribosome". *J. Phys. Chem. Lett. Letters*, 5(10): 1771–1782, 2014.

[69] Beauchamp, K. A., Lin, Y.-S., Das, R., and Pande, V. S. "Are protein force fields getting better? a systematic benchmark on 524 diverse nmr measurements". *J. Chem. Theory Comput.*, 8(4):1409–1414, 2012.

[70] Klauda, J. B., Venable, R. M., Freites, J. A., O' Connor, J. W., Tobias, D. J., Mondragon-Ramirez, C., Vorobyov, I., MacKerell, A. D., and Pastor, R. W. "Update of the charmm all-atom additive force field for lipids: Validation on six lipid types". *J. Phys. Chem. B*, 114(23):7830–7843, 2010.

[71] Dickson, C. J., Rosso, L., Betz, R. M., Walker, R. C., and Gould, I. R. "Gafflipid: a general amber force field for the accurate molecular dynamics simulation of phospholipid". *Soft Matter*, 8(37):9617–9627, 2012.

[72] Gould, I., A.A., S., Dickson, C., Madej, B., and Walker, R. "Lipid17: A comprehensive amber force field for the simulation of zwitterionic and anionic lipids". *Unpublished Work*, 2018.

[73] Dickson, C. J., Walker, R. C., and Gould, I. R. "Lipid21: complex lipid membrane simulations with amber". *J. Chem. Theory Comput*, 18(3):1726–1736, 2022.

[74] Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. "A second generation force field for the simulation of proteins, nucleic acids, and organic molecules". *J. Am. Chem. Soc.*, 117(19):5179–5197, 1995.

[75] Wang, J., Cieplak, P., and Kollman, P. A. "How well does a restrained electrostatic potential (resp) model perform in calculating conformational energies of organic and biological molecules?". *J. Comput. Chem.*, 21(12):1049–1074, 2000.

[76] Rosso, L. and Gould, I. R. "Structure and dynamics of phospholipid bilayers using recently developed general all-atom force fields". *J. Comput. Chem.*, 29(1):24–37, 2008.

[77] Skjevik, Å. A., Madej, B. D., Walker, R. C., and Teigen, K. "Lipid11: A modular framework for lipid simulations using amber". *J. Phys. Chem. B*, 116(36):11124–11136, 2012.

[78] Bayly, C. I., Cieplak, P., Cornell, W., and Kollman, P. A. "A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the resp model". *J. Phys. Chem. Lett.*, 97(40):10269–10280, 1993.

[79] Beiser, A. *Concepts of modern physics*. 2003.

[80] Kučerka, N., Liu, Y., Chu, N., Petrache, H. I., Tristram-Nagle, S., and Nagle, J. F. "Structure of fully hydrated fluid phase dmpc and dlpc lipid bilayers using x-ray scattering from oriented multilamellar arrays and from unilamellar vesicles". *Biophys. J*, 88(4):2626–2637, 2005.

[81] Kučerka, N., Nagle, J. F., Sachs, J. N., Feller, S. E., Pencer, J., Jackson, A., and Katsaras, J. "Lipid bilayer structure determined by the simultaneous analysis of neutron and x-ray scattering data". *Biophys. J.*, 95:2356–2367, 2008.

[82] Kucerka, N., Holland, B. W., Gray, C. G., Tomberli, B., and Katsaras, J. "Scattering density profile model of popg bilayers as determined by molecular dynamics simulations and small-angle neutron and x-ray scattering experiments". *J. Phys. Chem. B*, 116(1):232–239, 2012.

[83] Bengtsen, T., Holm, V. L., Kjølbye, L. R., Midtgaard, S. R., Johansen, N. T., Tesei, G., Bottaro, S., Schiøtt, B., Arleth, L., and Lindorff-Larsen, K. "Structure and dynamics of a nanodisc by integrating nmr, saxs and sans experiments with molecular dynamics simulations". *Elife*, 9:e56518, 2020.

[84] Doktorova, M., Kučerka, N., Kinnun, J. J., Pan, J., Marquardt, D., Scott, H. L., Venable, R. M., Pastor, R. W., Wassall, S. R., Katsaras, J., et al. "Molecular structure of sphingomyelin in fluid phase bilayers determined by the joint analysis of small-angle neutron and x-ray scattering data". *J. Phys. Chem. B*, 124(25):5186–5200, 2020.

[85] Penfold, J. and Thomas, R. K. "The application of the specular reflection of neutrons to the study of surfaces and interfaces". *J. Phys. Condens. Matter*, 2(6):1369, 1990.

[86] Abelès, F. and La, F. A. "théorie générale des couches minces". *J. Phys. Radium*, 11:307–309, 1950.

[87] Vineyard, G. H. "Grazing-incidence diffraction and the distorted-wave approximation for the study of surfaces". *Phys. Rev. B*, 26(8):4146, 1982.

[88] Kanduč, M., Schneck, E., and Stubenrauch, C. "Intersurfactant h-bonds between head groups of n-dodecyl-β-d-maltoside at the air-water interface". *J. Colloid Interface Sci.*, 586:588–595, 2021.

[89] Ibrahim, M., Wenzel, C., Lallemang, M., Balzer, B. N., and Schwierz, N. "Adsorbing dna to mica by cations: Influence of valency and ion type". *Langmuir*, 2023.

[90] Haes, A. J., Chang, L., Klein, W. L., and Van Duyne, R. P. "Detection of a biomarker for alzheimer's disease from synthetic and clinical samples using a nanoscale optical biosensor". *J. Am. Chem. Soc.*, 127(7):2264–2271, 2005.

[91] Rothemund, P. W. "Folding dna to create nanoscale shapes and patterns". *Nature*, 440(7082):297–302, 2006.

[92] Li, H., Cao, J., Zheng, W., Chen, Y., Wu, D., Dang, W., Wang, K., Peng, H., and Liu, Z. "Controlled synthesis of topological insulator nanoplate arrays on mica". *J. Am. Chem. Soc.*, 134(14):6132–6135, 2012.

[93] Riboh, J. C., Haes, A. J., McFarland, A. D., Ranjit Yonzon, C., and Van Duyne, R. P. "A nanoscale optical biosensor: real-time immunoassay in physiological buffer enabled by improved nanoparticle adhesion". *J. Phys. Chem. B*, 107(8):1772–1780, 2003.

[94] Woo, S. and Rothemund, P. W. "Self-assembly of two-dimensional dna origami lattices using cation-controlled surface diffusion". *Nat. Commun.*, 5:4889, 2014.

[95] Hansma, H., Vesenka, J., Siegerist, C., Kelderman, G., Morrett, H., Sinsheimer, R. L., Elings, V., Bustamante, C., and Hansma, P. "Reproducible imaging and dissection of plasmid dna under liquid with the atomic force microscope". *Science*, 256(5060):1180–1184, 1992.

[96] Hansma, H. G., Bezanilla, M., Zenhausern, F., Adrian, M., and Sinsheimer, R. L. "Atomic force microscopy of dna in aqueous solutions". *Nucleic Acids Res.*, 21(3):505–512, 1993.

[97] Guthold, M., Bezanilla, M., Erie, D. A., Jenkins, B., Hansma, H. G., and Bustamante, C. "Following the assembly of rna polymerase-dna complexes in aqueous solutions with the scanning force microscope.". *Proc. Natl. Acad. Sci. U.S.A*, 91(26):12927–12931, 1994.

[98] Hansma, H., Kim, K., Laney, D., Garcia, R., Argaman, M., Allen, M., and Parsons, S. "Properties of biomolecules measured from atomic force microscope images: a review". *J. Struct. Biol.*, 119(2):99–108, 1997.

[99] Shibata, M., Nishimasu, H., Kodera, N., Hirano, S., Ando, T., Uchihashi, T., and Nureki, O. "Real-space and real-time dynamics of crispr-cas9 visualized by high-speed atomic force microscopy". *Nat. Commun.*, 8(1):1430, 2017.

[100] Sorel, I., Piétrement, O., Hamon, L., Baconnais, S., Le Cam, E., and Pastré, D. "The ecori- dna complex as a model for investigating protein- dna interactions by atomic force microscopy". *Biochemistry*, 45(49):14675–14682, 2006.

[101] Crampton, N., Bonass, W. A., Kirkham, J., Rivetti, C., and Thomson, N. H. "Collision events between rna polymerases in convergent transcription studied by atomic force microscopy". *Nucleic Acids Res.*, 34(19):5416–5425, 2006.

[102] Hansma, H. G. "Surface biology of dna by atomic force microscopy". *Annu. Rev. Phys. Chem.*, 52(1):71–92, 2001.

[103] Bezanilla, M., Manned, S., Laney, D. E., Lyubchenko, Y. L., and Hansma, H. G. "Adsorption of dna to mica, silylated mica, and minerals: Characterization by atomic force microscopy". *Langmuir*, 11:655–659, 1995.

[104] Japaridze, A., Vobornik, D., Lipiec, E., Cerreta, A., Szczerbinski, J., Zenobi, R., and Dietler, G. "Toward an effective control of dna's submolecular conformation on a surface". *Macromolecules*, 49(2):643–652, 2016.

[105] "Polynucleotide adsorption to negatively charged surfaces in divalent salt solutions". *Biophys. J.*, 90:1164–1174, 2006.

[106] Pastré, D., Piétrement, O., Fusil, S., Landousy, F., Jeusset, J., David, M.-O., Hamon, L., Le Cam, E., and Zozime, A. "Adsorption of dna to mica mediated by divalent counterions: a theoretical and experimental study". *Biophys. J.*, 85(4):2507–2518, 2003.

[107] Rouzina, I. and Bloomfield, V. A. "Macroion attraction due to electrostatic correlation between screening counterions. 1. mobile surface-adsorbed ions and diffuse ion cloud". *J. Phys. Chem.*, 100(23):9977–9989, 1996.

[108] Netz, R. R. and Orland, H. "Beyond poisson-boltzmann: Fluctuation effects and correlation functions". *Eur. Phys. J. E*, 1:203–214, 2000.

[109] Kalcher, I., Schulz, J. C., and Dzubiella, J. "Electrolytes in a nanometer slab-confinement: Ion-specific structure and solvation forces". *J. Chem. Phys.*, 133(16):164511, 2010.

[110] Ray, J. and Manning, G. S. "An attractive force between two rodlike polyions mediated by the sharing of condensed counterions". *Langmuir*, 10(7):2450–2461, 1994.

[111] Izrailev, S., Stepaniants, S., Balsera, M., Oono, Y., and Schulten, K. "Molecular dynamics study of unbinding of the avidin-biotin complex". *Biophys. J.*, 72(4): 1568–1581, 1997.

[112] Grubmüller, H., Heymann, B., and Tavan, P. "Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force". *Science*, 271(5251):997–999, 1996.

[113] Schwierz, N., Horinek, D., Liese, S., Pirzer, T., Balzer, B. N., Hugel, T., and Netz, R. R. "On the relationship between peptide adsorption resistance and surface contact angle: A combined experimental and simulation single-molecule study". *J. Am. Chem. Soc.*, 134(48):19628–19638, 2012.

[114] Schwierz, N., Krysiak, S., Hugel, T., and Zacharias, M. "Mechanism of reversible peptide-bilayer attachment: combined simulation and experimental single-molecule study". *Langmuir*, 32(3):810–821, 2016.

[115] Horinek, D., Serr, A., Geisler, M., Pirzer, T., Slotta, U., Lud, S. Q., Garrido, J., Scheibel, T., Hugel, T., and Netz, R. R. "Peptide adsorption on a hydrophobic surface results from an interplay of solvation, surface, and intrapeptide forces". *Proc. Natl. Acad. Sci. U.S.A.*, 105(8):2842–2847, 2008.

[116] Rief, M., Oesterhelt, F., Heymann, B., and Gaub, H. E. "Single molecule force spectroscopy on polysaccharides by atomic force microscopy". *Science*, 275(5304): 1295–1297, 1997.

[117] Bano, F., Sluysmans, D., Wislez, A., and Duwez, A.-S. "Unraveling the complexity of the interactions of dna nucleotides with gold by single molecule force spectroscopy". *Nanoscale*, 7(46):19528–19533, 2015.

[118] Iliafar, S., Wagner, K., Manohar, S., Jagota, A., and Vezenov, D. "Quantifying interactions between dna oligomers and graphite surface using single molecule force spectroscopy". *J. Phys. Chem. C*, 116(26):13896–13903, 2012.

[119] Wei, G., Li, Q., Steckbeck, S., and Ciacchi, L. C. "Direct force measurements on peeling heteropolymer ssdna from a graphite surface using single-molecule force spectroscopy". *Phys. Chem. Chem. Phys.*, 16(9):3995–4001, 2014.

[120] Shlyakhtenko, L. S., Dutta, S., Banga, J., Li, M., Harris, R. S., and Lyubchenko, Y. L. "Apobec3g interacts with ssdna by two modes: Afm studies". *Sci. Rep.*, 5(1):1–12, 2015.

[121] Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. "Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers". *SoftwareX*, 1-2:19–25, 2015.

[122] Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. "Lincs: A linear constraint solver for molecular simulations". *J. Comput. Chem.*, 18(12):1463–1472, 1997.

[123] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. "Comparison of simple potential functions for simulating liquid water". *J. Chem. Phys.*, 79(2):926–935, 1983.

[124] Ivani, I., Dans, P. D., Noy, A., Pérez, A., Faustino, I., Hospital, A., Walther, J., Andrio, P., Goñi, R., Balaceanu, A., Portella, G., Battistini, F., Gelpí, J. L., González, C., Vendruscolo, M., Laughton, C. A., Harris, S. A., Case, D. A., and Orozco, M. "Parmbsc1: A refined force field for dna simulations". *Nat. Methods*, 13:55–58, 2015.

[125] Cygan, R. T., Liang, J. J., and Kalinichev, A. G. "Molecular models of hydroxide, oxyhydroxide, and clay phases and the development of a general force field". *J. Phys. Chem. B*, 108:1255–1266, 2004.

[126] Mamatkulov, S. and Schwierz, N. "Force fields for monovalent and divalent metal cations in tip3p water based on thermodynamic and kinetic properties". *J. Chem. Phys.*, 148, 2018.

[127] Grotz, K. K., Cruz-León, S., and Schwierz, N. "Optimized magnesium force field parameters for biomolecular simulations with accurate solvation, ion-binding, and water-exchange properties". *J. Chem. Theory Comput.*, 17(4):2530–2540, 2021.

[128] Cruz-León, S., Grotz, K. K., and Schwierz, N. "Extended magnesium and calcium force field parameters for accurate ion-nucleic acid interactions in biomolecular simulations". *J. Chem. Phys.*, 154(17):171102, 2021.

[129] Michaud-Agrawal, N., Denning, E. J., Woolf, T. B., and Beckstein, O. "Mdanalysis: A toolkit for the analysis of molecular dynamics simulations". *J. Comput. Chem.*, 32:2319–2327, 2011.

[130] Humphrey, W., Dalke, A., and Schulten, K. "Vmd: Visual molecular dynamics". *J. Mol. Graphics*, 14(1):33–38, 1996.

[131] Hummer, G. and Szabo, A. "Kinetics from nonequilibrium single-molecule pulling experiments". *Biophys. J.*, 85(1):5–15, 2003.

[132] Dudko, O. K., Hummer, G., and Szabo, A. "Intrinsic rates and activation free energies from single-molecule pulling experiments". *Phys. Rev. Lett.*, 96(10):108101, 2006.

[133] Bell, G. I. "Models for the specific adhesion of cells to cells: a theoretical framework for adhesion mediated by reversible bonds between cell surface molecules.". *Science*, 200(4342):618–627, 1978.

[134] Bullerjahn, J. T., Sturm, S., and Kroy, K. "Theory of rapid force spectroscopy". *Nat. Commun.*, 5(1):4463, 2014.

[135] Evans, E. and Ritchie, K. "Dynamic strength of molecular adhesion bonds". *Biophys. J.*, 72(4):1541–1555, 1997.

[136] Hummer, G. and Szabo, A. "Free energy profiles from single-molecule pulling experiments". *PNAS*, 107, 2010.

[137] Schwesinger, F., Ros, R., Strunz, T., Anselmetti, D., Güntherodt, H.-J., Honegger, A., Jermutus, L., Tiefenauer, L., and Plückthun, A. "Unbinding forces of single antibody-antigen complexes correlate with their thermal dissociation rates". *Proc. Natl. Acad. Sci. U.S.A.*, 97(18):9972–9977, 2000.

[138] Marcus, Y. *Ion properties.* Inc., New York, Basel, 1997.

[139] Hall, M. N., Gabay, J., Débarbouillé, M., and Schwartz, M. "A role for mrna secondary structure in the control of translation initiation". *Nature*, 295(5850):616–618, 1982.

[140] de Smit, M. H. and van Duin, J. "Control of translation by mrna secondary structure in escherichia coli: a quantitative analysis of literature data". *J. Mol. Biol.*, 244(2): 144–150, 1994.

[141] Chen, C., Zhang, H., Broitman, S. L., Reiche, M., Farrell, I., Cooperman, B. S., and Goldman, Y. E. "Dynamics of translation by single ribosomes through mrna secondary structures". *Nat. Struct. Mol. Biol.*, 20(5):582–588, 2013.

[142] Chemla, Y., Peeri, M., Heltberg, M. L., Eichler, J., Jensen, M. H., Tuller, T., and Alfonta, L. "A possible universal role for mrna secondary structure in bacterial translation revealed using a synthetic operon". *Nat. Commun.*, 11(1):4827, 2020.

[143] Stein, A. and Crothers, D. "Conformational changes of transfer rna. the role of magnesium (ii)". *Biochemistry*, 15(1):160–168, 1976.

[144] Sanchez de Groot, N., Armaos, A., Graña-Montes, R., Alriquet, M., Calloni, G., Vabulas, R. M., and Tartaglia, G. G. "Rna structure drives interaction with proteins". *Nat. Commun.*, 10(1):3246, 2019.

[145] Wassenaar, T. A., Ingólfsson, H. I., Bockmann, R. A., Tieleman, D. P., and Marrink, S. J. "Computational lipidomics with insane: a versatile tool for generating custom membranes for molecular simulations". *J. Chem. Theory Comput.*, 11(5):2144–2155, 2015.

[146] de Jong, D. H., Singh, G., Bennett, W. D., Arnarez, C., Wassenaar, T. A., Schafer, L. V., Periole, X., Tieleman, D. P., and Marrink, S. J. "Improved parameters for the martini coarse-grained protein force field". *J. Chem. Theory Comput.*, 9(1):687–697, 2013.

[147] Uusitalo, J. J., Ingólfsson, H. I., Marrink, S. J., and Faustino, I. "Martini coarse-grained force field: extension to rna". *Biophys. J.*, 113(2):246–256, 2017.

[148] Lorenz, R., Bernhart, S. H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F., and Hofacker, I. L. "Viennarna package 2.0". *Algorithms Mol Biol*, 6:1–14, 2011.

[149] Popenda, M., Szachniuk, M., Antczak, M., Purzycka, K. J., Lukasiak, P., Bartol, N., Blazewicz, J., and Adamiak, R. W. "Automated 3d structure composition for large rnas". *Nucleic Acids Res.*, 40(14):e112–e112, 2012.

[150] Díaz-Francés, E. and Rubio, F. J. "On the existence of a normal approximation to the distribution of the ratio of two independent normal random variables". *Stat Papers*, 54:309–323, 2013.

[151] Sedlak, S. M., Bruetzel, L. K., and Lipfert, J. "Quantitative evaluation of statistical errors in small-angle x-ray scattering measurements". *J. Appl. Crystallogr.*, 50(2): 621–630, 2017.

[152] Wassenaar, T. A., Pluhackova, K., Bockmann, R. A., Marrink, S. J., and Tieleman, D. P. "Going backward: a flexible geometric approach to reverse transformation from coarse grained to atomistic models". *J. Chem. Theory Comput.*, 10(2):676–690, 2014.

[153] López, C. A., Sovova, Z., van Eerden, F. J., de Vries, A. H., and Marrink, S. J. "Martini force field parameters for glycolipids". *J. Chem. Theory Comput.*, 9(3):1694–1708, 2013.

[154] Ibrahim, M., Gilbert, J., Heinz, M., Nylander, T., and Schwierz, N. "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.

[155] Mitchell, M. J., Billingsley, M. M., Haley, R. M., Wechsler, M. E., Peppas, N. A., and Langer, R. "Engineering precision nanoparticles for drug delivery". *Nat. Rev. Drug Discov*, 20(2):101–124, 2021.

[156] Cullis, P. R. and Hope, M. J. "Lipid nanoparticle systems for enabling gene therapies". *Mol Ther.*, 25(7):1467–1475, 2017.

[157] Adams, D., Gonzalez-Duarte, A., O' Riordan, W. D., Yang, C.-C., Ueda, M., Kristen, A. V., Tournev, I., Schmidt, H. H., Coelho, T., Berk, J. L., Lin, K.-P., Vita, G., Attarian, S., Planté-Bordeneuve, V., Mezei, M. M., Campistol, J. M., Buades, J., Brannagan, T. H., Kim, B. J., Oh, J., Parman, Y., Sekijima, Y., Hawkins, P. N., Solomon, S. D., Polydefkis, M., Dyck, P. J., Gandhi, P. J., Goyal, S., Chen, J., Strahs, A. L., Nochur, S. V., Sweetser, M. T., Garg, P. P., Vaishnaw, A. K., Gollob, J. A., and Suhr, O. B. "Patisiran, an rnai therapeutic, for hereditary transthyretin amyloidosis". *N. Engl. J. Med.*, 379(1):11–21, 2018.

[158] Kulkarni, J. A., Cullis, P. R., and van der Meel, R. "Lipid nanoparticles enabling gene therapies: From concepts to clinical utility". *Nucleic Acid Ther.*, 28(3):146–157, 2018.

[159] Ermilova, I. and Swenson, J. "DOPC versus DOPE as a helper lipid for gene-therapies: molecular dynamics simulations with DLin-MC3-DMA". *Phys Chem Chem Phys*, 22 (48):28256–28268, 2020.

[160] Park, S., Choi, Y. K., Kim, S., Lee, J., and Im, W. "Charmm-gui membrane builder for lipid nanoparticles with ionizable cationic lipids and pegylated lipids". *J. Chem. Inf. Model*, 61(10):5192–5202, 2021.

[161] Klauda, J. B., Kucerka, N., Brooks, B. R., Pastor, R. W., and Nagle, J. F. "Simulation-based methods for interpreting x-ray data from lipid bilayers". *Biophys. J.*, 90(8): 2796–2807, 2006.

[162] Smith, D. J., Klauda, J. B., and Sodt, A. J. "Simulation best practices for lipid membranes [article v1.0]". *Living J. Comp. Mol. Sci.*, 1(1):5966, 2019.

[163] Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J. A., Jr., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, Ö., Foresman, J. B., Ortiz, J. V., Cioslowski, J., and D. J. Fox", t. R. E. *Gaussian, Inc., Wallingford CT, 2009.* Gaussian Inc. Wallingford CT 2009.

[164] Dupradeau, F.-Y., Pigache, A., Zaffran, T., Savineau, C., Lelong, R., Grivel, N., Lelong, D., Rosanski, W., and Cieplak, P. "The r.e.d. tools: advances in resp and esp charge derivation and force field library building". *Phys. Chem. Chem. Phys.*, 12:7821–7839, 2010.

[165] Sousa da Silva, A. W. and Vranken, W. F. "Acpype - antechamber python parser interface". *BMC Res. Notes*, 5(1):367, 2012.

[166] Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. "Development and testing of a general amber force field". *J. Comput. Chem.*, 25(9):1157–1174, 2004.

[167] Shirts Michael R., Klein Christoph, Swails Jason M., Yin Jian, Gilson Michael K., Mobley David L., Case David A., and Zhong Ellen D. "Lessons learned from comparing molecular dynamics engines on the SAMPL5 dataset". *J Comput Aided Mol Des*, 31(1):147–161, 2017.

[168] Jo, S., Kim, T., Iyer, V. G., and Im, W. "Charmm-gui: A web-based graphical user interface for charmm". *J. Comput. Chem.*, 29(11):1859–1865, 2008.

[169] Knight, C. J. and Hub, J. S. "MemGen: a general web server for the setup of lipid membrane simulation systems". *Bioinformatics*, 31(17):2897–2899, 2015.

[170] Darré, L., Iglesias-Fernandez, J., Kohlmeyer, A., Wacklin, H., and Domene, C. "Molecular dynamics simulations and neutron reflectivity as an effective approach to characterize biological membranes and related macromolecular assemblies". *J. Chem. Theory Comput.*, 11(10):4875–4884, 2015.

[171] Koutsioubas, A. "Combined coarse-grained molecular dynamics and neutron reflectivity characterization of supported lipid membranes". *J. Phys. Chem. B*, 120(44): 11474–11483, 2016.

[172] Nelson, A. R. J. and Prescott, S. W. "refnx: neutron and x-ray reflectometry analysis in python". *J. Appl. Crystallogr.*, 52(Pt 1):193–200, 2019.

[173] Charitat, T., Bellet-Amalric, E., Fragneto, G., and Graner, F. "Adsorbed and free lipid bilayers at the solid-liquid interface". *Eur. Phys. J. B*, 8(4):583–593, 1999.

[174] Fragneto-Cusani, G. "Neutron reflectivity at the solid/liquid interface: examples of applications in biophysics". *J. Phys. Condens. Matter*, 13(21):4973, 2001.

[175] König, B. W., Krueger, S., Orts, W., Majkrzak, C. F., Berk, N. F., Silverton, J., and Gawrisch, K. "Neutron reflectivity and atomic force microscopy studies of a lipid bilayer in water adsorbed to the surface of a silicon single crystal". *Langmuir*, 12(5): 1343–1350, 1996.

[176] Tristram-Nagle, S., Petrache, H. I., and Nagle, J. F. "Structure and interactions of fully hydrated dioleoylphosphatidylcholine bilayers". *Biophys. J.*, 75(2):917–925, 1998.

[177] Kučerka, N., Nagle, J. F., Sachs, J. N., Feller, S. E., Pencer, J., Jackson, A., and Katsaras, J. "Lipid bilayer structure determined by the simultaneous analysis of neutron and x-ray scattering data". *Biophys. J.*, 95(5):2356–2367, 2008.

[178] "Structure of lipid bilayers". *Biochim. Biophys .Acta Biomembr*, 1469(3):159–195, 2000.

[179] Liu, Y. and Nagle, J. F. "Diffuse scattering provides material parameters and electron density profiles of biomembranes". *Phys. Rev. E*, 69:040901, 2004.

[180] Kulkarni, J. A., Darjuan, M. M., Mercer, J. E., Chen, S., van der Meel, R., Thewalt, J. L., Tam, Y. Y. C., and Cullis, P. R. "On the formation and morphology of lipid nanoparticles containing ionizable cationic lipids and sirna". *ACS Nano*, 12(5): 4787–4795, 2018.

[181] Kulkarni, J. A., Witzigmann, D., Leung, J., van der Meel, R., Zaifman, J., Darjuan, M. M., Grisch-Chan, H. M., Thöny, B., Tam, Y. Y. C., and Cullis, P. R. "Fusion-dependent formation of lipid nanoparticles containing macromolecular payloads". *Nanoscale*, 11:9023–9031, 2019.

[182] Ramezanpour, M., Schmidt, M. L., Bodnariuc, I., Kulkarni, J. A., Leung, S. S. W., Cullis, P. R., Thewalt, J. L., and Tieleman, D. P. "Ionizable amino lipid interactions with popc: implications for lipid nanoparticle function". *Nanoscale*, 11:14141–14146, 2019.

[183] Emami, F. S., Puddu, V., Berry, R. J., Varshney, V., Patwardhan, S. V., Perry, C. C., and Heinz, H. "Force field and a surface model database for silica to simulate interfacial properties in atomic resolution". *Chem. Mater.*, 26(8):2647–2658, 2014.

[184] Flyvbjerg, H. and Petersen, H. G. "Error estimates on averages of correlated data". *J. Chem. Phys.*, 91(1):461–466, 1989.

[185] Grossfield, A. and Zuckerman, D. M. "Quantifying uncertainty and sampling quality in biomolecular simulations". *Annu Rep Comput Chem.*, 5:23–48, 2009.

[186] Grava, M., Ibrahim, M., Sudarsan, A., Pusterla, J., Philipp, J., Rädler, J. O., Schwierz, N., and Schneck, E. "Combining molecular dynamics simulations and x-ray scattering techniques for the accurate treatment of protonation degree and packing of ionizable lipids in monolayers". *J. Chem. Phys.*, 159(15), 2023.

[187] Li, Z., Carter, J., Santos, L., Webster, C., van der Walle, C. F., Li, P., Rogers, S. E., and Lu, J. R. "Acidification-induced structure evolution of lipid nanoparticles correlates with their in vitro gene transfections". *ACS nano*, 17(2):979–990, 2023.

[188] Philipp, J., Dabkowska, A., Reiser, A., Frank, K., Krzysztoń, R., Brummer, C., Nickel, B., Blanchet, C. E., Sudarsan, A., Ibrahim, M., Johansson, S., Skantze, P., Skantze, U., Östman, S., Johansson, M., Henderson, N., Elvevold, K., Smedsrød, B., Schwierz, N., Lindfors, L., and Rädler, J. O. "ph-dependent structural transitions in cationic ionizable lipid mesophases are critical for lipid nanoparticle function". *Proc. Natl. Acad. Sci. U.S.A.*, 120(50):e2310491120, 2023.

[189] Carrasco, M. J., Alishetty, S., Alameh, M.-G., Said, H., Wright, L., Paige, M., Soliman, O., Weissman, D., Cleveland, T. E., Grishaev, A., and Buschmann, M. D. "Ionization and structural properties of mRNA lipid nanoparticles influence expression in intramuscular and intravascular administration". *Communications Biology*, 4(1):956, 2021.

[190] Donnini, S., Tegeler, F., Groenhof, G., and Grubmüller, H. "Constant ph molecular dynamics in explicit solvent with λ-dynamics". *J. Chem. Theory Comput.*, 7(6): 1962–1978, 2011.

[191] Bennett, W. D., Chen, A. W., Donnini, S., Groenhof, G., and Tieleman, D. P. "Constant ph simulations with the coarse-grained martini model—application to oleic acid aggregates". *Can. J. Chem.*, 91(9):839–846, 2013.

[192] Aho, N., Buslaev, P., Jansen, A., Bauer, P., Groenhof, G., and Hess, B. "Scalable constant ph molecular dynamics in gromacs". *J. Chem. Theory Comput.*, 18(10): 6148–6160, 2022.

[193] Grünewald, F., Souza, P. C., Abdizadeh, H., Barnoud, J., de Vries, A. H., and Marrink, S. J. "Titratable martini model for constant ph simulations". *J. Chem. Phys.*, 153 (2), 2020.

[194] Brezesinski, G. and Schneck, E. "Investigating ions at amphiphilic monolayers with x-ray fluorescence". *Langmuir*, 35(26):8531–8542, 2019.

[195] Schneck, E., Schubert, T., Konovalov, O. V., Quinn, B. E., Gutsmann, T., Brandenburg, K., Oliveira, R. G., Pink, D. A., and Tanaka, M. "Quantitative determination of ion distributions in bacterial lipopolysaccharide membranes by grazing-incidence x-ray fluorescence". *Proc. Natl. Acad. Sci. U.S.A.*, 107(20):9147–9151, 2010.

[196] Thompson, A., Vaughan, D., et al. *X-ray Data Booklet, Tables 1-3*. Lawrence Berkeley National Laboratory, University of California, CA, USA, 2001.

[197] Rowlinson, J. S. and Widom, B. *Molecular Theory of Capillarity*. Clarendon Press, Oxford, 1982.

[198] Vega, C. and de Miguel, E. "Surface tension of the most popular models of water by using the test-area simulation method". *J. Chem. Phys.*, 126(15), 2007.

[199] Helm, C. A., Möhwald, H., Kjaer, K., and Als-Nielsen, J. "Phospholipid monolayer density distribution perpendicular to the water surface. a synchrotron x-ray reflectivity study". *EPL*, 4(6):697, 1987.

[200] Schwierz, N., Horinek, D., and Netz, R. R. "Specific ion binding to carboxylic surface groups and the ph dependence of the hofmeister series". *Langmuir*, 31(1):215–225, 2015.

[201] Baoukina, S., Marrink, S. J., and Tieleman, D. P. *Biomembrane Frontiers: Nanostructures, Models, and the Design of Life*. Springer, 2009.

[202] Tempra, C., Ollila, O. S., and Javanainen, M. "Accurate simulations of lipid monolayers require a water model with correct surface tension". *J. Chem. Theory Comput.*, 18(3):1862–1869, 2022.

[203] Damjanović, A., Brooks, B. R., and García-Moreno E, B. "Conformational relaxation and water penetration coupled to ionization of internal groups in proteins". *J. Phys. Chem. A*, 115(16):4042–4053, 2011.

[204] Ullmann, R. T. and Ullmann, G. M. "Coupling of protonation, reduction, and conformational change in azurin from pseudomonas aeruginosa investigated with free energy measures of cooperativity". *J. Phys. Chem. B*, 115(34):10346–10359, 2011.

[205] Leung, A. K., Hafez, I. M., Baoukina, S., Belliveau, N. M., Zhigaltsev, I. V., Afshinmanesh, E., Tieleman, D. P., Hansen, C. L., Hope, M. J., and Cullis, P. R. "Lipid nanoparticles containing sirna synthesized by microfluidic mixing exhibit an electron-dense nanostructured core". *J. Phys. Chem. C*, 116(34):18440–18450, 2012.

[206] Yanez Arteta, M., Kjellman, T., Bartesaghi, S., Wallin, S., Wu, X., Kvist, A. J., Dabkowska, A., Székely, N., Radulescu, A., Bergenholtz, J., and Lindfors, L. "Successful reprogramming of cellular protein production through mrna delivered by functionalized lipid nanoparticles". *Proceedings of the National Academy of Sciences*, 115(15):E3351–E3360, 2018.

[207] Sebastiani, F., Yanez Arteta, M., Lerche, M., Porcar, L., Lang, C., Bragg, R. A., Elmore, C. S., Krishnamurthy, V. R., Russell, R. A., Darwish, T., et al. "Apolipoprotein e binding drives structural and compositional rearrangement of mrna-containing lipid nanoparticles". *ACS nano*, 15(4):6709–6722, 2021.

[208] Boyd, K. J. and May, E. R. "Bumpy: a model-independent tool for constructing lipid bilayers of varying curvature and composition". *J. Chem. Theory Comput.*, 14 (12):6642–6652, 2018.

[209] Marrink, S. J. and Mark, A. E. "Molecular dynamics simulation of the formation, structure, and dynamics of small phospholipid vesicles". *J. Am. Chem. Soc.*, 125 (49):15233–15242, 2003.

[210] Vögele, M., Köfinger, J., and Hummer, G. "Molecular dynamics simulations of carbon nanotube porins in lipid bilayers". *Faraday Discuss.*, 209:341–358, 2018.

[211] Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. "A second generation force field for the simulation of proteins, nucleic acids, and organic molecules j. am. chem. soc. 1995, 117, 5179- 5197". *J. Am. Chem. Soc.*, 118(9):2309–2309, 1996.

[212] Pérez, A., Marchán, I., Svozil, D., Sponer, J., Cheatham, T. E., Laughton, C. A., and Orozco, M. "Refinement of the amber force field for nucleic acids: improving the description of α/γ conformers". *Biophys. J.*, 92(11):3817–3829, 2007.

[213] Zgarbová, M., Otyepka, M., Šponer, J., Mládek, A., Banáš, P., Cheatham III, T. E., and Jurecka, P. "Refinement of the cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles". *J. Chem. Theory Comput.*, 7(9):2886–2902, 2011.

[214] Köfinger, J., Stelzl, L. S., Reuter, K., Allande, C., Reichel, K., and Hummer, G. "Efficient ensemble refinement by reweighting". *J. Chem. Theory Comput.*, 15(5): 3390–3401, 2019.

[215] Hilpert, C., Beranger, L., Souza, P. C., Vainikka, P. A., Nieto, V., Marrink, S. J., Monticelli, L., and Launay, G. "Facilitating cg simulations with mad: The martini database server". *J. Chem. Inf. Model.*, 63(3):702–710, 2023.

Appendices

Some Useful Jupyter Notebooks and Python Scripts

Here I provide a short description of the *git-hub* links containing some useful scripts and notebooks that were developed as a part of the thesis.

A.1   Neutron reflectivity profile from simulations

. The script and notebook provided on the following link calculates SLD profiles, obtains the substrate parameters, and optimizes the water patch fraction (γ) and water fraction at the silica-bilayer water interface (α). These scripts are used in chapter 5.

- https://github.com/ibrahim-mohd/Neutron-reflectivity-from-simulation.

Example files are also provided.

A.2   Simulation of silica surface at arbitrary pH

In chapter 5, we theoretically modeled the silica surface supporting the bilayer. To explicitly simulate silica we need the corresponding topology files for silica. The following script creates such files with arbitrary surface protonation degrees.

- https://github.com/ibrahim-mohd/Generate-silica-itp-file-Gromacs.

Example files are also provided.

A.3   GIXOS profiles from simulations

GIXOS curves in chapter 6 were obtained and compared with experiments using the following notebook.

- https://github.com/ibrahim-mohd/GIXOS-from-Simulations

Example files are also provided.

A.4   SAXS and SANS form factors from simulations

SAXS and SANS form factors presented in chapter 4 were obtained using the following notebook.

- https://github.com/ibrahim-mohd/Small-angle-X-ray-Form-factor-from-simulations

Example files are also provided.

A.5   MC3 force field files

The force field for MC3 in neutral and cationic form with example setup is available at the following link.

- https://github.com/bio-phys/ForceFieldsMC3

Acknowledgments

This thesis is a culmination of a long and arduous academic journey that would not have been possible without the beautiful people I am blessed to have in my life and the amazing and kindest of people I had the pleasure to come across as friends, teachers, colleagues, and collaborators. So many people have helped me throughout this journey, and it probably would not be possible to list all the names who directly or indirectly contributed to this journey.

I am deeply indebted to my supervisor Prof. Dr. Nadine Schwierz for giving me the opportunity to come here to Germany and carry out my Ph.D. Thank you Nadine for the trust, constant support, and encouragement throughout the years. I learned a lot of highly transferable skills both scientific and non-scientific which helped me grow as a researcher and will continue to be beneficial throughout my career. The constant scientific exchanges and your insightful advice during the group meetings, or outside of it always helped crystallize my otherwise amorphous, and vague ideas, and always provided a new perspective to look at the scientific problem at hand. Thank you also for establishing collaboration with many experimental groups that complemented all my work.

In all my projects, I had the great opportunity to work in close collaboration with experiments. I would like to thank Dr. Max Lallemang, Christiane Wenzel, and Dr. Bizan N. Balzer, at the University of Freiburg for the collaboration on DNA-Mica interactions. Thanks to Jennifer Gilbert and Prof. Dr. Tommy Nylander at Lund University, Sweden, for collaborating on the MC3 parametrization project. I would like to acknowledge Miriam Grava and Prof. Dr. Emanuel Schneck at the Technical University of Darmstadt for their collaboration on the monolayer protonation project. And thanks to Julian Philipp and Prof. Joachim O. Rädler at Ludwig-Maximilians-Universität, Munich, for involving us in the MC3/cholesterol mesophase project. I would like to thank Dr. Marcel Heinz and PD Dr. Jürgen Köfinger at Max Planck Institute for Biophysics, Frankfurt for collaborating on the MC3 parameterization project and RNA conformation project respectively.

As a part of the Emmy Noether group, I spent most of my Ph.D. time ($\approx 3\frac{1}{2}$ years) at the Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Frankfurt. The time spent there was an enlightening experience for me, I was exposed to a broad range of topics in computational sciences and closely witnessed the process of making of many cool works. For this, I thank the whole Theoretical Biophysics Department group members for the supportive work atmosphere. Within the Theoretical Biophysics department, I would like to especially thank PD Dr. Jürgen Köfinger for his constant feedback during the regular biweekly "PhD meetings", no matter how small or big the scientific problem at hand was, he always provided very detailed and insightful feedback. I also thank Prof. Dr. Gerhard Hummer for the opportunity to attend and present in his group meetings and for insightful feedback.

I thank past Emmy Noether group members, Sergio and Kara for their support and for sharing their expertise in MD simulations especially at the beginning of my Ph.D.