October 2021

MODELING STRUCTURES AND DYNAMICS OF POLYELECTROLYTES

Sadhana Chalise
University of Massachusetts Amherst

MODELING STRUCTURES AND DYNAMICS OF POLYELECTROLYTES

A Dissertation Presented

by

SADHANA CHALISE

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2021

Department of Polymer Science and Engineering

MODELING STRUCTURES AND DYNAMICS OF POLYELECTROLYTES

A Dissertation Presented

by

SADHANA CHALISE

Approved as to style and content by:

Murugappan Muthukumar, Chair

David A Hoagland, Member

Todd S Emrick, Member

Sarah Perry, Member

David A Hoagland, Department Head
Department of Polymer Science and Engineering

Dedication

TO MY FAMILY

ACKNOWLEDGEMENTS

I want to express my sincere gratitude to my advisor Professor Dr. Murugappan Muthukumar, for his guidance, encouragement, and support on this work throughout these years in his research lab. I am grateful for all the freedom and flexibility he provided during these years, which helped me learn at my own pace, especially during my journey of motherhood.

I want to thank my committee members Professor Dr. David Hoagland, Professor Dr. Sarah Perry, and Professor Dr. Todd Emrick, for their valuable time, feedback, and fruitful discussions to make my dissertation and presentation better.

I would like to thank my collaborator Dr. Jyoti Prakash Mahalik, for collaborating with me. I would like to thank National Human Genome Research Institute (NHGRI) for their financial support.

I am thrilled to be a part of the PSE community. Each member of the community is helpful and cooperative. I want to thank the PSE community: faculty, staff, students, and post-docs for welcoming me into the community and helping me to make my journey easier. I would like to express my special thanks to the custodial staff for making the Conte a safe place even during Covid. Because of that, I was able to come to the lab and work without any fear. With a toddler at home, it would not have been easy to work from home.

Lisa Groth has always been like my family, far from my family. I would like to thank her and Lisa McNamara for their emotional support during my journey of motherhood. Both of them made me feel proud and happy for being a mother and doing science simultaneously. Lisa McNamara even provided me a comfortable chair which became a blessing to me during my pregnancy.

I want to thank the class of 2015 for their support and friendship. My special thanks to Ashlin, Sarah, Konane, Alexa, Carolyn, Abhiram, and Dylan for helping me with chemistry and cumulative exams, helping with presentations, and being available for all support.

I am thankful to all past and present members of the Muthu group I overlapped during these years for helping me to make my journey easier. Each one is an amazing human being. There are many memories of hiking, bowling, biking, 5K run, picnic, bbq, game nights, universal studio visit, and birthday celebrations (before Covid) with Michael, Debasis, Hamid, Prabhat, Kiran, Zack, Amy, Di, Uma, and Khatcher. In addition, I would like to thank Michael for being my mentor and giving valuable suggestions with the cumulative exams, Zack for helping me improve my writings, and Kiran for assisting me with technical support.

I thank my close friend Sunita for all the love and support. I thank all my friends from Nepal who lived in Amherst area, whom I would like to call my family far from the family. I want to thank all of them for cultural gatherings, potlucks, bbq, rafting, picnic, and festivals celebrations. Our stay in North Pleasant Street was pleasant because of the presence of those friends.

My journey so far would not have been possible without the love and support of my family. I am grateful to my parents for their unconditional love, care, and support. My mother flew to the US from Nepal to help me with my baby, and my father could not come due to Covid as he was supposed to travel after a couple of months. Both of them stayed long-distance for nine long months for the first time in their life. I am indebted and feel lucky and blessed to have them as my parents. I am grateful to my parents-in-law for their love and support. I feel blessed to have them in my life. They flew to the US from Nepal, in the middle of Covid, to help me with our baby. Thank my siblings: brother and sister, for always being there for me whenever I needed them. The trust that my family and friends have shown towards me has

always motivated me to do better in life. Finally, I want to express my sincere thanks and appreciation to my best friend of my life, my husband Sabin, for love, care, and support. I am grateful to him for sharing a ride with me, helping me become a better version of myself, being an excellent companion, and being a wonderful father of our son Sanidh.

ABSTRACT

MODELING STRUCTURES AND DYNAMICS OF POLYELECTROLYTES

SEPTEMBER 2021

SADHANA CHALISE

B.Sc., TRIBHUVAN UNIVERSITY, NEPAL

M.Sc., TRIBHUVAN UNIVERSITY, NEPAL

M.Sc., UNIVERSITÀ DEGLI STUDI DI TRIESTE, ITALY

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Murugappan Muthukumar

Charged polymers exhibit interesting and complex structures, functions, and dynamics in both natural and synthetic environments. The equilibrium and dynamic properties of charged polymers are determined by the interplay of entropy and enthalpy from electrostatic interactions between charged polymers, counterions, salt ions, and short-ranged Van der Waals interactions. This work is mostly focused on understanding the equilibrium and dynamic behaviors of charged polymers in different environments. We use computer modeling, mostly coarse-grained Langevin dynamics simulations, to simulate complex environments having charged polymers, solvents, and charged ions. Our topics of interest in this work include a comparison of electrostatic potential across a nanopore using atomistic and coarse-grained

approaches, unfolding of RNA hairpins in response to applied external forces, chain conformation of a tagged polyelectrolyte chain inside a polyelectrolyte complex, and effects of charge density, temperature and salt concentration on the aggregated structure of oppositely charged polymers in semi-dilute solutions.

We model the electrostatic potential across charge decorated α-Hemolysin protein nanopores using both all-atom molecular dynamics simulations and coarse-grained Langevin dynamics simulations. We observe that the coarse-grained method gives a good approximation of the atomistic approach.

We also model simple RNA hairpin architectures and a nanopore using coarse-grained method. We monitor the mechanism of unfolding when different RNA hairpin architectures of an equal number of nucleotides are passing through a nanopore under the application of an electric field. We find that the RNA with longer hairpins requires more force to unfold and translocate through the nanopore. We also observe a distinct signature of unfolding time for the bases before and after unpaired bases in the RNA hairpin models.

Next, by using coarse-grained Langevin dynamics simulations of polyelectrolytes of symmetric and flexible polyelectrolytes of opposite charges alongside explicit counterions and salt ions, we study the role of charge density, polymer concentration, temperature, and salt concentration on the structure and dynamics of complexes. In a system of highly charged polyelectrolytes, the average radius of gyration ($\langle R_g \rangle$) of a labeled chain and the size-scaling exponent ν of a single isolated chain are maximum, and they shrink when two charged polymers of opposite charges come together forming a complex. The $\langle R_g \rangle$ of a labeled chain inside a polyelectrolyte complex increases with increasing the size of the complex reaches a plateau once the reasonable size of

the complex is formed. The value of ν also increases and reaches a plateau of 0.5, indicating that the labeled chain inside the complex shows Gaussian-like statistics. We observed that in semi-dilute solutions of polyelectrolyte complexes, the formation of complex structures is enhanced with an increase in charge density of polymers and with a decrease in temperature. The aggregates are de-complexed with an increase in salt concentration. We also observed that in the semi-dilute regime $\langle R_g \rangle$ and ν of a labeled chain is independent of the size of the complex formed, the charge density of the polymer, temperature, and salt concentration and chains show Gaussian-like conformations. Further, we observed that an isolated polyelectrolyte chain shows diffusive behavior, but the labeled chain in the complexes follows non-diffusive dynamical law as the chain becomes a part of the physical network due to the presence of other chains in the complex. The results of these studies complement experimental studies and provide a more in-depth understanding of already observed phenomena.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

CHAPTER 1

INTRODUCTION

Synthetic and biological charged polymers have been subjects of many scientific studies because of their applications in fields like pharmaceuticals [1–3], wastewater treatment [4], and the food industry [5, 6]. Furthermore, charged polymers are found in nature and play essential roles in biological processes such as transferring genetic information from DNA to RNA to proteins [7–12]. The effect of charges on either the systems containing DNA, RNA, and proteins in a different cellular environment or the systems containing polyelectrolytes, counterions, small electrolytes, salt ions, and water leads to interesting and complex shapes, sizes, structures, functions, and dynamics [13–16]. The general aim of this thesis is to study the equilibrium and dynamic behaviors of charged polymers in different environments.

The equilibrium and dynamic properties of charged macromolecules are determined by the interplay of entropy and enthalpy from electrostatic interactions between charged polymers, counterions, salt ions, short-ranged Van der Waals interactions. Two length scales, the Bjerrum length (l_B) and Debye screening length (ξ) are useful to study and rationalize the properties of charged systems (and will be frequently used in this thesis). To explain these two parameters more clearly, suppose two ions, carrying charges $z_i e$ and $z_j e$, with z_i being the valency of the i^{th} ion and e being the electronic charge, are in a medium of dielectric constant ϵ and separated by a distance r_{ij}. Coulomb's law describes the electrostatic interaction energy (U_{ij}) between the two ions as,

$$\frac{U_{ij}}{k_B T} = z_i z_j \frac{e^2}{4\pi\epsilon_0 \epsilon k_B T r_{ij}} \tag{1.1}$$

where k_B is the Boltzmann constant, T is absolute temperature, and ϵ_0 is the permittivity of free space. The quantity $\frac{e^2}{4\pi\epsilon_0\epsilon k_B T}$ is defined as the Bjerrum length (l_B), which when compared with r_{ij} gives us an idea of the strength of electrostatic energy in comparison with thermal energy. If r_{ij} is smaller than l_B for oppositely charged monovalent ions, the electrostatic energy between the pair is stronger than that of thermal energy. l_B is also inversely proportional to the dielectric constant of the medium, which is a temperature-dependent quantity [17]. For aqueous solution at room temperature ($25^0 C$), $l_B \simeq 0.7$ nm and for oil-like medium, $l_B \simeq 25$ nm at room temperature.

Furthermore, the presence of electrolyte solutions in a system screens the effect of Coulomb interactions as proposed by P. Debye and E. Hückel in their classic theory [18] of screening by point charges. The screened electrostatic energy, also known as the Debye-Hückel potential, is given by,

$$\frac{U_{ij}}{k_B T} = z_i z_j l_B \frac{\exp\left(-\kappa r_{ij}\right)}{r_{ij}} \tag{1.2}$$

where,

$$\kappa^2 = \frac{e^2}{\epsilon\epsilon_0 k_B T} \sum_i z_i^2 n_{i0} \tag{1.3}$$

and n_{i0} is the number density of ion type i. κ^{-1} is defined as the Debye length (ξ_D), which gives us the range of the electrostatic interaction. ξ_D for monovalent salt in water at room temperature ($25^0 C$) is given by,

$$\xi_D \simeq \frac{0.3}{\sqrt{c_s}} nm \tag{1.4}$$

where c_s is salt concentration in unit of moles/liter.

In this thesis, we employ computer modeling, mostly coarse-grained approach, to simulate complex environments comprising of charged polymers, solvent, and charged ions. We focus on the following specific problems in this thesis.

Electrostatic potential across a nanopore

In Chapter two, we model and compare the electrostatic potential across the alpha-Hemolysin (αHL) nanopore using both atomistic and coarse-grained methods. We observe that the coarse-grained method gives a good approximation of the atomistic approach. As the all-atom molecular dynamics simulations are computationally expensive and time-consuming for the available time and facility, we use coarse-grained approaches in other projects.

Unfolding of RNA hairpins in response to applied external forces

In Chapter three, we study the unfolding kinetics of RNA hairpin architectures in response to an applied external field. The primary goal of this project is to understand how the diverse conformations of RNA (secondary and tertiary structures of RNA) influence the dynamics of folding and unfolding, which is crucial to gain insight into fundamental biological processes. The exploration of unfolding kinetics of different RNA hairpin architectures made up of equal number of nucleotides is done in this chapter.

Polyelectrolyte complexation

Chapter four discusses the structures of polyelectrolyte complexes formed by flexible and symmetric polymers of opposite charges and the dynamics of a polyelectrolyte chain inside a complex. The goal of this project is to understand how the polyelectrolyte complexes respond to different variables. For example, how does the transition

from a homogeneous polyelectrolyte solution to a system of multiple aggregates occurs with the degree of ionization of the polymers? How do chains organize inside a complex in a different environment? When do they form ladder-like structures, and when do they form scrambled egg-like structures? Do the chains show worm-like models, coil-like conformations, or Gaussian-like statistics? Do the chains inside the complex show diffusive or non-diffusive dynamics? We present the systematic investigation of the roles of the degree of ionization, temperature, salt concentration, and polymer concentration on the structures and dynamics of polyelectrolyte complexes in this chapter.

Conclusions

Finally, Chapter five provides the conclusion and some possible future directions.

CHAPTER 2

ELECTROSTATIC POTENTIAL ACROSS A NANOPORE

2.1 Introduction

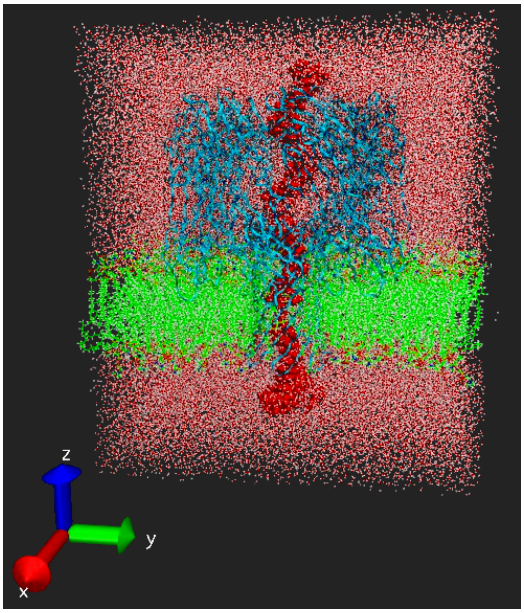
The translocation kinetics of a polyelectrolyte chain (charged molecule) through a nanopore is highly controlled by the electrostatic potential across the pore, along with other important factors such as chain structure, chain flexibility, pore-geometry, etc. So it is important to model the electrostatic potential across the pore correctly. In this project, we calculated the electrostatic potential across different charge-decorated alpha-hemolysin (αHL) nanopores by using an atomistic and coarse-grained model and show that the coarse-grained method gives a good approximation of the atomistic approach.

2.2 Methods and simulation details

Our model consists a single stranded DNA (ssDNA), α-hemolysin nanopore, lipid bilayer membrane, water, and ions as shown in Fig. 2.1.

We used all-atom MD simulations to study the biopolymer systems (proteins and nucleic acids) where the molecular structures are well known. The structure and conformations of ssDNA is generated using 3D-DART, which is a DNA structure modeling server [19]. The structures and conformations of α-hemolysin protein pore (entry 7AHL) is obtained from the available protein data bank "(http://www.rcsb.org)" [20]. To model lipid bilayer in α-hemolysin protein pores, membrane plugin for visual molecular dynamics (VMD) [21] is used. The membrane plugin for VMD uses

(a)



(b)

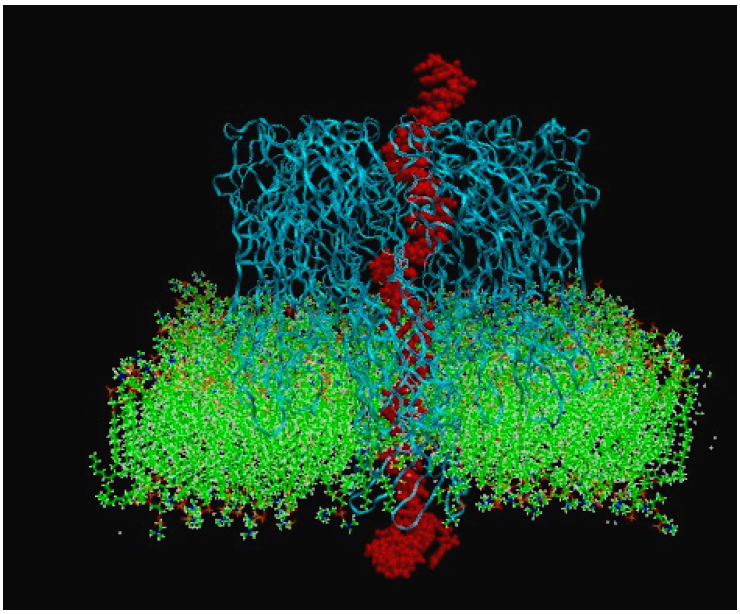


Figure 2.1. Snapshots of translocation of ssDNA through an α-hemolysin protein pore. Red beads represent ssDNA, cyan ribbons show α-hemolysin protein pore, green lines are lipid bilayer. Water and ions are removed from figure (b) for clarity.

pre-equilibrated patches of either POPC or POPE lipid bilayers, covering the desired size by tiling and trimming. The simulations are done in explicit water and salt ions using the Scalable Molecular Dynamics (NAMD) package [22].

The structure of double-stranded DNA made by 40 nucleotides was generated using 3D-DART. Two strands of DNA were separated and we were left with a strand of (poly$(dA)_{40}$. Thus obtained ssDNA was kept in a box of TIP3P water molecules removing overlapping water molecules with ssDNA. The system was ionized to get 1 molality (1 mole/kg) KCl solution. K^+ ions were added to ensure the charge neutrality since each nucleotide of DNA is negatively charged.

The system was energy-minimized. Minimization takes the system to the nearest local energy minimum so that the steric clashes and the possibility of high energy configuration of any system could be avoided. Thus modeled single-stranded DNA (poly$(dA)_{40}$) was relaxed in the presence of water molecules and 1 molality KCl salt solution. The system was heated to 300K at constant volume using a Langevin thermostat. The system was then equilibrated in an NPT ensemble at 295 K temperature and 1 atm pressure by applying Langevin forces to the atoms.

The structure of α-hemolysin protein (entry 7AHL) was obtained separately from the protein data bank. The lipid bilayer membrane was generated using membrane plugin of VMD. The protein and the membrane were solvated and equilibrated separately. The protein was then inserted into the membrane and overlapped molecules were removed from the system. The total system was solvated and ionized to achieve 1 molality NaCl solution. The system was again energy-minimized, heated, and equilibrated as in the case of a system of ssDNA at 295 K temperature in NPT ensemble.

The final system was prepared by combining the ssDNA and the pore membrane. The overlapped ions and water molecules were removed from the system to avoid clashes. Since we might have removed ions from the system, the system was re-neutralized. As before, the system was again energy-minimized, heated, and equilibrated.

The system was now ready for the translocation simulation. An external electric voltage bias equivalent to 120 mV was applied across the pore. For the simulation, open-pore ionic current and distribution of electrostatic potential were computed. Next, the system of ssDNA and α-hemolysin pore were simulated in an external electric field, which allowed us to observe the DNA translocation through the protein pore. The force calculated from the computed electrostatic potential distribution was applied to DNA to accelerate the process.

The equilibrium process and production simulations [23] were done in periodic boundary conditions using the particle mesh Ewald (PME) for the calculation of long-range electrostatic forces [24] in a grid spacing of $1\mathring{A}$. The simulations were performed using CHARMM27 forcefield [25].

2.3 Results

2.3.1 Comparison of electrostatic potential across the pore

We have calculated the average electrostatic potential across the αHL pore system by approximating every point charge by a spherical Gaussian as,

$$\rho_i(\mathbf{r}) = q_i \left(\frac{\beta}{\sqrt{\pi}}\right) \exp -\beta^2 |\mathbf{r} - \mathbf{r}_i|^2 \tag{2.1}$$

where q_i is the charge of the atom i and β gives the width of the Gaussian whose value is approximated to be 0.25 $\mathring{A}^{-1}$. We obtained the instantaneous electrostatic potential by solving Poisson equation as,

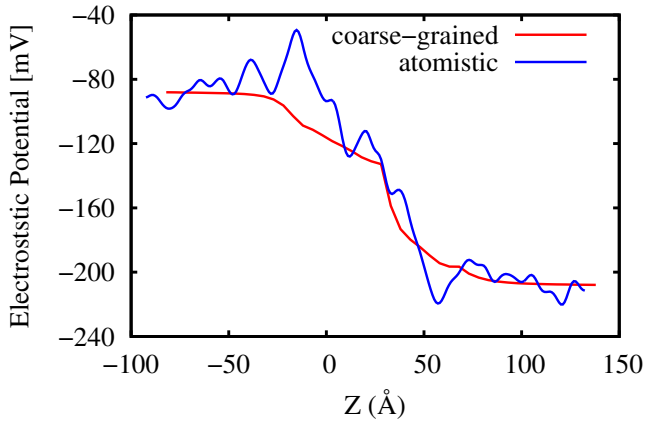
$$\nabla^2\phi(\mathbf{r}) = -4\pi \sum_i \rho_i(\mathbf{r}) \tag{2.2}$$

We averaged the instantaneous electrostatic potentials over the entire MD simulations to obtain the electrostatic potential across the pore. Above mentioned protocol of calculating electrostatic potential across the pore is adopted from already published paper by Aksimentiev et al. [26].

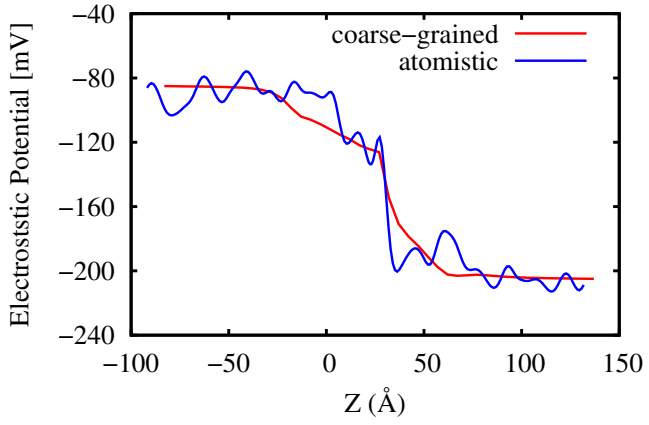
The comparison of electrostatic potential across the charge decorated nanopores calculated using the atomistic simulations and coarse-grained simulations are shown in Fig. 2.2. The coarse-grained simulation part is done by a postdoc from our group, Ining Jou [27]. In the Fig. 2.2, the y-axis is applied electrostatic potential and the x-axis is the z-position on the $\alpha-$hemolysin protein pore. We modeled three αHL nanopores, wild-type (WT), RL2 and RL2-M113R. Here, RL2 differs from WT due to mutations at five different amino acids, V124L, G130S, N139Q, I142L, and K8A, where, the original amino acid residue as represented by the first letter is replaced by the amino acid residue represented by the last letter on the location of polypeptide chain represented by the number. RL2-M113R has an extra arginine mutation at the position 113 in RL2 pore [28]. The positions of the charge decorations are shown in Fig. 2.3.

The qualitative behavior of the potential is in a good agreement in both the methods. The fluctuations in the curves from the atomistic calculations might be due to a lack of statistical samples, and due to electrostatic interactions between the charges in the pores and translocating ssDNA. We also observed that the fluctuations are

(a)



(b)



(c)

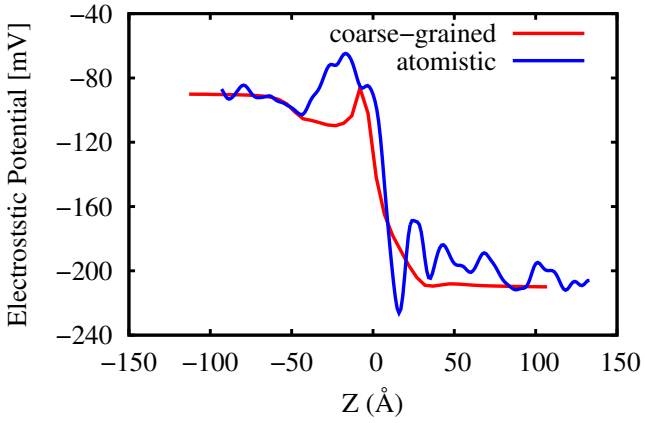


Figure 2.2. The comparison of electrostatic potential across the nanopore with atomistic simulations and coarse-grained simulations for (a) WT, (b) RL2, and (c) RL2-M113R charge decorated α-hemolysin nanopores.

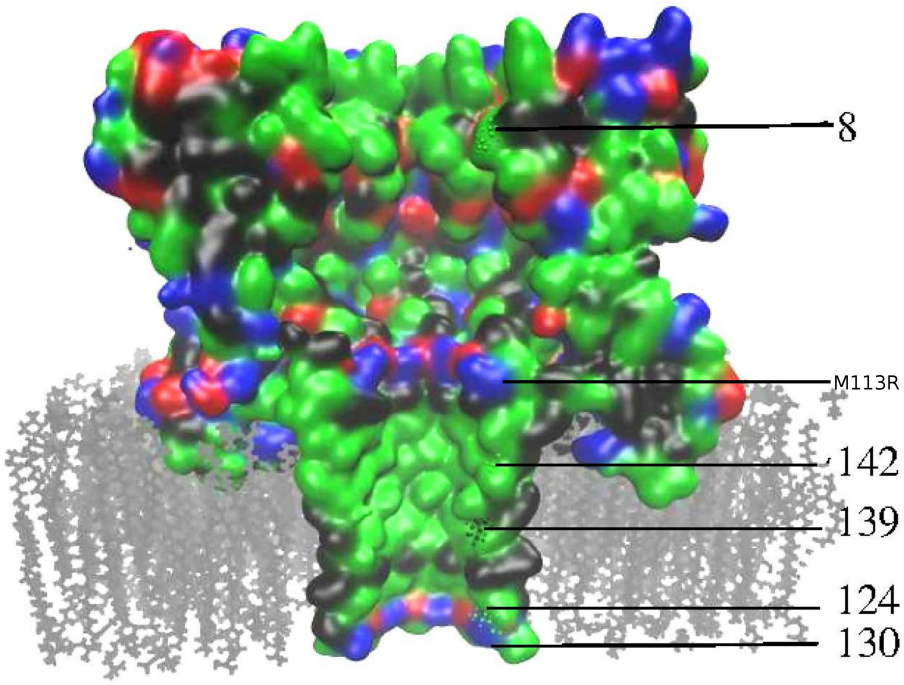


Figure 2.3. The α-hemolysin protein pore showing the location of mutation in RL2 and RL2-M113R from wild-type. Blue represent basic amino acids, Red represent acidic amino acids, green represent polar amino acids, and black represent location of neutral amino acids. Water and ions are not shown for clarity.

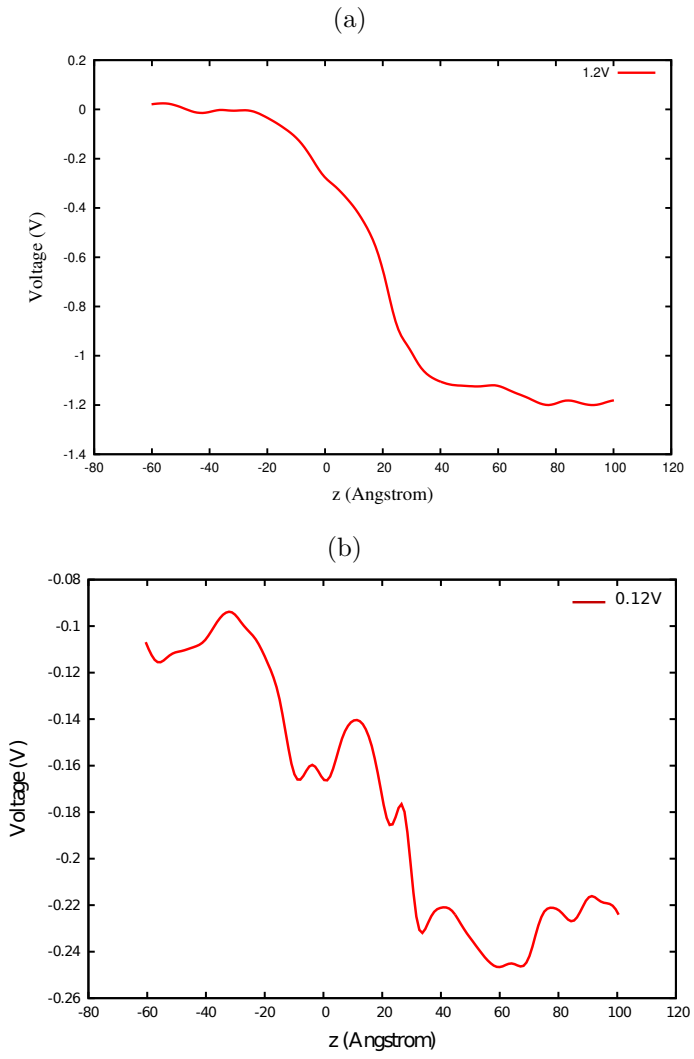


Figure 2.4. Electrostatic potential across the WT αHL nanopore when (a)1.2 V and (b) 0.12 V is applied.

high when an applied external voltage is low and vice versa as shown in Fig. 2.4.

We aimed to model and understand the translocation kinetics of ssDNA through charge decorated $\alpha-$hemolysin protein nanopore using all-atom molecular dynamics simulations. But the simulations were computationally expensive and time-consuming for the available time and facility as the translocation was very slow, and one single simulation took more than two months when we applied the laboratory used external

electric field ($\sim$ 120 mV). Therefore, we decided to stop the project and move-on to another project, where we use coarse-grained simulations.

CHAPTER 3

UNFOLDING OF RNA HAIRPINS IN RESPONSE TO APPLIED EXTERNAL FORCES

3.1 Introduction

Nucleic acids, which constitute some of the biopolymers present in living cells, are essential for the continuity of life because they make up the genetic information of living things [29, 30]. Nucleic acids are found in single stranded, secondary, and tertiary structures such as hairpins and pseudoknots, which play a crucial role in transforming genetic information from DNA to RNA to proteins.

The complex structure of RNA depends on the cellular environment during the transcription [7, 10], splicing [11], translation [8, 12] and protein synthesis [9]. The genetic information is transcripted from DNA to RNA through RNA polymerase. The transcripted information undergoes splicing which means some of the information is sliced out and removed from the sequences making new sequences. The newly formed sequences are then translated to synthesize one or more proteins. The above processes involves structural changes of secondary and tertiary structures of RNA including at the basic level folding and unfolding events and many also involve the transmission of nucleic acids through protein pores [7–12]. Therefore understanding how the diverse conformations of nucleic acids fold and unfold and how these complex structures move through nanopores is crucial to gain insight into mechanisms related to the transformation of genetic information from from DNA to RNA to proteins.

Taking inspiration from the biological applications, technologies involving synthetic and biological nanopores have been used in detection of genetic sequences.

This technology has been widely applied in DNA sequencing [31–35] and has been recently applied in RNA and protein sequencing as well [36–39]. The sequencing of DNA, RNA and proteins is essential in understanding human genomics, gene expression, health care, and the microbiome. Furthermore, such technologies can be applied to determine nucleic acid structures [37, 38, 40, 41]. With these motivations, the main goal of this project is to understand the role of secondary structures on RNA dynamics during translocation.

With a growing number of studies on DNA, RNA, and protein sequencing, it is very important to understand how the secondary and tertiary structures of RNA folds and unfolds in response to applied external forces. Similarly, it is crucial to study how RNA structure changes during translocation through nanopores. Systematic studies of a single molecule under applied mechanical forces have been done to understand the folding and unfolding mechanisms of complex nucleic acid structures [33, 42–46]. In typical experiments, a constant force is applied using optical tweezers to pull the two ends of DNA/RNA hairpins from which force extension curves are constructed. In these experiments, it is observed that at a critical force, RNA hairpins hop between folded and unfolded states [42, 46]. The critical forces were extracted from the force extension curves. Similarly, simulations have been done mimicking these experiments in order to find the folding and unfolding mechanisms of RNA hairpins in response to applied forces for a broad range of temperatures [43, 45]. The hoping between folded and unfolded states of polynucleotide hairpins was observed for hairpins with different stem lengths, loop lengths, and CG content [33]. In these studies, mechanical force is applied at the two ends of a polynucleotide hairpins, which might be different from the forces relevant to biological processes. However, in translocation, the force is distributed due to interactions between polynucleotides and the pore. Since the interaction of polynucleotides with different protein pores are observed, translocation might provide a better representation of real systems.

A number of studies have been done to understand the unfolding kinetics of polynucleotides during translocation through synthetic and biological nanopores. Translocation is driven by an applied voltage bias, which transmits a force onto the nucleic acid molecule (negative backbone charge), ultimately moving the molecule from one side of the pore to the other. The dependence of unfolding kinetics on pore diameter and threshold voltage have been studied for a wide range of voltages and lengths of hairpins [47–52]. These studies find that the unfolding process dominates the total translocation time.

Few experiments and little theoretical work has been done focusing on the the influence of the secondary structures of polynucleotides [53–56]. To obtain the signature of secondary structures of RNA, a new tool combining solid state nanopores and optical tweezers has been proposed to measure the net unfolding force for RNA structures [54]. However, no systematic studies have been done to understand the role of different architectures on the dynamics of translocation through nanopores.

In this project, we systematically study the role of hairpin architecture on the unfolding kinetics of RNA during translocation. To initiate studies on this, we employ coarse-grained simulations of a crude model of different RNA hairpins as they translocate through a geometric pore. We restrict our study to models containing 72 nucleotides of same types of base pairs in five different architectures. Our main focus is to find the influence of structural changes on the translocation dynamics. The results show that the threshold voltage to unfold different RNA architectures is higher when a longer hairpin segment enters the nanopore first. We also study the time evolution of probability of unfolding of all the models for all the applied voltages and extract the lag time as a function of applied voltages for all the architectures. The result show that the lag time is longer for lower applied voltages and for models with longer hairpin segment entering the nanopore first. However, the lag time shows no significant differences for the different models, when collapsed by the threshold

voltage. In simulations, we are able to access dynamical information during translocation which is not accessible through experiments, namely base by base unfolding dynamics. The results show that there is a distinct signature of the base unfolding time for bases near the unpaired bases. The signature is due to the lag between the unfolding of the basepair just before a group of unpaired bases and the basepair just after. This gives the foundation for the development of technologies that use translocation to predict secondary structures of polynucleotides.

This article is organized as follows. In Section 3.2 we describe the model and simulation details that we employ. In Section 3.3 the results of the simulations are presented, followed by conclusions and a discussion of future directions in Section 3.4.

3.2 Model and simulation details

3.2.1 Model

Our model consists of two parts, an RNA hairpin with a tail and two membrane walls containing a cylindrical pore inside it as shown in Fig. 3.1. The nucleic acid molecule was represented by a united atom model as described in our previous work [57]. Each nucleotide contains three spherical beads representing the phosphate, sugar, and base groups, respectively as shown in Fig. 3.1(c). The phosphate bead carries a charge of -e, where e is the fundamental unit charge. The beads representing the sugar and base do not carry charge. The membrane and the pore walls were represented by spherical beads that do not carry any charge. The length and diameter of the pore were chosen to mimic the β-barrel of the α-hemolysin protein pore. The stem length of β-barrel is $50\mathring{A}$ and the diameter is 14 $\mathring{A}$. For simplicity each bead was modeled with same diameter and mass.

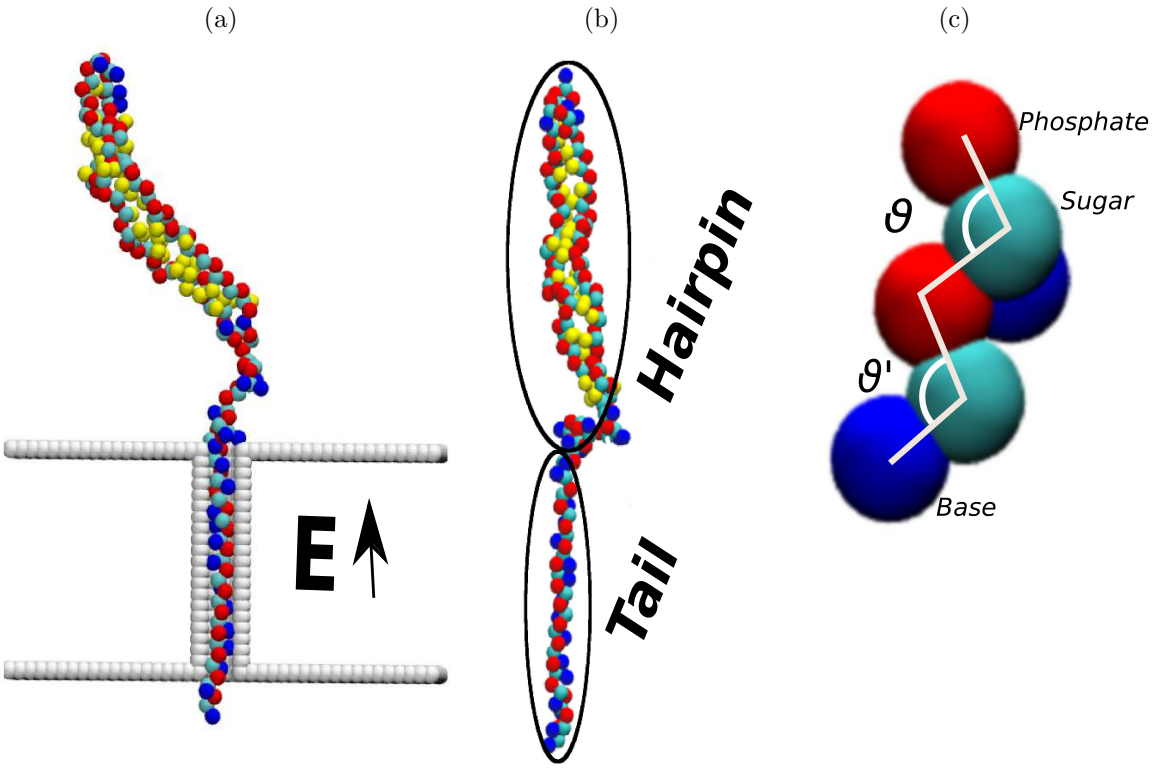


Figure 3.1. (a) The configuration before starting the simulation; the hairpin remains outside the membrane and the tail is inserted inside the pore. A voltage bias is applied across the membrane generating electric field **E**. (b) The tail and the hairpin section of one model is shown. (c) Representation of three bead unified model. θ is a representative backbone angle (phosphate-sugar-phosphate) and θ' is a representative side angle (phosphate-sugar-base). We take $\theta = \theta'$ in our simulations.

We consider five different architectures of the hairpin, each containing 72 nucleotides as shown in Fig. 3.2. The first model we consider consists of a simple hairpin with 22 base-pairs and a 4 unpaired base loop, also called a tetraloop as represented in Fig. 3.2 (a) (22HP). The second model consists of an interior loop connecting a stem of 11 base-pairs and a hairpin of 11 base-pairs, which is shown in Fig. 3.2 (b) (11IL11HP). The third model shown in Fig. 3.2(c) contains two hairpin domains containing 11 base-pairs in each hairpins and connected by 4 unpaired bases(11-11HP). The fourth and fifth models are similar to the 11-11HP model, however, both models contain two hairpins with 16 and 6 base-pairs. In the fourth model,

the tail is connected to the 16 base-paired hairpin loop (16-6HP) shown by Fig. 3.2 (d) and in the fifth model, it is connected to the 6 base-paired hairpin loop (6-16HP) as represented in Fig. 3.2(e). Each of the hairpin loops has an unpaired tetraloop.

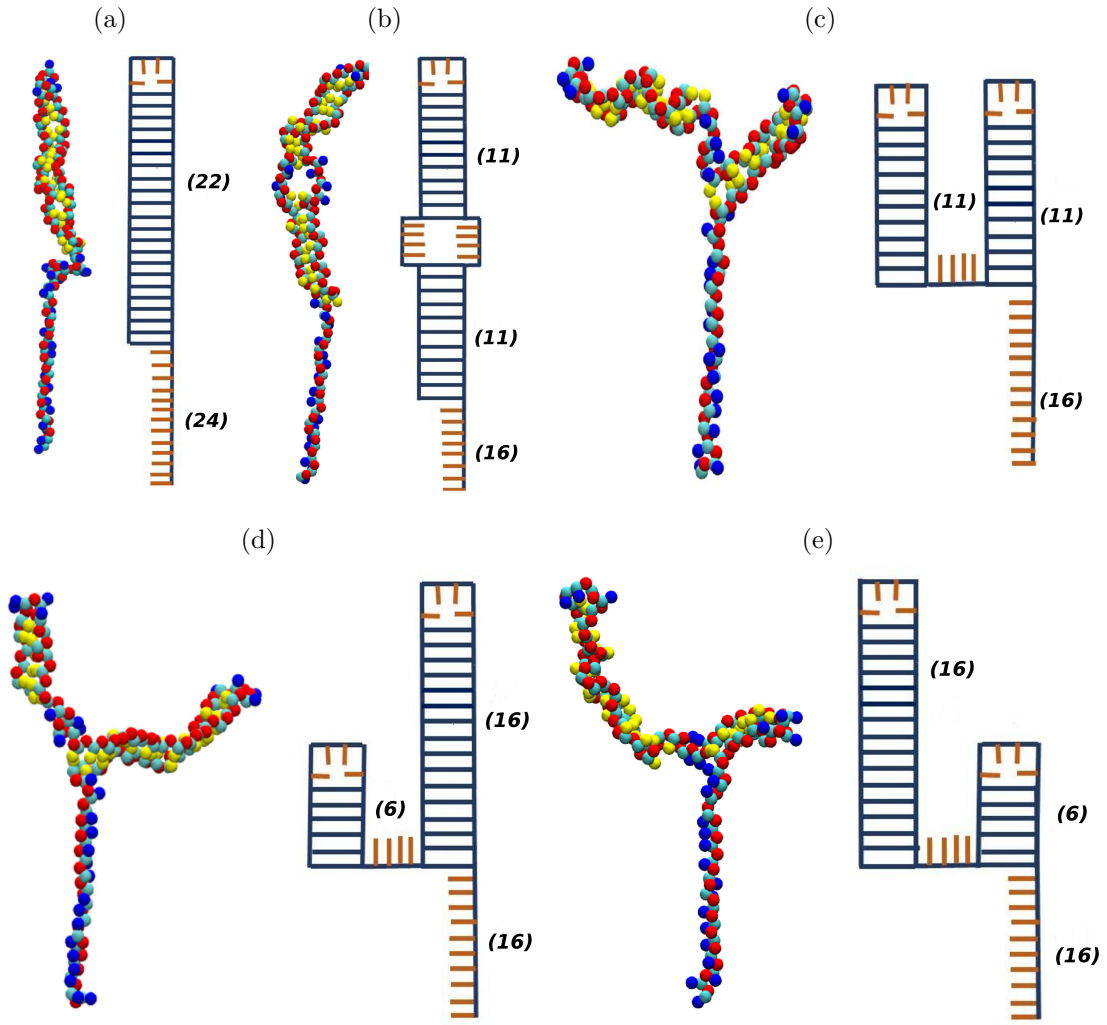


Figure 3.2. Different architectures of an RNA hairpins with tails. The left image in each frame shows the equilibrated model and the right image shows a 2D structure corresponding to the model. Red beads represent negatively charged phosphate groups, cyan beads represent sugar, yellow beads are paired bases and blue beads are unpaired bases. The models are (a) 22HP, (b) 11IL11HP, (c) 11-11HP, (d) 16-6HP and (e) 6-16HP.

3.2.2 Simulation detail

Langevin dynamics simulations were used to observe the unzipping trajectory of the hairpins under the influence of an applied electric field. Simulations were done using the Large-Scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) package [58]. The unzipping trajectory was computed by solving the Langevin equation of the i^{th} bead of the molecule:

$$m\frac{d^2\mathbf{r}_i}{dt^2} = -\xi\frac{d\mathbf{r}_i}{dt} - \nabla_i U(\mathbf{r}_i) + \mathbf{F}_i(t) + \mathbf{F}_{ext}, \tag{3.1}$$

where m is the mass of the bead, $\mathbf{r}_i$ is the position of the i^{th} bead, ξ is the friction coefficient, U is the total interaction potential acting on the i^{th} bead, $\mathbf{F}_{ext}$ is the force acting on the bead due to the applied external electric field $\mathbf{E}$, and $\mathbf{F}_i(t)$ is the random force acting on the i^{th} bead due to solvent molecules at time t and temperature T. This random force term satisfies the fluctuation dissipation theorem:

$$\langle \mathbf{F}_i(t) \cdot \mathbf{F}_j(t') \rangle = 6k_B T\xi\delta_{ij}\delta(t-t'), \tag{3.2}$$

where k_B is Boltzmann's constant, and $\delta(t)$ is the Dirac delta function.

Simulations were done in dimensionless Lennard-Jones (LJ) units, in agreement with LAMMPS. In defining units, $\sqrt{(m\sigma^2/\epsilon)}$ is taken as a unit time and $(4\pi\epsilon_0\sigma\epsilon)^{1/2}$ is the unit charge. Each beads has a unit mass, $m = 1$ which is equal to 96Da. The beads has diameter of 1 σ which is equal to 3.0 $\mathring{A}$. The phosphate bead has charge equal to $e = 1.602 \times 10^{-19}$ C and unit ϵ is taken approximately equal to 0.2 $kcal/mol$. For simplicity, the value of the friction coefficient in Eq (3.1) is choosen arbitrarily to be 1 LJ units. In the simulation, we used an integration time step of $dt = 0.0002\tau$ LJ units, where, the characteristic time ($\tau = \sqrt{(m\sigma^2/\epsilon)} = 1$) in our simulations.

There are bonded and non-bonded interaction potentials acting on each bead.

$$U = U_{bonded} + U_{non-bonded} \tag{3.3}$$

where $U_{bonded} = U_b + U_a$ and $U_{non-bonded} = U_{LJ} + U_C + U_{hb}$ are the different potentials acting in the system, which are described below. The bonded potential includes contributions from a spring-like bond (U_b) and an angle potential (U_a). In our system, the bond potential between connected beads (i and i+1) is represented by a harmonic potential and the angle potential between three successive beads (i, i+1, and i+2) is represented by a cosine-squared potential as:

$$\begin{aligned} U_b &= k_b(r - r_0)^2, \\ U_a &= k_a[cos(\theta) - cos(\theta_0)]^2 \end{aligned} \tag{3.4}$$

where, $k_b = 5344$ units (≈ 171 kcal/mol $\mathring{A}^2$) is the spring constant [57], $k_a = 250$ units (≈ 60 kcal/mol) is the angle constant [59], r is the distance between connected beads, $r_0 = 1$ is the equilibrium distance, θ is angle of backbone and side groups as shown in Fig 3.1 (c) and $\theta_0 = 105^0$ is the equilibrium angle [59]. The values of the parameters are choosen close to realistic values to promote RNA folding.

The non-bonded potential acting on each bead includes three contributions: excluded-volume interactions, a screened Coulomb potential and a Gauss potential to mimic hydrogen bonding.

To model excluded volume interactions, a truncated Lennard Jones (LJ) potential

$$\begin{aligned} U_{LJ} &= 4\epsilon[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^6] + \epsilon \quad for \quad r \leq r_c \\ &= 0 \qquad for \quad r > r_c \end{aligned} \tag{3.5}$$

is used where $\epsilon = 1$ (≈ 0.2 kcal/mol) is the depth of potential well which is taken as $k_BT/3$ for temperature $T = 300K$. Additionally, $\sigma = 3.0\mathring{A}$ is the effective bead diameter, and the potential is truncated at $r = r_c = 1.12\sigma$.

The screened Coulomb potential is modeled by using the truncated Debye-H*ü*ckel potential as

$$\begin{aligned} U_c &= (\frac{q_i q_j}{4\pi\varepsilon})\frac{\exp(-\kappa r)}{r} \quad for \quad r \leq r_{coul} \\ &= 0 \qquad for \quad r > r_{coul} \end{aligned} \tag{3.6}$$

where, q_i denotes the electric charge on bead i, $\kappa = 0.639$ units is the inverse Debye length for a monovalent salt concentration of 0.6 M, $\varepsilon = \varepsilon_0\varepsilon_r$ is the permittivity with ε_0 being permittivity in free space and ε_r being the dielectric constant, and $r_{coul} = 10/\kappa$ is the cut-off distance for tructation of the electrostatic potential. Although the effective dielectric constant inside the pore is unknown, our value of the dielectric constant is 80 which is the dielectric constant of water at room temperature. No beads other than the phosphates carry charge. Therefore, the Coulomb potential is only computed between phosphates.

In order to model hydrogen bonding between the bases in the nucleic acid, a Gauss potential,

$$\begin{aligned} U_{hb} &= -A\exp(-Br^2) \quad for \quad r \leq r_g \\ &= 0 \qquad for \quad r > r_g \end{aligned} \tag{3.7}$$

is applied between paired bases where A and B are parameters with units of energy and distance^{-2}, respectively. In RNA, the strength of hydrodgen bonding ranges

from 2 to 12 k_BT. For simplicity, we have used A = $10k_BT$ for the amplitude of the potential. In order to form the desired folded structures we have employed one-to-one interaction, U_{hb}, applied pairwise between the two specific bases that form a base pair in the desired folded state. Although this model is highly idealized, we are trying to understand the qualitative physics of RNA hairpin unfolding using a crude model.

3.2.3 Simulation process

To start the simulation, an equilibrated hairpin as shown in Fig. (3.2) was required. Using a flat initialized configuration, the tail was inserted into the pore and the last bead of the chain was kept fixed. The configuration was allowed to equilibrate for 2000τ LJ time units by solving Eq. (3.1), with no applied field. A Verlet algorithm in LAMMPS [58] was used to solve the equations of motion. Initially, a random velocity drawn from Boltzmann distribution at a given temperature was also assigned to all moving beads.

The configuration at the end of this equilibration process was taken as the initial conformation for a simulation of unfolding dynamics under the applied electric field. A uniform electric field E was applied across the pore. The pore and the membrane were fixed and only the nucleic acid was allowed to move throughout the simulation. Eq. (3.1) was integrated in time using the Verlet algorithm until the hairpin was either completely unfolded and translocated to the other side or until 5×10^7 timesteps ($10,000\tau$ LJ time units) was elapsed. The positions of the beads were recorded after every 1000 timesteps. The diameter of the pore was such that only a single nucleotide was allowed to enter the pore. Therefore, the translocation of the nucleotide was possible only after unfolding has happened. This allowed us to study the unfolding dynamics of the RNA hairpin. For a given applied voltage bias, 500 to 1500 runs were

performed for each architecture. The number of runs were varied such that we have about 500 successful runs for the analysis of each architecture and applied voltage. For the analysis, only successful unfolding events were taken into consideration.

3.3 Results

3.3.1 Fraction of successful unfolding events

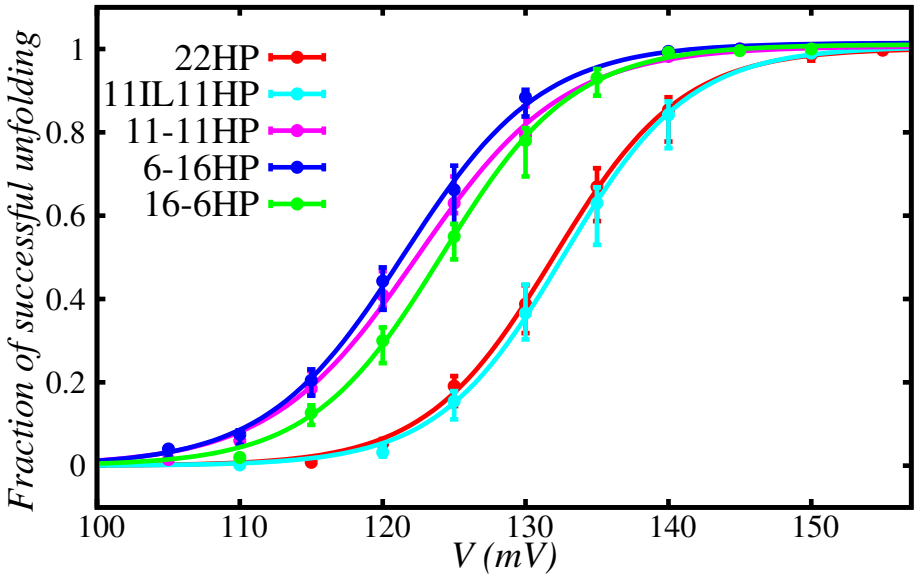


Figure 3.3. Fraction of successful unfolding events for all models as a function of applied voltage. The curves from top to bottom represent models 6-16HP, 11-11HP, 16-6HP, 22HP and 11IL11HP respectively.

For all the models and all applied voltages, we first checked if the hairpin completely unfolded and translocated at the end of the simulation, $10,000\tau$ LJ time units. The ratio of the number of unfolded events to the total number of simulations was calculated and fitted with a Sigmoid function:

$$Frac = \frac{1}{1 + \exp(-\frac{V - V_0}{W})}, \tag{3.8}$$

where, V_0 is the voltage at which the inflection point of the Sigmoid curve is observed, and W is a measure of the sensitivity of the unfolding fraction to changes in voltage. Fig. (3.3) shows that the nature of unfolding is similar in all the models where the primary difference is the threshold voltage, defined by V_0. This qualitative behavior is consistent with single molecule experiment on mechanical pulling of RNA hairpin using optical tweezers [42].

Table 3.1 shows the values of fit-constants we obtained. The models 22HP and 11IL11HP show similar threshold voltages. This shows that the internal loop has very little effect on the unfolding mechanism. Among the two domain hairpins (11-11HP, 6-16HP, and 16-6HP), it shows that the threshold voltage depends on the length of the hairpin attached to the tail. When a longer hairpin is attached to the tail, the threshold voltage for unfolding is higher. While the primary differences between the different models is the threshold voltages, the W or sensitivity does slightly vary between the models as well. Table 3.1 shows that the two domain cases seem to have slightly larger values of W. It is clear that the different structures play a role and further study is needed to understand this.

Model	V_0 (mV)	$Width$ (mV)
22HP	131.97	4.49
11IL11HP	132.64	4.41
11-11HP	122.45	4.90
6-16HP	121.49	4.83
16-6HP	124.15	4.58

Table 3.1. The threshold voltage (V_0) and the W of the unfolding fraction for different models, when fit to Equation (3.8).

Further, we studied how the probability of unfolding (P_{unfold}) evolves with time. We have defined P_{unfold} as the ratio of unfolded events to the total number of events at given time for each applied voltage bias and architectures. P_{unfold} at long time is equal to the fraction of unfolding at different voltages as shown in Fig. 3.3. Fig. 3.4

shows how the unfolding of RNA hairpins occurs as a function of time for different architectures and at different applied voltages. Fig. 3.4 (a) shows the time evolution of P_{unfold} for the 22HP architecture; it shows that for a given voltage, the unfolding process starts only after a certain lag time. The lag time is shorter for higher voltages. The nature of unfolding of other architectures as shown in Fig. 3.4 (b-e) is similar to Fig. 3.4 (a).

We extracted the lag time from the intersection of the baseline and fitted a line of the evolution of P_{unfold} as shown in Fig. 3.5(a). The example is for the case of 155mV applied voltage and 22HP model. Using the same method for all the applied voltages and all the structures, we get the lag time as a function of applied voltages as shown in Fig. 3.5(b). The red 'plus' data points represent the lag time for the 22HP architecture. The figure shows that on increasing the applied voltage, the lag time decreases, as it is easier to unfold the hairpin when more energy is used to the system. Blue 'cross' data points represent the values for the 11IL11HP model. The data points in the curve show that the presence of an internal loop in the model seems to have a minimal effect on the lag time. Other data points show similar behavior for all other architectures but with different time scales. The figure shows that the lag time increases with the increasing length of the hairpin closer to the tail for a constant voltage. Therefore, lag time might be a way to capture the secondary structures of RNA. Moreover, we subtract the threshold voltages of different models as shown in Table 3.1 from the applied voltages ($V - V_0$) and replot the lag time as a function of $V - V_0$ for all the architectures as shown in Fig. 3.6. The figure shows no significant difference in the lag time between models as a function of $V - V_0$ for all the models. This result implies that the threshold voltage of a model carries the information about the structures of RNA.

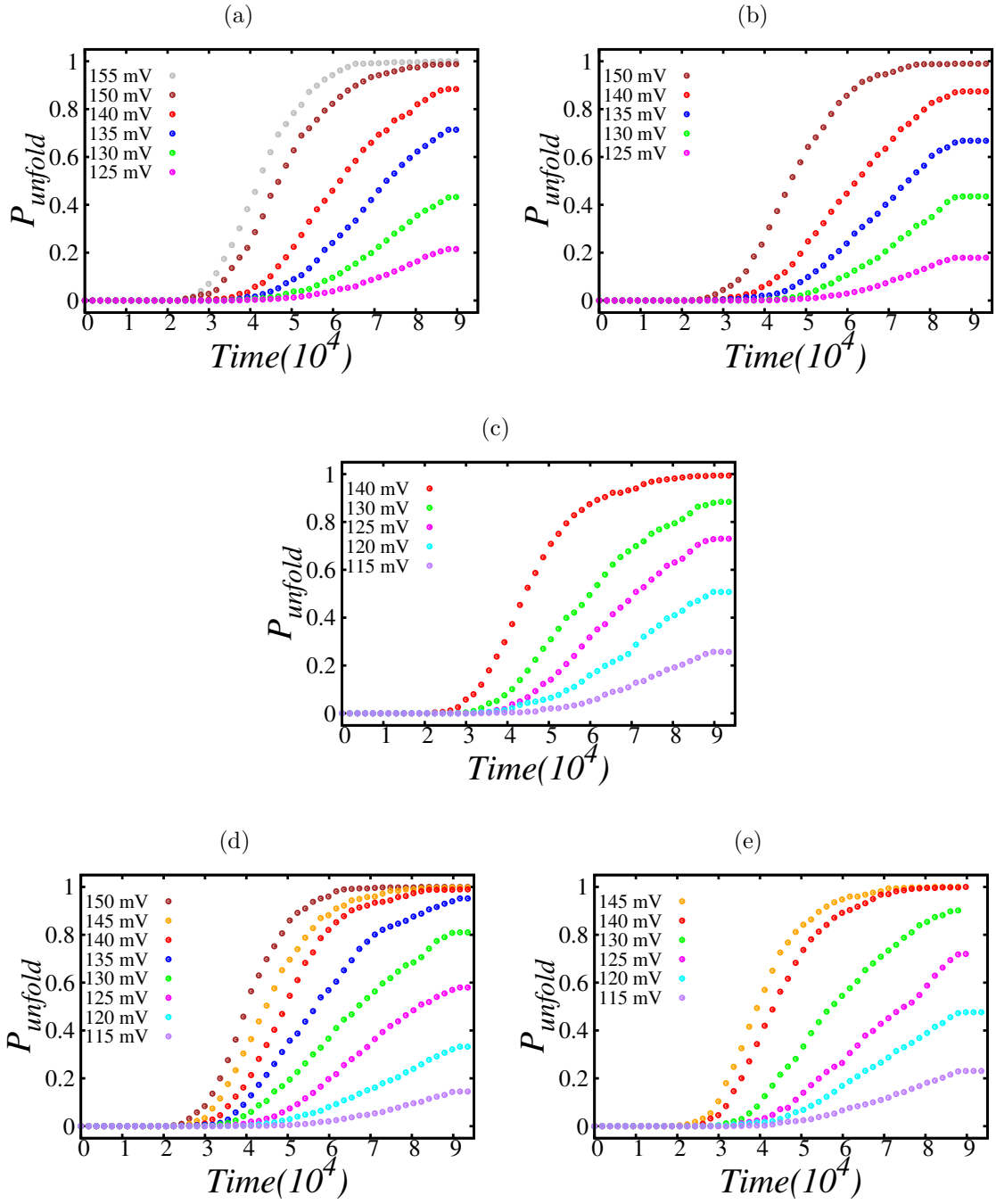


Figure 3.4. Probability of unfolding as a function of time for models (a) 22HP, (b) 11IL11HP, (c) 11-11HP, (d) 16-6HP and (e) 6-16HP. In all the cases the left-most curve has the highest voltage and the right-most curve has the lowest voltage.

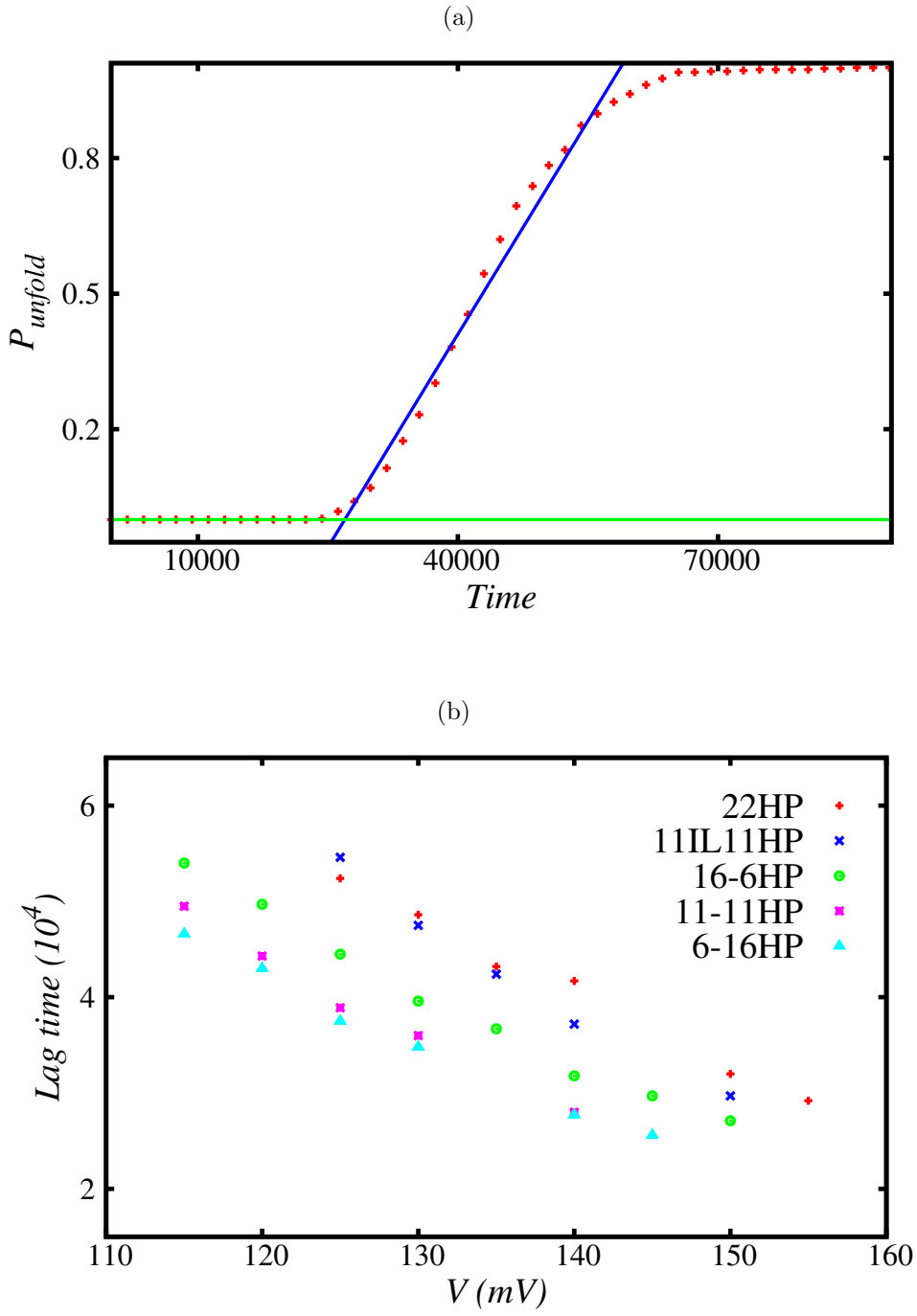


Figure 3.5. (a) Example of lag time extraction. This is the case of 155 mV for 22HP model. (b) Lag time as a function of applied voltage for different architectures. Lag time is longer for smaller applied voltages and for the structures with longer hairpin attached to the tail.

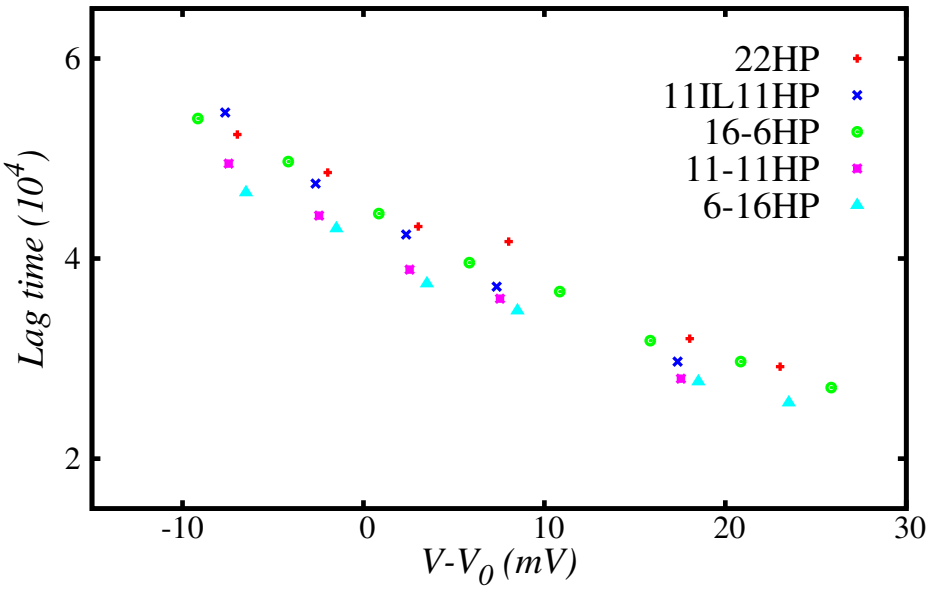


Figure 3.6. Lag time as a function of difference between applied voltage and threshold voltages for all the architectures. Lag time shows no significant differences for different models.

3.3.2 Unfolding dynamics of individual base-pairs

Here we show how unfolding behavior is observed for each base-pairs in the various models. We define unfolding time of each base-pair as the time it takes a base-pair to permanently unfold; after which it never returns to a folded state. Fig. 3.7 shows the unfolding time of each base-pair at different applied voltages and for different architectures. Base-pair 1 is the one closest to the tail and base-pair 22 is the one that unfolds the last.

From Fig. 3.7 it is observed that for a hairpin of any length, the unfolding time increases with base-pair number, unfolding occurs sequentially along with the RNA. Additionally, the unzipping is very fast towards the end of the hairpin. We also observed that for two domain architectures (11-11HP, 16-6HP, and 6-16HP), the unfolding of the first few base-pairs of the second hairpin is very fast. After the complete unfolding of the first hairpin loop, the tetraloop can translocate with less

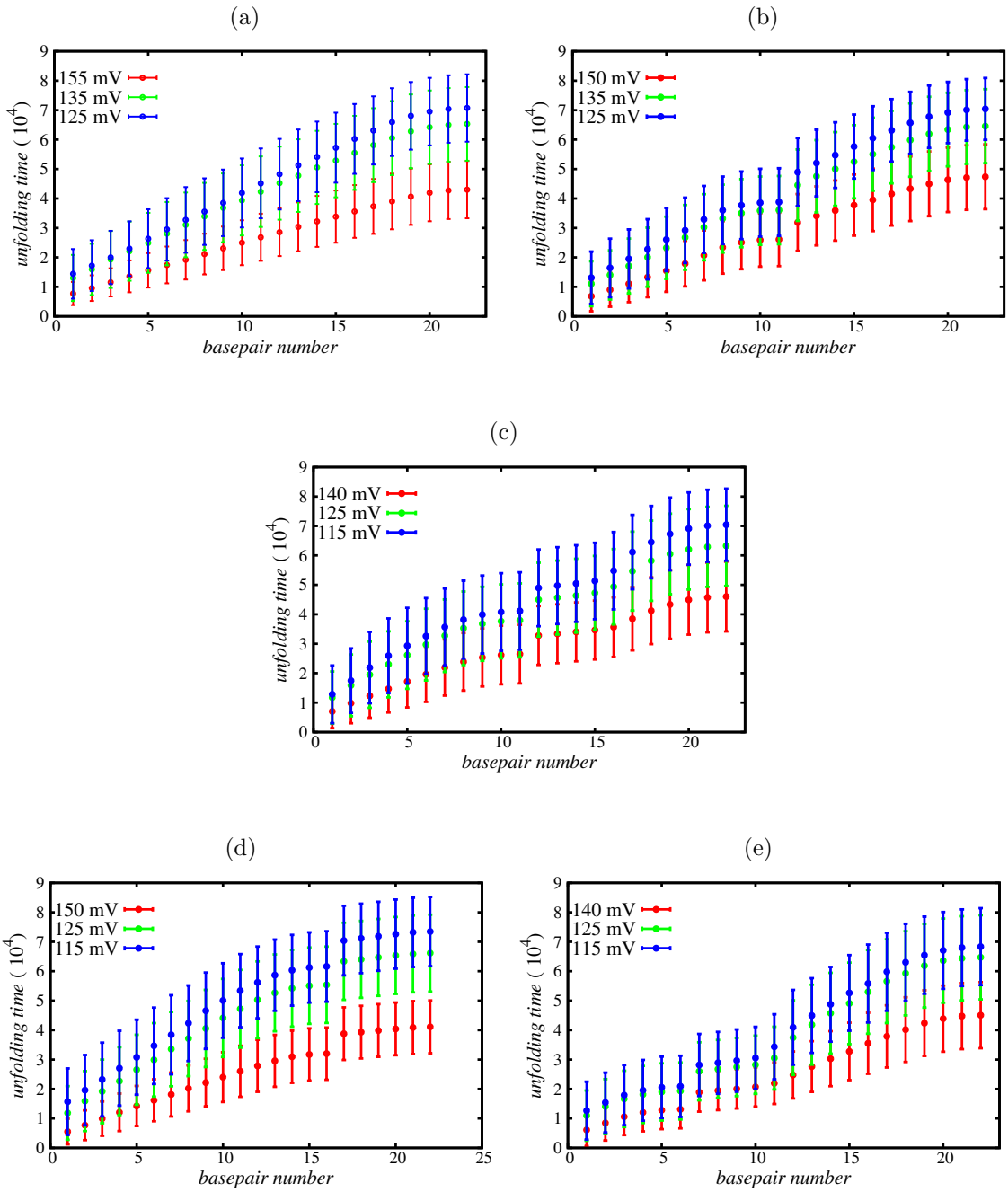


Figure 3.7. Unfolding time per base-pair for models (a) 22HP, (b) 11IL11HP, (c) 11-11HP, (d) 16-6HP, and (e) 6-16HP. The points are averages from at least 500 simulations and the error bars represent the standard deviations. In all cases the voltages increases moving from the top curve to the bottom one.

force because there is no force (hydrogen bonding in base-pairs) acting against the applied voltage bias. This might have assisted the unfolding of the first few base-pairs of the second hairpin. However, further studies are needed to verify this.

Fig. 3.8 further clarifies this effect where we plot the time difference of unfolding of consecutive base-pairs versus base-pair number for different architectures. In Fig. 3.8(a) (22HP architecture), the first data point shows the time taken to unfold the first base-pair, the second data point shows the time between the unfolding of the first and the second base-pairs, and so on. The results show that, the time taken to unfold the first base-pair is higher than that for rest of the base-pairs. This is because of the lag time taken by the hairpin to start unfolding. Fig. 3.8 also shows that, the unzipping of the last few base-pairs is faster compared to other base-pairs. The 11IL11HP architecture shows similar behavior to that of the 22HP architecture, except it shows the unfolding of both the stem and the hairpin domains, as indicated by the two repeated regions as shown in Fig. 3.8 (b). Fig. 3.8 (c) and (e) are for the 11-11HP and 6-16HP architectures, respectively. Both of them consists of two hairpin domains. The variation of time difference with respect to base-pair numbers of the first hairpin in both cases shows similar behavior to that of the 22HP architecture. The unfolding of base-pairs of the second hairpins however, show qualitatively different behaviors. Once the unfolding of the first hairpin domain is completed, the unfolding of the second hairpin domain begins after a certain lag time. This lag time is due to the existance of the tetraloop in the hairpin. After the unfolding of first base-pair in the second hairpin domain, the unfolding of next few base-pairs is faster and then it slows down for a few base-pairs. Finally, unfolding gets faster for the remaining last few base-pairs. The faster unfolding of the first few base-pairs of the second hairpin domain is due to the assisted force from tetraloop. Fig. 3.8 (d) shows results for the 16-6HP architecture. This architecture also consists of two hairpin domains. The variation of time difference with respect to base-pair number of the first hairpin

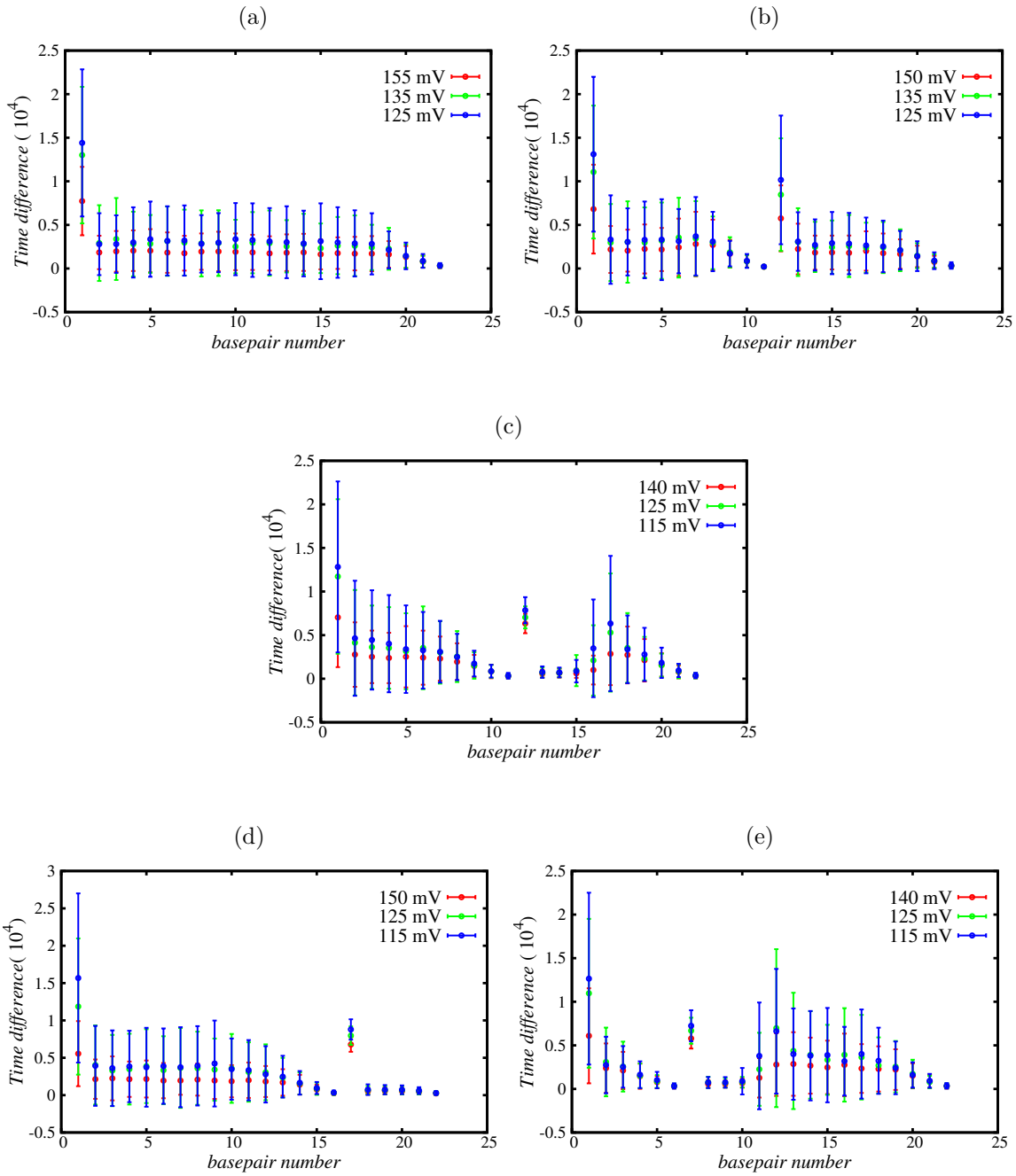


Figure 3.8. Unfolding time difference per base-pair for models (a) 22HP, (b) 11IL11HP, (c) 11-11HP, (d) 16-6HP, and (e) 6-16HP. The points are averages from at least 500 simulations and the error bars represent the standard deviations. Results for all voltages overlap within the error bars.

shows similar behavior to the rest of the architectures, as explained above. Once the unfolding of the first hairpin is completed, all base-pairs of the second hairpin unfold fast. This is because the hairpin has only six base-pairs.

3.4 Conclusions

In this study, we considered simple models of various RNA hairpins, and employed Langevin dynamics simulations to study their unfolding kinetics. We have used the united atom model to model nucleotides and a simple cylindrical pore to model the nanopore. We have considered five different architectures of the hairpin, each containing 72 nucleotides and 22 base-pairs. The first model is 22HP representing a simple hairpin with 22 base-pairs. The second model is 11IL11HP, made up of a stem of 11 base-pairs connected to a hairpin of 11 base-pairs by an interior loop. The other three models are 11-11HP, 16-6HP, and 6-16HP. They consists of two hairpin domains connected by four unpaired bases. The numbers in the name of the model are the numbers of base-pairs in each hairpin domains. Each of the hairpin domain has an unpaired tetraloop.

Our results showed that the threshold voltage of unfolding depends on the length of the hairpin attached to the tail. The longer the hairpin is attached to the tail, the higher the threshold voltage. In addition, the lag time to unfold is longer for lower applied voltages and the models with longer hairpins attached to the tail. However, the lag time collapses when the threshold voltage adjusts the voltage. Moreover, the base by base unfolding dynamics shows a distinct signature of base unfolding time for the bases before and after the unpaired bases in all the RNA hairpin models. The above-discussed structure dependence of translocation kinetics behavior of different RNA hairpins considered in this study, although obtained from simple models, demonstrates the potential use of translocation to predict the secondary structures

of polynucleotides.

In the future, more complex models of secondary and tertiary structures of polynucleotides and different geometries of pores can be used to understand the role of RNA architectures on the dynamics of translocation through nanopores and to predict the secondary and tertiary structures of polynucleotides.

CHAPTER 4

POLYELECTROLYTE COMPLEXATION

4.1 Introduction

Polyelectrolyte complexes (PECs) have attracted significant scientific interest in recent years because of their potential applications in gene therapy [1], drug delivery [2, 3], wastewater treatment [4], and food industry [5]. Numerous theoretical [60–65], and experimental [66–75] studies have focused on understanding the phase behavior of the systems which readily phase separate upon mixing of oppositely charged polyelectrolytes in aqueous solutions, as summarized in recent reviews [14, 13, 15, 16]. These studies have shown that the driving force of complexation is mainly entropic, which comes from the release of counterions from the polyelectrolytes into the solvent [76, 71]. In general, the phase behavior of such multi-component systems containing polyelectrolytes of opposite charges, salt, and water depends on various variables such as temperature, and salt concentration, together with molecular weight, charge density, and composition asymmetry of polyelectrolytes. It has been established in the theories and experiments that the polymer chains have to be ionized above a certain degree and also the ions concentration has to be below a certain threshold for stable complexes to form. Polyelectrolytes with higher charge density, or larger molecular weight, or both, have a larger tendency of formation of complexes and form broader two-phase regions. The critical ion concentration below which the system phase separates increases with an increase in charge density [60, 61, 68, 73].

The first experimental observation of polyelectrolyte complexes or complex coacervates by Bungenberg de Jong and Kruyt in 1929 [66] opened the field of polyelectrolyte complexes. The first theory of the phase behavior of complex coacervates is the Vorn-Overbeek (VO) model [60],which is a crude model based on Flory-Huggins theory to account for the entropy of mixing of chains and Debye-Hückel theory of simple electrolytes to account for the electrostatic contribution of charged monomers. According to the theory, the high molecular weight and high charge densities of polyelectrolytes enhance phase separation due to increasing electrostatic strength. The addition of salt suppresses phase separation by screening the electrostatic interaction [60]. The main drawback of this theory is the decomposition of polyelectrolytes into their charged monomers upon complexation. Despite its approximations, the theory is in good agreement with some experiments with no or some adjustable parameters [61, 66–68]. The reason for its ability to explain some of the experimental observations is the cancellation of the effect of chain connectivity and excluded volume [65, 77]. After VO theroy, numerous theoretical and simulation studies have been done and have enriched understandings of coacervation: including the effect of charge connectivity, excluded volume interactions, the finite size of ions, chain asymmetry, charge sequence on the polyelectrolyte, role of temperature-dependent dielectric constant and solvent-polymer interaction parameter (χ) [65, 77–84]. The main approaches to describe the phenomenon in these studies were mainly based on the random phase approximation (RPA), density functional theory, liquid-state theory, and field theoretic approximations.

The structures of polyelectrolyte complexes are highly dependent on many factors such as charge density [74, 85], charge sequence [81, 86], chain length [87], temperature [74], polymer concentration [85], pH [73, 88], salt concentration [71–73, 88, 89], chain topology [87, 90], stoichiometry [71, 85, 88], chain geometry [64], and chain stiffness

[91, 92]. Marciel et al. performed small-angle X-ray scattering (SAXS) experiments and Fourier-transform infrared spectroscopy on polyelectrolyte complexes made with polylysine and polyglutamic acid and reported that only the solid polyelectrolyte complexes contain stiff-ladder like structures due to hydrogen bonding interaction in the complexes. On the other hand, the formation of the stiff-ladder like structures due to hydrogen-bonding is not observed in the liquid polyelectrolyte complexes [93].

In asymmetric polyelectrolyte complexes, as in protein-polyelectrolyte and micelle-polyelectrolyte complex-coacervates, the overall structure of the complex is dominated by the longer polyelectrolyte molecules and these complexes have in-homogeneous density fluctuations which are observed by the presence of high scattering intensity upturn at low q [94–98], as reviewed by Spruijt et al. in ref [89].

Schlenoff and coworkers have performed numbers of experiments; extrusion, doping level measurements, conductivity measurements, thermogravimetric analysis (TGA), and electron microscopy on polyelectrolyte complexes. They formed the complexes with different pairs of polyanions and polycations and observed that the association strength of the polyelectrolyte complexes determines whether the formed complexes are glassy (very strong), rubbery (intermediate) or soft (weak) [75]. They have also reported that there can be a clear transformation of glassy, rubbery and liquid forms of polyelectrolyte complexes based only on the amount of salt in the system [71]. The transformation of the complexes is found to be a reversible process on salt doping and the doping level is affected by the type of anion present in the salt [99]. The association strength of the complexation is found to be dependent on the types of ions, amount of water content and the mobility of ions within the complexes [75]. They have proposed that Donnan equilibrium can predict the association strength of polyelectrolyte pairs and the salt content inside the polyelectrolyte complex. How-

ever, the estimation is only proposed to be accurate for low levels of salt doping [100].

As evident from the above studies, the properties of polyelectrolyte complexes are controlled by myriad of variables. This work focuses on investigation of structures of the complexes formed and the chains themselves, dynamics of a chain inside a complex of different sizes, for different charged density in dilute and semi-dilute solutions of polyelectrolyte complexes. For example, how does the transition from a homogeneous polyelectrolyte solution to a system of multiple complexes takes place on slowly increasing the degree of ionization? How do chains organize themselves inside a complex? When do they form ladder-like structures and when do they form scrambled egg structures? Do the chains show worm-like model, coil-like conformation or Gaussian-chain statistics? Do the chains inside complex show diffusive or non-diffusive dynamics?

We perform Langevin dynamics simulations of flexible and symmetric polyelectrolytes of opposite charges with explicit counterions at different polymer concentrations to answer these questions. In this work, we use coarse-grained (CG) models to study the role of charge density (α), temperature (T^*), monomer density (ρ), and salt concentration (c_s) on structures of polyelectrolyte complexes formed. Above overlap concentration, for small α, no or few numbers of polymer chains form a cluster of polyelectrolyte complexes. For high α, all or most of the polymer chains form an aggregation of polyelectrolyte complexes. The average biggest cluster size as a function of α shows Sigmoid behavior. The critical charge density (α_c) as defined by the point of inflection of the Sigmoid curve decreases with a decrease in temperature. It is independent of the polymer concentration considered in our simulations. Also, we found that the complexes are de-complexed with an increase in c_s. The transition salt concentration decreases with a decrease in α. Moreover, the average radius of

gyration ($\langle R_g \rangle$) of a labeled chain is independent of the size of the cluster formed, α, ρ, T^*, and c_s and follow Gaussian-chain conformation. The labeled chain shows non-diffusive dynamical law for high α and is diffusive for low α.

4.2 Model and simulation details

We modeled a polyelectrolyte chain as a freely jointed chain of N spherical beads of diameter σ, with αN charged beads ($0 \leq \alpha \leq 1$), connected linearly by harmonic spring, where, α is the charge density. Each charged bead is assigned a charge of $+1e$ for polycations and $-1e$ for polyanions (e is an unit electronic charge). The position of charged beads are located randomly along the chain, at the center of excluded volume spheres. A counterion and a salt ion are modeled as a sphere with a point charge of $-1e$ (or $+1e$). Total n number of polyanion chains and $\alpha(n)N$ counterions together with an equal number of polycation chains and their counterions were placed in a cubic box of edge length L as shown in Fig. 4.1. Langevin dynamics simulations were done to study the complexation of polyelectrolyte chains. The simulations were done using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) package [58]. The coarse-graining scheme and the relevant parameters were adopted from published papers of our group for polyelectrolyte simulations [101, 76, 102, 103].

The Langevin equation for the j^{th} component of the i^{th} particle is given by

$$m_i \frac{d^2 r_{ij}}{dt^2} = -\xi_i \frac{dr_{ij}}{dt} - \nabla_j U_i + f_{ij}, \tag{4.1}$$

where m_i is the mass, r_{ij} is the j^{th} position, ξ_i is the friction coefficient of the i^{th} bead, U_i is the total interaction potential acting on the i^{th} bead, and f_{ij} is the j^{th} component of the random force acting on the i^{th} bead due to solvent molecules at time t and

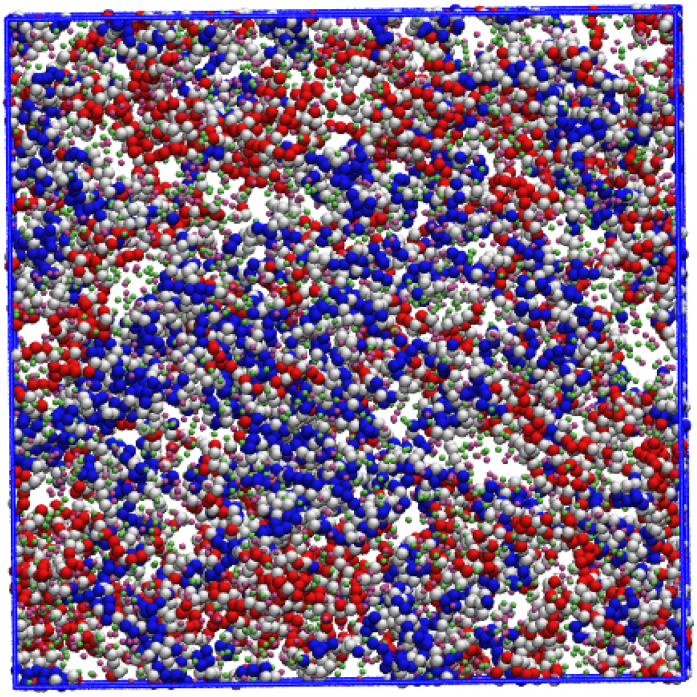


Figure 4.1. The configuration before starting the simulation. Equal number of polyanions and polycations together with counterions are placed in a simulation box of length L. Red beads are anions, blue beads are cations, silver beads are neutral monomers, pink beads are negative counterions and green beads are positive counterions.

absolute temperature T. This random force term satisfies the fluctuation-dissipation theorem:

$$\langle \mathbf{f}_i(t) \cdot \mathbf{f}_j(t') \rangle = 6k_B T \xi \delta_{ij} \delta(t - t'), \tag{4.2}$$

where k_B is Boltzmann's constant, and $\delta(t)$ is the Dirac delta function.

The total interaction potential U comprises of three independent interactions:

$$U = U_{LJ} + U_C + U_b \tag{4.3}$$

where U_{LJ} is the excluded volume interaction, U_C is the electrostatic interaction, and U_b is the bond potential. The charged beads representing monomers experience all three potentials, neutral beads representing monomers experience U_{LJ} and U_b, while the beads representing counterions and salt ions experience U_{LJ} and U_C.

The bond potential between connected beads (i and i+1) is represented by a harmonic potential,

$$U_b = K_b(l - l_0)^2 \tag{4.4}$$

where K_b is the spring constant, l is the bond length and l_0 is the equilibrium bond length.

To model excluded volume interactions, a truncated Lennard Jones (LJ) potential,

$$\begin{aligned} U_{LJ} &= 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] - 4\epsilon \left[\left(\frac{\sigma}{r_c}\right)^{12} - \left(\frac{\sigma}{r_c}\right)^{6} \right] \quad for \quad r \leq r_c \\ &= 0 \qquad for \quad r > r_c \end{aligned} \tag{4.5}$$

is used where ϵ is the interaction strength and the potential is truncated at $r = r_c = 1.12\sigma$.

The electrostatic interaction between the i^{th} bead carrying a charge of q_i and j^{th} bead carrying a charge q_j separated by distance r is modeled by using the Coulomb potential as

$$U_c = \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon r} \tag{4.6}$$

where $\varepsilon = \varepsilon_0\varepsilon_r$ is the permittivity with ε_0 being permittivity of free space and ε being the relative dielectric constant of the medium. The parameters in our simulations are presented in dimensionless Lennard-Jones (LJ) units, in agreement with LAMMPS.

4.2.1 Parameters

Coarse-grain model parameters of the polymers were chosen from the flexible polyelectrolyte chains such as sodium polystyrene sulfonate [76, 103], with $\sigma = l_0 = 0.25$ nm (l_0 represents the charge separation distance between the consecutive monomers), the magnitude of charge per charged-monomer or ion was e, and the excluded volume interaction strength (ϵ) was approximated to be 1 k_BT at room temperature. The electrolyte ions were modeled to be smaller than l_0. The value of σ were $1.0l_0$, $0.8l_0$, and $0.6l_0$ for beads representing polymer-polymer, polymer-ion and ion-ion pairs, respectively [76, 103]. ε_r was set at 80 for all the simulations, to represent the dielectric constant of water at room temperature, although some simulations were done at a temperature different than the room temperature.

All the parameters were converted from real units to dimensionless Lennard-Jones (LJ) units, fully consistent with LAMMPS [58]. The simulations were then performed in the LJ units, followed by analysis. The summary of the model parameters in real units and their corresponding LJ unit values are presented below

Parameter	Real Units	LJ Units
Distance	$l_0 = 0.25$ nm	1 LJ unit
Energy	$\epsilon = 1k_B$T	1 LJ unit
Temperature	$T^* = 298$ K	1 LJ unit
Charge	1 e	14.97 LJ units

Table 4.1. Few important parameters and their real and LJ unit values

m_i and ξ_i were assumed to be the same for all the coarse grained beads (m_i=1 LJ unit), and we set the damping parameter (m_i/ξ_i) at 10 LJ units in order to monitor dynamics and equilibrium properties of the system in a reasonable computational time frame (20 days). With smaller damping parameter values, the system did not reach equilibrium in a reasonable computational time frame and with larger damping parameter, the dynamic evolution of the system could not be monitored. Harmonic spring constant, $K_b = 5000$ LJ unit was used to keep the bond fluctuations within 10% of the equilibrium bond length.

We used particle-particle-particle-mesh (pppm) solver to compute the long-range Coulombic interactions [104]. Two additional parameters, desired relative error in forces for the iterative solver, and a cutoff distance were needed for pppm solver. Coulomb interaction between two charged beads is computed in the real space by using equation 4.6, if their separation distance is within the cutoff. If the separation distance is more than the cutoff distance, it is computed in the reciprocal space. We chose the relative error in forces to be 10^{-5}. The value of cutoff distance was determined by trial-error in some sample simulations, where, maximum computation efficiency was observed when cutoff distance was set at 20 LJ units, hence a cutoff distance = 20 LJ units was used.

The equation 4.1 is integrated in time using the velocity Verlet algorithm. The integration time step was 0.001τ LJ units, with a characteristic time $\tau = \sqrt{\frac{m\sigma^2}{\epsilon}} = 1$

LJ units.

4.2.2 Initial configuration and simulation details

We model the initial state of n polycations and n polyanions having N number of beads in each chain, with αN being the number of charged beads in a chain at random. $2\alpha nN$ number of counterions are randomly generated inside the simulation box of L×L×L LJ unit^3 (without overlap). To equilibrate the initial configuration, the system energy is minimized by adjusting bead positions, using LAMMPS's build-in minimize command. We follow this energy minimization with a dynamic simulations for $2 \times 10^3\tau$ LJ time units. At this stage, charge on the beads are not assigned yet to disperse the polymer chains and counterions homogeneously in the system. After the equilibration, we assigned charges on the polymers and counterions. $+e$ charges were assigned to αnN free beads (counterions) and $-e$ charges were assigned on remaining αnN free beads. A polymer chain was selected at random, and αN beads of this chain were assigned a charge of $+e$. This process was repeated for all n polycation chains. Similar protocol was followed to assign charge of $-e$ to n polyanion chains.

For systems with additional monovalent salt, the total number of salt electrolytes were determined from the salt molarity (input parameter) and the simulation box volume (L^3 in nm^3). The same protocol was followed for equilibration and charge assignment, only difference is more number of electrolytes are present in the system with salt compared to the salt free systems. After charge assignment, the simulation was conducted at a temperature of 3 k_BT and ramped to the desired temperature (around 1 k_BT) in $2 \times 10^3\tau$ LJ time units. The system was then equilibrated for $2 \times 10^3\tau$ at this temperature, followed by additional $2 \times 10^3\tau$ for equilibrium data analysis. Such ramping was was done for systems with high α, when the system could get stuck in a metastable state near the initial state due to high electrostatic

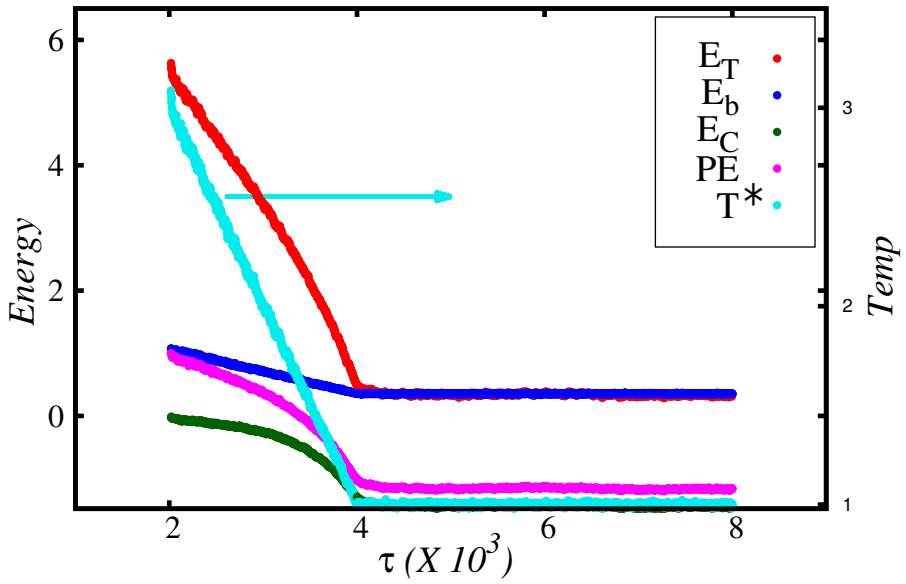


Figure 4.2. Left y-axis is the energy per particle and the right y-axis is temperature. x-axis is time steps. This is the energy profile for $\rho = 0.04$, $\alpha = 0.4$, and T^* =1.0. E_T is total energy, E_b is bond energy, E_C is Coulomb energy, and PE is potential energy per particle.

strength, Γ ($\Gamma = e^2/(4\pi\varepsilon_0\varepsilon_r l_0 k_B T = 2.8)$). A typical energy profile is shown in Fig. 4.2. The Fig. shows that the electrostatic energy is low at low temperature as the complexation between opposite chains is more favorable at low temperature.

We have performed two sets of simulations to understand the role of different variables. All the units are in LJ units unless otherwise mentioned.

- To study the structure and dynamics of a labeled chain inside a complex, we modeled systems with high charge density (α=1.0). This high charge density is chosen to facilitate the complexation. The parameters are chain length (N) and total number of chains ($2n$. In our study, N = 30, 45, 60, 75, and 90, and $2n$ = 2, 4, 8, 12, 16, 24, 48, 72, 96, and 120.

  For reference, we have performed simulations with single isolated polyelectrolyte chain having N = 30, 45, 60, 75, 90. We chose simulation box size to be L = 50 LJ units and L = 160 LJ units depending on the chain length such that the

length of a box is more than 10 times the average radius of gyration (R_g) of a chain. The monomer density (ρ) of the system with $N = 60$, $2n = 48$, and $L = 50$ is 0.023.

- To study the role of charge density, temperature, and salt concentration on the structure of a structure of a complex, our parameters are
  - $L = 62$ LJ units, $2n = 160$, and $N = 60$ (equivalent to a monomer density (ρ) of 0.04), charge density (α) = 0.3, 0.35, 0.4, 0.45, 0.5, 0.6, and 0.7, temperature (T^*) = 0.95 LJ units, 1.0 LJ units, and 1.2 LJ units, and salt concentration (c_s) = 0.25, 0.5 M, 1 M, 1.5 M, 2.0 M, 2.5 M and 3 M.
  - $L = 62$ LJ units, $2n = 400$, and $N = 60$ (equivalent to ρ of 0.1), α are 0.2, 0.3, 0.4, 0.5, 0.6, and $T^* = 1.0$ LJ units.

4.2.2.1 Is our system at equilibrium?

To test if the system is in equilibrium, we checked if the potential energy and the total energy of the system were constant for at least 2000 τ time steps as shown in Fig. 4.2. We have calculated and compared energies and average radius of gyration $\langle R_g \rangle$ of a labeled chain of different independent trajectories. Table 4.2 shows the values of energies and $\langle R_g \rangle$ of a labeled chain for $2n = 48$, $N = 60$, $L = 50$, $T^* = 1.0$, and $\alpha = 1$ for independent trajectories. We observed that they are within one standard deviation.

Trajectory	Potential energy	Total energy	$\langle R_g \rangle$
1	-0.567 ± 0.011	0.937 ± 0.016	5.082 ± 0.748
2	-0.576 ± 0.010	0.925 ± 0.017	4.981 ± 0.462
3	-0.568 ± 0.010	0.938 ± 0.014	4.871 ± 0.543
4	-0.575 ± 0.010	0.928 ± 0.014	5.001 ± 0.455

Table 4.2. Comparison of energies and $\langle R_g \rangle$ of $2n = 48$, $N = 60$, $L = 50$, $T^* = 1.0$, and $\alpha = 1$ for different trajectories.

Further, we ran another 1000τ time units and calculated energies and $\langle R_g \rangle$ for the same system. We observed that the values are within one standard deviation as shown in Table 4.3.

Trajectory	Potential energy	Total energy	$\langle R_g \rangle$
1	-0.566 ± 0.012	0.937 ± 0.016	4.824 ± 0.521
2	-0.572 ± 0.011	0.929 ± 0.019	4.890 ± 0.512
3	-0.571 ± 0.013	0.939 ± 0.018	5.075± 0.550
4	-0.571 ± 0.010	0.929 ± 0.014	4.861 ± 0.740

Table 4.3. Comparison of energy and $\langle R_g \rangle$ of $2n = 48$, $N = 60$, $L = 50$, $T^* = 1.0$, and $\alpha = 1$ for different trajectories.

For the case of $\alpha = 1.0$, and $T^* = 1.0$, randomly for many samples, we have repeated the equilibration process. The protocol is heating the system by ramping from $k_BT/\epsilon = 1$ to $k_BT/\epsilon = 3$ for 1000τ time units, then cooling it by ramping from $k_BT/\epsilon = 3$ to $k_BT/\epsilon = 1$ for another 1000τ time units, and running the simulations at $k_BT/\epsilon = 1$ for 1000τ time units as shown in Fig. 4.3. This is one way to check the consistency of our simulations.

We have followed this protocol for the trajectory 1 mentioned in Tables 4.2 and 4.3. The calculated values of energies and $\langle R_g \rangle$ are listed in Table 4.4.

Trajectory	Potential energy	Total energy	$\langle R_g \rangle$
1	-0.576 ± 0.013	0.919 ± 0.024	4.926 ± 0.565

Table 4.4. The potential energy, total energy and $\langle R_g \rangle$ of $2n = 48$, $N = 60$, $L = 50$ ($\rho = 0.023$), $T^* = 1.0$, and $\alpha = 1$ of one trajectory.

We believe that our systems are in equilibrium and are not kinetically trapped by running the number of time steps mentioned above because,

- Potential energy and total energy are constant for at least the last 2000 τ time steps.

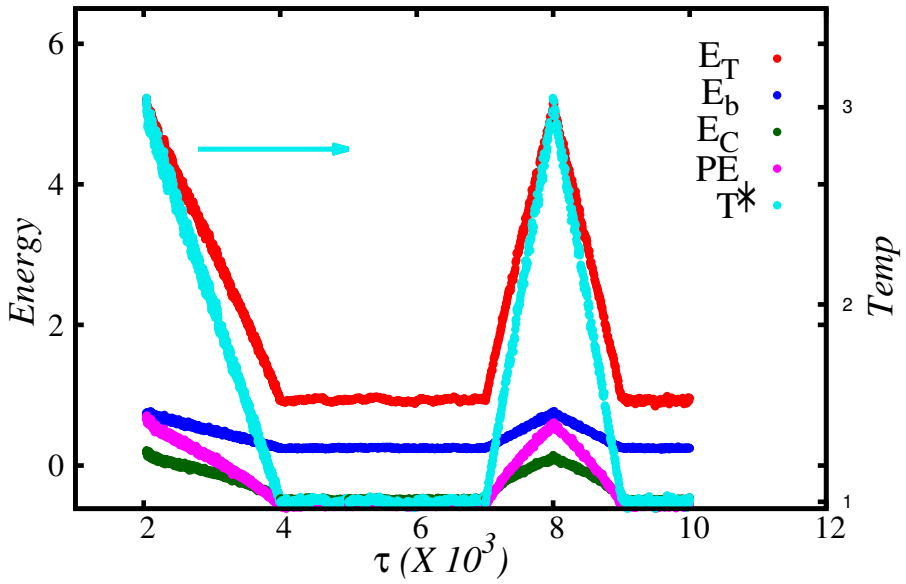


Figure 4.3. Left y-axis is the energy per particle and the right y-axis is temperature. x-axis is time steps. This is the energy profile for $\alpha = 1.0$, $T^* = 1.0$, $2n = 48$, $N = 60$, and $L = 50$. E_T is total energy, E_b is bond energy, and E_C is Coulomb energy per particle.

- $\langle R_g \rangle$ of a labeled chain for four independent runs are within a standard deviation.
- $\langle R_g \rangle$ of a labeled chain is within standard deviation when more 1000 τ time steps simulation is done.
- The potential energy, total energy, and $\langle R_g \rangle$ of a labeled chain for one trajectory are within standard deviation when re-ramping is performed.

4.3 Results

4.3.1 Overlap concentration

We first calculated the radius of gyration (R_g) of a single isolated chain having variable charge density (α). α is the ratio of charged beads to the total number of

beads of a chain. The systems contain explicit counterions. The overlap concentration is defined as,

$$c^* = \frac{nN}{V_{overlap}} = \frac{nN}{n\frac{4}{3}\pi R_g^3} \tag{4.7}$$

where, N is degree of polymerization.

α	R_g	c^*
0.2	6.87 ± 1.13	0.044
0.3	7.67 ± 1.13	0.032
0.4	7.88 ± 1.13	0.029
0.5	8.25 ± 0.88	0.026
1.0	8.52 ± 1.14	0.023

Table 4.5. Radius of gyration (R_g) of a single isolated chain (mean ± sd) and the overlap concentration (c^*) for systems with different charge density.

Table 4.5 shows the R_g of a single isolated chain and the overlap concentration (c^*) for different α. As the α is increased, the chain swells leading to a larger R_g and lower overlap concentration.

4.3.2 Structure and dynamics of a single chain within a complex

We define the monomer density (ρ) as

$$\rho = \frac{nN}{V_{box}} = \frac{nN}{L^3} \tag{4.8}$$

Table 4.6 shows the total number of polyelectrolyte chains and corresponding monomer density of our simulation systems where, $\alpha = 1.0$ and $T^* = 1.0$.

$2n$	2	8	12	24	48	72	96	120
ρ	0.001	0.004	0.006	0.012	0.023	0.035	0.046	0.058

Table 4.6. Total number of chains and corresponding monomer density of our system.

4.3.2.1 Snapshots

Figs. 4.4 and 4.5 show the snapshots complexation of polyanions and polycations of $\alpha = 1.0$ at $T^* = 1.0$ after the system is in equilibrium. The Figs. show that oppositely charged polyelectrolytes form strong complexes with the release of counterions.

4.3.2.2 Size of a labeled chain in a complex

We calculated the average radius of gyration ($\langle R_g \rangle$) of a labeled chain in all the complexes as shown in Fig. 4.6, defined by [105]

$$R_g^2 = \frac{1}{N} \Sigma_{i=0}^{N} < R_i^2 > \tag{4.9}$$

We examined the structural response in polymer chain conformations with an increase in the size of the complex. Previous work from our group showed that an isolated chain shows rod-like conformation in a salt-free medium; when a pair of oppositely charged polymers are placed in contact with each other, they collapse into a globule after releasing all the counterions [76]. We took a step further and determined how the ($\langle R_g \rangle$) of a lableled chain responds to gradual increase in polymer concentrations with explicit counterion. Fig. 4.6 shows that the $\langle R_g \rangle$ of a labeled chain varies non-monotonically with the number of polyelectrolyte chains in the complex. As expected, the $\langle R_g \rangle$ of a single chain is maximum and it shrinks when two chains (one polycation and one polyanion) come together forming a complex. On increasing the concentration of polymer in a complex, the value of $\langle R_g \rangle$ increases and saturates for higher polymer concentration. This high value of $\langle R_g \rangle$ of a single polyelectrolyte chain is due to the electrostatic repulsion among the charged beads at the backbone [106–109]. When the system contains two polyelectrolyte chains of opposite charges, the chains collapse due to electrostatic attraction between oppositely charged backbones and all the counterions are released from the complex. This release of counterions

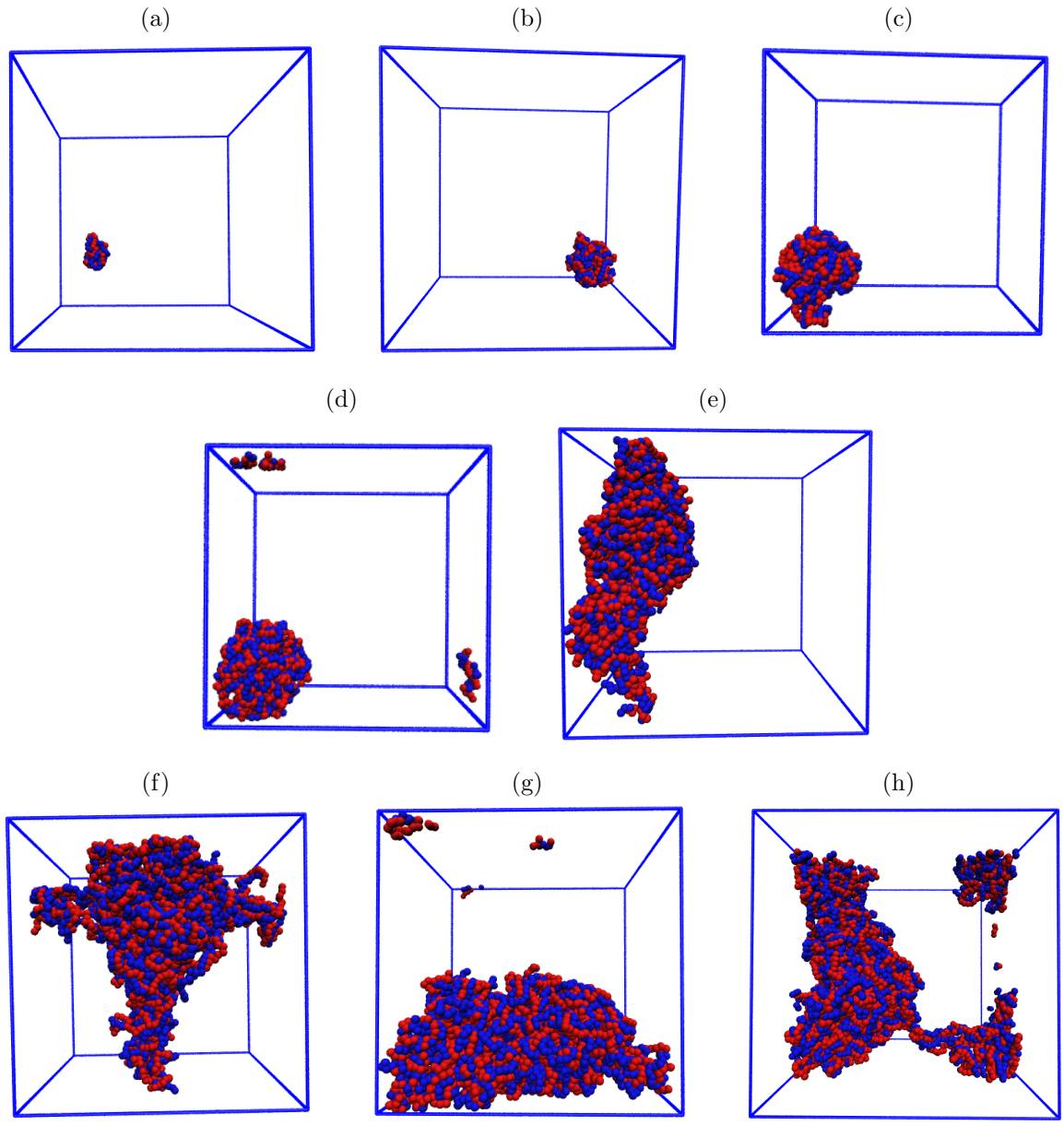


Figure 4.4. Simulation snapshots of complexation formed by complexation of $2n$ number of chains made up of equal numbers of polycations and polyanions with $\alpha = 1.0$ and at $T^* = 1.0$. Each chain has $N = 60$ beads. Blue beads are polycations and red beads are polyanions. Counterions are not shown for clarity. (a) $2n = 2$, (b) $2n = 8$, (c) $2n = 12$, (d) $2n = 24$, (e) $2n = 48$, (f) $2n = 72$, (g) $2n = 96$, and (h) $2n = 120$.

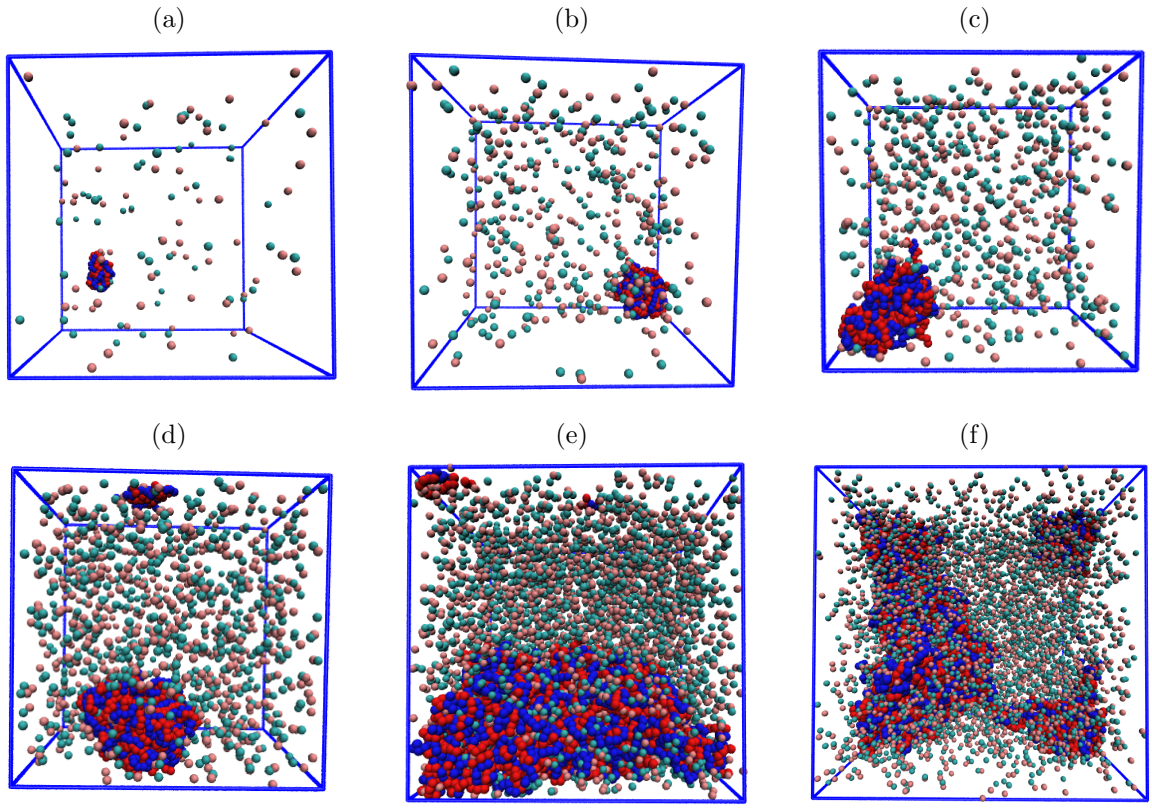


Figure 4.5. Simulation snapshots of complexation formed by complexation of $2n$ number of chains made up of equal numbers of polycations and polyanions with $\alpha = 1.0$ and at $T^* = 1.0$. Each chain has $N = 60$ beads. Blue beads are polycations, red beads are polyanions, pink beads are negative counterions and cyan beads are positive counterions. (a) $2n = 2$, (b) $2n = 8$, (c) $2n = 12$, (d) $2n = 24$, (e) $2n = 96$, and (f) $2n = 120$.

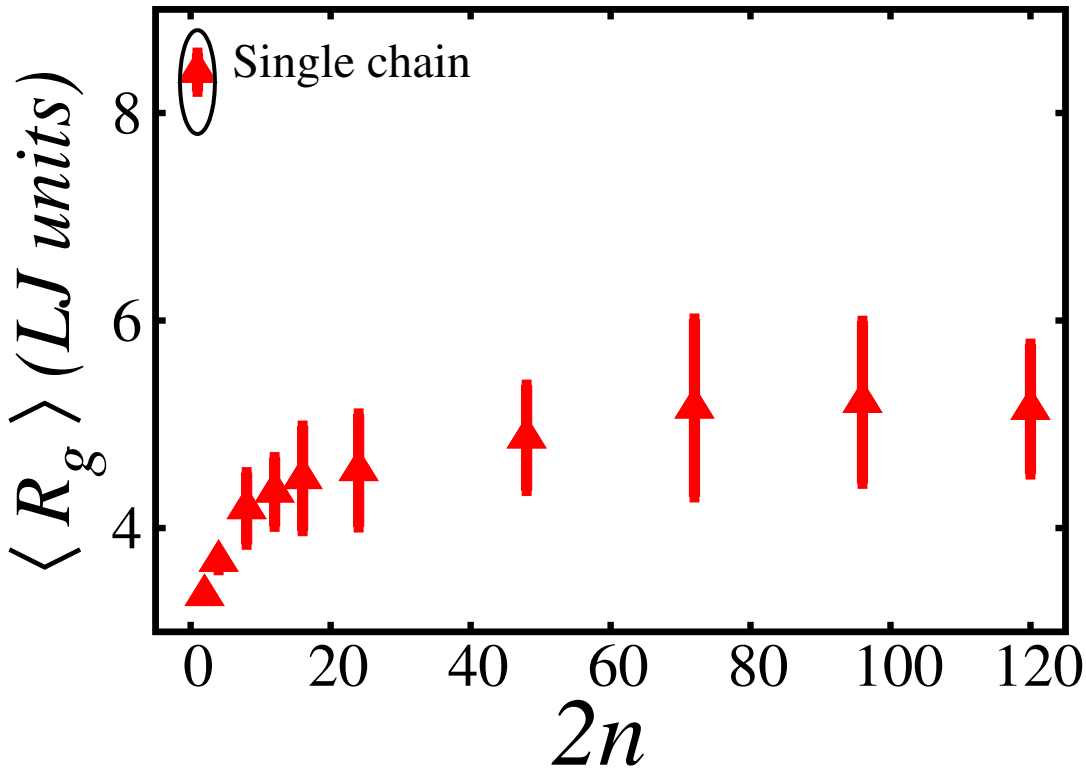


Figure 4.6. $\langle R_g \rangle$ of a labeled chain in the complexes. The first data point is for single polyelectrolyte chain. Here, $N = 60$ for all the cases.

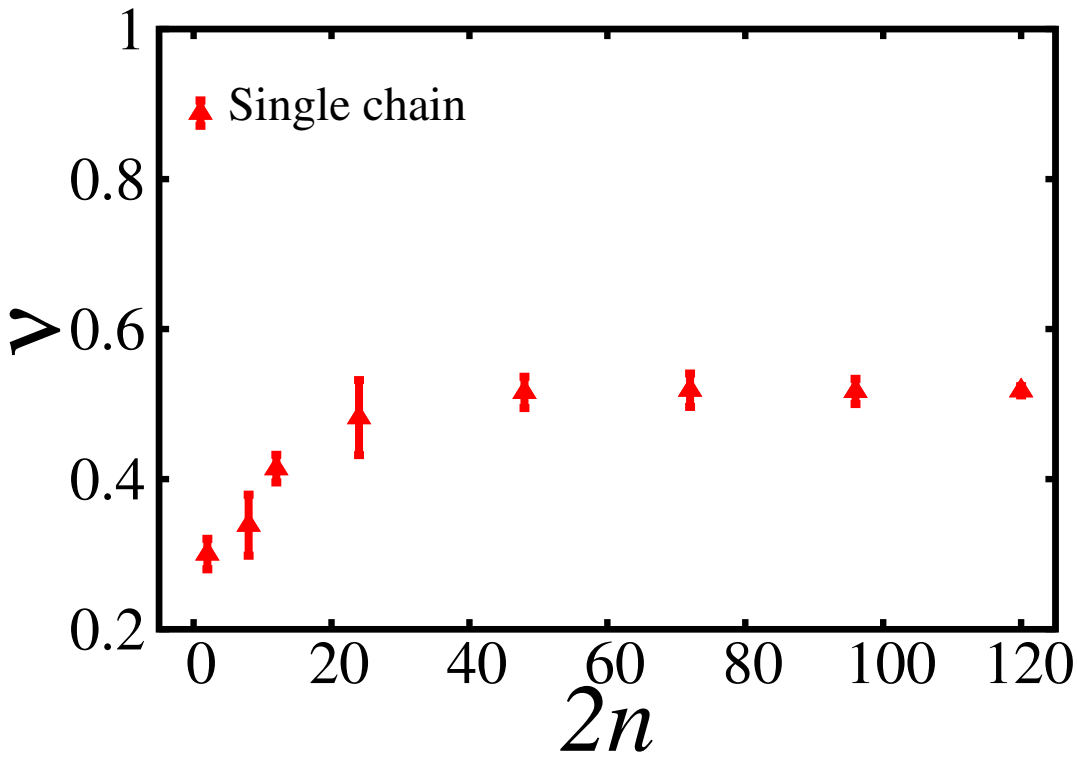


Figure 4.7. Size exponent of a labelled chain in the complexes. Size exponent is non-monotonic with the number of chains in the system.

on complexation of oppositely charged polyelectrolytes is well known from previous studies [76, 103, 110]. On increasing the concentration of polyanion and polycation chains in the system, the size of the complexes increases and so does the $\langle R_g \rangle$ of a labeled chain in the system. On further increasing the polymer concentration in the system, the $\langle R_g \rangle$ of a labeled chain in the complexes saturates. This particular case of polyelectrolyte complexes seems to indicate that, although the interaction between the beads are long range, the interaction is screened by the presence of other chains of opposite charges in its environment resulting in smaller $\langle R_g \rangle$ than that of a single isolated chain.

To further understand this saturation effect, we investigated the size-scaling exponents (ν) of the labeled chain which are shown in Fig. 4.7. The ν for a single

chain is found to be 0.88, representing a rod-like chain [111]. The ν of two chains of opposite charges is very small as they collapse once they are close to each other and all the counterions are released from the system. In this case, the size exponent is 0.3, representing a globule. We found that the ν of a polyelectrolyte chain changes from globular-like conformation to Gaussian-like conformation with increase in the number of polyelectrolytes forming the complex as shown in Fig. 4.7. ν of a chain inside the complex shows globular-like conformation when the complex is made with few chains. It is because the electrostatic attraction between the chains is very high leading to the very compact structure forming globular-like conformation. The chain conformation opens-up on increasing the size of the complex, as the attractive electrostatic interaction is screened by the presence of other chains of opposite charges. Once the complex of reasonable size is formed, a chain inside the complex shows Gaussian-like conformation. Similar behavior is observed from the form factor of a chain inside the complex as a function of scattering wave vector as shown in Fig. 4.8. Different colored lines show the numbers of chains in a complex. The form factor of a single polymer chain is given by [89]:

$$P(q) = \frac{1}{N_b^2} \sum_{n=1}^{N_b} \sum_{m=1}^{N_b} < \exp(iq(\mathbf{r_n} - \mathbf{r_m})) > \tag{4.10}$$

where, q is the scattering wave vector, b is the Kuhn length, N_b is the number of Kuhn monomers. At the limit of $\frac{1}{R_g} < q < \frac{1}{b}$, the form factor scales as, $P(q) \propto (qb)^{-1/\nu}$, where, ν is the size-scaling exponent.

This result agrees with neutron scattering experiments by Spruijt et al. [89]. They used neutron, X-ray and light scattering experiments to study the structure of deuterated polymers inside the complex coacervates made up of flexible polyelectrolytes, poly(acrylic acid) and poly(N,N-dimethylaminoethyl methacrylate). They found that both polyelectrolytes of opposite charges show Gaussian chain conforma-

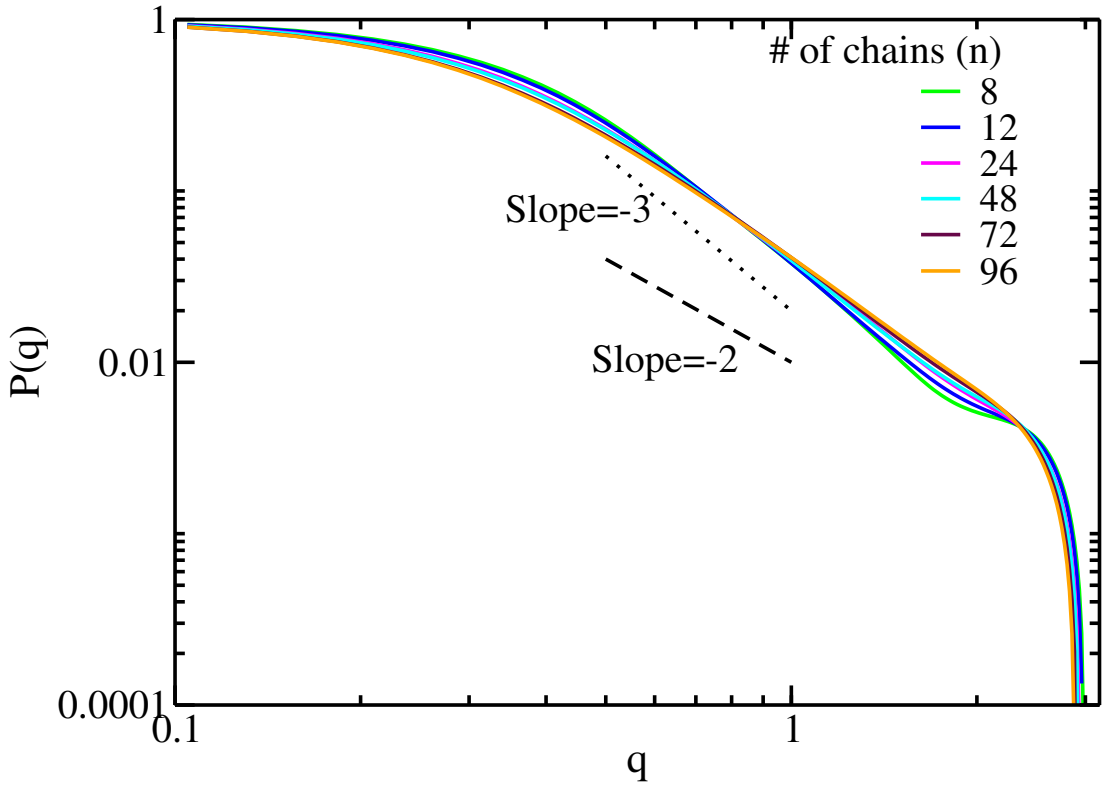


Figure 4.8. Form factor of a chain inside a complex as a function of scattering wave vector. Different colored lines are the number of chains making the complex.

tion inside the complex.

In addition, we calculated how the radius of gyration of a complex ($\langle R_{gA} \rangle$) itself scales with the number of monomers present in the complex as shown in Fig. 4.9. But we are focusing in the region where, both R_g and ν are not saturated i.e. for the region where $2n < 20$ in Figs. 4.6 and 4.7. The slopes in the Fig. 4.9 show that size-scaling exponent of the complex is globule-like when smaller complexes are formed. This is because of the strong electrostatic attraction between opposite charges. On increasing the number of monomers in the complex, the size-scaling exponent increases showing that the chains expand because of presence of other chains of both the charges in the surrounding. This shows both the $\langle R_g \rangle$ and $\langle R_{gA} \rangle$ scale in a similar manner with number of monomers.

4.3.2.3 Mean square displacement of a center of mass of a chain inside a complex

We tracked how a labeled chain inside a complex moves. For this the monomer density of the system we chose is $\rho = 0.046$ which has $2n = 96$, and $N = 60$. The movement of the yellow chain as shown in Fig. 4.10 is observed throughout the simulations.

Fig. 4.11 shows the mean square displacement (MSD) of a center of mass (CM) of a chain versus simulation time. The red line is the MSD of a CM of a single isolated polyelectrolyte chain and blue line represents the MSD of a CM of a labeled chain inside a complex made by $2n = 96$ and $N = 60$, where, $\rho = 0.046$. We found that MSD of a CM of a tagged chain inside a complex follow non-diffusive law given by,

$$< [R_{CM}(t) - R_{CM}(0)]^2 > \sim t^{\frac{4}{5}} \tag{4.11}$$

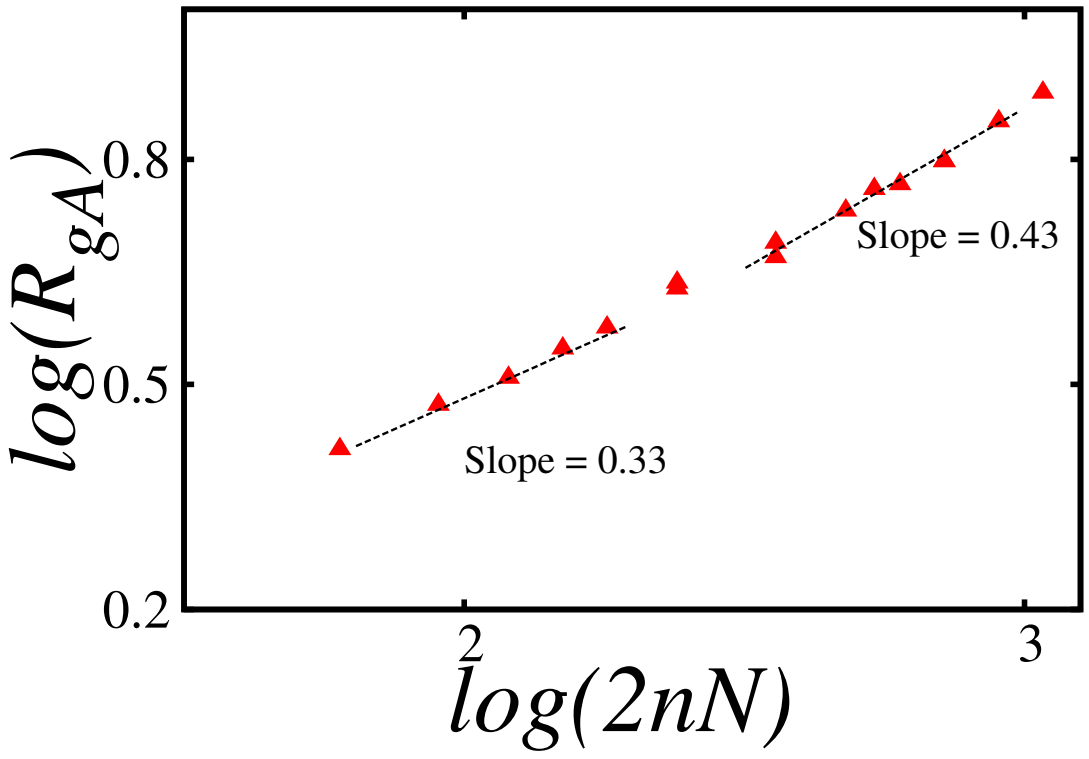


Figure 4.9. Form factor of a chain inside a complex of different chains. Different colored lines are the number of chains that make the complex.

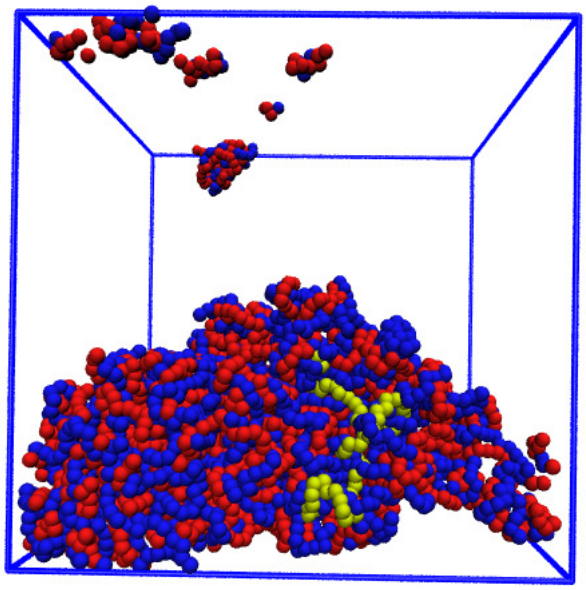


Figure 4.10. Snapshot of a complex formed by $2n = 96$ and $N = 60$. Blue beads are cations and red are anions. A chain with yellow color is a labeled chain. Counterions are not shown for clarity.

This result agrees with the results from the Monte-Carlo simulations by Srivastava et al. [112]. This is because the chain is a part of a physical network formed by the chains of opposite charges.

4.3.2.4 Effect of salt on $\langle R_g \rangle$ of a chain

Fig. 4.12 shows the effect of salt on the $\langle R_g \rangle$ of a labeled chain as a function of the number of polyelectrolyte chains in the complex. The $\langle R_g \rangle$ of a labeled chain in the complex at low polymer concentration were found to increase with the addition of salt on the system. It is because the screening of the electrostatic interaction due to salt ions makes the chain swollen. When a complex is made by more number of chains i.e. higher polymer concentration, addition of 1 M salt does not affect on the $\langle R_g \rangle$ of the labeled chain because the charges on the backbone of polyanions and polycations are already screened due to presence of other chains of opposite charges. This is again shown by the size-scaling exponent in Fig. 4.13.

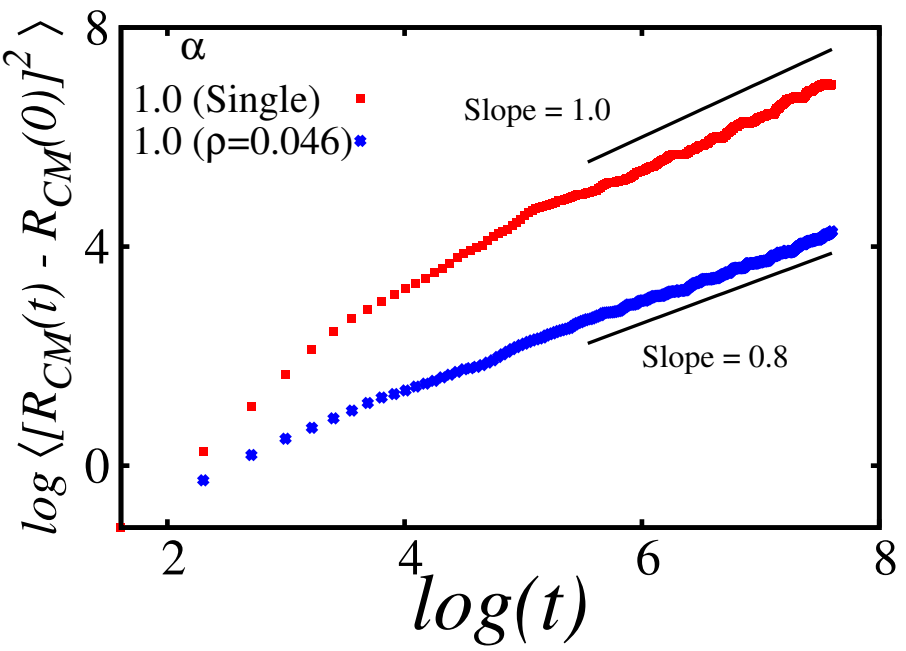


Figure 4.11. MSD of center of mass (CM) of a labeled chain inside a complex of $\rho = 0.046$ at $T^* = 1.0$. Red line is MSD of CM of a single isolated chain.

A single polyelectrolyte chain changes from a rod-like conformation to a coil-like conformation due to screening of backbone charge on addition of salt. At low polymer concentration forming complex, the size-scaling exponent of labeled chain were found to be increased as swollen of chain is observed due to screening of charges. On further increasing the polymer concentration in the system, the exponent does not change due to salt concentration (1 M) as the charges on the polymer backbone is already screened because of the presence of polymers of opposite charges.

To understand this phenomenon, we monitored the effect of salt concentration on the $\langle R_g \rangle$ of labeled chain inside the complex for low and moderate polymer concentration as shown in Fig. 4.14. In the Fig. $2n = 8$ corresponds to monomer density of 0.004 which is shown by red triangle points and $2n = 48$ corresponds to monomer density of 0.023 represented by blue square points. At low polymer concentration, the $\langle R_g \rangle$ increases slightly with salt concentration and for moderate polymer concentra-

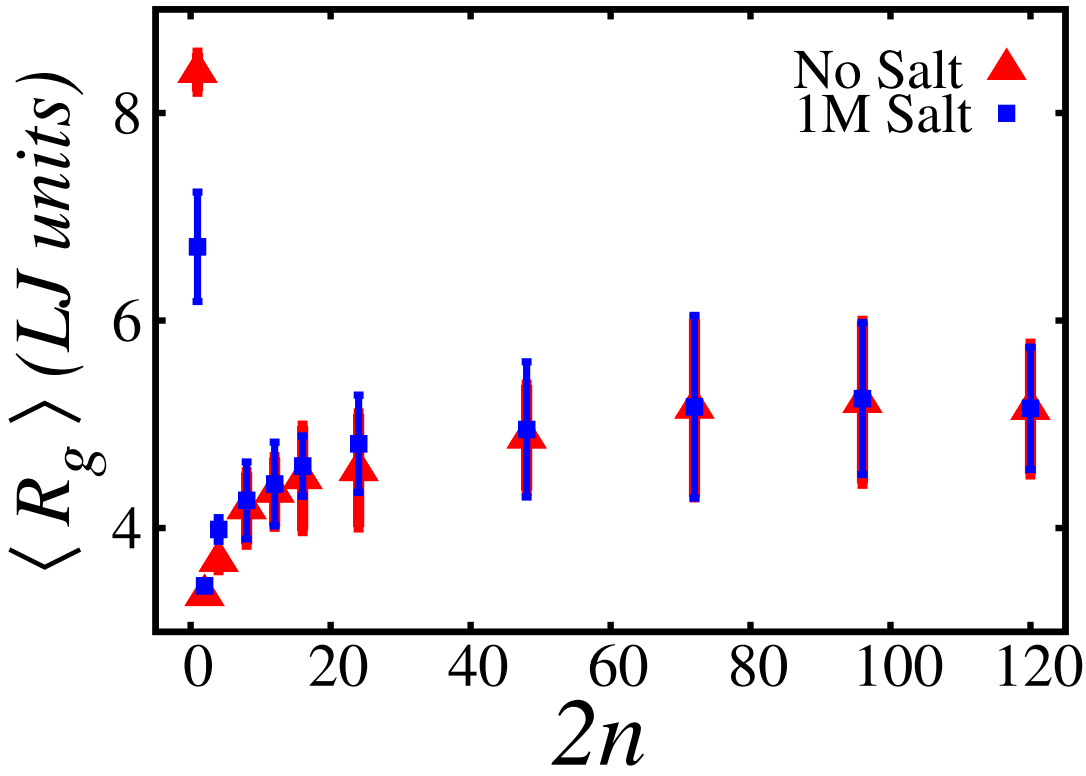


Figure 4.12. Comparison of $\langle R_g \rangle$ of a labeled chain in the complexes in salt free and with 1M salt concentration. Here, $N = 60$ for all the cases.

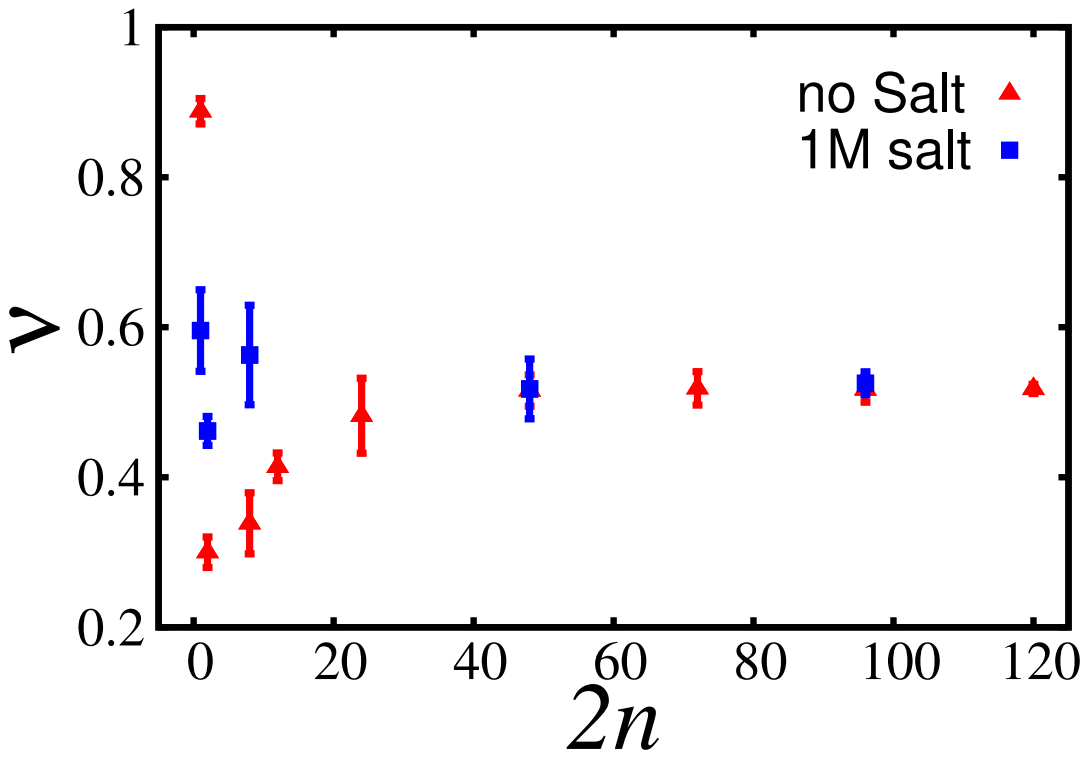


Figure 4.13. Comparison of size-scaling exponent of a labeled chain in complexes. Red triangle points represent salt-free case and blue square points are at 1M salt concentration.

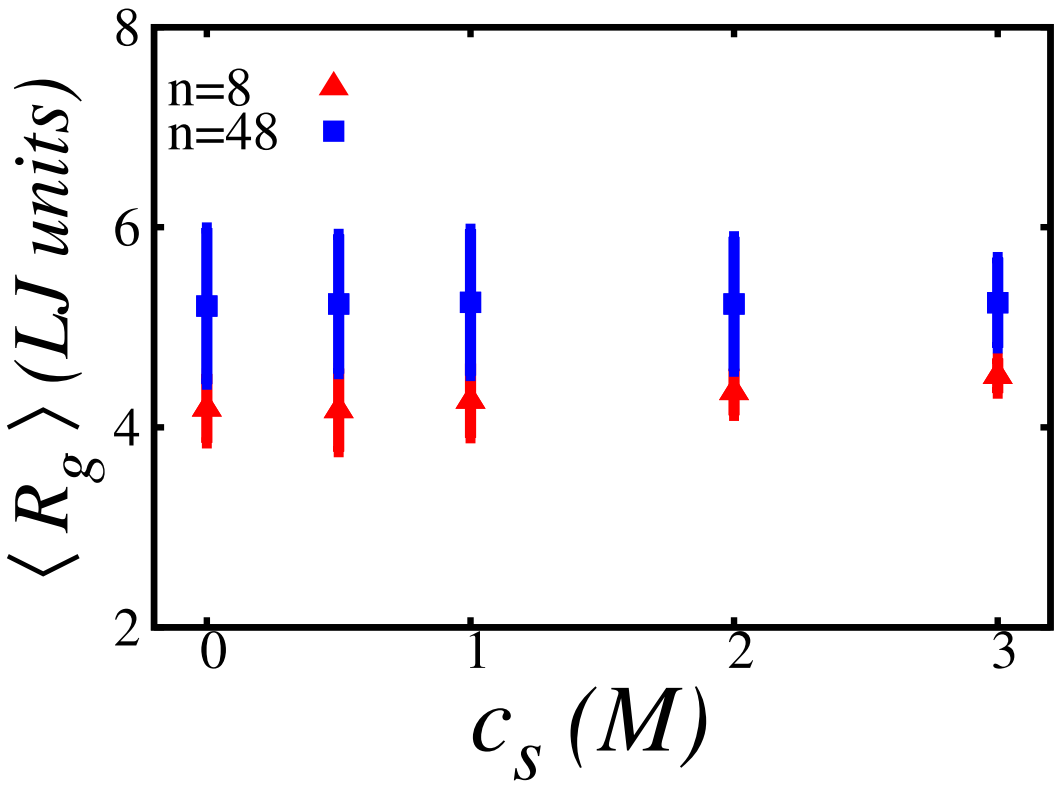


Figure 4.14. $\langle R_g \rangle$ of a labeled chain in two different polymer concentration per salt concentration. Red triangle points represent $\langle R_g \rangle$ of labeled chain when $2n$ is 8 ($\rho = 0.004$) and blue square points are that of when $2n$ is 48 ($\rho = 0.023$).

tion, the $\langle R_g \rangle$ doesnot change at all. At moderate and higher polymer concentration, we did not observe any structural changes at higher $\alpha = 1$ (not shown) in response to salt; aggregate (all chains assembled into one tightly packed large structure) did not respond at all to salt concentration as high as 6 M, indicating formation of glassy structure at very high α. Fares et al. have performed neutron scattering experiment to observe the effect of salt on the size of a deuterated poly(styrenesulfonate) inside the polyelectrolyte complex [113]. They found that, the R_g of the deuterated chain remain constant upto 1.4 M KB_r and then decreases on increasing the salt concentration. Our result shows that the $\langle R_g \rangle$ remains constant even for salt concentration higher than 1.4 M. This might be because we have not considered the effect of excess salt on the dielectric constant of water in our model [114].

Another reason for not observing the effect of salt concentration in our simulations might be the high charge density of polyelectrolyte chains. The complexation is too strong that the salt concentration we have used might not be sufficient to screen the electrostatic interaction between chains.

Therefore, how the polyelectrolytes of lower charge density organize inside a complex is another problem we explored. We expect that compared with the structure of highly charged polyelectrolyte complexes, they might show qualitatively different behaviors.

4.3.3 Role of charge density, temperature, and salt concentration on the structure of complex

Refering to the overlap concentration obtained from $\langle R_g \rangle$ of single isolated chain for different α listed in Table 4.5, monomer density (ρ) in our simulations are $\rho = 0.04$ and $\rho = 0.1$, for different values of α, which are above the overlap concentration. We

investigated the effect of different parameters: charge density of polymers (α), temperature (T^*), and salt conentration (c_s) on the structure and dynamics of polyelectrolyte complexes (PECs).

4.3.3.1 Snapshots

Figs. 4.15 and 4.16 are the representative snapshots of systems in equilibrium. We observed qualitatively similar behavior for different temperatures and monomer densities. After scanning through all the simulations, we found that qualitatively there are three distinct classes of aggregates as shown in Figs. 4.15 and 4.16. It is observed that at low α, only few chains complex together showing larger number of smaller sized complexes which is more like homogeneous distribution. With increasing α, number of chains forming complex together also increases. At intermediate α, smaller and larger complexes coexist. At high α, we observed the formation of one big complex.

4.3.3.2 Cluster size distribution

We calculated the cluster size distribution of a complex formed for different α, T^*, and ρ. The criteria we chose to find the cluster are as follows: We first calculated the distance between every charged bead of a polyanion and a polycation. If at least one of the distance between pairs is less than the threshold value (here, the threshold is 1.12σ), then these two chains are considered to be directly connected. We calculated the distance between all the possible pairs of chains to find which of the chains are directly connected. A chain is a member of a cluster if chains are either directly or indirectly connected. Lets take a example to describe what we mean by indirectly connected. If chain A is connected with chain B, and chain B is connected with chain C, then, chain A and chain C are indirectly connected. From the list of all directly connected chain pairs, we found the members of clusters. For example, in a system of chains A B, C, and D, let [A, B], [A, D], and [B, C], be directly connected pairs.

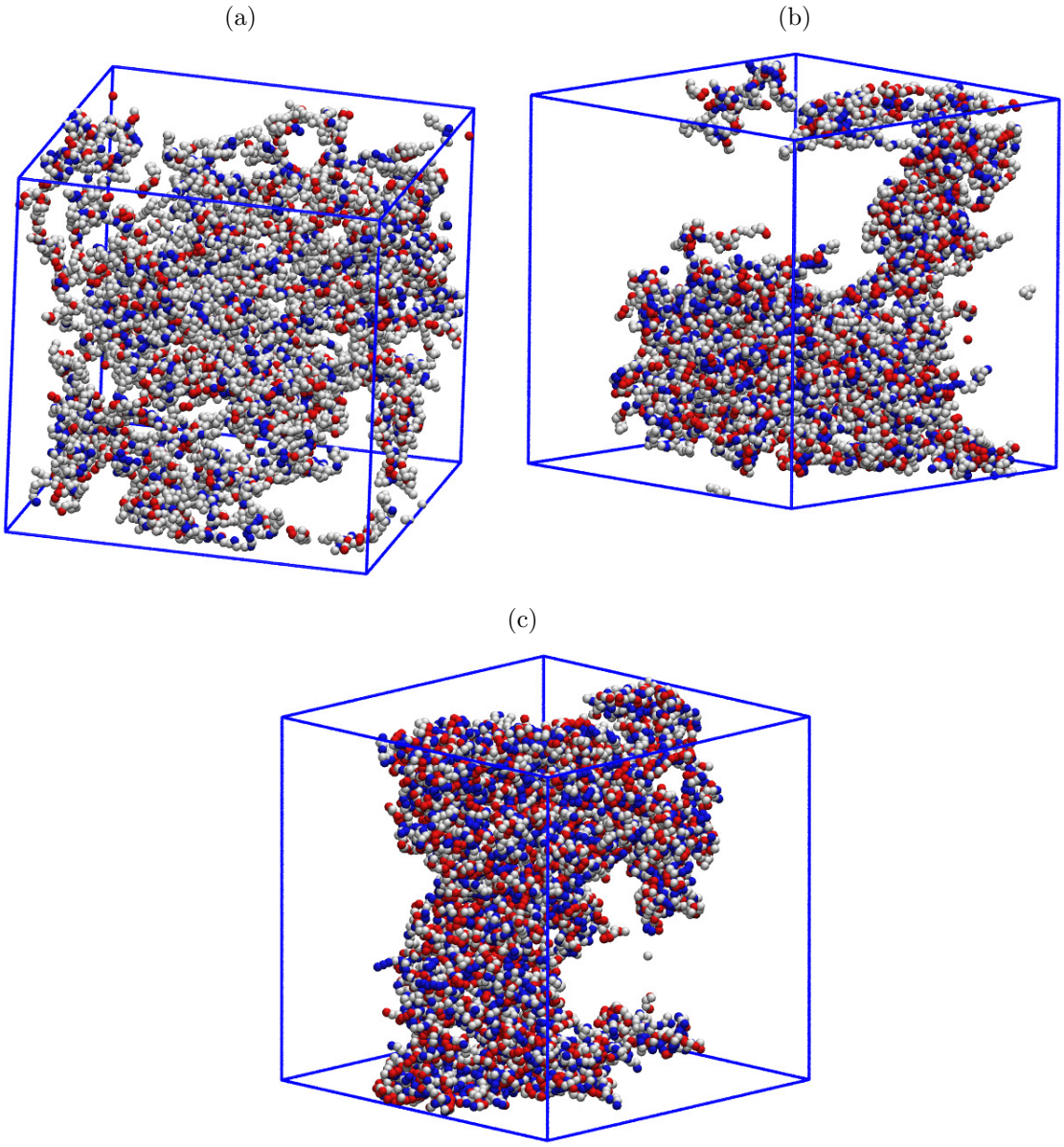


Figure 4.15. Snapshots of complex structures formed by the complexation of flexible polycations and flexible polyanions of same chain length in the absence of added salt. Snapshots are at $\rho = 0.04$, $T^* = 1.0$ and (a) $\alpha = 0.3$, (b) $\alpha = 0.45$, and (c) $\alpha = 0.6$. Red beads are negative monomers, blue beads are positive monomers, white beads are neutral monomers. Counterions are not shown for clarity.

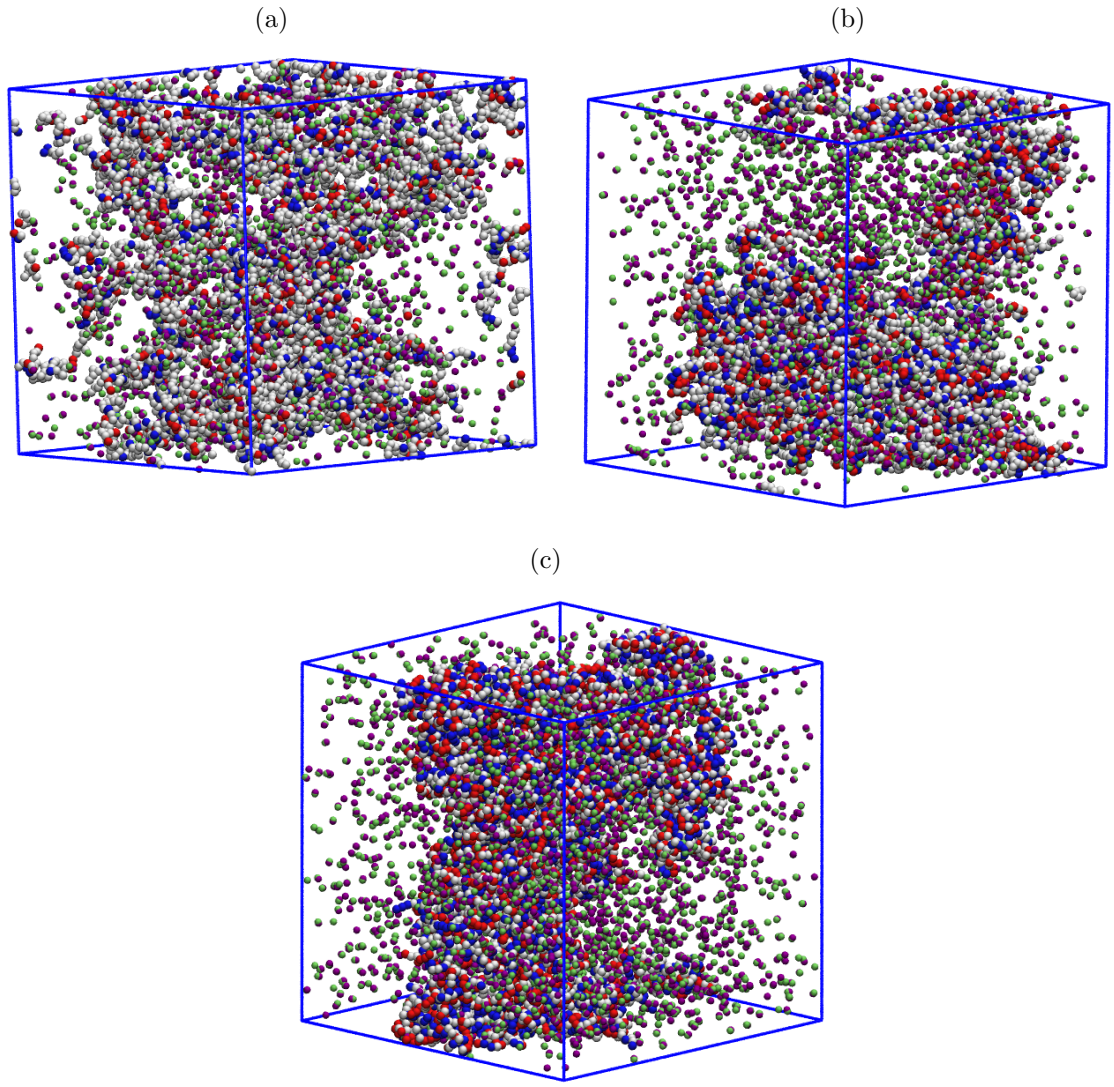


Figure 4.16. Snapshots of complex structures formed for a system at $\rho = 0.04$, $T^* = 1.0$ and (a) $\alpha = 0.3$, (b) $\alpha = 0.45$, and (c) $\alpha = 0.6$. Red beads are negative monomers, blue beads are positive monomers, white beads are neutral monomers, green beads are positive counterions, and magenta beads are negative counterions.

(a)

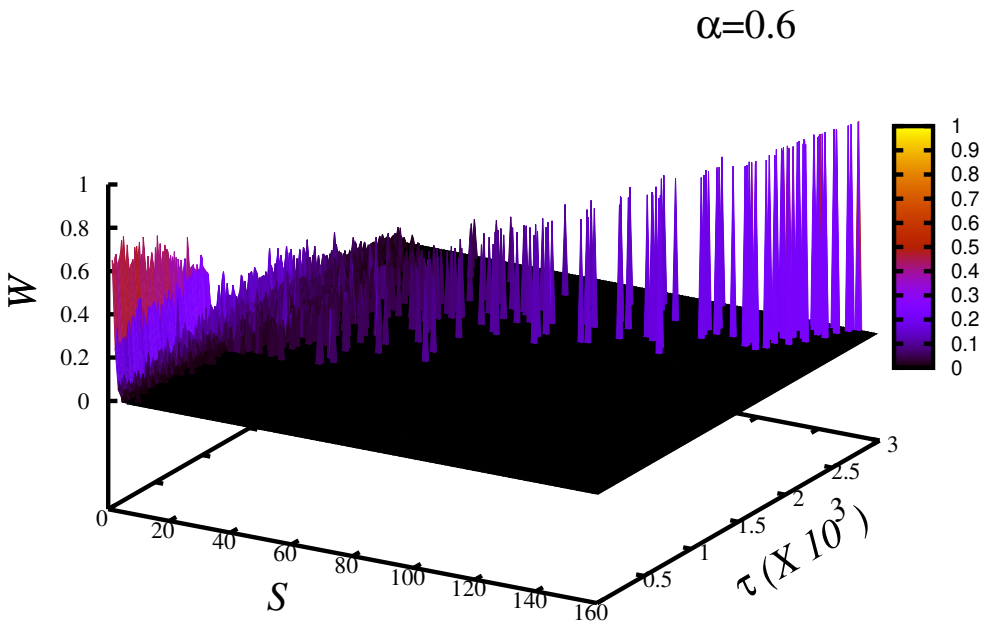


Figure 4.17. Time evolution of mass fraction of cluster size (s) for $\rho = 0.04$, $\alpha = 0.6$, and $T^* = 1.0$ before equilibrium.

Then chains [A, B, C, D] make a cluster.

We also defined the mass fraction of cluster size as $W = sf/2n$. Here, s is the number of chains that make a cluster, f is the frequency of a cluster of size s, and n is the total number of polycations or polyanions ($2n$ is the total number of chains in the system).

Fig. 4.17 shows the evolution of mass fraction of cluster size s for $\rho = 0.04$, $\alpha = 0.6$, and at $T^* = 1.0$ before equilibrium, which is a representative figure of all simulations. The figure shows the time evolution of bigger clusters at high α with increasing time. Fig. 4.18 shows the time evolution of mass fraction of cluster size for intermediate α, and high α. The Fig. shows that, at intermediate α, both the bigger

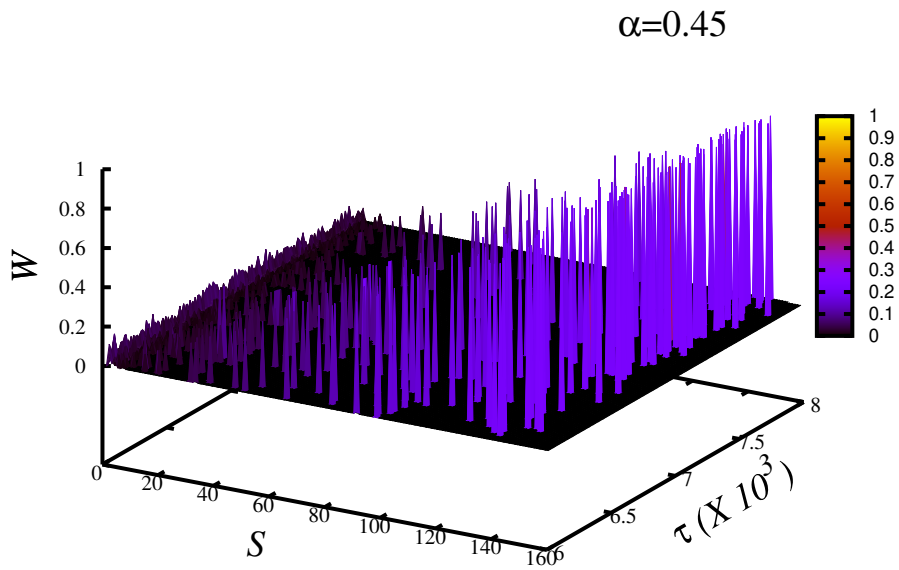
and smaller aggregates are present, while at high α, only the bigger clusters are visible.

Once the system reached equilibrium, the cluster size distribution did not get one particular distribution. There was a fluctuation in the size of the cluster formed, as shown in Fig. 4.18. The fluctuations in the size of the aggregates are more clearly seen in Fig. 4.19. The fluctuations in the size of the clusters formed are minor at low and high α, while they are more significant at intermediate α.

It was observed that the cluster size distribution was not static even after the equilibrium. After the system reached equilibrium, we compared the cluster size distribution between 0τ - $10^3\tau$ and $10^3\tau$ - $2\times10^3\tau$, they were very similar, indicating a dynamic equilibrium system. When we compared 100 observation time frames in a simulation time of $10^3\tau$, we found that the statistics were sufficient for determining the equilibrium cluster size and we chose 200 observation time frames in a simulation time of $2\times10^3\tau$ for cluster size distribution analysis.

In figs. 4.20 to 4.23, the y-axes are mass fractions (W) of the cluster size s, and x-axes are size of the clusters formed for different ρ, α and T^*. We calculated the cluster size distribution by averaging over 200 time-frames after the system reached equilibrium. The Figs. show a clear transition of formation of larger clusters from smaller clusters in going from small α to large α as well as from high temperature to low temperature. As observed in Figs., at higher α, almost all chains of the system formed a single complex. At smaller α, α =0.3, α =0.35, and α =0.4, the cluster size distribution varies narrow to wider distribution, where narrow distribution is observed for smaller α. At intermediate α, α =0.45 and $\alpha = 0.5$, the aggregate size distribution is extensive, indicating the formation of both smaller and larger aggregates together.

(a)



(b)

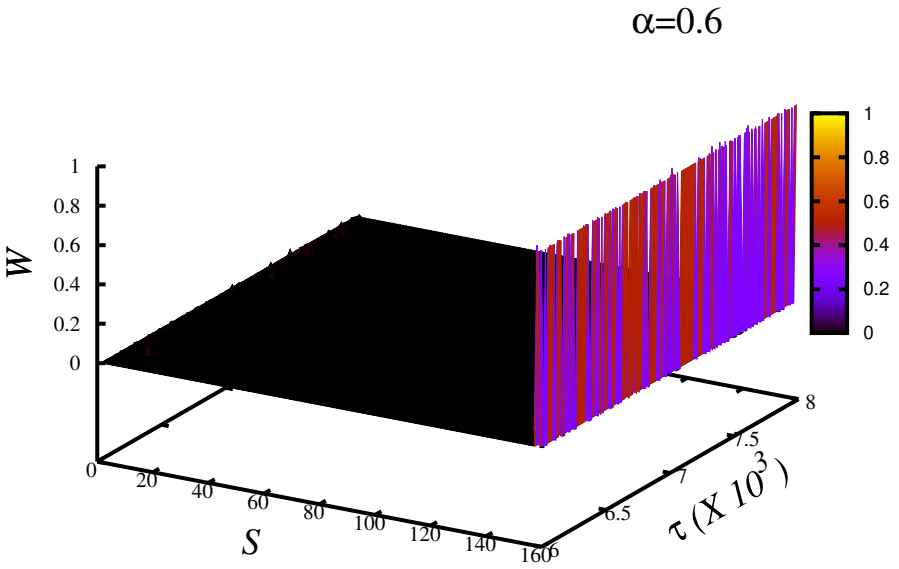


Figure 4.18. Time evolution of mass fraction of cluster size (s) for $\rho = 0.04$, and $T^* = 1.0$ (a) $\alpha = 0.45$ (b) $\alpha = 0.6$. Both of these are after equilibrium

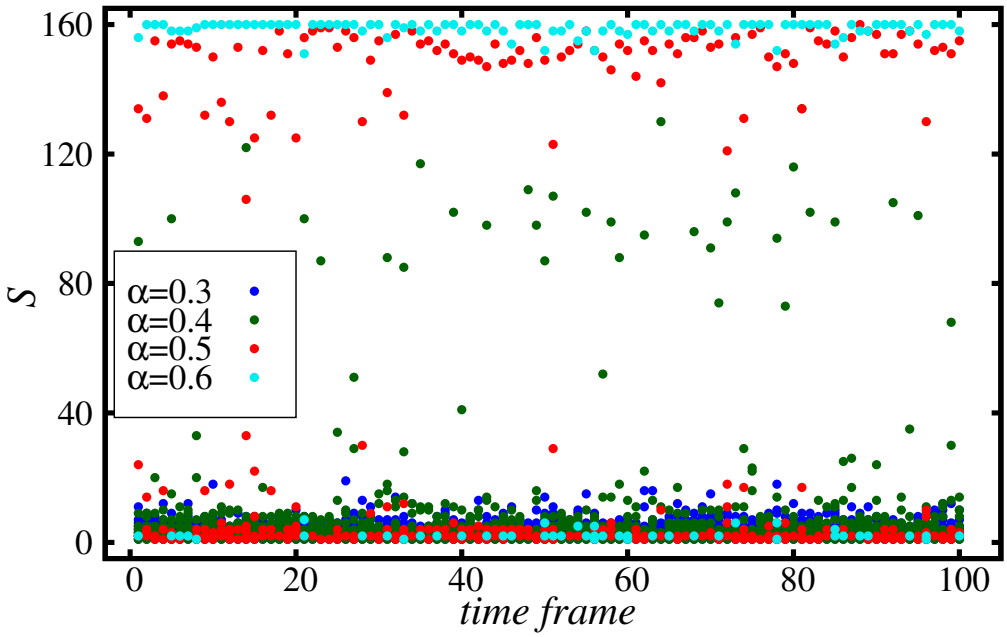


Figure 4.19. Distribution of size of the clusters formed for $\rho = 0.04$ and at $T^* = 1.0$. Different colored data-points are for for different α. The cluster size distribution is tracked for 100 time-frames after the system reached equilibrium.

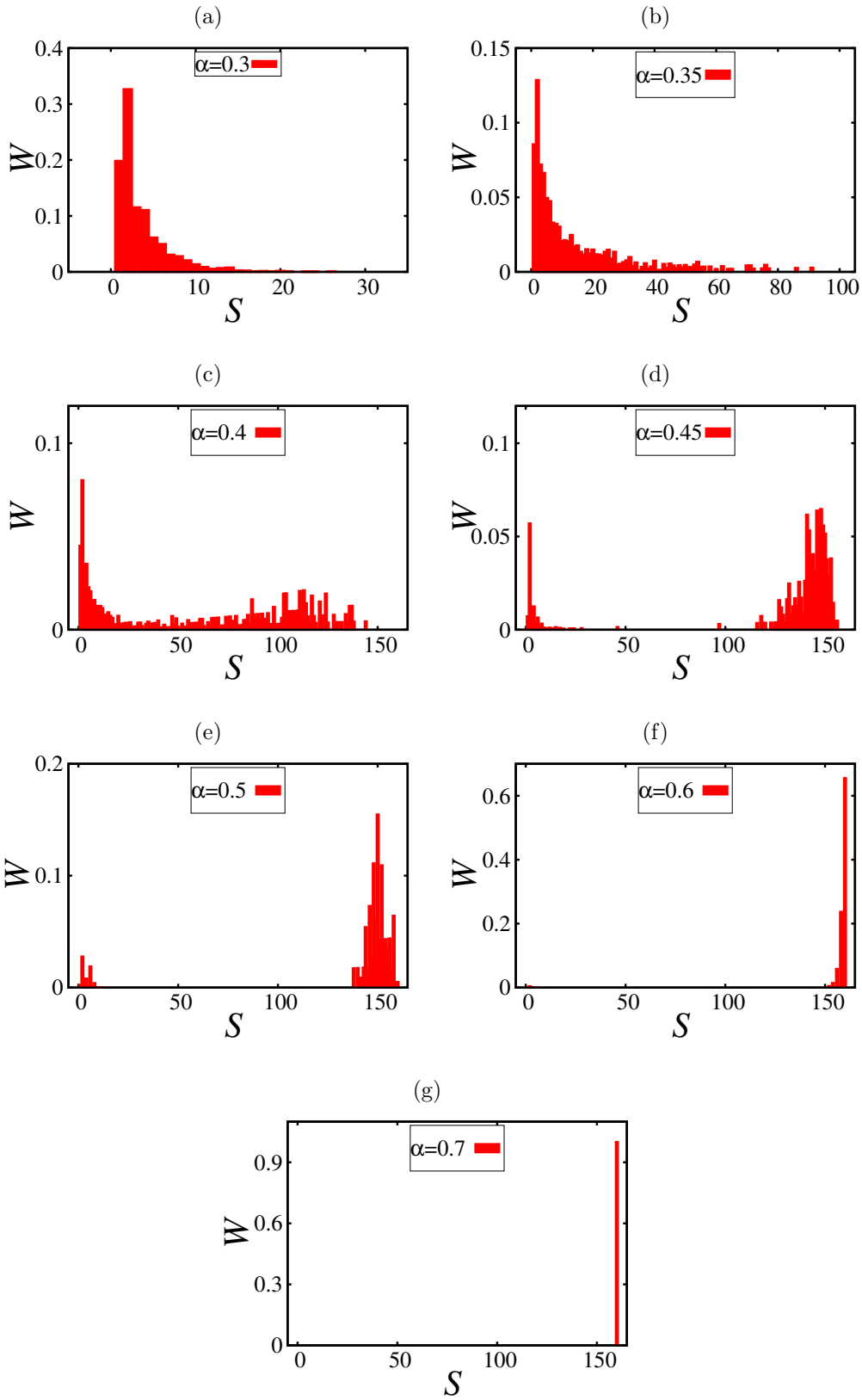


Figure 4.20. Distribution of the mass fraction of the clusters for $\rho = 0.04$, at $T^* = 0.95$ and different α. The x-axes are size of the clusters formed and y-axes are their mass fraction.

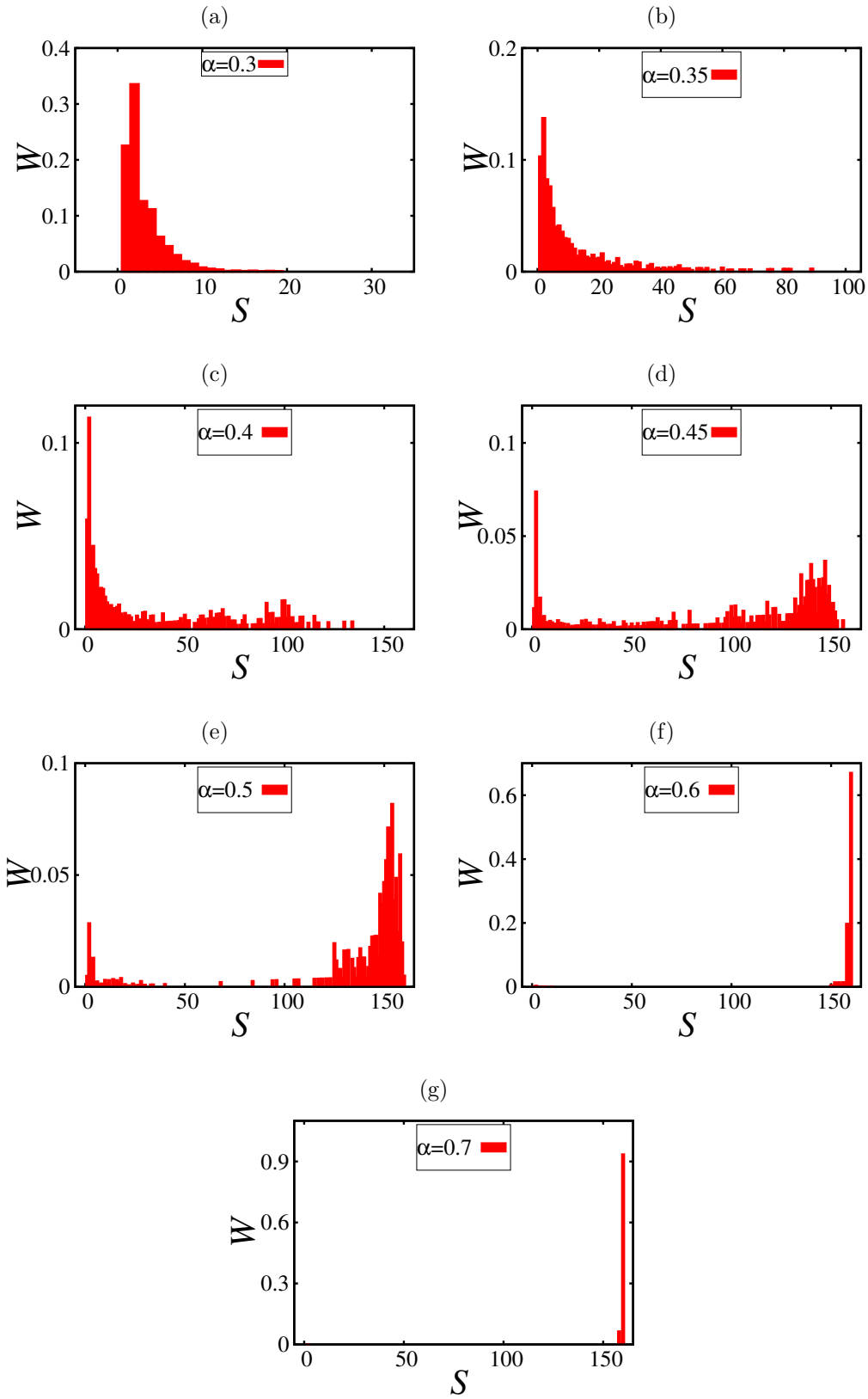


Figure 4.21. Distribution of the mass fraction of the clusters for $\rho = 0.04$, at $T^* = 1.0$ and different α. The x-axes are size of the clusters formed and y-axes are their mass fraction.

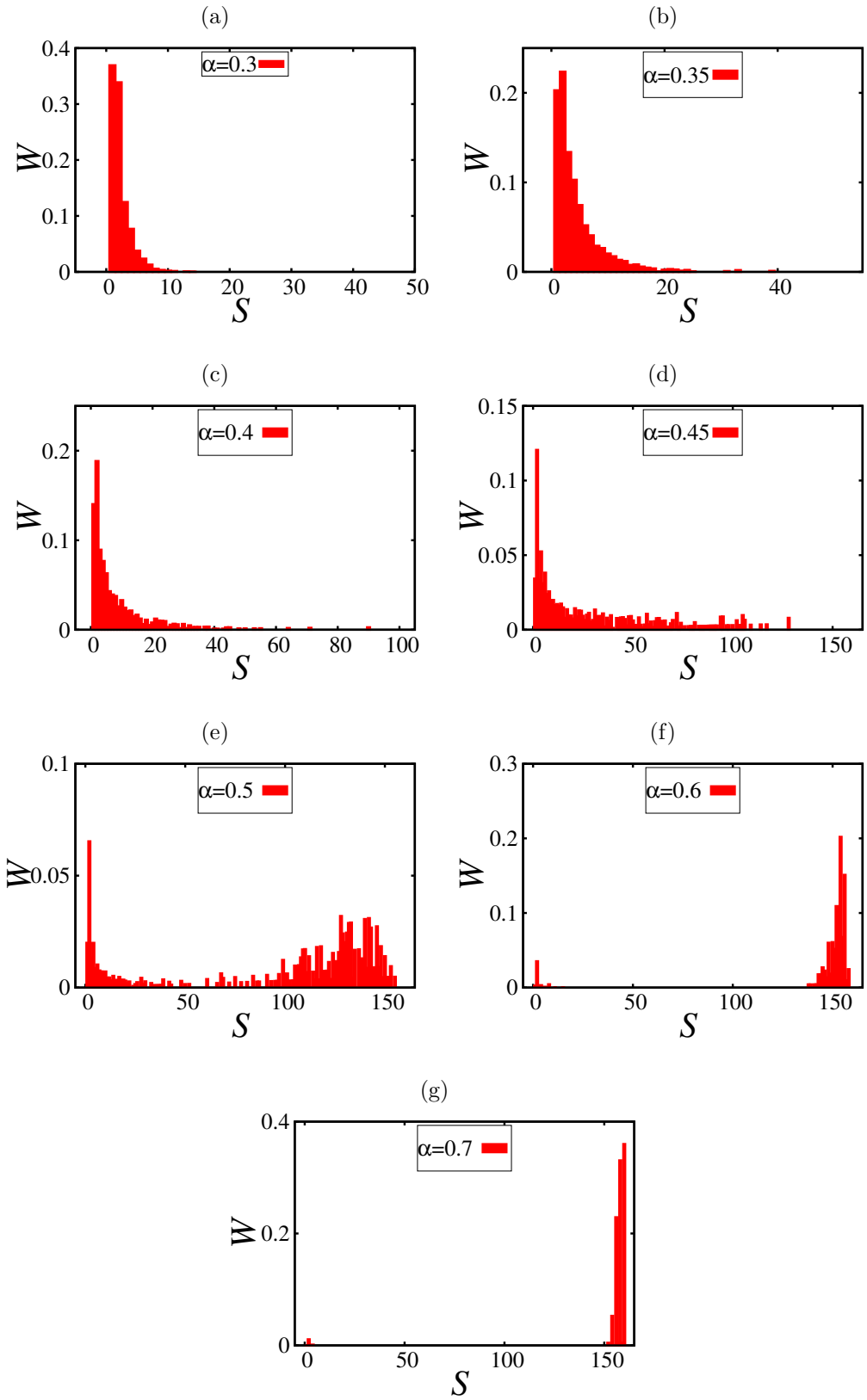


Figure 4.22. Distribution of the mass fraction of the clusters for $\rho = 0.04$, at $T^* = 1.2$ and different α. The x-axes are size of the clusters formed and y-axes are their mass fraction.

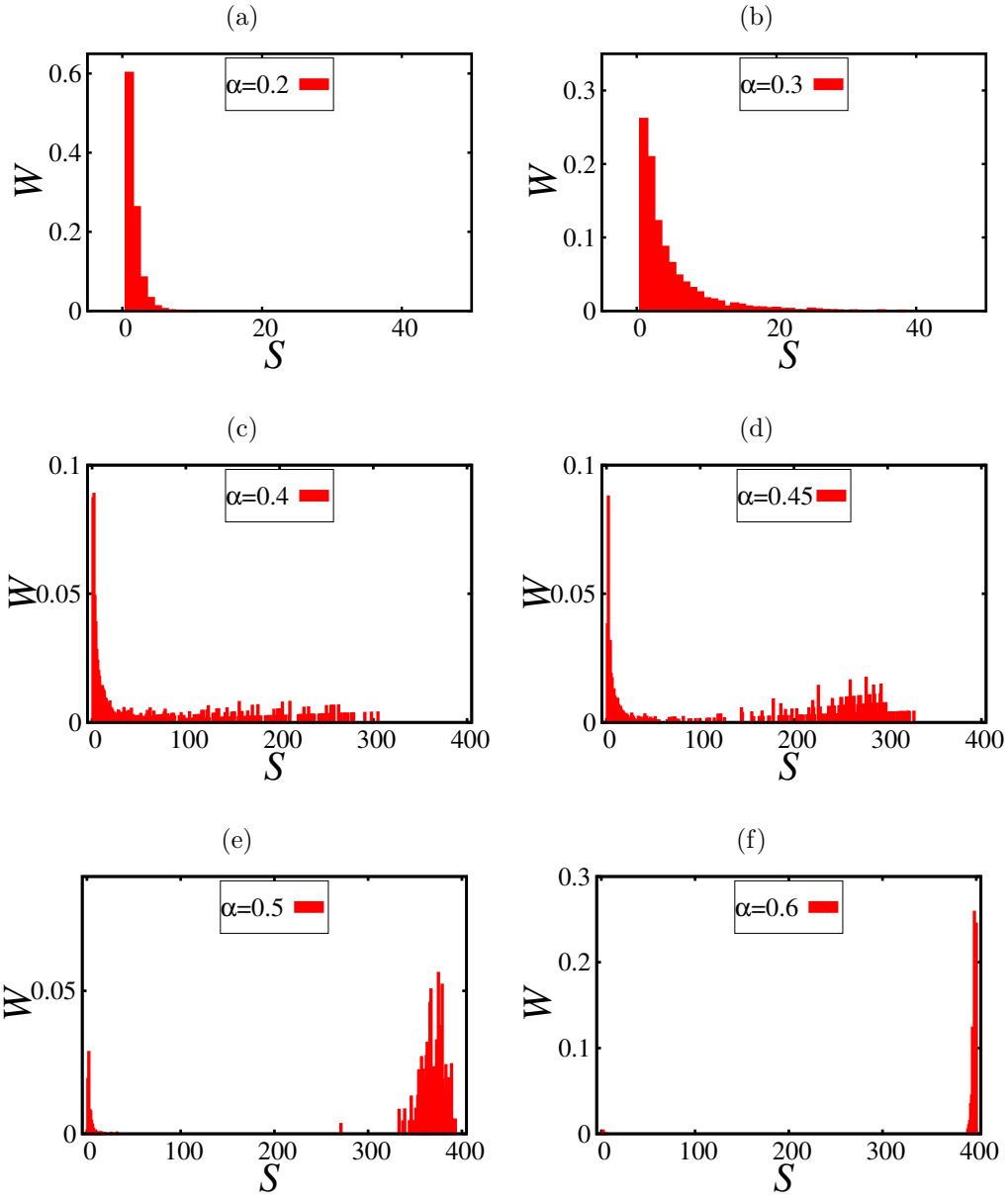


Figure 4.23. Distribution of the mass fraction of the clusters for $\rho = 0.1$, at $T^* = 1.0$ and different α. The x-axes are size of the clusters formed and y-axes are their mass fraction.

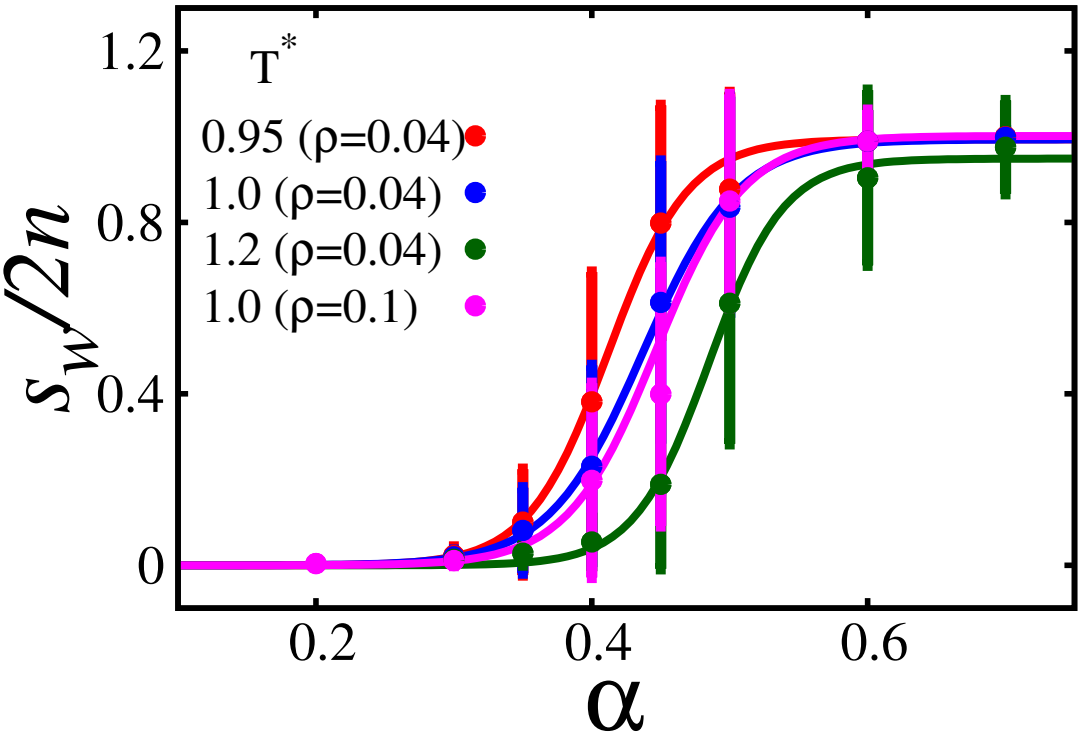


Figure 4.24. Weighted mean and variance of the size of the cluster formed (normalized by the total number chains) as a function of α fitted by the Sigmoid function shown by smooth lines.

4.3.3.3 Weighted mean

One way of summarizing the above observation from figs. 4.20 to 4.23 is with the weighted mean. The reason for choosing the weighted mean over non-weighted mean is for giving more weight to the large clusters formed because, larger clusters occupy more volume than the small clusters.

Weighted mean $(\bar{S_w}) = \sum_{i=1}^{N} S_i w_i$,

where, $w_i = \frac{s_i f_i}{\sum S_i f_i}$, S is the size of the cluster and f is the frequency of the cluster of size S.

Variance $(\sigma_w^2) = \sum_{i=1}^{N} \frac{w_i (S_i - \bar{S_w})^2}{V}$,

where, $V = \sum_{i=1}^{N} w_i = 1$ (in our case).

In Fig. 4.24, y-axis is normalized $\bar{s_w}$ and x-axis is α fitted by the function $f(x) = \frac{1}{1+exp(-w*(x-\alpha_c))}$, where, α_c is the critical charge density which is the inflection point of the fitted sigmoid function, and w is a measure of sensitivity of S_w to $\Delta\alpha$. Different colored plots are for different temperatures and ρ. The Fig. shows that the formation of no or smaller complexes at small α and big complexes at high α, and there is a clear transition from the smaller to higher complexes at intermediate α. The larger standard deviation at intermediate α represents the distribution of different sized clusters. The presence of single chains floating around bigger clusters also contributes towards the larger standard deviation, especially at the intermediate α. Since the weighted mean does not give the complete picture of the cluster size distribution and there is more weight to the bigger clusters formed, similar to the analysis of the biggest clusters formed. Hence, to make a quantitative comparison of equilibrium aggregation behavior, we calculated the biggest cluster size normalized by the total number of chains in the system ($s_b/2n$) followed by Sigmoid fitting to determine the critical charge density.

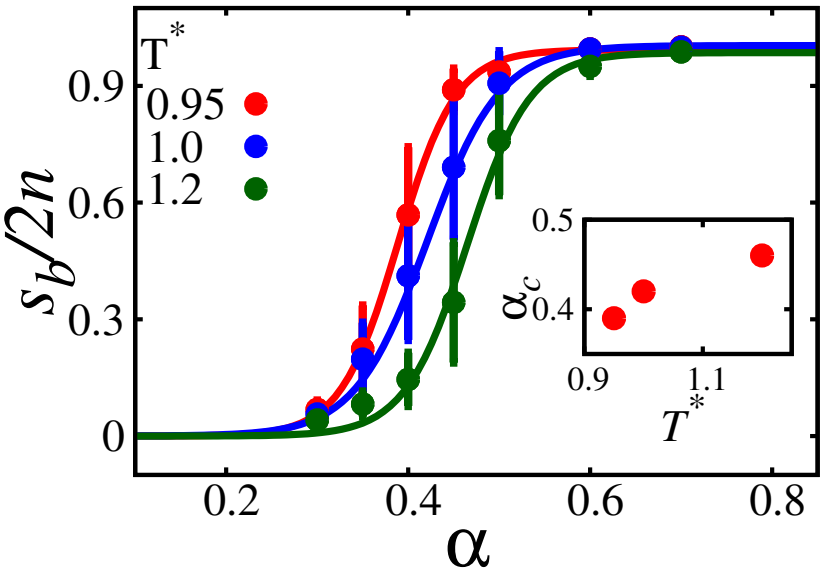
4.3.3.4 Biggest cluster size

From the distribution of different ρ, T^*, and α, we calculated the mean and standard deviation of the biggest cluster (s_b) formed as a function of α. The system showed Sigmoidal aggregation behavior as a function of α for the three temperatures and two polymer concentrations. In Fig. 4.25(a), the y-axis is the normalized s_b and the x-axis is α. We fit the results with a Sigmoid function given by,

$$f(x) = \frac{1}{1 + exp(-w(x - \alpha_c))} \tag{4.12}$$

where α_c is the critical charge density, which is the inflection point of the fitted Sigmoid function, and w is a measure of the sensitivity of S_b to $\Delta\alpha$. Different colored

(a)



(b)

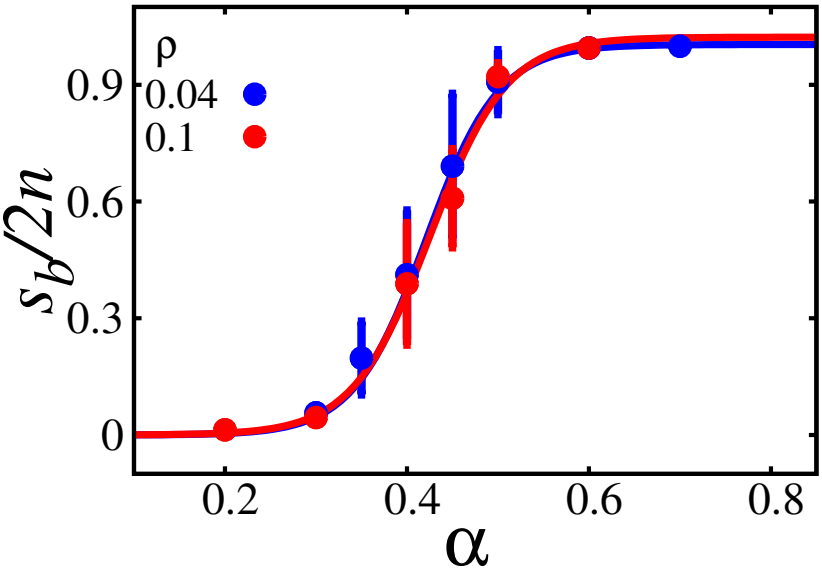


Figure 4.25. (a) Mean and standard deviation of biggest cluster formed (normalized by the total number chains) as a function of α fitted by a Sigmoid function for the system with $\rho = 0.04$ for different T^*. The values of critical charge density (α_c) as a function of temperature (T^*) is shown in the inset. α_c is the point of inflection of the Sigmoid curve. (b) Mean and standard deviation of biggest cluster formed (normalized by total number of chains) as a function of α fitted by a Sigmoid function for systems at $T^* = 1.0$ for $\rho = 0.04$ and $\rho = 0.1$.

plots are for the system with different T^* for $\rho = 0.04$. When converted to real units, the dimensionless temperature ranges from 285 K, 300 K, and 360 K, corresponding to $T^* = 0.95$, 1.0, and 1.2, respectively. In this temperature range, the Sigmoid curve shifted as a function of temperature, with α_c increasing with temperature. α_c represents a crossover from un-clustered structure to clustered structure. The inset in the Fig. shows the values of α_c as a function of temperature. The numerical values of α_c are listed in Table 4.7. We observed that the α_c increases with an increase in temperature. With the increase in temperature, the thermal fluctuation increases, acting against the formation of clusters resulting in higher α_c. Moreover, we found that α_c is independent of ρ as considered in our simulations, which are both above the overlap concentration, as shown in Fig. 4.25(b).

System	α_c
$\rho = 0.04$, $T = 0.95$	0.389
$\rho = 0.04$, $T = 1.0$	0.4210
$\rho = 0.04$, $T = 1.2$	0.465
$\rho = 0.1$, $T = 1.0$	0.4251

Table 4.7. Values of α_c from Fig. 4.25.

To understand the effect of high salt concentration on the structure of complexes formed, we added monovalent salt in the system of $\rho = 0.04$ for $\alpha = 0.5$, and $\alpha = 0.6$ at $T^* = 1.0$. These are the systems where almost all chains form a cluster with only a few chains floating around a big cluster when no salt is present. Figs. 4.26 and 4.27 represent the scatter plot of cluster size distribution with time after the system is in equilibrium. Figs. show that the bigger clusters are dissociated with the increase in salt concentration.

As shown in Figs. 4.26, 4.27 and 4.28, increase in c_s leads to decrease in the average size of the complexes. For smaller α=0.5, the aggregate size decreased with increase in salt concentration and it fully breaks down at 2M salt. For higher α=0.6,

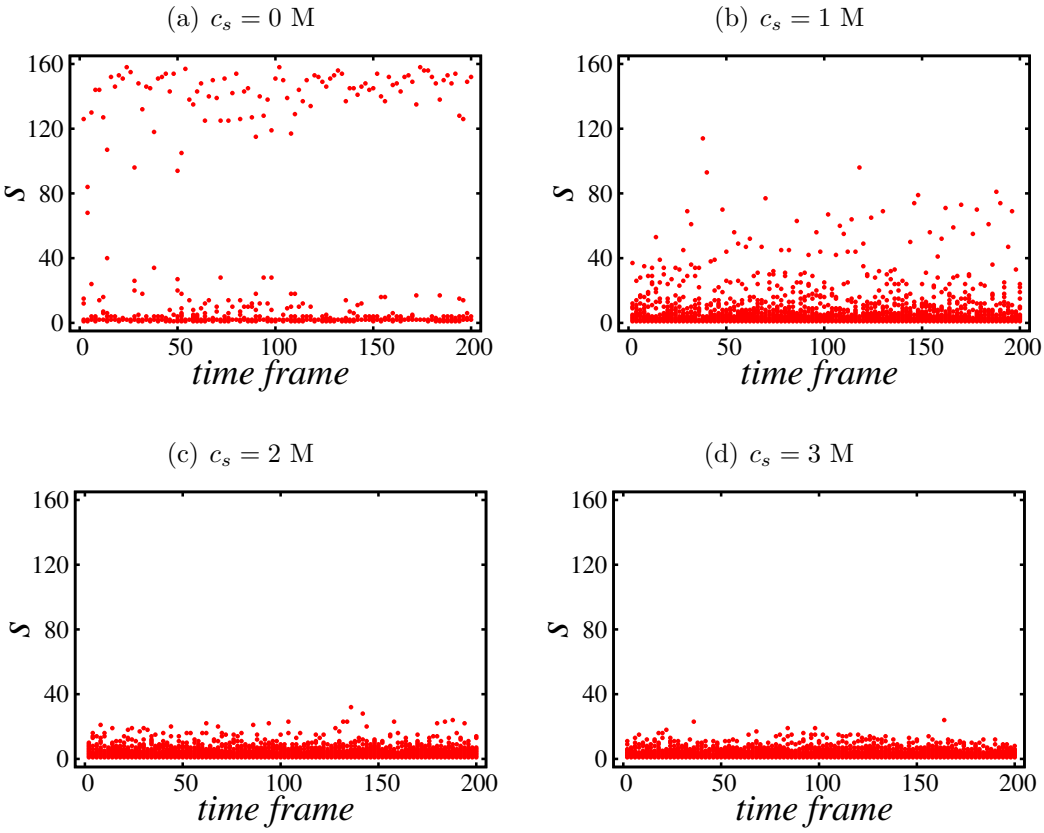


Figure 4.26. Scatter plot of cluster size distribution with time after the system is in equilibrium for $\rho = 0.04$, and $\alpha = 0.5$ at $T^* = 1.0$ with diferent c_s.

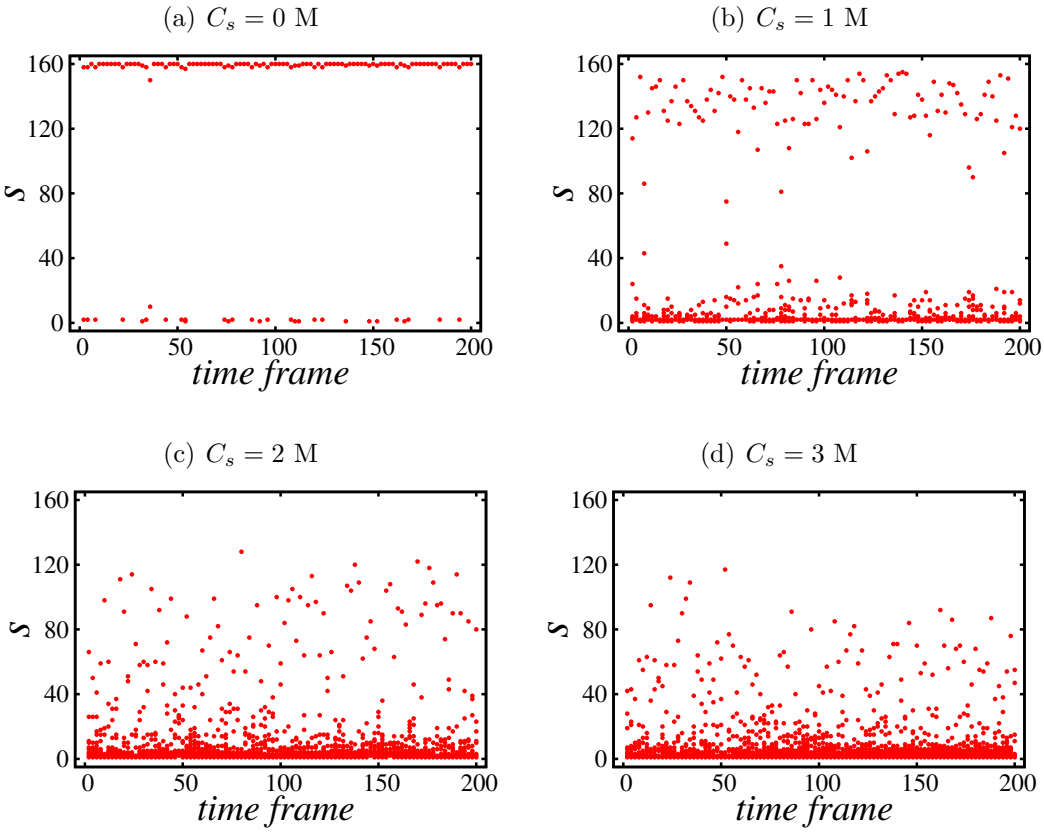


Figure 4.27. Scatter plot of cluster size distribution with time after the system is in equilibrium for $\rho = 0.04$, and $\alpha = 0.6$ at $T^* = 1.0$ with diferent c_s.

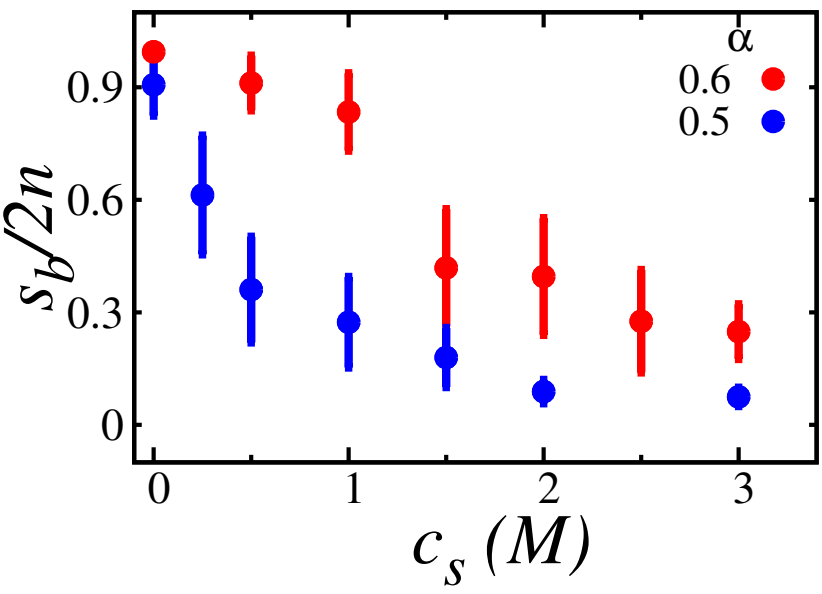


Figure 4.28. Mean and standard deviation of biggest cluster formed (normalized by total number of chains) as a function of salt concentration for $\rho = 0.04$, $\alpha = 0.5$ and $\alpha = 0.6$ at $T^* = 1.0$ with explicit salt concentration.

the aggregate size reduced with increase in salt concentration but it did not fully break down even at a high salt concentration of 3M. In summary, the polymers with smaller α are more responsive to salt, due to smaller number of electrostatic contact points between oppositely charged polyelectrolytes, the rest of them being excluded volume interactions.

We made qualitative comparisons of our results with the literature related to charge density and salt effects on polyelectrolyte complexes. Using optical turbidity measurement on model acrylamide copolymers with varying charge density and non-ionic comonomer hydrophobicity, Huang et al. [115] showed that with an increase in charge density of the polymers, the transition salt concentration (from 2 phases to single-phase), as well as the concentration of polymers in coacervates, increased. The transition salt concentration increased almost linearly from 0.4-0.6 M with an increase in mole fraction of charged monomers from 0.65 to 1, respectively, independent of the nature of the non-ionic component (whether hydrophilic or hydrophobic).

Another system, Hyaluronic acid - chitosan compl [116] showed qualitatively similar behavior; the strong electrostatic association was observed at very low pH, which can prevent if a large amount of salt (1.5 M NaCl or more) is introduced. Qualitatively similar behavior was observed for Gum arabic -whey protein [94, 117]. A system containing a strong ionic group such as poly(sodium styrene sulfonate) (PSS) and poly(methacryloxyethyltrimethylammonium) chloride (polyMETAC) showed a more drastic behavior. They precipitate out of solution when exact stoichiometry is achieved [85]. Using molecular dynamics simulations, Shakya et al. [118] showed disassembly of the polyelectrolyte complexes for a dilute polymer solution corresponding to weak electrostatic interaction ($\Gamma = 0.7$, Bjerrum length over σ). Complete disassembly was observed at a low salt concentration of 150 mM. The common outcome of these studies and our simulation study is that increasing the charge density of the monomers can enhance the aggregation, and it can be reduced by decreasing the charge density of monomers or increasing salt concentration in the system. This kind of behavior has been observed for a large number of polycation-polyanion pairs with different types of salt ions [75]. A more quantitative treatment is discussed later in this Section.

4.3.3.5 $\langle R_g \rangle$ of a labeled chain

Similar to the investigation of labeled chain inside of complex of different sizes as explained in Section 4.3.2.2, we investigated the influence of α, ρ, T^*, and c_s on $\langle R_g \rangle$ of individual polymer chains in semi-dilute polymer solutions. For this, we calculated the average radius of a gyration ($\langle R_g \rangle$) of a labeled chain for each size (s) of clusters formed,. In Fig. 4.29, the y-axis is the $\langle R_g \rangle$ of a labeled chain, and the x-axis is the size of the cluster in which a labeled chain belongs to. Different colored data are for different ρ and α. The Fig. shows that the $\langle R_g \rangle$ is independent of cluster size, α, ρ,

T^*, and c_s at semi-dilute regime. The $\langle R_g \rangle$ is averaged from 200 time-frames after the system is in equilibrium.

In addition, we also calculated the form factor of a labeled chain in semi-dilute polymer solutions as a function of α as shown in Fig. 4.30. The slope of $P(q)$ versus q (scattering wave vector in units of inverse length LJ units) was -2, suggesting a Gaussian conformation. The size-scaling exponent (ν) for the system at $\rho = 0.04$ at different α is listed in Table 4.8. The values of $\nu \sim 0.5$ is independent of the charge density. The result indicates that the chains show Gaussian-like conformation for all the charge density of the system. This might be due to the screening of coulomb interactions at high concentration of oppositely charged monomers for higher values of α. The same slope of -2 was observed for all T^*, c_s, and ρ.

α	ν
0.7	0.498
0.6	0.496
0.5	0.501
0.4	0.499
0.3	0.502

Table 4.8. Size-scaling exponent of a chain in the system of different charge density.

In experimental literature, the conformation of labeled chains in the polyelectrolyte complexes has been determined using small angle neutron scattering (SANS) of samples containing a small fraction of deuterated chains of either polycation or polyanion [89, 119, 113]. The constituent polyelectrolytes either take Gaussian chain conformation [89, 119] or self-avoiding random walk conformation [113]. Using poly(acrylic acid) and poly(N,N-dimethylaminoethylmethacrylate) Spruijt et al. [89] showed that the overall structure is similar to semidilute polymer solutions, with polycations and polyanions strongly overlapping to form a network with a mesh size that is much smaller than the radius of gyration of the polymers. The mesh size and the radius of gyration of labeled chains increased with increasing salt concentration. For

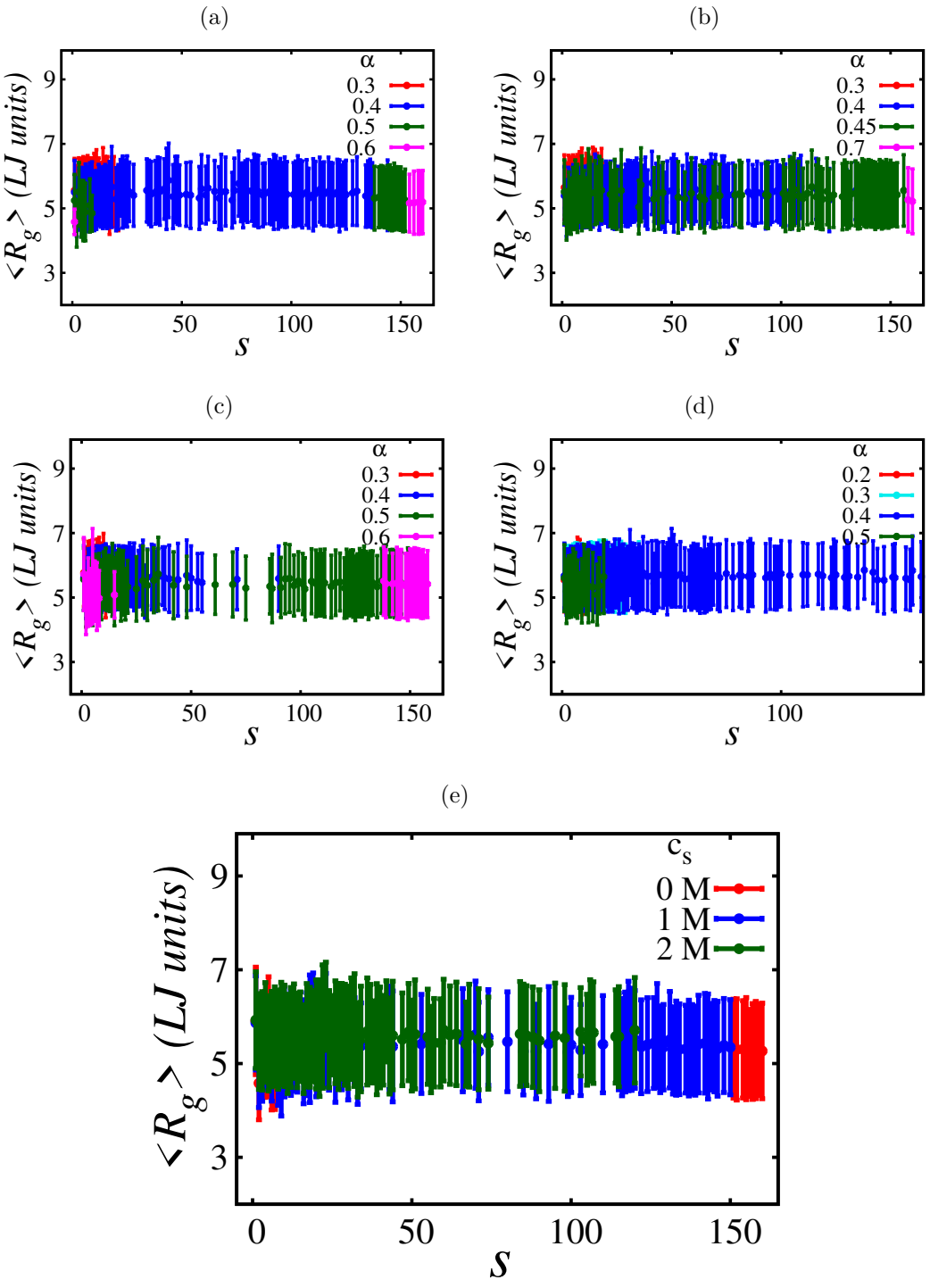


Figure 4.29. $\langle R_g \rangle$ of a chain in a cluster of size s for (a) $\rho = 0.04$ at T^*=0.95, (b) $\rho = 0.04$ at T^*=1.0, (c) $\rho = 0.04$ at T^*=1.2, and (d) $\rho = 0.1$ at T^*=1.0. Different colored data are for different α. (e) $\langle R_g \rangle$ of a chain in a cluster of size s for $\rho = 0.04$, $\alpha = 0.6$, at T^*1.0. Different colored data are for different salt concentration.

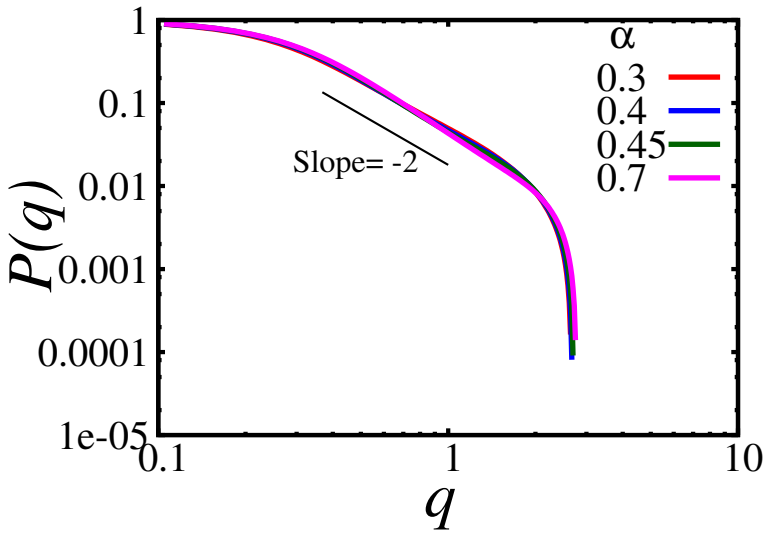


Figure 4.30. Form factor of a single polymer chain inside a complex for $\rho = 0.04$, at $T^* = 1.0$, and $c_s = 0$ M. Different colored curves are for the system of different charge density.

an aqueous solution of poly(styrene sulfonate) and poly(diallyldimethylammonium), Markarian et al. [119] showed that the labeled chains follow Gaussian conformation and the radius of gyration of the labeled chains were insensitive to addition of salt up to 1.5 M NaCl. For the same system with a different salt (KBr instead of NaCl), they behaved differently where labeled chains followed statistical random walk model and its radius of gyration was sensitive to salt above certain concentration [113]. According to them, the system retained its solid complex form upto about 1.2 M KBr, beyond which it took liquid like form as evident from the decrease in radius of gyration of labeled chains with further increase in salt concentration. We simulated symmetric oppositely charged polyelectrolytes in semidilute polymer solutions and for strong electrostatic interactions (Γ=2.8) the labeled chains take Gaussian conformation, qualitatively agreeing with the equilibrium behavior reported for solid like complexes [89, 119]. Similar to the observations of Markarian et al. [119] we also find that the $\langle R_g \rangle$ was insensitive to salt, although the aggregate structures respond to salt addition by opening up more as in Fig. 4.28.

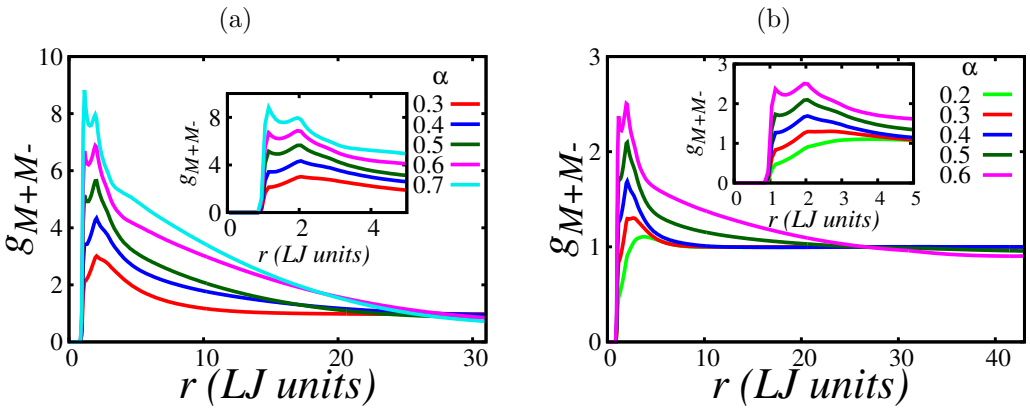


Figure 4.31. Radial distribution function between monomers of polyanions and polycations for (a)$\rho = 0.04$, at $T^* = 1.0$ and (b) $\rho = 0.1$ at $T^* = 1.0$ (here, $dr = 0.1$). Different colored plots represent systems for different α and no salt.

4.3.3.6 Monomer arrangement in polyelectrolyte complexes

Finally, we focused on the smallest length scale of the system: monomer level. We compute the radial distribution function ($g(r)$) as given by [120],

$$g(r) = \lim_{dr \to 0} \frac{p(r)}{4\pi(N_{pairs}/V)r^2 dr} \tag{4.13}$$

where, r is the distance between pair particles, $p(r)$ is the average number of atom pairs found at distance between r and $r+dr$, N_{pairs} is number of unique pairs between two types of atom selections.

We compute the $g(r)$ from 200 time frames after the system is in equilibrium. In our case,

$$p(r) = \frac{1}{N_{frame}} \sum_{i=1}^{N_{frame}} \sum_{j=selection1} \sum_{k=selection2} \delta(r - r_{jk}) \tag{4.14}$$

where, N_{frame} is the total number of frames, r_{jk} is the distance between atom j and atom k at frame i.

Fig. 4.31 shows the radial distribution functions (RDFs) between monomers of polyanions and polycations ($g_{M_+M_-}$) for the system with different ρ, α and at

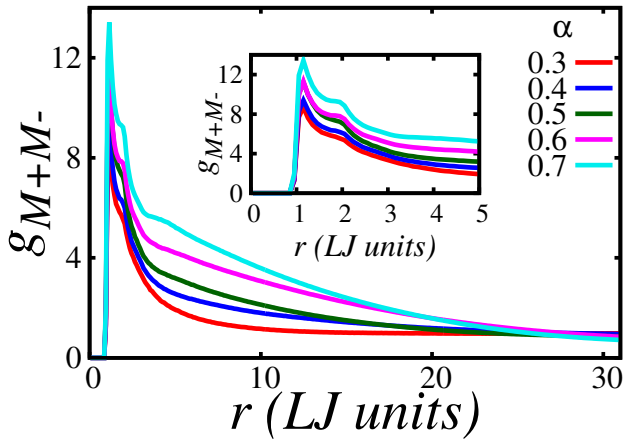


Figure 4.32. Radial distribution function between charged monomers of polyanions and polycations for $\rho = 0.04$ at $T^* = 1.0$ (here, $dr = 0.1$). Different colored plots represent systems for different α and no salt.

$T^* = 1.0$. In general, $g_{M_+M_-}$ shows a well defined correlation peak at 2σ, (which measures the typical nearest neighbor monomer-monomer separation in the system). Another peak at 1.12σ emerges and is more prominent with increase in α of the system corresponding to the association of oppositely charged monomers of polyanions and polycations.

To look into this more closely, we calculate the $g_{M_+M_-}$ between charged beads of monomers of polyanions and polycations as shown in Fig. 4.32. The Fig. shows peaks at 1.12σ corresponding to the association of oppositely charged monomer beads.

We determined the effect of T^*, ρ, and c_s on the RDF of oppositely charged monomers (g_{M+M-}) and charged polycation monomer and its counterion (g_{M+I-}). The RDF for the charged polyanion monomer and it's counterion was the same as g_{M+I-} and g_{M+M-}. Moreover, RDF was insensitive to ρ as shown in Fig. 4.33.

Effect of T^* on g_{M+M-} in Fig. 4.34(a) shows a large peak at $r = 1.12$ LJ units and a smaller peak at $r = 2$ LJ units; presence of the long-range correlation peak (second peak) indicates long-range ordering in the assembled structure. The g_{M+M-} decays to unity over a length scale of 25 LJ units. The large peak value decreased from

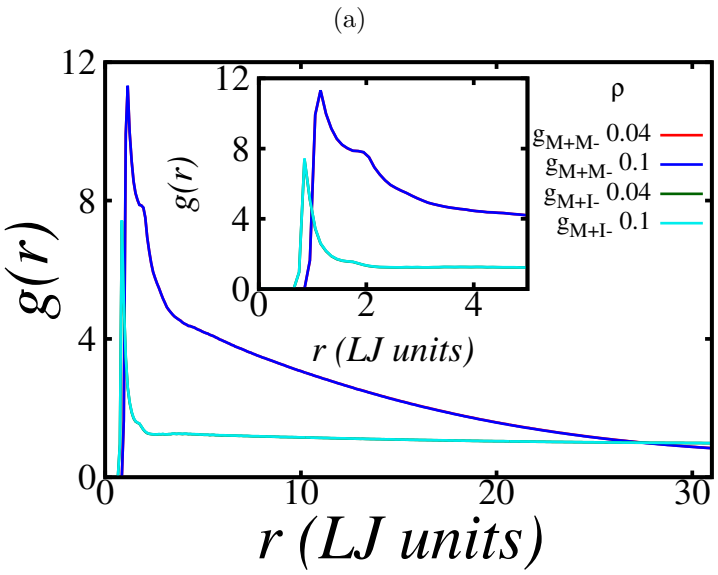


Figure 4.33. Comparison of g_{M+M-} and g_{M+I-} for $\alpha = 0.6$ at $T^* = 1.0$ for different monomer density ρ and no salt.

12 to 9 with increase in T^* from 0.95 to 1.2, indicating that the charged monomer which was closely surrounded by 12 oppositely charged monomers (relative to bulk) at T^*=0.95 is closely surrounded by 9 oppositely charged monomers at T^*=1.2. This is expected, as we increase the temperature, we are effectively reducing the electrostatic interaction strength (reduced Bjerrum length) leading to less electrostatic associations. Next, we determined the effect of salt on g_{M+M-} as shown in Fig. 4.34(b). Increasing salt resulted in a qualitatively similar trend as increasing temperature, but the quantitative effect is more pronounced with reasonable amount of salt addition (3M monovalent salt) compared to the reasonable limits of temperature change for an aqueous medium (ranges between 283 K to 358 K in real units in our simulation). The peak height decreased from 11 to 6 with increase in salt concentration from 0M to 3M monovalent salt, and g_{M+M-} decayed to unity at a smaller r for higher salt concentration. g_{M+M-} reached unity at about half the r for 3M salt as compared to 0M salt. This shows that the monomer distribution becomes more uniform with increase in salt concentration, which is consistent with the aggregate size distribution

in response to salt concentration in Fig. 4.28.

Next, we followed g_{M+I-} as a function of radial distance r for different T^* as shown in Fig. 4.34(c). It showed two peaks, one at 1 LJ units and another at about 1.7 LJ units (Note: The σ for monomers is 1.0 LJ units but it is 0.6 LJ units for the ions). g_{M+I-} decayed to unity at r =2, which is an order of magnitude smaller compared to r at which g_{M+M-} reached unity. This shows that the ions are more uniformly distributed as compared to the monomers, although a small but noticeable trend in g_{M+I-} was observed as a function of T^*. Increasing T^* resulted in a slight increase of the peak heights, which is opposite to the qualitative trend observed for g_{M+M-} dependence on T^*. With an increase in T^* the association between oppositely charged monomers weakened, leading to slightly more access of counterions to these monomers. Qualitatively similar features were observed for the effect of c_s on g_{M+I-} as shown in Fig. 4.34(d). As salt concentration is increased from 0M to 1M, the peak height decreased from 7.5 to 5 and remained at 5 with further increase in c_s. As the bulk salt concentration is increased, the number of counterions surrounding a charged polycation monomer increased too, such that the ratio of the number of the counterions surrounding the monomer is five times that of the bulk. As more counterions surround a charged monomer, its electrostatic association with oppositely charged monomer is reduced as evident from Fig. 4.34(b).

We compared our RDF with that reported for flexible-flexible oppositely charged polyelectrolyte complexes by Shakya et al. [118]. Although they studied dilute polymer solution for a very asymmetric system ($N_+ = 240$ and $N_- = 22$) but qualitative features at monomer level can be compared. There are some similarities between our g_{M+M-} and theirs and some dissimilarities too. Their peak height was about three times higher than our peak value. The large difference between the peak height can be explained in terms of number of charged monomers in the system. Corresponding

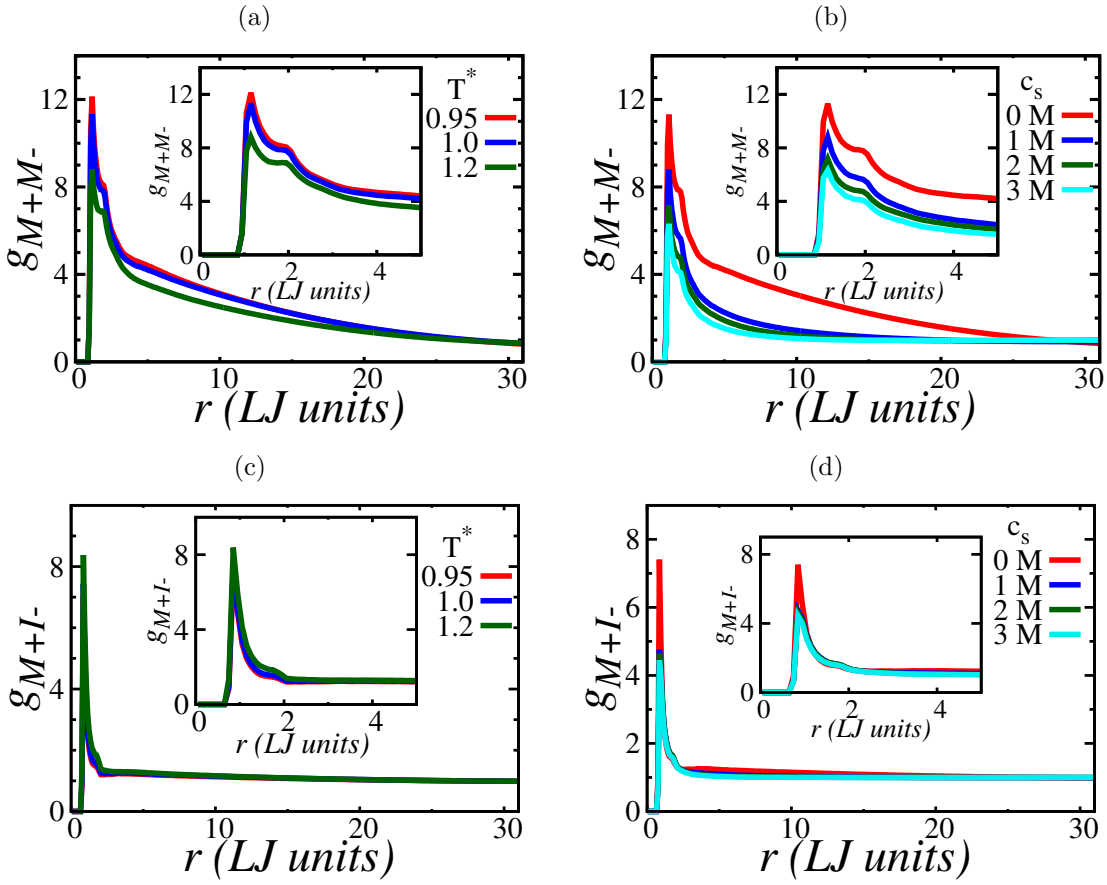


Figure 4.34. Effect of T^* and c_s on the pairwise radial distribution function of charged monomers (g_{M+M-}), and charged monomer and it's counterion (g_{M+I-}) for $\rho = 0.04$, $\alpha = 0.6$. (a) g_{M+M-} for $c_s = 0$ M, and different T^*, (b) g_{M+M-} for $T^* = 1.0$ and different c_s.

to Fig. 4.34, only 0.6 fraction of the monomers are charged but in their system all the monomers are charged (we are showing correlation between charged monomers only). They observed two peaks only for salt free case, with addition of salt as little as 25 mM the long-range correlation peak vanished. We observed two peaks for salt concentrations as high as 3M, because their electrostatic interaction strength parameter, which is measured as the ratio of Bjerrum length over σ is 0.7 is 4 times weaker than ours. Their g_{M+M-} decayed to unity at a much larger length scale (an order of magnitude larger than their monomer size) similar to our observation for salt-free case. With an increase in salt from 0 to 150 mM, the peak of the g_{M+M-} peak decreased by a quarter and g_{M+M-} decayed to unity over a length scale of 10 nm, in contrast to 40 nm for the salt free case where we observed qualitatively similar behavior. Lastly, they also monitored g_{M+I-} as a function of r, which showed two peaks at 1 nm and 2 nm; decayed to unity over a very small length scale (an order of magnitude smaller than that observed for salt-free g_{M+M-}) similar to our observation. Quantitatively there were differences in the peak heights, their peak height was five times smaller than ours in salt-free case, due to weaker electrostatic associations. In summary, we observed many qualitatively similar features in RDF as reported by Shakya et al. [118], but stronger electrostatic interaction strength in our system introduces more ordering in the structure and offers more immunity against salt.

4.4 Conclusion

In this project, we have studied structures and dynamics of complexes of flexible and symmetric polyelectrolytes of opposite charges by employing coarse-grained Langevin dynamics simulations. We have studied the role of polyelectrolyte concentration, charge density, temperature, and salt concentration on the structure and dynamics of complexes.

For systems of high charge density, The results show that the $\langle R_g \rangle$ of a labeled chain inside a complex and the size-scaling exponent (ν) of a single polyelectrolyte chain inside a complex varies non-monotonically with the polymer concentration in the complex. $\langle R_g \rangle$ and ν of a single isolated chain is maximum, and they shrink when two chains of opposite charge come together, forming a complex. On increasing the polymer concentration, $\langle R_g \rangle$ saturates for higher polymer concentration. The value of ν reaches a plateau of 0.5, showing that the labeled chain in the complex shows Gaussian-like chain conformation once the charged polymers form an aggregate of a reasonable size.

Moreover, in semi-dilute solutions of polyelectrolyte complexes, we observed that the formation of aggregates of polyelectrolytes of opposite charges is enhanced with an increase in charge density and decrease in temperature. In addition, the formation of complexes of such polymers is decreased with an increase in temperature and salt concentration. Although drastic changes in the large length scale aggregate structure were observed, the constituent polymers were insensitive to these parameters. The calculation of $\langle R_g \rangle$ and form factors indicate that the chains behave Gaussian-like conformation irrespective of charge density of polymers, temperature, salt concentration, and monomer density. The structural changes in the aggregate were reflected in the lowest length scale (monomers), as determined from their radial distribution function(RDF). The strong association between oppositely charged monomers can be reduced by increasing temperature or adding salt, but there was no effect of polymer concentration on the RDF.

Finally, we determined the dynamics of polymer chains by monitoring the mean square displacement (MSD) of the center of mass of a labeled chain versus time (t). An isolated polyelectrolyte chain shows diffusive behavior ($MSD \sim t$), but the

labeled chains in the complexes are slowed down significantly, and the chain follows non-diffusive dynamical law as the chain becomes a part of the physical network due to the presence of other chains in the complex.

For further investigation in this topics, one possible direction would be to explore the role of branched structures and asymmetries, such as in polycation-polyanion chain length and charge density, in the equilibrium and dynamic behaviors of polyelectrolyte complexes. In addition, it would be interesting to explore possible transition from a solution of oppositely charged polymers to a percolated structure.

CHAPTER 5

CONCLUSIONS

In this thesis, we studied the equilibrium and dynamic behaviors of charged polymers in biological and synthetic environments by employing computational modeling to simulate complex environments consisting of charged polymers, solvents, and charged ions. Below, we summarize our main findings and discuss the possible directions for future research.

In Chapter 2, we calculated and compared the electrostatic potential across a α-hemolysin protein nanopore by using atomistic and coarse-grained methods. We showed that the coarse-grained approach gives a good approximation of the atomistic method for the problems under consideration. Since all-atom molecular dynamics simulations are computationally expensive and time-consuming for the available time and facility. Therefore, we chose a coarse-grained approach in other projects.

In Chapter 3, we studied the unfolding kinetics of five different architectures of simple RNA hairpins of equal numbers of nucleotides and base pairs using a simple coarse-grained model of homo-nucleotide RNA hairpins. We observed that the threshold voltage of unfolding depends on the length of the hairpin attached to the tail. The threshold voltage is higher for the models with longer hairpins attached to the tail. The qualitative behavior of fraction of unfolding as a function of the applied voltage is consistent with the single-molecule experiment on mechanical pulling of RNA hairpin using optical tweezers [42]. We showed a distinct signature of base unfolding time for the bases before and after the unpaired bases in the RNA hairpin models

considered in the project. The above-discussed structure dependence of translocation kinetics behavior of different RNA hairpins considered in this study, although obtained from simple models, demonstrates the potential use of translocation to predict the secondary structures of polynucleotides. In the future, more complex models of secondary and tertiary structures of polynucleotides and different geometries of pores can be used to understand the role of RNA architectures on the dynamics of translocation through nanopores and to predict the secondary and tertiary structures of polynucleotides.

In Chapter 4 we studied the role of polymer concentration, charge density, temperature, and salt concentration on the structure and dynamics of complexes of flexible and symmetric polyelectrolytes of opposite charges by employing coarse-grained Langevin dynamics simulations. The underlying physics of charged macromolecules are determined by the interplay of entropy and enthalpy from long-ranged electrostatic interactions between charged polymers, counterions, salt ions, short-ranged Van der Waals interactions. In semi-dilute solutions of polyelectrolyte complexes, we observed that the formation of aggregates of polyelectrolytes of opposite charges is enhanced with an increase in charge density of polymers and decreased temperature. The aggregates are de-complexed with the increase in salt concentration. We also studied the structure and dynamics of a labeled chain inside a complex of highly charged polyelectrolytes at different concentrations. The results show that the radius of gyration of a labeled chain ($\langle R_g \rangle$) and size-scaling exponent (ν) of a single isolated chain are maximum, and they shrink when two chains of opposite charge come together, forming a complex. On increasing the polymer concentration, $\langle R_g \rangle$ saturates for higher polymer concentration. The value of ν reaches a plateau of 0.5, showing that the labeled chain on the complex shows Gaussian-like chain conformation once a complex of reasonable size is formed. We also observed that in a semi-dilute

regime of polyelectrolyte complexes $\langle R_g \rangle$ and ν of a labeled chain is independent of the size of the complex formed, the charge density of the polymer, temperature, and salt concentration, where chains show Gaussian-like conformations. We also observed that an isolated polyelectrolyte chain shows diffusive behavior but the labeled chain in the complexes follows non-diffusive dynamical law as the chain becomes a part of physical network due to presence of other chains in the complex. At much larger time-scales, we expect diffusive behavior. Most of the observations in this study are in agreement with the experiments and contribute to a better understanding of the experimentally observed phenomena. For further investigation in this topics, one possible direction would be to explore the role of branched structures and asymmetries, such as in polycation-polyanion chain length and charge density, in the equilibrium and dynamic behaviors of polyelectrolyte complexes. In addition, it would be interesting to explore possible transition from a solution of oppositely charged polymers to a percolated structure.

APPENDIX

A.1 Basics of Langevin dynamics

Langevin equation used in LAMMPS is given by

$$F = F_c + F_f + F_r \tag{A.1}$$

where, F_c is conservative force, F_f is frictional drag force, and F_r is random force.

In LAMMPS, for a i^{th} particle,

$$F_c = -\nabla U_i \tag{A.2}$$

$$F_f = -\frac{m_i}{\gamma_d} v_i \tag{A.3}$$

$$F_r = \sqrt{\frac{k_B T}{dt} \frac{m_i}{\gamma_d}} \xi(t) \tag{A.4}$$

where, U_i is the total interaction potential acting on the i^th bead, m_i is the mass, γ_d is a damp parameter in LAMMPS language, v_i is the velocity of i^{th} particle, $\xi(t)$ is unitless random distribution.

For our system,

$$U = U_{LJ} + U_b + U_c \tag{A.5}$$

where, U_{LJ} is Lennard Jones potential, U_b is bond potential and U_c is long-range electrostatic interaction.

To model excluded volume interactions, a truncated Lennard Jones (LJ) potential,

$$\begin{aligned} U_{LJ} &= 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12}-\left(\frac{\sigma}{r}\right)^{6}\right]-4\epsilon\left[\left(\frac{\sigma}{r_c}\right)^{12}-\left(\frac{\sigma}{r_c}\right)^{6}\right] \quad for \quad r \le r_c \\ &= 0 \qquad for \quad r > r_c \end{aligned} \tag{A.6}$$

is used where ϵ is the interaction strength and the potential is truncated at $r = r_c = 1.12\sigma$. The value of σ is taken to be $1.0l_0$, $0.8l_0$, and $0.6l_0$ for beads representing polymer-polymer, polymer-ion and ion-ion pairs respectively, where l_0 is the equilibrium bond length.

The bond potential between connected beads (i and i+1) is represented by a harmonic potential,

$$U_b = K_b(l - l_0)^2 \tag{A.7}$$

where K_b is the spring constant, l is the bond length and l_0 is the equilibrium bond length.

The long-range electrostatic interaction between i^{th} bead carrying a charge of q_i and j^{th} bead carrying a charge q_j separated by distance r is modeled by using the Coulomb potential as

$$U_c = \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_r r} \tag{A.8}$$

where ε_0 is permittivity of free space and ε_r is the relative dielectric constant of the medium.

Additionally, we define the electrostatic strength parameter between beads as,

$$\Gamma = \frac{l_B}{l_0}, \tag{A.9}$$

where l_B is the Bjerrum length defined by,

$$l_B = \frac{e^2}{4\pi\varepsilon_0\varepsilon_r k_B T}, \tag{A.10}$$

where $k_B T$ is the Boltzmann constant times the absolute temperature T.

Let's write A.1 in its full form as,

$$m_i \frac{d^2 r_i}{dt^2} = -\nabla U_i - \frac{m_i}{\gamma_d}\frac{dr_i}{dt} + \sqrt{\frac{k_B T}{dt}\frac{m_i}{\gamma_d}}\xi(t) \tag{A.11}$$

What are the characteristic distance, time, etc of the system?

distance = σ

mass = m

Dimensions: $\frac{1}{\gamma_d} velocity = acceleration$

γ_d is something which has to be measured experimentally. Here, In LAMMPS, this is one of the parameter and has dimension of time.

Now, lets make equation A.11 dimensionless

Let, $\widetilde{m} = m/m_s$, $\widetilde{r} = r/\sigma$, $\widetilde{t} = t/\tau$, $\widetilde{q} = q_i/q_s$, $\widetilde{l} = l/\sigma$, where, m_s, σ, τ, and q_s are scaled mass, scaled distance, characteristic time, and characteristic charge respectively. We also define $\widetilde{\gamma_d} = \gamma_d/\tau$, as the unitless damp parameter, $\widetilde{T} = k_B T/\epsilon$, is the reduced temperature and $\tau = \sqrt{\frac{m_s\sigma^2}{\epsilon}}$. Therefore, we get,

$$\widetilde{m_i}\frac{d^2\widetilde{r_i}}{d\widetilde{t}^2} = -(\frac{\tau^2}{m_s\sigma}\frac{\epsilon}{\sigma})\widetilde{\nabla}(\widetilde{U_i})_{LJ} - (\frac{\tau^2}{m_s\sigma}\frac{\epsilon}{\sigma})\widetilde{\nabla}(\widetilde{U_i})_b - (\frac{\tau^2}{m_s\sigma}\frac{q_s^2}{4\pi\varepsilon_0\varepsilon_r\sigma}\frac{1}{\sigma})\widetilde{\nabla}(\widetilde{U_i})_c - \frac{\widetilde{m}_i}{\widetilde{\gamma}_d}\frac{d\widetilde{r_i}}{d\widetilde{t}} + (\frac{\tau^2}{m_s\sigma}\frac{\sqrt{m_s\epsilon}}{\tau})\sqrt{\frac{\widetilde{T}}{d\widetilde{t}}\frac{\widetilde{m}_i}{\widetilde{\gamma}_d}}\xi(t) \tag{A.12}$$

Therefore, the following terms in equation A.12 are dimensionless,

- $\frac{\tau^2}{m_s\sigma}\frac{\epsilon}{\sigma}$ where, $\tau = \sqrt{m_s\sigma^2/\epsilon}$
- $\frac{\tau^2}{m_s\sigma}\frac{\epsilon}{\sigma}$
- $\frac{\tau^2}{m_s\sigma^2}\frac{q_s^2}{4\pi\varepsilon_0\varepsilon_r\sigma} \rightarrow$ where, $q_s = \sqrt{4\pi\varepsilon_0\epsilon\sigma}$
- $\frac{\tau^2}{m_s\sigma}\frac{\sqrt{m_s\epsilon}}{\tau}$

Therefore equation A.12 becomes,

$$\widetilde{m_i}\frac{d^2\widetilde{r_i}}{d\widetilde{t}^2} = -\widetilde{\nabla}(\widetilde{U_i})_{LJ} - \widetilde{\nabla}(\widetilde{U_i})_b - (\frac{1}{\varepsilon_r})\widetilde{\nabla}(\widetilde{U_i})_c - \frac{\widetilde{m}_i}{\widetilde{\gamma}_d}\frac{d\widetilde{r_i}}{d\widetilde{t}} + \sqrt{\frac{\widetilde{T}}{d\widetilde{t}}\frac{\widetilde{m}_i}{\widetilde{\gamma}_d}}\xi(t) \tag{A.13}$$

So, for our system in LAMMPS, we need to assign, m_s, σ, ϵ, $\widetilde{\gamma}_d = \gamma_d/\tau$, $\widetilde{T} = k_BT/\epsilon$, and $\widetilde{q} = \frac{q_i}{q_s}$.

BIBLIOGRAPHY

[1] A. Bertin, Polyelectrolyte complexes of dna and polycations as gene delivery vectors, in *Polyelectrolyte Complexes in the Dispersed and Solid State II* (Springer, 2013), pp. 103–195.

[2] K. Roy, H.-Q. Mao, S.-K. Huang, and K. W. Leong, Oral gene delivery with chitosan–dna nanoparticles generates immunologic protection in a murine model of peanut allergy, *Nature medicine* **5**(4), 387–391 (1999).

[3] I. Genta, P. Perugini, T. Modena, F. Pavanetto, F. Castelli, R. Muzzarelli, C. Muzzarelli, and B. Conti, Miconazole-loaded 6-oxychitin–chitosan microcapsules, *Carbohydrate polymers* **52**(1), 11–18 (2003).

[4] R. Zemaitaitiene, J. Barkauskaite, E. Mazoniene, and E. Zliobaite, Cationic polymers and their complexes in wastewater treatment, in *Chemical Water and Wastewater Treatment V* (Springer, 1998), pp. 163–172.

[5] C. Schmitt and S. L. Turgeon, Protein/polysaccharide complexes and coacervates in food systems, *Advances in colloid and interface science* **167**(1-2), 63–70 (2011).

[6] S. Turgeon, C. Schmitt, and C. Sanchez, Protein–polysaccharide complexes and coacervates, *Current Opinion in Colloid & Interface Science* **12**(4-5), 166–178 (2007).

[7] T. Pan and T. Sosnick, Rna folding during transcription, *Annual Review of Biophysics and Biomolecular Structure* **35**(1), 161–175 (2006).

[8] M. Kozak, Regulation of translation via mrna structure in prokaryotes and eukaryotes, *Gene* **361**, 13 – 37 (2005).

[9] P. Svoboda and A. D. Cara, Hairpin rna: a secondary structure of primary importance, *Cellular and Molecular Life Sciences CMLS* **63**(7), 901–908 (Apr 2006).

[10] J. S. Pedersen, G. Bejerano, A. Siepel, K. Rosenbloom, K. Lindblad-Toh, E. S. Lander, J. Kent, W. Miller, and D. Haussler, Identification and classification of conserved rna secondary structures in the human genome, *PLOS Computational Biology* **2**, 1–12 (04 2006).

[11] M. B. Warf and J. A. Berglund, Role of rna structure in regulating pre-mrna splicing, *Trends in Biochemical Sciences* **35**(3), 169 – 178 (2010).

[12] V. Ramakrishnan, Ribosome structure and the mechanism of translation, *Cell* **108**(4), 557 – 572 (2002).

[13] M. Muthukumar, 50th anniversary perspective: A perspective on polyelectrolyte solutions, *Macromolecules* **50**(24), 9528–9560 (2017).

[14] S. Srivastava and M. V. Tirrell, Polyelectrolyte complexation, *Adv. Chem. Phys* **161**, 499–544 (2016).

[15] C. E. Sing, Development of the modern theory of polymeric complex coacervation, *Advances in colloid and interface science* **239**, 2–16 (2017).

[16] C. E. Sing and S. L. Perry, Recent progress in the science of complex coacervation, *Soft Matter* (2020).

[17] J. Hasted, D. Ritson, and C. Collie, Dielectric properties of aqueous ionic solutions. parts i and ii, *The Journal of Chemical Physics* **16**(1), 1–21 (1948).

[18] E. Hückel and P. Debye, The theory of electrolytes: I. lowering of freezing point and related phenomena, *Phys. Z* **24**(185-206), 1 (1923).

[19] M. van Dijk and A. M. Bonvin, 3d-dart: a dna structure modelling server, *Nucleic acids research* **37**(suppl_2), W235–W239 (2009).

[20] H. Berman, K. Henrick, and H. Nakamura, Announcing the worldwide protein data bank, *Nature Structural & Molecular Biology* **10**(12), 980–980 (2003).

[21] W. Humphrey, A. Dalke, and K. Schulten, VMD – Visual Molecular Dynamics, *Journal of Molecular Graphics* **14**, 33–38 (1996).

[22] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, Scalable molecular dynamics with namd, *Journal of computational chemistry* **26**(16), 1781–1802 (2005).

[23] J. R. Comer, D. B. Wells, and A. Aksimentiev, Modeling nanopores for sequencing dna, in *DNA Nanotechnology* (Springer, 2011), pp. 317–358.

[24] P. F. Batcho, D. A. Case, and T. Schlick, Optimized particle-mesh ewald/multiple-time step integration for molecular dynamics simulations, *The Journal of Chemical Physics* **115**(9), 4003–4018 (2001).

[25] A. D. MacKerell Jr, D. Bashford, M. Bellott, R. L. Dunbrack Jr, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, *et al.*, All-atom empirical potential for molecular modeling and dynamics studies of proteins, *The journal of physical chemistry B* **102**(18), 3586–3616 (1998).

[26] A. Aksimentiev and K. Schulten, Imaging α-hemolysin with molecular dynamics: ionic conductance, osmotic permeability, and the electrostatic potential map, *Biophysical journal* **88**(6), 3745–3761 (2005).

[27] I. Jou and M. Muthukumar, Effects of nanopore charge decorations on the translocation dynamics of dna, *Biophysical journal* **113**(8), 1664–1672 (2017).

[28] G. Maglia, M. R. Restrepo, E. Mikhailova, and H. Bayley, Enhanced translocation of single dna molecules through α-hemolysin nanopores by manipulation of internal charge, *Proceedings of the National Academy of Sciences* **105**(50), 19720–19725 (2008).

[29] A. Blanco and G. Blanco, Chapter 6 - nucleic acids, in *Medical Biochemistry*, edited by A. Blanco and G. Blanco, (Academic Press, 2017), pp. 121 – 140.

[30] K. E. van Holde and J. Zlatanova, Chapter 9 - the central dogma, in *The Evolution of Molecular Biology*, edited by K. E. van Holde and J. Zlatanova, (Academic Press, 2018), pp. 87 – 94.

[31] M. Wanunu, Nanopores: A journey towards dna sequencing, *Physics of Life Reviews* **9**(2), 125 – 158 (2012).

[32] J. B. Heng, C. Ho, T. Kim, R. Timp, A. Aksimentiev, Y. V. Grinkova, S. Sligar, K. Schulten, and G. Timp, Sizing dna using a nanometer-diameter pore, *Biophysical Journal* **87**(4), 2905 – 2911 (2004).

[33] M. T. Woodside, W. M. Behnke-Parks, K. Larizadeh, K. Travers, D. Herschlag, and S. M. Block, Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins, *Proceedings of the National Academy of Sciences* **103**(16), 6190–6195 (2006).

[34] B. M. Venkatesan and R. Bashir, Nanopore sensors for nucleic acid analysis., *Nature nanotechnology* **6 10**, 615–24 (2011).

[35] E. A. Manrao, I. M. Derrington, A. H. Laszlo, K. W. Langford, M. K. Hopper, N. Gillgren, M. Pavlenok, M. Niederweis, and J. H. Gundlach, Reading dna at single-nucleotide resolution with a mutant mspa nanopore and phi29 dna polymerase, *Nature biotechnology* **30**(4), 349 (2012).

[36] A. M. Smith, M. Jain, L. Mulroney, D. R. Garalde, and M. Akeson, Reading canonical and modified nucleobases in 16s ribosomal rna using nanopore native rna sequencing, *PloS one* **14**(5), e0216709 (2019).

[37] C. Soneson, Y. Yao, A. Bratus-Neuenschwander, A. Patrignani, M. D. Robinson, and S. Hussain, A comprehensive examination of nanopore native rna sequencing for characterization of complex transcriptomes, *Nature Communications* **10**, 3359–3372 (2019).

[38] W. L. Xie, J. G. Chipman, D. L. Robertson, R. L. Erikson, and D. L. Simmons, Expression of a mitogen-responsive gene encoding prostaglandin synthase is regulated by mrna splicing., *Proceedings of the National Academy of Sciences* **88**(7), 2692–2696 (1991).

[39] D. P. Depledge, K. P. Srinivas, T. Sadaoka, D. Bready, Y. Mori, D. G. Placantonakis, I. Mohr, and A. C. Wilson, Direct rna sequencing on nanopore arrays redefines the transcriptional complexity of a viral pathogen, *Nature communications* **10**(1), 754 (2019).

[40] D. J. Lockhart and E. A. Winzeler, Genomics, gene expression and dna arrays, *Nature* **405**(6788), 827–36 (Jun 15 2000).

[41] N. Kono and K. Arakawa, Nanopore sequencing: Review of potential applications in functional genomics, *Development, growth & differentiation* **61**(5), 316–326 (2019).

[42] J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, and C. Bustamante, Reversible unfolding of single rna molecules by mechanical force, *Science* **292**(5517), 733–737 (2001).

[43] C. Hyeon and D. Thirumalai, Mechanical unfolding of rna hairpins, *Proceedings of the National Academy of Sciences* **102**(19), 6789–6794 (2005).

[44] C. Hyeon and D. Thirumalai, Forced-unfolding and force-quench refolding of rna hairpins, *Biophysical Journal* **90**(10), 3410 – 3427 (2006).

[45] F. Liu and Z. can Ou-Yang, Monte carlo simulation for single rna unfolding by force, *Biophysical Journal* **88**(1), 76 – 84 (2005).

[46] S. Harlepp, T. Marchal, J. Robert, J.-F. Léger, A. Xayaphoummine, H. Isambert, and D. Chatenay, Probing complex rna structures by mechanical force, *The European Physical Journal E* **12**(4), 605–615 (Nov 2003).

[47] Q. Zhao, J. Comer, V. Dimitrov, S. Yemenicioglu, A. Aksimentiev, and G. Timp, Stretching and unzipping nucleic acid hairpins using a synthetic nanopore, *Nucleic Acids Research* **36**(5), 1532–1541 (01 2008).

[48] J. Mathé, H. Visram, V. Viasnoff, Y. Rabin, and A. Meller, Nanopore unzipping of individual dna hairpin molecules, *Biophysical Journal* **87**(5), 3205 – 3212 (2004).

[49] B. McNally, M. Wanunu, and A. Meller, Electromechanical unzipping of individual dna molecules using synthetic sub-2 nm pores, *Nano Letters* **8**(10), 3418–3422 (2008).

[50] O. K. Dudko, J. Mathé, A. Szabo, A. Meller, and G. Hummer, Extracting kinetics from single-molecule force spectroscopy: Nanopore unzipping of dna hairpins, *Biophysical Journal* **92**(12), 4188 – 4195 (2007).

[51] A. F. Sauer-Budge, J. A. Nyamwanda, D. K. Lubensky, and D. Branton, Unzipping kinetics of double-stranded dna in a nanopore, *Phys. Rev. Lett.* **90**, 238101 (2003).

[52] A. Stachiewicz and A. Molski, Diffusive dynamics of dna unzipping in a nanopore, *Journal of computational chemistry* **37**(5), 467–476 (2016).

[53] X. Wang, Y. Li, T. Li, L. Liu, and H.-C. Wu, The effect of secondary structures on the generation of characteristic events during the translocation of dna hybrid through α-hemolysin, *Science China Chemistry* **59**(1), 135–141 (2016).

[54] M. van den Hout, I. D. Vilfan, S. Hage, and N. H. Dekker, Direct force measurements on double-stranded rna in solid-state nanopores, *Nano Letters* **10**(2), 701–707 (2010).

[55] U. Gerland, R. Bundschuh, and T. Hwa, Translocation of structured polynucleotides through nanopores, *Physical Biology* **1**(1), 19–26 (feb 2004).

[56] S. Schink, S. Renner, K. Alim, V. Arnaut, F. C. Simmel, and U. Gerland, Quantitative analysis of the nanopore translocation dynamics of simple structured polynucleotides, *Biophysical journal* **102**(1), 85–95 (2012).

[57] M. Muthukumar and C. Y. Kong, Simulation of polymer translocation through protein channels, *Proceedings of the National Academy of Sciences* **103**(14), 5273–5278 (2006).

[58] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, *Journal of Computational Physics* **117**(1), 1 – 19 (1995), http://lammps.sandia.gov.

[59] D. Klimov, M. Betancourt, and D. Thirumalai, Virtual atom representation of hydrogen bonds in minimal off-lattice models of α helices: effect on stability, cooperativity and kinetics, *Folding and Design* **3**(6), 481 – 496 (1998).

[60] J. T. G. Overbeek and M. Voorn, Phase separation in polyelectrolyte solutions. theory of complex coacervation, *Journal of Cellular and Comparative Physiology* **49**(S1), 7–26 (1957).

[61] I. Michaeli, J. T. G. Overbeek, and M. Voorn, Phase separation of polyelectrolyte solutions, *Journal of Polymer Science* **23**(103), 443–450 (1957).

[62] M. Muthukumar, Phase diagram of polyelectrolyte solutions: weak polymer effect, *Macromolecules* **35**(24), 9142–9145 (2002).

[63] C.-L. Lee and M. Muthukumar, Phase behavior of polyelectrolyte solutions with salt, *The Journal of chemical physics* **130**(2), 01B608 (2009).

[64] R. de Vries and M. C. Stuart, Theory and simulations of macroion complexation, *Current opinion in colloid & interface science* **11**(5), 295–301 (2006).

[65] S. L. Perry and C. E. Sing, Prism-based theory of complex coacervation: Excluded volume versus chain correlation, *Macromolecules* **48**(14), 5040–5053 (2015).

[66] B. de Jong, Coacervation, in *Proc. Royal Acad. Amsterdam* (1929), vol. 32, pp. 849–856.

[67] A. Veis and C. Aranyi, Phase separation in polyelectrolyte systems. i. complex coacervates of gelatin, *The Journal of Physical Chemistry* **64**(9), 1203–1210 (1960).

[68] E. Spruijt, A. H. Westphal, J. W. Borst, M. A. Cohen Stuart, and J. van der Gucht, Binodal compositions of polyelectrolyte complexes, *Macromolecules* **43**(15), 6476–6484 (2010).

[69] D. Priftis and M. Tirrell, Phase behaviour and complex coacervation of aqueous polypeptide solutions, *Soft Matter* **8**(36), 9396–9405 (2012).

[70] R. Chollakup, J. B. Beck, K. Dirnberger, M. Tirrell, and C. D. Eisenbach, Polyelectrolyte molecular weight and salt effects on the phase behavior and coacervation of aqueous solutions of poly (acrylic acid) sodium salt and poly (allylamine) hydrochloride, *Macromolecules* **46**(6), 2376–2390 (2013).

[71] Q. Wang and J. B. Schlenoff, The polyelectrolyte complex/coacervate continuum, *Macromolecules* **47**(9), 3108–3116 (2014).

[72] S. L. Perry, Y. Li, D. Priftis, L. Leon, and M. Tirrell, The effect of salt on the complex coacervation of vinyl polyelectrolytes, *Polymers* **6**(6), 1756–1772 (2014).

[73] P. K. Jha, P. S. Desai, J. Li, and R. G. Larson, ph and salt effects on the associative phase separation of oppositely charged polyelectrolytes, *Polymers* **6**(5), 1414–1436 (2014).

[74] M. Zhao, J. Zhou, C. Su, L. Niu, D. Liang, and B. Li, Complexation behavior of oppositely charged polyelectrolytes: Effect of charge distribution, *The Journal of Chemical Physics* **142**(20), 204902 (2015).

[75] J. Fu, H. M. Fares, and J. B. Schlenoff, Ion-pairing strength in polyelectrolyte complexes, *Macromolecules* **50**(3), 1066–1074 (2017).

[76] Z. Ou and M. Muthukumar, Entropy and enthalpy of polyelectrolyte complexation: Langevin dynamics simulations, *The Journal of Chemical Physics* **124**(15), 154902 (2006).

[77] M. Radhakrishna, K. Basu, Y. Liu, R. Shamsi, S. L. Perry, and C. E. Sing, Molecular connectivity and correlation effects on polymer coacervation, *Macromolecules* **50**(7), 3030–3037 (2017).

[78] J. Qin and J. J. de Pablo, Criticality and connectivity in macromolecular charge complexation, *Macromolecules* **49**(22), 8789–8800 (2016).

[79] M. Andreev, A. Chremos, J. de Pablo, and J. F. Douglas, Coarse-grained model of the dynamics of electrolyte solutions, *The Journal of Physical Chemistry B* **121**(34), 8195–8202 (2017).

[80] S. Adhikari, M. A. Leaf, and M. Muthukumar, Polyelectrolyte complex coacervation by electrostatic dipolar interactions, *The Journal of chemical physics* **149**(16), 163308 (2018).

[81] T. K. Lytle, L.-W. Chang, N. Markiewicz, S. L. Perry, and C. E. Sing, Designing electrostatic interactions via polyelectrolyte monomer sequence, *ACS Central Science* **5**(4), 709–718 (2019).

[82] J. J. Madinya, L.-W. Chang, S. L. Perry, and C. E. Sing, Sequence-dependent self-coacervation in high charge-density polyampholytes, *Molecular Systems Design & Engineering* (2020).

[83] S. Adhikari, V. M. Prabhu, and M. Muthukumar, Lower critical solution temperature behavior in polyelectrolyte complex coacervates, *Macromolecules* **52**(18), 6998–7004 (2019).

[84] J. Lee, Y. O. Popov, and G. H. Fredrickson, Complex coacervation: A field theoretic simulation study of polyelectrolyte complexation, *The Journal of chemical physics* **128**(22), 224908 (2008).

[85] A. Shovsky, I. Varga, R. Makuska, and P. M. Claesson, Formation and stability of water-soluble, molecular polyelectrolyte complexes: effects of charge density, mixing ratio, and polyelectrolyte concentration, *Langmuir* **25**(11), 6113–6121 (2009).

[86] L.-W. Chang, T. K. Lytle, M. Radhakrishna, J. J. Madinya, J. Vélez, C. E. Sing, and S. L. Perry, Sequence and entropy-based control of complex coacervates, *Nature communications* **8**(1), 1–8 (2017).

[87] Z. Dai and C. Wu, How does dna complex with polyethylenimine with different chain lengths and topologies in their aqueous solution mixtures?, *Macromolecules* **45**(10), 4346–4353 (2012).

[88] D. Priftis, X. Xia, K. O. Margossian, S. L. Perry, L. Leon, J. Qin, J. J. de Pablo, and M. Tirrell, Ternary, tunable polyelectrolyte complex fluids driven by complex coacervation, *Macromolecules* **47**(9), 3076–3085 (2014).

[89] E. Spruijt, F. A. Leermakers, R. Fokkink, R. Schweins, A. A. van Well, M. A. Cohen Stuart, and J. van der Gucht, Structure and dynamics of polyelectrolyte complex coacervates studied by scattering of neutrons, x-rays, and light, *Macromolecules* **46**(11), 4596–4605 (2013).

[90] M. Lemmers, I. K. Voets, M. A. C. Stuart, and J. van der Gucht, Transient network topology of interconnected polyelectrolyte complex micelles, *Soft Matter* **7**(4), 1378–1389 (2011).

[91] C. Narambuena, E. Leiva, M. Chávez-Páez, and E. Pérez, Effect of chain stiffness on the morphology of polyelectrolyte complexes. a monte carlo simulation study, *Polymer* **51**(14), 3293–3302 (2010).

[92] A. A. Lazutin, A. N. Semenov, and V. V. Vasilevskaya, Polyelectrolyte complexes consisting of macromolecules with varied stiffness: Computer simulation, *Macromolecular theory and simulations* **21**(5), 328–339 (2012).

[93] A. B. Marciel, S. Srivastava, and M. V. Tirrell, Structure and rheology of polyelectrolyte complex coacervates, *Soft Matter* **14**(13), 2454–2464 (2018).

[94] F. Weinbreck, H. S. Rollema, R. H. Tromp, and C. G. de Kruif, Diffusivity of whey protein and gum arabic in their coacervates, *Langmuir* **20**(15), 6389–6395 (2004).

[95] A. B. Kayitmazer, S. P. Strand, C. Tribet, W. Jaeger, and P. L. Dubin, Effect of polyelectrolyte structure on protein- polyelectrolyte coacervates: Coacervates of bovine serum albumin with poly (diallyldimethylammonium chloride) versus chitosan, *Biomacromolecules* **8**(11), 3568–3577 (2007).

[96] X. Wang, Y. Li, Y.-W. Wang, J. Lal, and Q. Huang, Microstructure of β-lactoglobulin/pectin coacervates studied by small-angle neutron scattering, *The Journal of Physical Chemistry B* **111**(3), 515–520 (2007).

[97] S. S. Singh, V. Aswal, and H. Bohidar, Structural evolution of aging agar-gelatin co-hydrogels, *Polymer* **50**(23), 5589–5597 (2009).

[98] S. Chodankar, V. Aswal, J. Kohlbrecher, R. Vavrin, and A. Wagh, Structural study of coacervation in protein-polyelectrolyte complexes, *Physical Review E* **78**(3), 031913 (2008).

[99] R. A. Ghostine, R. F. Shamoun, and J. B. Schlenoff, Doping and diffusion in an extruded saloplastic polyelectrolyte complex, *Macromolecules* **46**(10), 4089–4094 (2013).

[100] J. B. Schlenoff, M. Yang, Z. A. Digby, and Q. Wang, Ion content of polyelectrolyte complex coacervates and the donnan equilibrium, *Macromolecules* **52**(23), 9149–9159 (2019).

[101] S. Liu and M. Muthukumar, Langevin dynamics simulation of counterion distribution around isolated flexible polyelectrolyte chains, *The Journal of Chemical Physics* **116**(22), 9975–9982 (2002).

[102] S. Liu, K. Ghosh, and M. Muthukumar, Polyelectrolyte solutions with added salt: A simulation study, *The Journal of Chemical Physics* **119**(3), 1813–1823 (2003).

[103] B. Peng and M. Muthukumar, Modeling competitive substitution in a polyelectrolyte complex, *The Journal of Chemical Physics* **143**(24), 243133 (2015).

[104] R. W. Hockney and J. W. Eastwood, *Computer simulation using particles* (crc Press, 1988).

[105] M. Muthukumar, *Polymer translocation* (CRC press, 2011).

[106] A. Katchalsky, O. Künzle, and W. Kuhn, Behavior of polyvalent polymeric ions in solution, *Journal of Polymer Science* **5**(3), 283–300 (1950).

[107] M. Beer, M. Schmidt, and M. Muthukumar, The electrostatic expansion of linear polyelectrolytes: Effects of gegenions, co-ions, and hydrophobicity, *Macromolecules* **30**(26), 8375–8385 (1997).

[108] M. J. Stevens and K. Kremer, The nature of flexible linear polyelectrolytes in salt free solution: A molecular dynamics study, *The Journal of Chemical Physics* **103**(4), 1669–1690 (1995).

[109] M. Muthukumar, Double screening in polyelectrolyte solutions: Limiting laws and crossover formulas, *The Journal of Chemical Physics* **105**(12), 5183–5199 (1996).

[110] J. Fu and J. B. Schlenoff, Driving forces for oppositely charged polyion association in aqueous solutions: enthalpic, entropic, but not electrostatic, *Journal of the American Chemical Society* **138**(3), 980–990 (2016).

[111] R. G. Winkler, M. Gold, and P. Reineker, Collapse of polyelectrolyte macromolecules by counterion condensation and ion pair formation: a molecular dynamics simulation study, *Physical review letters* **80**(17), 3731 (1998).

[112] D. Srivastava and M. Muthukumar, Interpenetration of interacting polyelectrolytes, *Macromolecules* **27**(6), 1461–1465 (1994).

[113] H. M. Fares, Y. E. Ghoussoub, J. D. Delgado, J. Fu, V. S. Urban, and J. B. Schlenoff, Scattering neutrons along the polyelectrolyte complex/coacervate continuum, *Macromolecules* **51**(13), 4945–4955 (2018).

[114] N. Gavish and K. Promislow, Dependence of the dielectric constant of electrolyte solutions on ionic concentration: A microfield approach, *Phys. Rev. E* **94**, 012611 (Jul 2016).

[115] J. Huang, F. J. Morin, and J. E. Laaser, Charge-density-dominated phase behavior and viscoelasticity of polyelectrolyte complex coacervates, *Macromolecules* **52**(13), 4957–4967 (2019), ¡Go to ISI¿://WOS:000475408800017.

[116] A. B. Kayitmazer, A. F. Koksal, and E. K. Iyilik, Complex coacervation of hyaluronic acid and chitosan: effects of ph, ionic strength, charge density, chain length and the charge ratio, *Soft Matter* **11**(44), 8605–8612 (2015), ¡Go to ISI¿://WOS:000364824600006.

[117] F. Weinbreck, R. de Vries, P. Schrooyen, and C. G. de Kruif, Complex coacervation of whey proteins and gum arabic, *Biomacromolecules* **4**(2), 293–303 (2003), ¡Go to ISI¿://WOS:000181496300014.

[118] A. Shakya, M. Girard, J. T. King, and M. O. de la Cruz, Role of chain flexibility in asymmetric polyelectrolyte complexation in salt solutions, *Macromolecules* **53**(4), 1258–1269 (2020), ¡Go to ISI¿://WOS:000517351100016.

[119] M. Z. Markarian, H. H. Hariri, A. Reisch, V. S. Urban, and J. B. Schlenoff, A small-angle neutron scattering study of the equilibrium conformation of polyelectrolytes in stoichiometric saloplastic polyelectrolyte complexes, *Macromolecules* **45**(2), 1016–1024 (2012).

[120] B. G. Levine, J. E. Stone, and A. Kohlmeyer, Fast analysis of molecular dynamics trajectories with graphics processing units—radial distribution function histogramming, *Journal of computational physics* **230**(9), 3556–3569 (2011).