11-20-2019

Modeling and Simulations of Peptoids

Pu Du
Louisiana State University and Agricultural and Mechanical College

MODELING AND SIMULATIONS OF PEPTOIDS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Chemistry

by
Pu Du
B.S., Nicholls State University, 2014
May 2020

To my mother Hong Jiang, who made my world magical
I miss you so much.

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere gratitude to my advisor, Prof. Revati Kumar for her motivation, passion, enthusiasm, inspiration, and support over the last five years. I am deeply indebted for having the opportunity to be a student under her advising. I could not thank her enough for being a fantastic mentor, advisor, and friend during my Ph.D. study.

My sincere gratitude extends to my committee members, Prof. Donghui Zhang, Prof. Bin Chen, and Prof. Jeffery Blackmon for their valuable suggestions, insightful comments, constructive discussions, and support as my committee members.

I am also grateful for Kumar research group members, Dr. Kiara Taylor-Edinbyrd, Dr. Caitlin Bresnahan, Visal Subasinghege Don, and Ke Li, for working together and good friendship. Also, I thank collaborators I have been working with, Prof. Isiah Warner, Prof. Steven Rick, Prof. Rendy Kartika, Prof. Semin Lee, Prof. Daniel Kuroda, and Dr. Xin Li, for providing me great opportunities to work on diverse and exciting research projects.

Last but not least, I would like to express my sincere appreciation to my parents, for bringing me into this world and giving me their unconditional love. A special thank you to my fiancée, Mi. Without her love and support, this dissertation would not have been possible. She is the lighthouse in the night, guiding me home safely.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Polypeptoids, or poly-N-substituted glycines, are a class of sequence defined polymers that are structural mimics of polypeptides. Polypeptoids currently have received a growing interest due to their improved thermal stability, larger chemical diversity, and easier synthetic pathways as compared to peptides. Their lack of backbone hydrogen bonding and stereochemistry coupled with their easily tunability make them an ideal prototypical model system to study the effect of secondary/non-covalent interactions on self-assembly in solution. In order to develop a molecular level understanding of the effect of secondary interactions on polypeptoid self-assembly, systematic studies were carried out using molecular dynamics simulations on several polypeptoid solutions.

Firstly, atomistic molecular dynamics simulations were performed to study self-aggregation of a cyclic polypeptoid and its linear analog in a low dielectric solvent, namely methanol. The cyclic polypeptoids bearing zwitterionic end-groups showed small cluster formation experimentally, ranging from a single polypeptoid chain to small oligomers in methanol solution, while the experiments showed no aggregation in the linear case. Atomistic molecular simulation results revealed that the aggregation in the cyclic case was as results of several different interactions. During the initial approach of the two cyclic polypeptoids, the attractive dipole-dipole interaction dominates. At closer distance, the attractive solvophobic effect takes over, while the effective repulsive interaction resulting from solvation of the dipoles dramatically reduces the attractive dipole-dipole component.

A second set of studies was carried out to investigate the micelle formation of sequence-defined singly charged ionic peptoid block copolymers consisting of a hydrophobic and a hydrophilic segment in aqueous solution. Results from this study demonstrated that water molecules mediate the structure and shape of micelles by interacting with the polymer chains and ionic monomers on the backbone. Specifically, a key interaction is that of the charged moiety with water that results in the charged group exposed to the solvent, even when it is placed next to the hydrophobic segment of the polypeptoid chain. This study contributed to our understanding of structure-property relationships in peptoid based soft matter systems.

Finally, a coarse-grained (CG) model of N, N-dimethylacetamide, the backbone of the polypeptoid system, has been developed that can reproduce the solvation environment around the DMA molecule, as observed from the reference all-atom simulations. These results suggest a promising approach to develop CG models of complex peptoid with various side chains that will enable better sampling and longer simulation length scales.

CHAPTER 1
INTRODUCTION

1.1 Polypeptoids

Polypeptoids, or poly-N-substituted glycines, are a class of polymers that are structural mimics of polypeptides with side chains attached to nitrogen atoms instead of α-carbons. Chemical structures of polypeptides and polypeptoids are illustrated in Figure 1.1. Due to their ease of synthesis, stability, and structural similarity to polypeptides, polypeptoids are perfectly suitable for combinatorial approaches to self-assembly, nanoscience, protein mimicry, and drug discovery.[1–10]

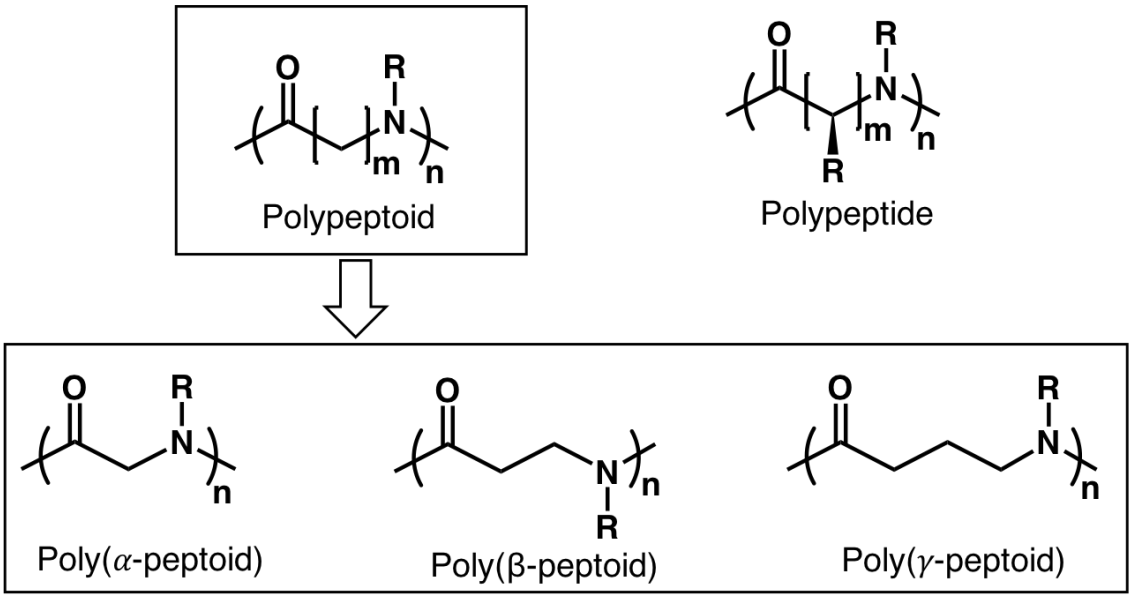

Figure 1.1. Chemical structures of polypeptides and polypeptoids.

From a synthetic perspective, polypeptoids are particularly appealing since the chemical sequence can be controlled resulting in a model system that allows for great tunability of non-covalent secondary interactions. These interactions include electrostatics, van der Waals interactions, hydrogen bonding, as well as hydrophobicity and hydrophilicity. Two fundamentally different approaches have been commonly used for synthesis of polypeptoids including solid-phase synthesis and solution polymerization.[11] An outline of these two techniques are shown in Figure 1.2. Similar to solid-phase peptide synthesis, peptoid synthesis uses a stepwise monomer addition cycle, which involves the use of primary amines (R-NH_2) with side group R (Figure 1.2a).[2,12] The

key advantage of solid-phase synthesis approach is that the sequence definition of a polymer can be precisely controlled. Using such synthesis approach can yield the degree of polymerization up to fifty monomeric units.[11] On the other hand, solution polymerization can be used to access those polypeptoid with higher molecular weight (Figure 1.2b). Although less sequence control is available with this approach, a large array of commercially available primary amines can be incorporated, providing access to a large diversity in terms of primary structure.

Aggregation or self-assembly of block copolymers in solution is key for a wide range of applications including biomedical uses.[13–18] Decoupling the effect of different secondary/non-covalent interactions (electrostatic, van der Waals, hydrophobic effect etc.) on self-assembly can be challenging given the complexity in macromolecular systems. The lack of stereochemistry and hydrogen bonding in the backbone of peptoids combined with the programmable synthesis and tunability of their side chains makes them an ideal platform to study the effect of different side group chemistries and hence secondary interactions on self-assembly.[19–27] The focus of this dissertation is to use molecular simulations to gain insight into the molecular mechanisms that govern self-aggregation in polypeptoids. A brief introduction to molecular simulations is presented in the proceeding sections.

1.2 Molecular Dynamics Simulation

Molecular simulation, specifically molecular dynamics, is a powerful technique for molecular level investigations of the structural and dynamic properties of soft matter systems including macromolecular systems. The growing importance of molecular dynamics simulations in modern research efforts lead to the 2013 Nobel Prize in Chemistry for Martin Karplus, Michael Levitt, and Ariel Warshel "for the development of multiscale models for complex chemical systems." In Figure 1.3, the number of publications per year that include molecular dynamics simulations shows an exponential increase from a few in 1972 to more than 12000 in 2018. It is clear that molecular dynamics simulations have become substantially more popular and visible in recent years. There are countless articles in the literature where molecular dynamics simulations provide the underlying molecular interpretation to observed physicochemical properties.[28] For example,

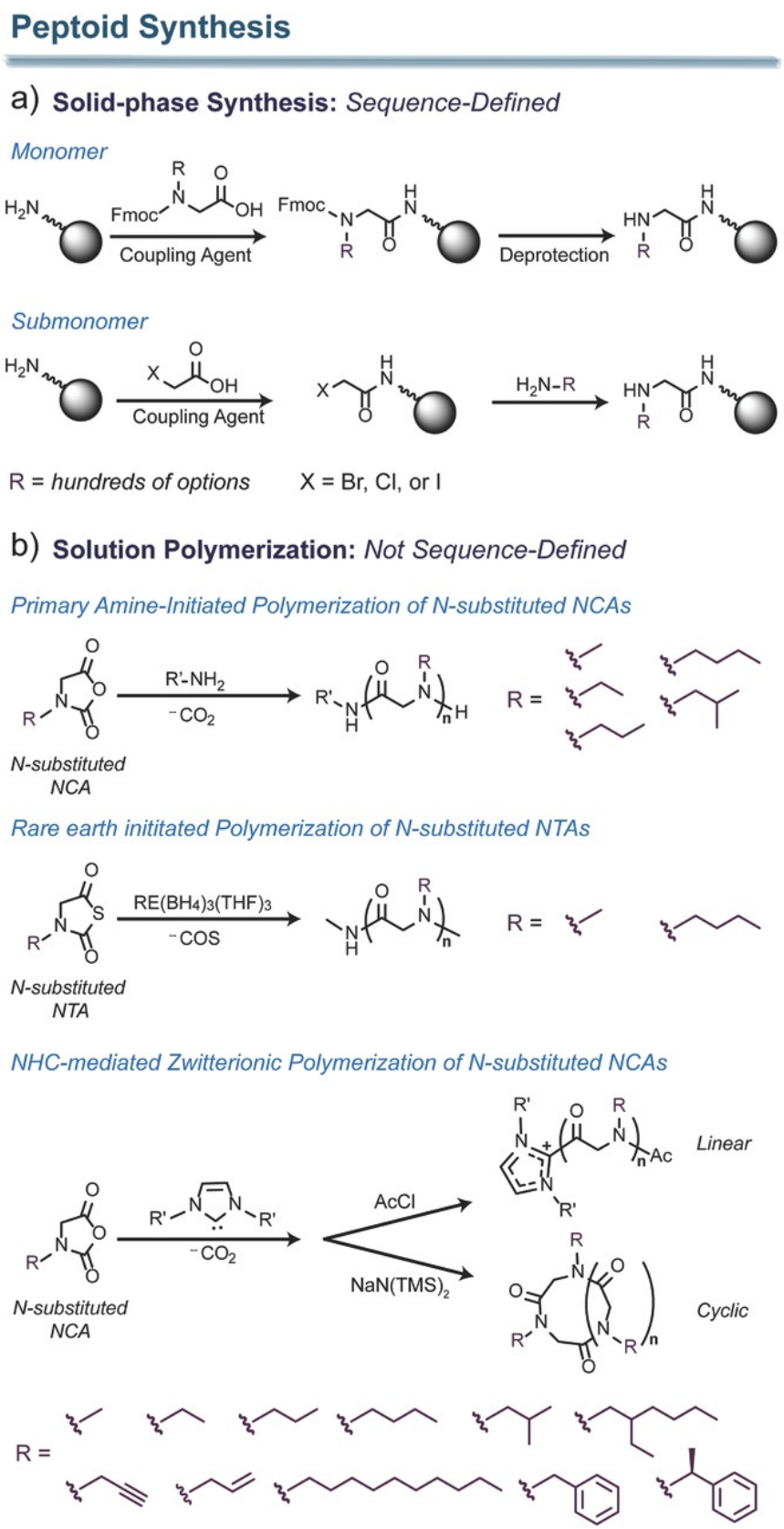

Figure 1.2. Two common methods to synthesize polypeptoids: a) solid-phase synthesis and (b) solution polymerization. This figure is reproduced from the studies of knight *et al.*[11] with permission.

molecular simulations have greatly improved our understanding of glassy polymers, including different relaxation processes, effect of confinement, conformation changes etc.[29] Atomistic simulations reveal the role of solvent in protein dynamics.[30,31] Specialized simulations can help us learn the mechanism of protein folding.[32–34] With continuing advances in the methodology, molecular dynamics studies are being extended to larger systems and longer time scales. This makes it possible for molecular dynamics simulations to obtain information that is not accessible from experiments.

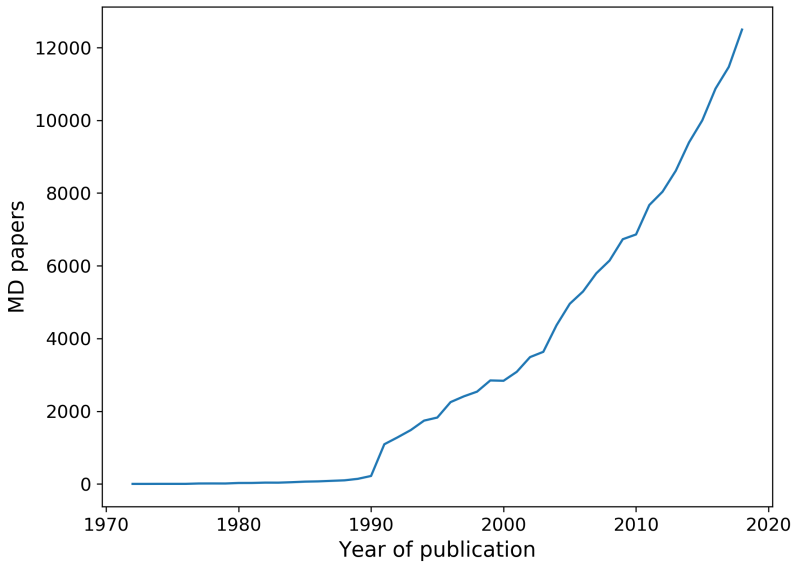

Figure 1.3. Growth of molecular dynamics Simulations. This analysis was performed via Web of Science Core Collection (https://www.webofknowledge.com), which provides access to publications from world's leading scholarly journals, books, and proceedings in the sciences, social sciences, and humanities.

Molecular dynamics is based on solving Newtons equation of motion numerically. Newton's equation of motion is given by:

$$F_i = m_i a_i \tag{1–1}$$

where F_i is force on particle i, m_i is mass of the particle i, and a_i is the particle i's acceleration. The force acting on each particle can be expressed as the gradient of the potential energy:

$$F_i = -\nabla_i U \tag{1–2}$$

where U is the potential energy of the system. Since velocity is the derivative of position and acceleration is the derivative of velocity, we can thus write equations of motion as:

$$\frac{-dU}{dr_i} = m_i \frac{d^2 r_i}{dt^2} \tag{1–3}$$

Newton's equation of motion can then relate the derivative of the potential energy to the changes in position as a function of time. The above differential equation is the basis of molecular dynamics simulations.

1.2.1 Force Field

To solve Equation 1–3, one needs the gradient of the potential energy of the system. Within the Born-Oppenheimer approximation, one could solve the time-independent Schrödinger Equation using various electronic structure methods and determine the atomic forces. Indeed, this can be done at the Density Functional Theory level and is the basis of the so-called *ab initio* molecular dynamics simulations.[35,36] However, these methods are extremely computationally expensive and hence one often uses force fields that are essentially modeled the potential energy of the system under consideration.[37–40] Unlike quantum mechanical methods, force fields can be evaluated rapidly and, if parameterized correctly, can accurately reproduce sufficient details of the properties of interest of the system studied.[41–43] The parameters of a force field can be derived from experiments such as neutron, X-ray and electron diffraction, NMR, infrared, Raman and neutron spectroscopy, or quantum mechanics, or both. Many force fields employ similar functional forms but differ in the parameters used in the equations. A typical expression for a force field can be written as:

$$U_{total} = U_{bonded} + U_{nonbonded} \tag{1–4}$$

where the components of the covalent and noncovalent contributions can be broken down as shown below:

$$U_{bonded} = U_{bond} + U_{angle} + U_{dihedral} \tag{1–5}$$

$$U_{nonbonded} = U_{electrostatic} + U_{vdw} \tag{1–6}$$

Bond stretching and angle bending are usually represented by a harmonic potential, which does not allow bond breaking. Thus, no chemical processes can be studied. Normal dihedral interactions can usually be represented by a cosine function such as Fourier expansion. A small number of terms is applied, generally three, in the Fourier expansion of dihedral potentials. The functional forms of bond, angle, and dihedral potentials are shown as follows:

$$U_{bond}(r_{ij}) = \frac{k_{ij}^{l}}{2}(r_{ij} - r_{ij}^{0})^2 \tag{1–7}$$

$$U_{angle}(\theta_{ijk}) = \frac{k_{ijk}^{\theta}}{2}(\theta_{ijk} - \theta_{ijk}^{0})^2 \tag{1–8}$$

$$U_{dihedral}(\phi_{ijkl}) = \sum_{n} \frac{V_n}{2}[1 + cos(n\phi_{ijkl} - \gamma_n)] \tag{1–9}$$

In the preceding equations, k^l, r_{ij}, and r_{ij}^0, respectively, are the bond force constant, bond length between particle i and j, and equilibrium bond distance, k_{ijk}^{θ}, θ_{ijk}, and θ_{ijk}^0 are, respectively, the angle force constant, angle between particle i, j, k, and equilibrium angle, while V_n, n, θ_{ijkl}, and γ_n are the dihedral force constant, dihedral periodicity, dihedral angle between particle i, j, k, l, and a phase of the dihedral angle θ_{ijkl}. The bond and angle parameters can be obtained from gas phase spectroscopy and small molecule crystal structures. The dihedral parameters are usually derived from gas phase quantum mechanics calculations. It is worth noting that the functional forms of each term in equations 1–7, 1–8, and 1–9 are highly variable, the details of which are beyond the scope of this dissertation.

Non-bonded interactions of classical force fields are usually divided in two: electrostatic interactions and van der Waals interactions. The electrostatic interactions can be generally modelled by a Coulomb potential:

$$U_{Coulomb}(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \tag{1–10}$$

where q_i and q_j are the charges on the particle i and j, r_{ij} is their separation distance, and ϵ_0 is the vacuum electric permittivity. The van der Waals interaction contains a repulsion and a

dispersion term. One of the most common forms for the combination of repulsion and dispersion interactions is the Lennard-Jones (LJ) potential, which is expressed as:

$$U_{LJ}(r_{ij}) = 4\epsilon_{ij}[(\frac{\sigma_{ij}}{r_{ij}})^{12} - (\frac{\sigma_{ij}}{r_{ij}})^{6}] \tag{1–11}$$

where ϵ_{ij} is the strength of dispersion interactions between particle i and j, σ_{ij} is the contact distance between particle i and j, and r_{ij} is the distance between particle i and j. The LJ potential depends only on the distance between pairs of atoms. The first part of the equation, $(\frac{\sigma}{r})^{12}$ describes the repulsive forces between particles while the latter part of the equation, $(\frac{\sigma}{r})^{6}$ denotes the attractive forces. A simple schematic view of force field interactions is shown in Figure 1.4. Numerous community-developed force fields have been developed over the

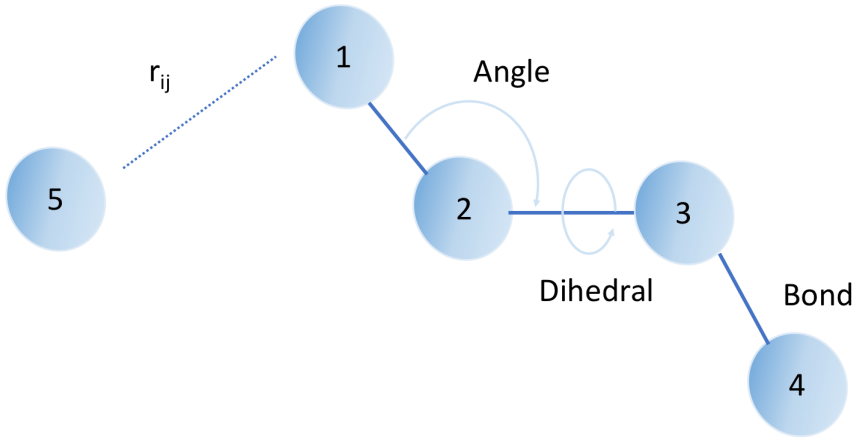

Figure 1.4. A schematic view of force field interactions. Covalent bonds are indicated by solid blue lines, non-bonded by a dash line.

decades for simulations, and it is not the intent of this dissertation to provide any sort of comprehensive review. It is worth noting that four force fields are wildly used in the literature, and these include Amber,[44] Chemistry at HARvard Macromolecular Mechanics (CHARMM),[45,46] GROningen MOlecular Simultion (GROMOS),[47] and the all-atom Optimized Potential for Liquid

Simulations (OPLS-AA).[48,49] The comparison between these four commonly used force fields is summarized in Table 1.1.

Table 1.1. Comparison between several force fields.

Force field	Chemical reactions	Atomic charge	Type of systems
Amber	No	Static	Proteins, lipids, nucleic acids, carbohydrates
CHARMM	No	Static	Proteins, lipids, nucleic acids, organics
GROMOS	No	Static	Proteins, sugars, nucleic acids, organics
OPLS	No	Static	Many liquids

1.2.2 Integration Algorithms

As mentioned above, the equations of motion can only be solved numerically and there is no analytical solution. In molecular dynamics, one of most commonly used integrators is Verlet algorithm.[50,51] All the integration algorithms assume that the positions, velocities and accelerations can be approximated by a Taylor series expansion:

$$r(t+\delta t) = r(t) + v(t)\delta t + \frac{1}{2}a(t)\delta t^2 + \ldots \tag{1–12}$$

$$v(t+\delta t) = r(t) + a(t)\delta t + \frac{1}{2}b(t)\delta t^2 + \ldots \tag{1–13}$$

$$a(t+\delta t) = a(t) + b(t)\delta t + \ldots \tag{1–14}$$

where r is the position, v is the velocity, and a is the acceleration. To derive the Verlet Algorithm, one has:

$$r(t+\delta t) = r(t) + v(t)\delta t + \frac{1}{2}a(t)\delta t^2 \tag{1–15}$$

$$r(t-\delta t) = r(t) - v(t)\delta t + \frac{1}{2}a(t)\delta t^2 \tag{1–16}$$

Adding the two expressions together gives:

$$r(t+\delta t) = 2r(t) - r(t-\delta t) + a(t)\delta t^2 \tag{1–17}$$

This is the basic form of the standard Verlet algorithm, which uses positions at time t and $t-\delta t$, and acceleration at time t to obtain new positions at time $t+\delta t$. The acceleration, $a(t)$, which is the second derivative of the position $r(t)$ as a function of time is determined from Equation 1–3 (Newton's equation of motion). A problem with this version of the Verlet algorithm is that velocities are not explicitly solved. Although velocities are not needed for time evolution, this information is sometimes necessary, such as to determine the kinetic energy of the system. One could compute the velocities from first order central difference:

$$v(t) = \frac{r(t+\delta t) - r(t-\delta t)}{2\delta t} \tag{1–18}$$

The standard Verlet algorithm is not a self-starting method since position vector at $t+\delta t$ requires positions previous two time steps. The advantages of the standard Verlet algorithm are, good stability, great simplicity, and modest requirement for memories. The disadvantage is moderate precision. To improve the accuracy compared to standard Verlet, some variants of the Verlet algorithm have been developed.[52] For instance, the Leap-Frog algorithm[53] is one of such variants where velocities are handled better:

$$r(t+\delta t) = r(t) + v(t+\frac{1}{2}\delta t)\delta t \tag{1–19}$$

$$v(t+\frac{1}{2}\delta t) = v(t-\frac{1}{2}\delta t) + a(t)\delta t \tag{1–20}$$

In the Leap-Frog algorithm, velocities are explicitly solved. The algorithm is visualized in Figure 1.5. The algorithm is called Leap-Frog because the positions and velocities are at interleaved time points, staggered in such a way that they are leaping like frogs over each other's back. The disadvantage of this algorithm is that velocities are not calculated at the same time as the positions. The velocities at time t can be calculated by:

$$v(t) = \frac{1}{2}[v(t-\frac{1}{2}\delta t) + v(t+\frac{1}{2}\delta t)] \tag{1–21}$$

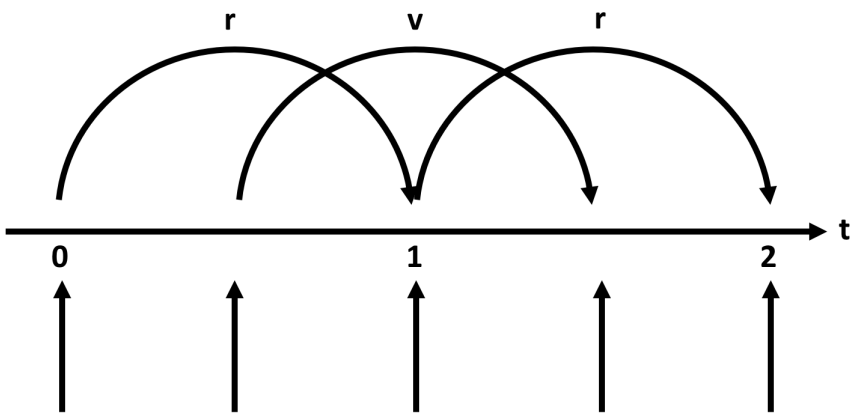

Figure 1.5. The Leap-Frog algorithm.

An alternative implementation of the same basic algorithm is the velocity Verlet algorithm.[54] The velocity Verlet algorithm requires updates of both positions and velocities:

$$r(t+\delta t) = r(t) + v(t)\delta(t) + \frac{1}{2}a(t)\delta t^2 \tag{1–22}$$

$$v(t+\delta t) = v(t) + \frac{1}{2}[a(t) + a(t+\delta t)]\delta t \tag{1–23}$$

Since the velocity Verlet algorithm synchronizes the calculations of positions and velocities, there is no compromise on precision. All of these commonly used integration algorithms are time reversible. Computational demands of using any particular integration scheme is important. Choosing the reasonable time step is another critical consideration when running molecular dynamics simulations. If the time step is too small, the simulation trajectories cover only a limited part of the phase space. If the time step is too large, the particles are moving too much between time intervals and simulation is therefore unstable. The chosen time step should be approximately one order of magnitude smaller than the shortest motion time scale. In general, most of atomistic molecular dynamics simulation studies often employ 1 or 2 fs as time steps for

numerical stability. If using a coarse-grained potential, a larger time step can be used to speed up simulations.

Simplified workflow of a typical molecular dynamics simulation is depicted in Figure 1.6. The first step of a molecular dynamics simulation normally requires initializing the calculation by providing the initial coordinates and velocities of all the particles in the system as well as the specification of the force fields, the potential energy functions of the system that determine how the particles interact with each other. The loop in the Figure 1.6 refers to the main computational routine of a molecular dynamics simulation. In this loop, the forces on each atom will be calculated repeatedly and then these forces will be used to update the position and velocity of each atom using one of the integrators mentioned above. The essential output of a simulation is a three-dimensional trajectory, which consists of consecutive snapshots of the system coordinates at every point during the simulated time interval. Such information is very difficult to obtain with any experimental technique. The microscopic level of information generated from the simulation is then analyzed using statistical mechanics to calculate a wide range of macroscopic properties, such as pressure, density, energy, temperature, enthalpy, entropy and heat capacities, etc. Researchers these days do not have to write computer code from scratch to perform molecular dynamics simulations. There are many simulation software/engines that can be used to study different chemical systems. Common choices include LAMMPS,[55] GROMACS,[56] AMBER,[57] CHARMM,[45] NAMD,[58] Tinker,[59] and OpenMM.[60]

1.3 Previous Molecular Dynamics Studies of Peptoids

Over the past several decades, molecular dynamics simulations have played a prominent role in understanding peptoid systems.[27,61] The first all-atom simulations of simple peptoid were performed in the late 1990s. Möhle *et al.* used CHARMM22 force field to study the conformation change of peptoid monomers in comparison to the corresponding peptides in aqueous solution.[62] The results of these simulations reproduced the essential features of *ab initio* data. In 1997, Armmand *et al.*[20] used molecular mechanics calculations and AMBER force field to show that the presence of the Nα chiral center in the peptoid side chains has a dramatic impact on the

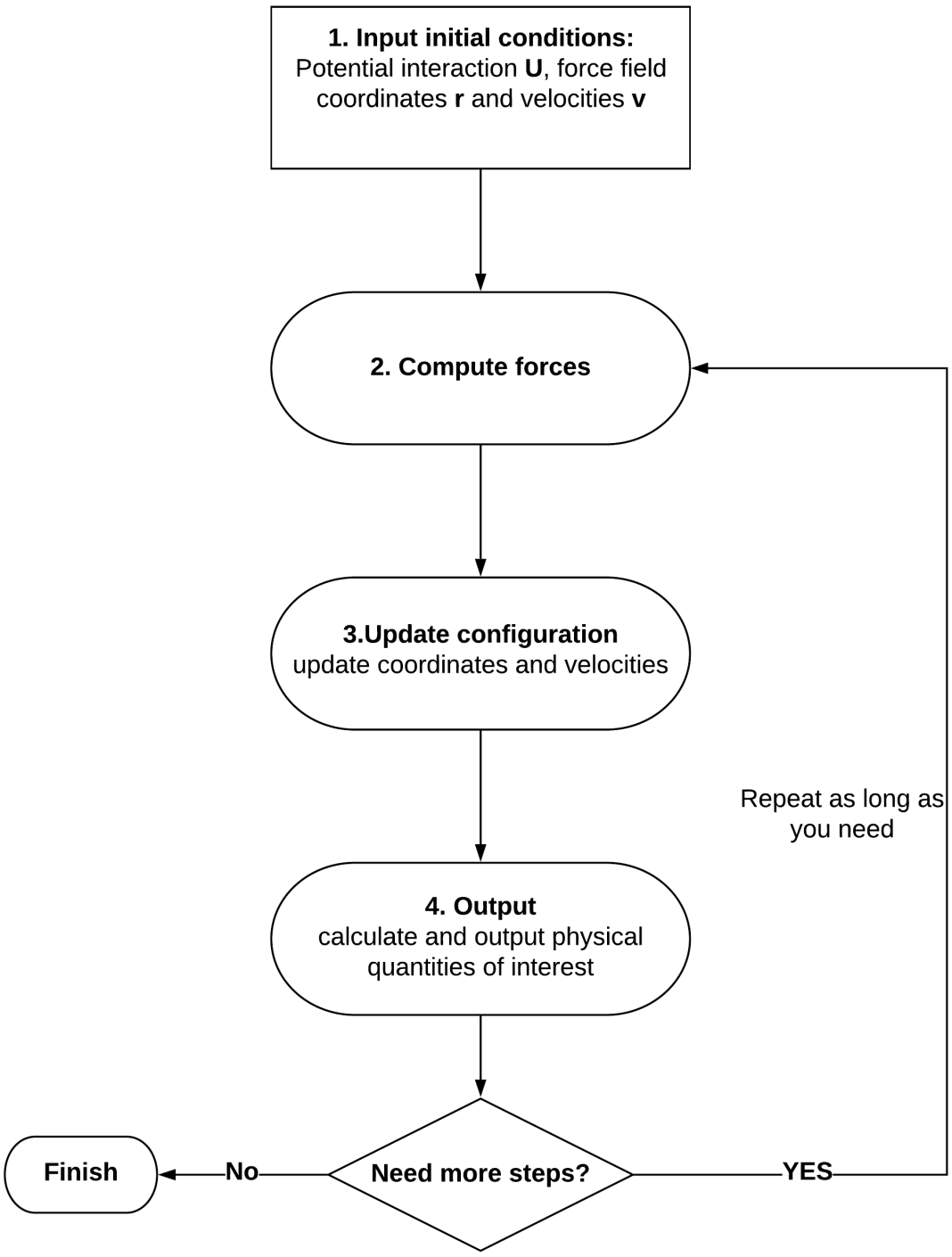

Figure 1.6. Flowchart of a typical molecular dynamics simulation.

available backbone conformations. Park and coworkers carried out all-atom molecular dynamics simulations on peptoid oligomers with methyl and methoxyethyl side chains in aqueous phase with explicit water molecules using OPLS force field.[63] They found peptoids have unique backbone dihedral angle distributions, which are significantly different from the peptide counterparts.

However, the force fields used in the previous studies mentioned above are not primarily optimized for peptoids, the results may not represent peptoid structures accurately.[64] Recently, researchers have developed new all-atom force fields to address this concern. Mirijanian *et al.* developed MFTOID force field for peptoid, based on the CHARMM22[65,66] peptide force field.[67] Notably, their parameterization for peptoid mainly focused on the methyl side chains, thus leading to potential transferability issues. Mukherjee *et al.* reported that general AMBER force filed[68] (GAFF) with new parameters could successfully predict a range of experimental structures for peptoid systems in implicit solvent but worked poorly in explicit solvent.[69] Weiser *et al.* developed new CHARMM general force field[70] (CGenFF) parameters that successfully reproduce structures of both *cis* and *trans* conformers.[71]

1.4 Scope of Dissertation

The primary goal of this dissertation is to provide nanoscale insight into the self-assembly of polypeptoids in solution by using computational approaches. In chapter 2, we used simulation techniques to study aggregation behavior of cyclic polypeptoids bearing zwitterionic and end-groups in methanol. Small-angle neutron scattering and cryo-TEM data indicate that these cyclic polypeptoids form small clusters, ranging from a single polypeptoid chain to small oligomers. Atomistic molecular dynamics simulations reveal that the driving force for this clustering behavior is attributed to the interplay between the effective repulsion due to the solvation of the dipoles formed by the charged end-groups in each peptoid chain and the attractive forces due to dipole-dipole interactions and the solvophobic effect.

In Chapter 3, atomistic molecular dynamics simulations were performed on sequence-defined ionic peptoid block copolymers consisting of a hydrophobic and a hydrophilic segment (with one singly charged and the remaining uncharged polar groups) in aqueous solution. Results reveal

that water molecules were highly acknowledged for playing an important role in the stability and structure of self-assembled charged polypeptoids involving strong inter/intramolecular secondary interactions. Besides, charge-dipole secondary interactions dictate micelle shape, water penetration, and other micellar properties. These results shed light on the physics of self-assembly of weakly charged soft matter and advance our understanding of the precise design of self-assembled micelles by controlling the position of charge on the backbone of peptoid block copolymers.

In Chapter 4, a coarse-grained (CG) model for N, N-dimethylacetamide (DMA), which represents the polypeptoid backbone, was developed as a step towards establishing a CG model of the complex polypeptoid system. The DMA CG model is parameterized to reproduce structural properties of DMA liquid and solution using a reference all-atom model, namely the OPLS-AA force field. The intermolecular forces are represented by the Stillinger-Weber potential, that consists of both two- and three-body terms that are very short-ranged. The model is validated on thermodynamic properties of liquid and aqueous DMA, the vapor-liquid interface of liquid DMA, and the structure of a concentrated aqueous solution of DMA in water as well as a simple peptoid in water. Without long-ranged interactions and the absence of interaction sites on hydrogen atoms, the CG DMA model is an order of magnitude faster than the higher resolution all-atom (AA) model. Finally, chapter 5 presents the conclusion of this dissertation and offers suggestions about future work.

1.5 References

[1] Ronald N Zuckermann, "Peptoid origins", *Peptide Science* **96**(5), pp. 545–555 (2011).

[2] Reyna J Simon, Robert S Kania, Ronald N Zuckermann, Verena D Huebner, David A Jewell, Steven Banville, Simon Ng, Liang Wang, Steven Rosenberg, and Charles K Marlowe, "Peptoids: a modular approach to drug discovery.", *Proceedings of the National Academy of Sciences* **89**(20), pp. 9367–9371 (1992).

[3] Ronald N Zuckermann and Thomas Kodadek, "Peptoids as potential therapeutics", *Curr. Opin. Mol. Ther* **11**(3), pp. 299–307 (2009).

[4] Ronald N Zuckermann, "The chemical synthesis of peptidomimetic libraries: Current opinion in structural biology 1993, 3: 580–584", *Current Opinion in Structural Biology* **3**(4), pp. 580–584 (1993).

[5] Kent Kirshenbaum, Ronald N Zuckermann, and Ken A Dill, "Designing polymers that mimic biomolecules", *Current opinion in structural biology* **9**(4), pp. 530–535 (1999).

[6] Annelise E Barron and Ronald N Zuckerman, "Bioinspired polymeric materials: in-between proteins and plastics", *Current opinion in chemical biology* **3**(6), pp. 681–687 (1999).

[7] John E Murphy, Tetsuo Uno, Janice D Hamer, Fred E Cohen, Varavani Dwarki, and Ronald N Zuckermann, "A combinatorial approach to the discovery of efficient cationic peptoid reagents for gene delivery", *Proceedings of the National Academy of Sciences* **95**(4), pp. 1517–1522 (1998).

[8] Byoung-Chul Lee, Michael D Connolly, and Ronald N Zuckermann, "Bioinspired polymers for nanoscience research", *ChemInform* **40**(32), pp. i (2009).

[9] Donghui Zhang, Samuel H Lahasky, Li Guo, Chang-Uk Lee, and Monika Lavan, "Polypeptoid materials: current status and future perspectives", *Macromolecules* **45**(15), pp. 5833–5841 (2012).

[10] Robert Luxenhofer, Corinna Fetsch, and Arlett Grossmann, "Polypeptoids: A perfect match for molecular definition and macromolecular engineering?", *Journal of Polymer Science Part A: Polymer Chemistry* **51**(13), pp. 2731–2752 (2013).

[11] Abigail S Knight, Effie Y Zhou, Matthew B Francis, and Ronald N Zuckermann, "Sequence programmable peptoid polymers for diverse materials applications", *Advanced Materials* **27**(38), pp. 5665–5691 (2015).

[12] Ronald N Zuckermann, Janice M Kerr, Stephen BH Kent, and Walter H Moos, "Efficient method for the preparation of peptoids [oligo (n-substituted glycines)] by submonomer solid-phase synthesis", *Journal of the American Chemical Society* **114**(26), pp. 10646–10647 (1992).

[13] Pallabi Samaddar, Akash Deep, and Ki-Hyun Kim, "An engineering insight into block copolymer self-assembly: Contemporary application from biomedical research to nanotechnology", *Chemical Engineering Journal* **342**, pp. 71–89 (2018).

[14] Hongbo Feng, Xinyi Lu, Weiyu Wang, Nam-Goo Kang, and Jimmy W Mays, "Block copolymers: Synthesis, self-assembly, and applications", *Polymers* **9**(10), pp. 494 (2017).

[15] Martin Hrubỳ, Sergey K Filippov, and Petr Štěpánek, "Biomedical application of block copolymers", *Macromolecular Self-assembly* , pp. 231–250 (2016).

[16] Dali Wang, Gangsheng Tong, Ruijiao Dong, Yongfeng Zhou, Jian Shen, and Xinyuan Zhu, "Self-assembly of supramolecularly engineered polymers and their biomedical applications", *Chemical Communications* **50**(81), pp. 11994–12017 (2014).

[17] Yoshinori Kakizawa and Kazunori Kataoka, "Block copolymer self-assembly into monodispersive nanoparticles with hybrid core of antisense dna and calcium phosphate", *Langmuir* **18**(12), pp. 4539–4543 (2002).

[18] Saziye Yorulmaz Avsar, Myrto Kyropoulou, Stefano Di Leone, Cora-Ann Schoenenberger, Wolfgang P Meier, and Cornelia G Palivan, "Biomolecules turn self-assembling amphiphilic block co-polymer platforms into biomimetic interfaces", *Frontiers in chemistry* **6** (2018).

[19] Cindy W Wu, Tracy J Sanborn, Ronald N Zuckermann, and Annelise E Barron, "Peptoid oligomers with α-chiral, aromatic side chains: effects of chain length on secondary structure", *Journal of the American Chemical Society* **123**(13), pp. 2958–2963 (2001).

[20] Philippe Armand, Kent Kirshenbaum, Alexis Falicov, Roland L Dunbrack Jr, Ken A Dill, Ronald N Zuckermann, and Fred E Cohen, "Chiral n-substituted glycines can form stable helical conformations", *Folding and Design* **2**(6), pp. 369–375 (1997).

[21] Cindy W Wu, Tracy J Sanborn, Kai Huang, Ronald N Zuckermann, and Annelise E Barron, "Peptoid oligomers with α-chiral, aromatic side chains: sequence requirements for the formation of stable peptoid helices", *Journal of the American Chemical Society* **123**(28), pp. 6778–6784 (2001).

[22] Philippe Armand, Kent Kirshenbaum, Richard A Goldsmith, Shauna Farr-Jones, Annelise E Barron, Kiet TV Truong, Ken A Dill, Dale F Mierke, Fred E Cohen, Ronald N Zuckermann, et al., "Nmr determination of the major solution conformation of a peptoid pentamer with chiral side chains", *Proceedings of the National Academy of Sciences* **95**(8), pp. 4309–4314 (1998).

[23] Cindy W Wu, Kent Kirshenbaum, Tracy J Sanborn, James A Patch, Kai Huang, Ken A Dill, Ronald N Zuckermann, and Annelise E Barron, "Structural and spectroscopic studies of peptoid oligomers with α-chiral aliphatic side chains", *Journal of the American Chemical Society* **125**(44), pp. 13525–13530 (2003).

[24] Kai Huang, Cindy W Wu, Tracy J Sanborn, James A Patch, Kent Kirshenbaum, Ronald N Zuckermann, Annelise E Barron, and Ishwar Radhakrishnan, "A threaded loop conformation adopted by a family of peptoid nonamers", *Journal of the American Chemical Society* **128**(5), pp. 1733–1738 (2006).

[25] J Aaron Crapster, Ilia A Guzei, and Helen E Blackwell, "A peptoid ribbon secondary structure", *Angewandte Chemie International Edition* **52**(19), pp. 5079–5084 (2013).

[26] Benjamin C Gorske, Emily M Mumford, Charles G Gerrity, and Imelda Ko, "A peptoid square helix via synergistic control of backbone dihedral angles", *Journal of the American Chemical Society* **139**(24), pp. 8070–8073 (2017).

[27] Ranjan V Mannige, Thomas K Haxton, Caroline Proulx, Ellen J Robertson, Alessia Battigelli, Glenn L Butterfoss, Ronald N Zuckermann, and Stephen Whitelam, "Peptoid nanosheets exhibit a new secondary-structure motif", *Nature* **526**(7573), pp. 415 (2015).

[28] Martin Karplus and J Andrew McCammon, "Molecular dynamics simulations of biomolecules", *Nature Structural & Molecular Biology* **9**(9), pp. 646 (2002).

[29] Jean-Louis Barrat, Jörg Baschnagel, and Alexey Lyulin, "Molecular dynamics simulations of glassy polymers", *Soft Matter* **6**(15), pp. 3430–3446 (2010).

[30] Dennis Vitkup, Dagmar Ringe, Gregory A Petsko, and Martin Karplus, "Solvent mobility and the protein'glass' transition", *Nature Structural & Molecular Biology* **7**(1), pp. 34 (2000).

[31] Stephen J Hagen, James Hofrichter, and William A Eaton, "Protein reaction kinetics in a room-temperature glass", *Science* **269**(5226), pp. 959–962 (1995).

[32] Christopher M Dobson, Andrej Šali, and Martin Karplus, "Protein folding: a perspective from theory and experiment", *Angewandte Chemie International Edition* **37**(7), pp. 868–893 (1998).

[33] Yaoqi Zhou and Martin Karplus, "Interpreting the folding kinetics of helical proteins", *Nature* **401**(6751), pp. 400 (1999).

[34] Andrea Cavalli, Philippe Ferrara, and Amedeo Caflisch, "Weak temperature dependence of the free energy surface and folding pathways of structured peptides", *Proteins: Structure, Function, and Bioinformatics* **47**(3), pp. 305–314 (2002).

[35] Robert G Parr, "Density functional theory of atoms and molecules", In *Horizons of Quantum Chemistry* , pp. 5–15. Springer (1980).

[36] Richard Car and Mark Parrinello, "Unified approach for molecular dynamics and density-functional theory", *Physical review letters* **55**(22), pp. 2471 (1985).

[37] Aron J Cohen, Paula Mori-Sánchez, and Weitao Yang, "Challenges for density functional theory", *Chemical reviews* **112**(1), pp. 289–320 (2011).

[38] Christopher J Cramer and Donald G Truhlar, "Density functional theory for transition metals and transition metal chemistry", *Physical Chemistry Chemical Physics* **11**(46), pp. 10757–10816 (2009).

[39] Edoardo Aprà, Eric J Bylaska, David J Dean, Alessandro Fortunelli, Fei Gao, Predrag S Krstić, Jack C Wells, and Theresa L Windus, "Nwchem for materials science", *Computational Materials Science* **28**(2), pp. 209–221 (2003).

[40] Frank Jensen, *Introduction to computational chemistry*, John wiley & sons (2017).

[41] Mike P Allen and Dominic J Tildesley, *Computer simulation in chemical physics* volume 397, Springer Science & Business Media (2012).

[42] Yalin Dong, Qunyang Li, and Ashlie Martini, "Molecular dynamics simulation of atomic friction: A review and guide", *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films* **31**(3), pp. 030801 (2013).

[43] Judith A Harrison, J David Schall, M Todd Knippenberg, Guangtu Gao, and Paul T Mikulski, "Elucidating atomic-scale friction using molecular dynamics and specialized analysis techniques", *Journal of Physics: Condensed Matter* **20**(35), pp. 354009 (2008).

[44] Paul K Weiner and Peter A Kollman, "Amber: Assisted model building with energy refinement. a general program for modeling molecules and their interactions", *Journal of Computational Chemistry* **2**(3), pp. 287–303 (1981).

[45] Bernard R Brooks, Charles L Brooks III, Alexander D Mackerell Jr, Lennart Nilsson, Robert J Petrella, Benoît Roux, Youngdo Won, Georgios Archontis, Christian Bartels, Stefan Boresch, et al., "Charmm: the biomolecular simulation program", *Journal of computational chemistry* **30**(10), pp. 1545–1614 (2009).

[46] Bernard R Brooks, Robert E Bruccoleri, Barry D Olafson, David J States, S a Swaminathan, and Martin Karplus, "Charmm: a program for macromolecular energy, minimization, and dynamics calculations", *Journal of computational chemistry* **4**(2), pp. 187–217 (1983).

[47] Xavier Daura, Alan E Mark, and Wilfred F Van Gunsteren, "Parametrization of aliphatic chn united atoms of gromos96 force field", *Journal of computational chemistry* **19**(5), pp. 535–547 (1998).

[48] William L. Jorgensen, Jeffry D. Madura, and Carol J. Swenson, "Optimized intermolecular potential functions for liquid hydrocarbons", *Journal of the American Chemical Society* **106**(22), pp. 6638–6646 (1984).

[49] William L. Jorgensen, David S. Maxwell, and Julian Tirado-Rives, "Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids", *Journal of the American Chemical Society* **118**(45), pp. 11225–11236 (1996).

[50] Loup Verlet, "Computer" experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules", *Physical review* **159**(1), pp. 98 (1967).

[51] Loup Verlet, "Computer" experiments" on classical fluids. ii. equilibrium correlation functions", *Physical Review* **165**(1), pp. 201 (1968).

[52] Stewart A Adcock and J Andrew McCammon, "Molecular dynamics: survey of methods for simulating the activity of proteins", *Chemical reviews* **106**(5), pp. 1589–1615 (2006).

[53] Roger W Hockney, "The potential calculation and some applications", *Methods Comput. Phys.* **9**, pp. 136 (1970).

[54] William C Swope, Hans C Andersen, Peter H Berens, and Kent R Wilson, "A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters", *The Journal of Chemical Physics* **76**(1), pp. 637–649 (1982).

[55] Steve Plimpton, "Fast parallel algorithms for short-range molecular dynamics", *Journal of computational physics* **117**(1), pp. 1–19 (1995).

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

[57] D.S. Cerutti T.E. Cheatham III T.A. Darden R.E. Duke T.J. Giese H. Gohlke A.W. Goetz N. Homeyer S. Izadi P. Janowski J. Kaus A. Kovalenko T.S. Lee S. LeGrand P. Li C. Lin T. Luchko R. Luo B. Madej D. Mermelstein K.M. Merz G. Monard H. Nguyen H.T. Nguyen I. Omelyan A. Onufriev D.R. Roe A. Roitberg C. Sagui C.L. Simmerling W.M. Botello-Smith J. Swails R.C. Walker J. Wang R.M. Wolf X. Wu L. Xiao D.A. Case, R.M. Betz and P.A. Kollman, "Amber 2016", University of California, San Francisco.

[58] James C Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D Skeel, Laxmikant Kale, and Klaus Schulten, "Scalable molecular dynamics with namd", *Journal of computational chemistry* **26**(16), pp. 1781–1802 (2005).

[59] Jay W Ponder and Frederic M Richards, "An efficient newton-like method for molecular mechanics energy minimization of large molecules", *Journal of Computational Chemistry* **8**(7), pp. 1016–1024 (1987).

[60] Peter Eastman, Mark S Friedrichs, John D Chodera, Randall J Radmer, Christopher M Bruns, Joy P Ku, Kyle A Beauchamp, Thomas J Lane, Lee-Ping Wang, Diwakar Shukla, et al., "Openmm 4: a reusable, extensible, hardware independent library for high performance molecular simulation", *Journal of chemical theory and computation* **9**(1), pp. 461–469 (2012).

[61] Glenn L Butterfoss, Barney Yoo, Jonathan N Jaworski, Ilya Chorny, Ken A Dill, Ronald N Zuckermann, Richard Bonneau, Kent Kirshenbaum, and Vincent A Voelz, "De novo structure prediction and experimental characterization of folded peptoid oligomers", *Proceedings of the National Academy of Sciences* **109**(36), pp. 14320–14325 (2012).

[62] Kerstin Möhle and Hans-Jörg Hofmann, "Peptides and peptoids-a systematic structure comparison", In *Molecular modeling annual* volume 2 , pp. 307–311. Springer (1996).

[63] Sung Hyun Park and Igal Szleifer, "Structural and dynamical characteristics of peptoid oligomers with achiral aliphatic side chains studied by molecular dynamics simulation", *The Journal of Physical Chemistry B* **115**(37), pp. 10967–10975 (2011).

[64] Laura J Weiser and Erik E Santiso, "Molecular modeling studies of peptoid polymers", *AIMS MATERIALS SCIENCE* **4**(5), pp. 1029–1051 (2017).

[65] Alexander D Mackerell Jr, Michael Feig, and Charles L Brooks III, "Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations", *Journal of computational chemistry* **25**(11), pp. 1400–1415 (2004).

[66] Robert B Best, Xiao Zhu, Jihyun Shim, Pedro EM Lopes, Jeetain Mittal, Michael Feig, and Alexander D MacKerell Jr, "Optimization of the additive charmm all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ1 and χ2 dihedral angles", *Journal of chemical theory and computation* **8**(9), pp. 3257–3273 (2012).

[67] Dina T Mirijanian, Ranjan V Mannige, Ronald N Zuckermann, and Stephen Whitelam, "Development and use of an atomistic charmm-based forcefield for peptoid simulation", *Journal of computational chemistry* **35**(5), pp. 360–370 (2014).

[68] Junmei Wang, Romain M Wolf, James W Caldwell, Peter A Kollman, and David A Case, "Development and testing of a general amber force field", *Journal of computational chemistry* **25**(9), pp. 1157–1174 (2004).

[69] Sudipto Mukherjee, Guangfeng Zhou, Chris Michel, and Vincent A Voelz, "Insights into peptoid helix folding cooperativity from an improved backbone potential", *The Journal of Physical Chemistry B* **119**(50), pp. 15407–15417 (2015).

[70] Kenno Vanommeslaeghe, Elizabeth Hatcher, Chayan Acharya, Sibsankar Kundu, Shijun Zhong, Jihyun Shim, Eva Darian, Olgun Guvench, P Lopes, Igor Vorobyov, et al., "Charmm general force field: A force field for drug-like molecules compatible with the charmm all-atom additive biological force fields", *Journal of computational chemistry* **31**(4), pp. 671–690 (2010).

[71] Laura J Weiser and Erik E Santiso, "A cgenff-based force field for simulations of peptoids with both cis and trans peptide bonds", *Journal of computational chemistry* (2019).

CHAPTER 2
AGGREGATION OF CYCLIC POLYPEPTOIDS BEARING ZWITTERIONIC END-GROUPS WITH ATTRACTIVE DIPOLE-DIPOLE AND SOLVOPHOBIC INTERACTIONS: A STUDY BY MOLECULAR DYNAMICS SIMULATION

2.1 Introduction

Cluster formation is a common phenomenon that has been widely observed for solutions containing colloidal particles[2,3] and biomacromolecules (e.g., proteins[4] and DNAs[5]) as well as synthetic polymers. Clusters can exist in various morphologies and are often precursors to glass or gel formation[3,6–8]. Thus, cluster formation has significant implications for the practical uses of these solutions in the areas of food products, agricultural products and biopharmaceuticals.

The mechanism of cluster formation often varies and depends on the nature of the specific solutes in the solution and the interplay of various contributing forces. To form stable finite clusters, the presence of competing interactions is often required.[3,7,9] The attractive interaction promotes the formation and growth of clusters, whereas the repulsive interaction will arrest and prevent indefinite cluster growth and further aggregation. The balance of these forces will lead to stable clusters consisting of finite number of particles. Colloidal particles (CPs) with short range attractive interaction and long range repulsive interaction (SALR) have been shown to form stable clusters with specific size and size distribution determined by the thermodynamics.[8–11] Interestingly, it has also been shown that CPs with only attractive interaction (e.g., depletion force) can also form stable clusters, suggesting that details of attractive mechanism are important.[12] With low attractive potentials, the clusters can bump into each other but quickly diffuse away from each other before they can reorganize and merge into larger clusters, thus allowing stable small clusters to exist. Even more intriguingly, theoretical work has indicated that particles with

Reproduced from P. Du, A. Li, X. Li, Y. Zhang, C. Do, L. He, S. W. Rick, V. John, R. Kumar and D. Zhang, *Phys. Chem. Chem. Phys.*, 2017, **19**, 14388-14400 (https://www.doi.org/10.1039/C7CP01602F)[1] with permission from the PCCP Owner Societies.

purely repulsive interactions can also form finite clusters, stabilized by a free energy competition due to the intra-cluster and inter-cluster interactions.[13]

The majority of studies on clustering are focused on the CPs. Several globular proteins have been recently shown to have characteristics of CPs with short-range attractive interaction and long-range repulsive interactions, giving rise to liquidliquid phase transitions and some critical phenomenon.[4,14] Synthetic polymers including neutral polymers,[15] polyelectrolytes[16] and dendrimers/star polymers[17–19] have been described as soft colloids and were shown to form clusters driven by the interplay of various interactions.[17,19–22] For examples, high molecular weight poly(ethylene oxide) (PEOs) have been shown to form clusters driven by attractive solvophobic interactions of chain ends with the backbone.[23,24] The effect of counter ions in the case of charged polyelectrolytes has also been explored,[25] with larger counter ions giving rise to finite clusters due to steric and short-range electrostatic effects.[26] The interplay between hydrophobic and electrostatic forces in the aggregations of ionomers, another class of charged polymers, has received attention due to its implication in PEM materials.[27]

Investigations of cluster formation in macromolecular systems have been carried out using a number of experimental tools including small angle neutron and X-ray scattering (SANS and SAXS), optical microscopy as well as dielectric spectroscopy.[4,8,12,23,28] Features in the low Q region in neutron and X-ray scattering present a signature of clustering in macromolecular systems. Large aggregates that extend to the micron region can be observed in optical microscopy experiments, and presence of slow modes in DLS suggests the formation of aggregates. While these are extremely useful in analyzing the mechanism of clustering, one has to resort to theoretical and computational studies to tease out the subtle forces involved in the formation of clusters in various macromolecular systems. The bulk of these computational studies involve molecular dynamics or Monte Carlo simulations of coarse-grained models of the polymer that lack atomistic level detail.[2,10,13,15,16,29] The use of coarse-grained models arises due to the slow dynamics of aggregation in these systems, which gives rise to inadequate sampling when using atomistic simulation techniques. A number of coarse-grained Monte Carlo simulations of clustering of

colloidal particles describe the inter-particle interactions in terms of an attractive component described by a Lennard-Jones type interaction and a longer-range repulsive Yukawa interacti on.[2,3] In addition, mean field type approaches as well as classical density functional theory based methods wherein a free energy functional is constructed and minimized with respect to component densities have also been explored.[13,21,26,27,30]

Polypeptoids, a class of polymers with N-substituted poly-glycine backbones, are structural mimics of polypeptides.[31,32] Because of the N-substitution, polypeptoids lack extensive hydrogen bonding or stereogenic centers on the main chain that are characteristic of polypeptides. As a result, polypeptoids tend to have more flexible backbone conformation[33,34] and enhanced proteolytic stability and solubility[35] in various solvents relative to polypeptides.[31,32] Early studies have shown that the physicochemical properties of polypeptoids (e.g., conformation,[36–38] solubility and crystallinity[39–41]) are strongly dependent on the side chain structures.[31,32]

In this study, we investigated the cluster formation of cyclic polypeptoids bearing zwitterionic end-groups (see Figure 2.1b) in solution using molecular dynamics (MD) simulations combined with umbrella sampling (US), an enhanced sampling method. The simulations reveal that the subtle interplay between the attractive dipoledipole and the solvophobic interactions along with the repulsive forces due to solvation of the dipoles provide the driving force for the cluster formation of the cyclic polypeptoids.

2.2 Methods

Computational studies were carried out on the linear and the cyclic counterpart of the $PNMG_{100}$ polypeptoid (Figure 2.1b) in methanol solvent. In the cyclic case, the two ends are repre- sented by oppositely charged groups, namely $-O^-$ and $-NH_3^+$ whereas in the linear case the end groups are -OH and $-NH_3^+$. While these are not the end groups in the actual peptoids (Figure 2.1a), they carry the appropriate charge and hence present a reasonable representation of the relevant end group interactions.

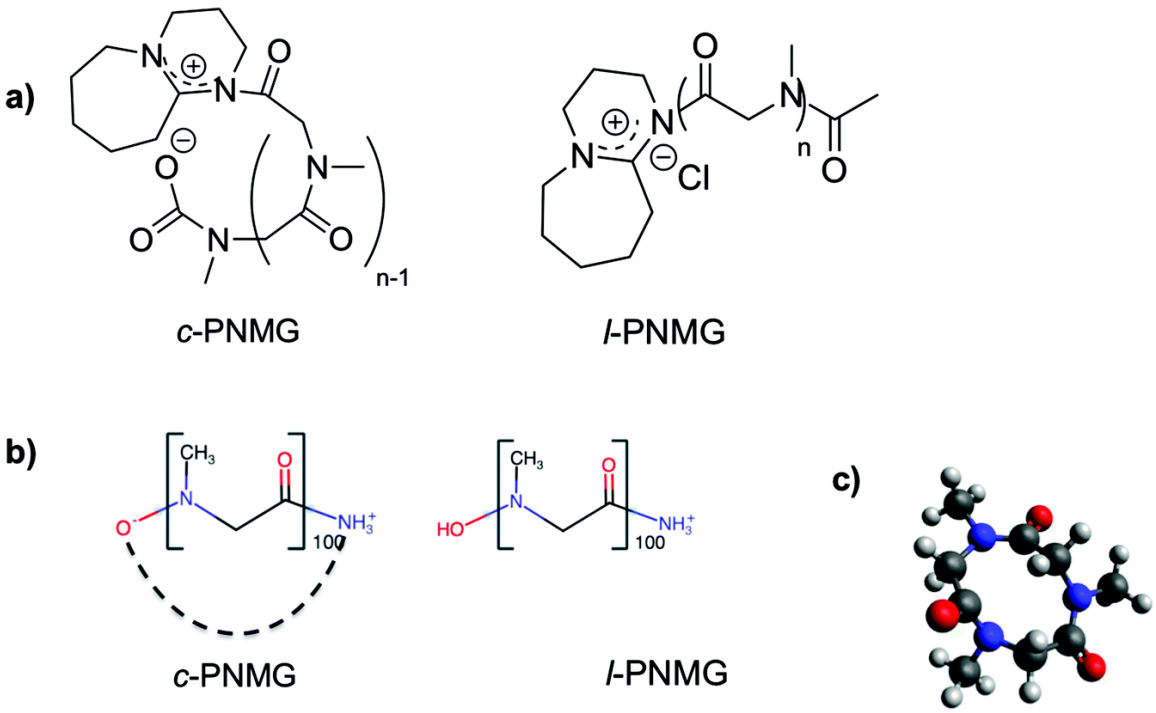

Figure 2.1. (a) Sketch of the experimental cyclic polypeptoid structure c-PNMG (DP=100) (left) and the linear polypeptoid structure l-PNMG (DP=100) (right). (b) Sketch of the simplified cyclic and linear peptoids simulated in this work. The cyclic polypeptoid is zwitterion with negatively charged -O group at one end and positively charged -NH_3 at the other that very close together forming a cyclic peptoid in methanol solution, while the linear case has a single positively charged -NH_3 group. (c) The representative peptoid cluster used to determine the partial charges. Red spheres are used for oxygen atoms, blue for nitrogen, grey for carbon and white for hydrogen.

2.2.1 Electronic Structure Calculations

The partial charges on the atomic sites of the polymer and methanol were determined by carrying out electronic structure calculations on two representative clusters: the first consisting of three monomeric units (Figure 2.1c) in order to represent the polymer and the second consisting of a single methanol molecule to determine the charges for the methanol. Specifically, the CHELPG method was used,[42] which determines the charges by fitting to the *ab initio* electrostatic potential at specific grid points around the cluster. The final charges on the polymer were obtained by averaging over the three monomeric units in the cluster. These calculations were carried out using Density Functional Theory approach with the B3LYP functional and the 6-31G basis set.

These calculations were performed using the GAUSSIAN09 suite of programs.[43] The final charges are tabulated in table 4.1.

Table 2.1. The partial charges obtained from the CHELPG electronic structure methods using the B3LYP functional and the 6-31G basis set. Note that the terminal charges were not calculated using this method but were directly assigned values of 1 and -1.

Type	Charge (e)	Type	Charge (e)
Methanol O	-0.72	Backbone N	-0.22
Methanol H	0.44	Backbone C	-.02
Methanol CH_3	0.28	Side chain C	-.26
Amide C	0.5	Terminal NH_{3+}	1.0
Amide O	-.05	Terminal O^-	-1.0
Hydrogen	0.1	Chloride ion	-1.0

2.2.2 Molecular Dynamics Simulations

Single polypeptoid chain. The initial configurations for each of the two types of polypeptoids were generated by connecting 100 monomers, each with a methyl side chain. In the case of the cyclic polypeptoid, the starting configuration had the two charged ends in proximity to each other, mimicking the compound obtained from synthesis while for the linear counterpart, the starting structure consisted of the monomeric units arranged in a line.[44] The OPLSAA force field parameters for the Lennard-Jones and intramolecular interactions were used along with the partial charges determined from electronic structure calculations for the peptoids.[45,46] The same dihedral parameters that are used in the peptide case for the planarity of the amide bond were used here. The OPLS based force-field has previously been used successfully to model peptoids with .[47] A short 50 ps simulation of the l-$PNMG_{100}$ was carried out in vacuum at a temperature of 300 K in the NVT ensemble in a cubic box of length 100 Å, and similarly for the c-$PNMG_{100}$ polymer. The resulting l-$PNMG_{100}$ with a chloride counter anion and c-$PNMG_{100}$ were each solvated, using the packmol program,[48] in a periodic cubic box of length 100 Å, consisting of the same number (14500) of methanol (represented using the united atom OPLS force-field)[45] molecules in each case. Each system was equilibrated for 10 ns, with a 1 fs timestep, in the canonical (NVT) ensemble at a temperature of 300 K using a NoseHoover thermostat. Production runs of 50 ns

each were then carried out for each peptoid system. A snapshot from the simulation of the cyclic polymer case is shown in Figure 2.2a.

Two polypeptoid chains. The initial configuration for the linear and cyclic cases were generated by replicating the respective single chain boxes in the x-direction resulting in a periodic box with two polypeptoid chains, and 29000 methanol molecules in a box of length 200 Å, X 100 Å, X 100 Å. Each system was once again equilibrated for 10 ns, with a 1 fs time step in the NVT ensemble followed by a 50 ns production run. A snapshot from the simulation of the cyclic polymer case is shown in Figure 2.2b.

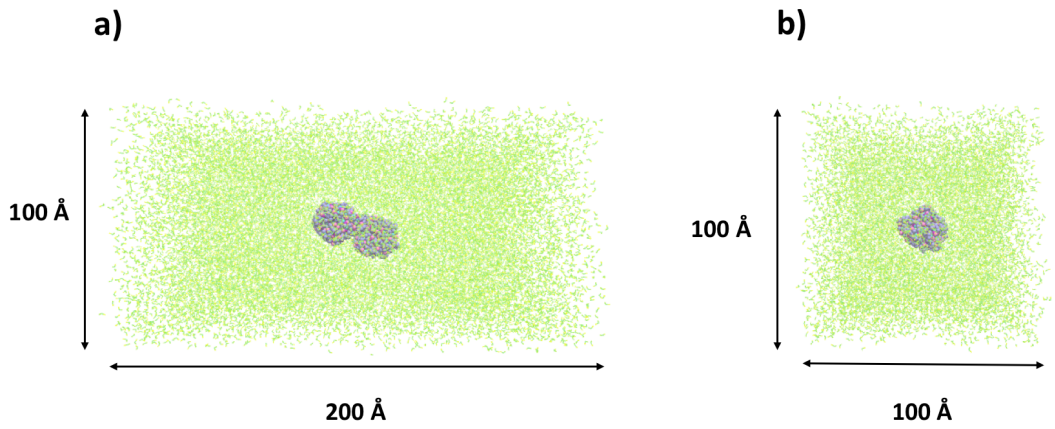

Figure 2.2. Representative snapshot of the simulation box with (a) just one cyclic polypeptoid (c-PNMG) in a box of methanol solvent (b) two c-PNMG in methanol. The methanol solvent is shown in green.

2.2.3 Umbrella Sampling Simulations for the Two Polypeptoid Chain Case

Umbrella sampling along with the weighted histogram analysis method (WHAM) was used to determine the free energy profile or potential of mean force (PMF) of bringing the center of mass of one c-PNMG$_{100}$ /l-PNMG$_{100}$ monomer to the center of mass of another c-PNMG$_{100}$ /l-PNMG$_{100}$ monomer in methanol solution.[49,50] A total of sixteen umbrella sampling windows ranging from an inter-polymer distance of 50 *Å* to 22 *Å*, were carried out for the linear case, each with a ten nanosecond equilibration run followed by a 40 ns production run. A harmonic force constant of 3.0 kcal mol^{-1} was used for the 22 and 23 Å, windows and 2.0 kcal mol^{-1} for the remaining windows. For the linear case a total of eighteen umbrella sampling windows were

used with a harmonic force constant of 3.0 kcal mol^{-1} for the 20, 22 and 23 Å, windows and 2.0 kcal mol^{-1} for the remaining windows The errors bars for the PMFs were obtained using block averaging, with each block from a 10 ns run.

All simulations, both canonical and umbrella sampling, were carried using the LAMMPS package.[51]

2.3 Results and Discussion

To further investigate the mechanism and driving force of the cluster formation, we performed MD simulations on the cyclic and linear PNMG molecules in methanol solution with both DP = 100. First, as shown in the left panel of Figure 2.3, the peptoid form factors of a cyclic (green line) and a linear (red line) PNMG molecule were calculated from the simulation trajectory, with each atom weighted by its neutron scattering length.[52] It can be seen that the Guinier region of the scattering profile for a cyclic molecule extends more towards high Q than a linear one, which indicates a smaller overall size than its linear analog, and it agrees with the theoretical expectation.[53] In the intermediate and high Q region ($Q > 0.1$ $Å^{-1}$), the scattering profile of a cyclic homopolymer decays faster than a linear one, which means the scatterers in a cyclic polymer are more densely packed like compact colloids rather than loose coils given by a linear polymer. The bump in the scattering profile of a cyclic molecule around $Q = 0.4$ $Å^{-1}$ suggests that the interface between the polymeric materials and the solvent is relatively well-defined and spherical, compared to the irregular interface for a linear one. The scattering function of a dimer of two cyclic polymers is also calculated, and its Guinier region shifts towards low Q, corresponding to a larger size as expected. Compared to the experimental data in the right panel of Figure 2.3, where the SANS profile of the cyclic polymers has the Guinier region shifted towards low Q, it confirms that the solution of cyclic polymers may actually contain dimers or even oligomers. This then raises the question as to what the driving force is that causes the aggregation in the case of the cyclic peptoid and not the linear case. The following sections address this important question.

The process to form a dimer in the cyclic case can be described using the free energy profile, also called the potential of mean force (PMF), between two interacting molecules as a function

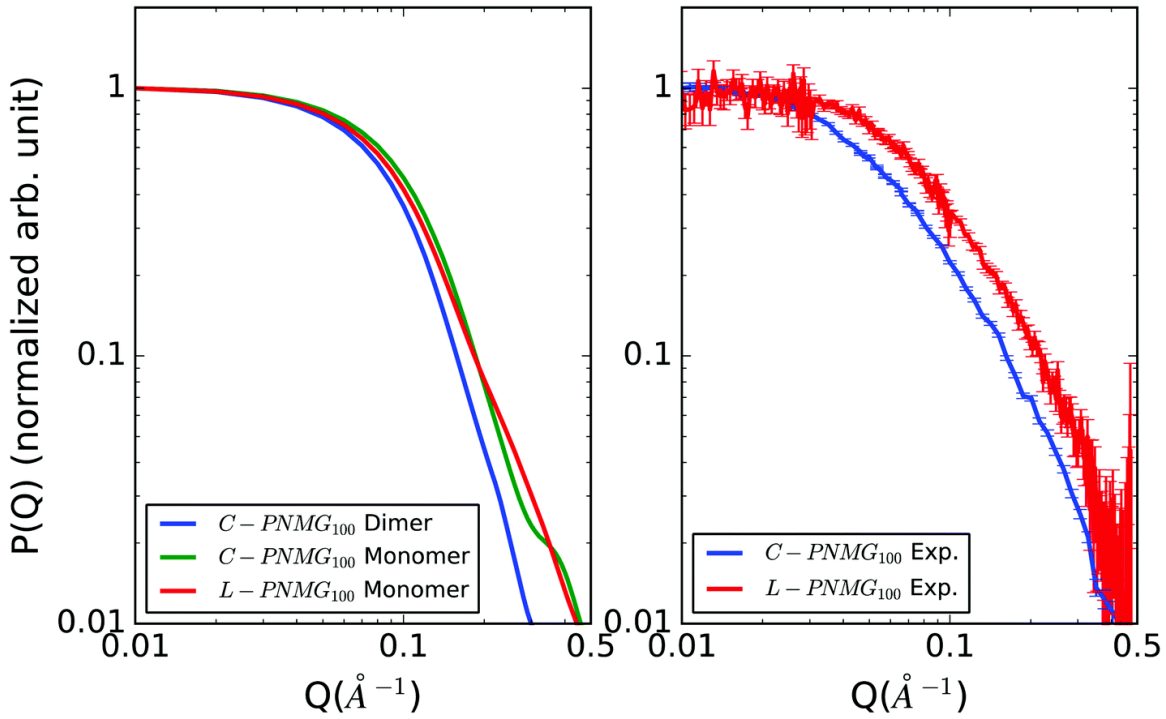

Figure 2.3. Form factors of the PNMG polypeptoids (targeted DP=100) from simulation (left) compared to the experimental small-angle neutron scattering (SANS) data (right). In the left panel (simulation) the linear PNMG polypeptoid is shown in red, the cyclic monomeric version is shown in green and the cyclic dimer in blue. In the right panel (experimental data) the linear PNMG is shown in red and the cyclic PNMG in blue. Note that the curves in the right panel were obtained from the measured intensity in experiments from which the incoherent scattering has been subtracted, and the coherent scattering is normalized by its extrapolated zero-angle scattering, to make it easier to compare with the calculated form factor from MD simulation.

of the center of mass separation. The use of the center of mass distance as the collective variable is appropriate as it allows for the study of the two polymers as they approach each other without biasing the polymers in any specific orientation. Using umbrella sampling, we obtained the PMF of two cyclic PNMG polymers with DP = 100 in methanol solution, as shown in Figure 2.4a. At distances greater than 30 Å, the PMF is essentially constant, within the statistical error. At a separation of about 24 Å, there is a free energy minimum of around 0.5 kcal mol^{-1} , indicating an attractive region, which may come from the dipoledipole interaction or solvophobic interaction, and which in turn drives the formation of a dimer. Overall, the statistical error of the PMF is comparable to the energy barrier and the depth of the relative free energy and is of the order

of k_BT, which means both the formation and dissociation of dimers can occur at a fast rate and there exists an equilibrium between the two. Therefore, in solution, monomer, dimers, and oligomers should exist, and indeed this agrees with our experimental measurements. On the other hand, the PMF or free energy profile for association of the linear counterpart as a function of the distance between the center of masses of the two polymers (see Figure 2.4b) shows no minimum and in fact the free energy increases as the two come closer. This suggests that the linear polymer can at best form transient aggregate species in dilute solution.

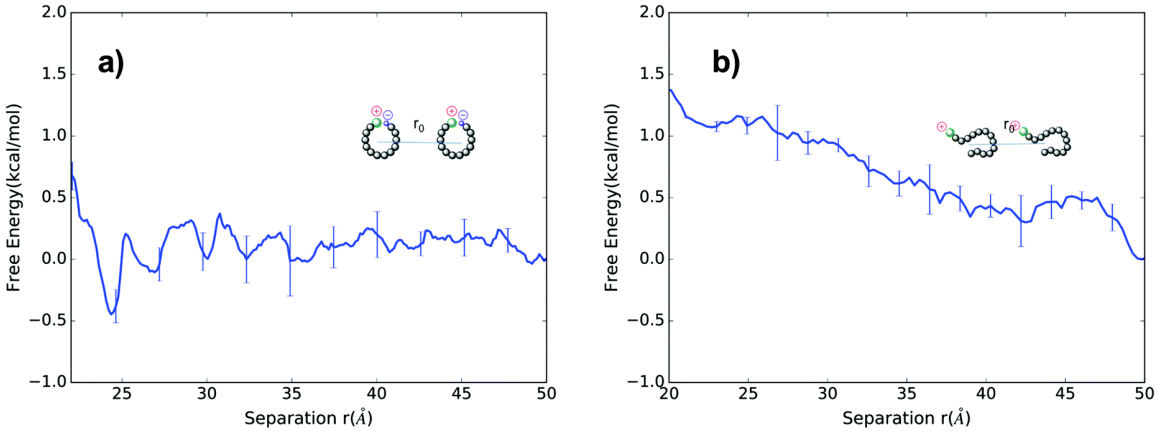

Figure 2.4. The free energy profile as a function of the center of mass separation between (a) two c-PNMG polymer chains (DP = 100) and (b) two l-PNMG polymer chains (DP = 100).

To further study the dissociation of solvent molecules from the cyclic polymers, we calculated the radial distribution function of the oxygen atom of methanol to the oxygen of the carbonyl group on the PNMG backbone, in Figure 2.5. The very sharp peak with a minimum at r = 3.3 Å, corresponds to the first solvation shell and the small peak at r = 7 Å, to the second solvation shell. Hence the first solvation shell coordination number is defined as the number of solvent molecules with the O atom of its hydroxyl group within this cuto distance of r = 3.3 Å, from the O atom of the corresponding carbonyl group of the polymer. On the basis of the spatial distribution of solvent molecules around the polymer, we can estimate the solvation level of cyclic PNMG molecules during the dimer formation, which is later used to demonstrate the solvophobic interaction. We define a monomer in the PNMG polymer as "coordinated" to

solvent molecules if there is at least one oxygen atom from the hydroxyl group in methanol located within 3.3 Å, from the oxygen in the carbonyl group of the monomer. The percentage of the coordinated monomers in the whole polymer chain indicates the level of solvation of the PNMG. This coordination percentage is plotted as a function of the center of mass separation distance during dimer formation, as shown in Figure 2.6. When two PNMG molecules are approaching each other but still separated ($r > 35$ Å), the coordination percentage gradually and smoothly decreases, which may be due to the rearrangement of the first solvation shell influenced by the hydrodynamic interaction through the excluded solvent molecules. In the range of 22 Å, $< r < 35$ Å, the coordination percentage shows a plateau between 25 Å, and 35 Å, and then a drop from $r = 25$ Å, to 22 Å. Combined with the PMF, a clear picture emerges regarding the aggregation process. When two cyclic PNMG molecules approach each other within 35 Å, the orientation of methanol molecules located in the first solvation shell between the two polymers are rearranged, to minimize the free energy, which give the constant coordination percentage. Until their separation distance reaches $r = 25$ Å, the first solvation shell between the two PNMG polymers starts to get fully excluded and the coordination percentage approaches its minimum value. It is worth noting that, even at the shortest separation distance, more than half of the monomers in PNMG located in the outer region of the dimer are still exposed to the surrounding solvent molecules, and only the monomers in the middle of the two PNMG molecules are solvent-excluded. Even in the form of separated molecules, a fraction of the monomers may be shielded by other ones from being solvated. Therefore, it is reasonable that the coordination percentage is reducing from 80.5% to 67.5%.

In our zwitterionic cyclic polymer system, solvophobic interactions, demonstrated in the coordination percentage curve (Figure 2.6), plays an important role in the final structure of the aggregate, in addition to the dipoledipole interaction, with the dipole created by the oppositely charged ends of the cyclic polymer. To further reveal the mechanism of dimer formation and the interplay between these forces, we calculated the probability distribution of the dipoledipole distance at different separation distances between the two PNMG molecules, presented in Figure

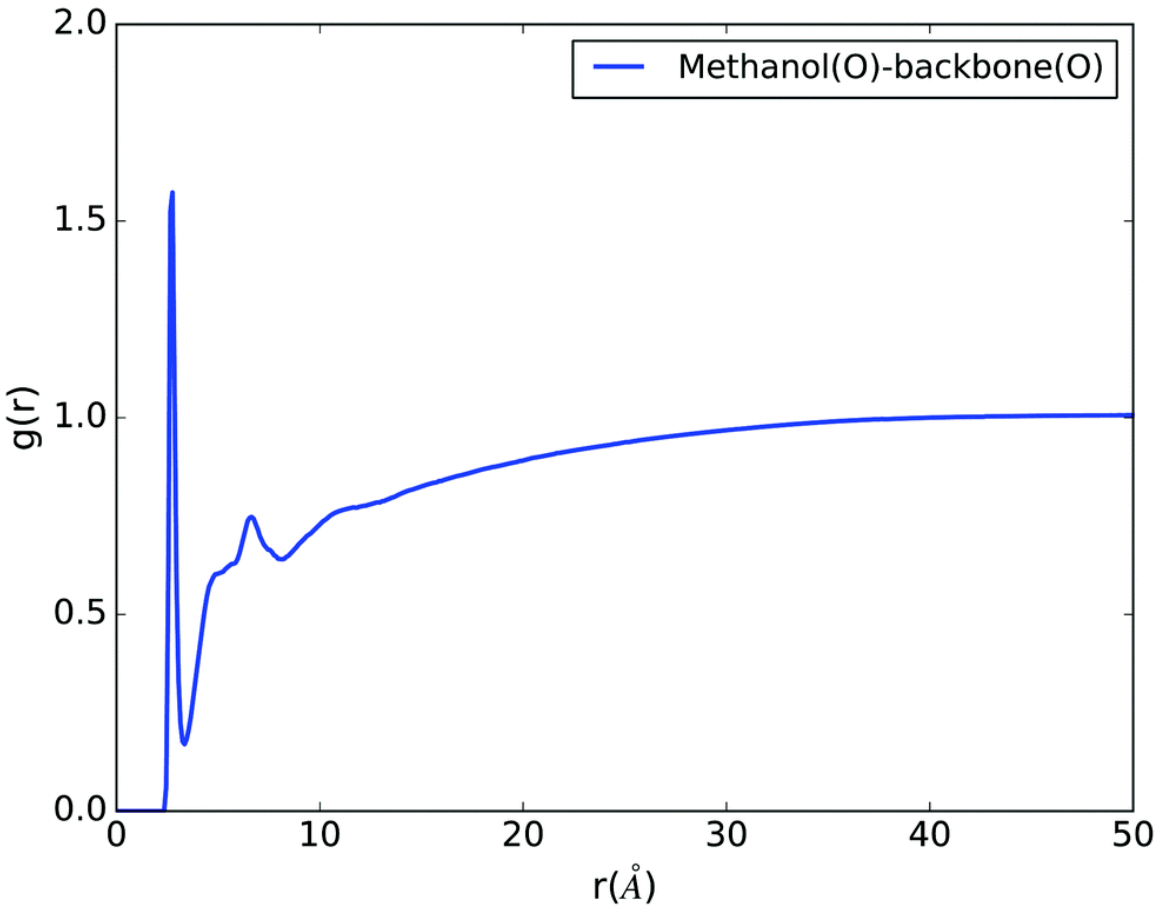

Figure 2.5. The radial distribution function, g(r), where r is the distance between the oxygen atoms of the polymer backbone and the oxygen atoms of the methanol solvent from MD simulation at T = 300 K. The first minimum is at 3.3 Å.

2.7. When the dipoledipole interaction is stronger than the solvophobic interaction, the dipoles prefer to minimize their distance to reduce the free energy, and as a result, the two molecules will rotate with the dipoles towards each other. However, the dipole shows much greater preference to be solvated than the backbone. When the dipoles are rotated close to each other, they will also tend to bring solvent molecules between the two PNMG molecules. This will counteract the solvophobic interaction, resulting in an increase of the free energy due to the excluded volume effect. Therefore, the conformation of the dimer is determined by the competition between the attractive dipoledipole and solvophobic interaction and the effectively repulsive interaction arising from the solvation of the dipoles, and it is apparent in the dipoledipole distance distribution. A configuration with the dipoledipole distance smaller than the center of mass separation distance

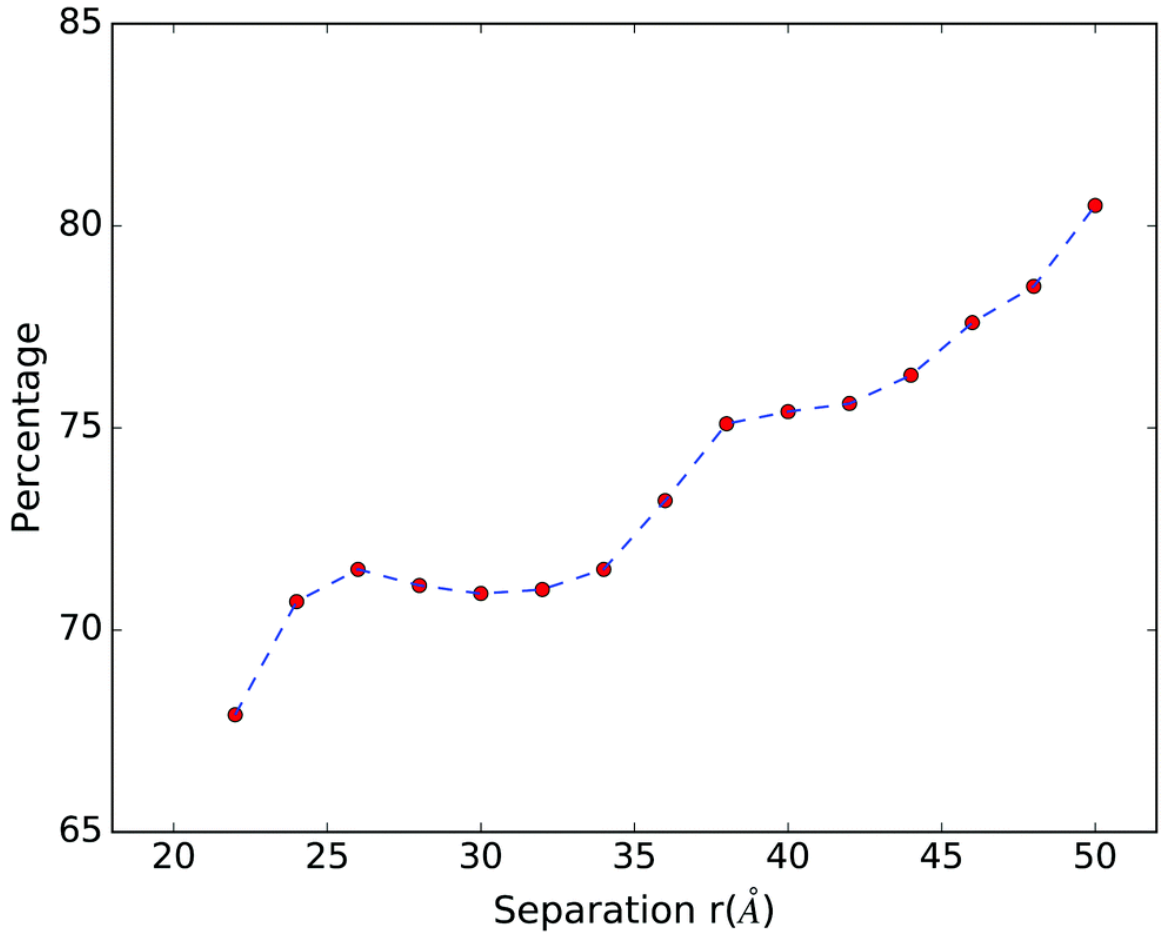

Figure 2.6. The percentage of O atoms on the polymer backbones that is coordinated to methanol as a function of the distance between the center of mass of the two c-PNMG monomers (DP=100).

indicates that the dipoledipole interaction dominates in this conformation. On the other hand, when the dipoledipole distance is larger than the center of mass distance, the solvophobic interaction coupled with dipole solvation is the dominant force in determining the conformation of the dimer. The dipoledipole distance probability distributions at four representative center of mass separation distances are presented in Figure 2.7. Representative snapshots of the dimer conformation, from umbrella sampling simulations, corresponding to each probability peak in Figure 2.7 are shown in Figure 2.8, with the dipoles highlighted. It can be seen that for the snapshot from the umbrella sampling window with $r_0 = 36$ Å, (cyan curve) a majority rests in the range above r_0 and only a small part appears in the range below r_0, which indicates that at

this separation distance dipoledipole interactions do not play an important role. However, at r_0 = 32 Å, (red curve), the majority is below r_0, which means that dipoledipole interactions dominate in determining the conformation. At r_0 = 28 Å, (green curve), the conclusion is similar. At r_0 = 24 Å, (blue curve), we see that all events are located in the region above r_0 , which implies that the solvophobic interaction dominates and the methanol between the PNMG molecules has been fully excluded, with the result that the dipoles are rotated opposite to each other in order to achieve a better solvation environment. The dipoles, open to the solvent at this inter-polymer distance, along with the hydrophobic and dipoledipole interactions will all contribute to further clustering to form trimers and other oligomers. From the inter-polymer distance dependence of these probability distribution functions, we can directly observe the competition between these driving forces during dimer formation and their effect on the conformation. Lastly, the need for the dipoles to be solvated suggests a reason for the formation of only small oligomers. With increasing cluster size, some of the dipoles will inevitably be buried in the interior region of the cluster, thus losing their solvation shell, which is unfavorable and will thereby restrict the aggregation to fairly small cluster sizes. To reiterate, the formation and structure of the cyclic dimer is driven by both the solvophobic effect and the dipoledipole interactions and opposed by the solvation of the dipoles, with the different interactions dominating at different inter-polymer distances. In the case of the linear counterpart there is no dipoledipole interaction to form a stable dimer and furthermore the charged ends prefer to be solvated, resulting in the free energy increasing continuously with decreasing distance.

At this point it is worth noting that the experimental data for the cyclic polypeptoids can be thought of as arising from association of several macro zwitterions *i.e.* the positive end of one polymer aggregates with the negative end of another and so on. The simulations were carried out for the cyclic polymer with a starting configuration with the two charged ends of the polypeptoid in proximity to each other, since this mimics the polypeptoid architecture that was obtained from the synthesis. The center of mass coordinate based collective variable should not bias the simulations against the opening of the cyclic polypeptoid chain. We do not see any ring

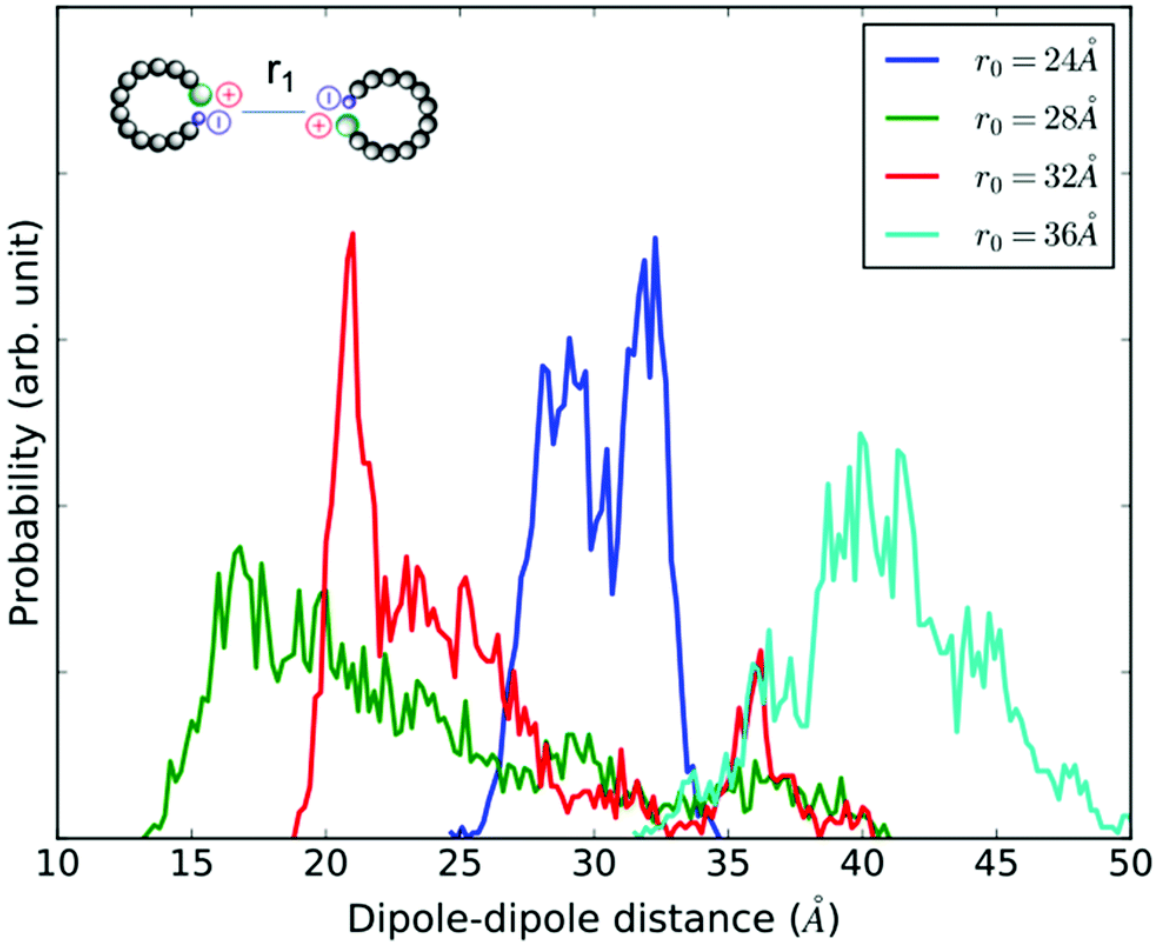

Figure 2.7. Dipole-dipole distance distribution as a function of the distance between the center of mass of the two c-PNMG monomers (DP=100). The curves are normalized to have unit area.

opening events in the simulations. This behavior of the cyclic polypeptoids in methanol arises because the comparatively low dielectric constant of methanol does not allow for the opening of the dipole in the cyclic zwitterionic polypeptoids. This, in turn, prevents the charged ends of each zwitterionic polypeptoid from aggregating with the oppositely charged ends of other zwitterionic polypeptoids in solution. In future work, the decomposition of the free energy into enthalpic and entropic components as well as the further decomposition of the enthalpic term into dierent intermolecular interactions will be carried out.

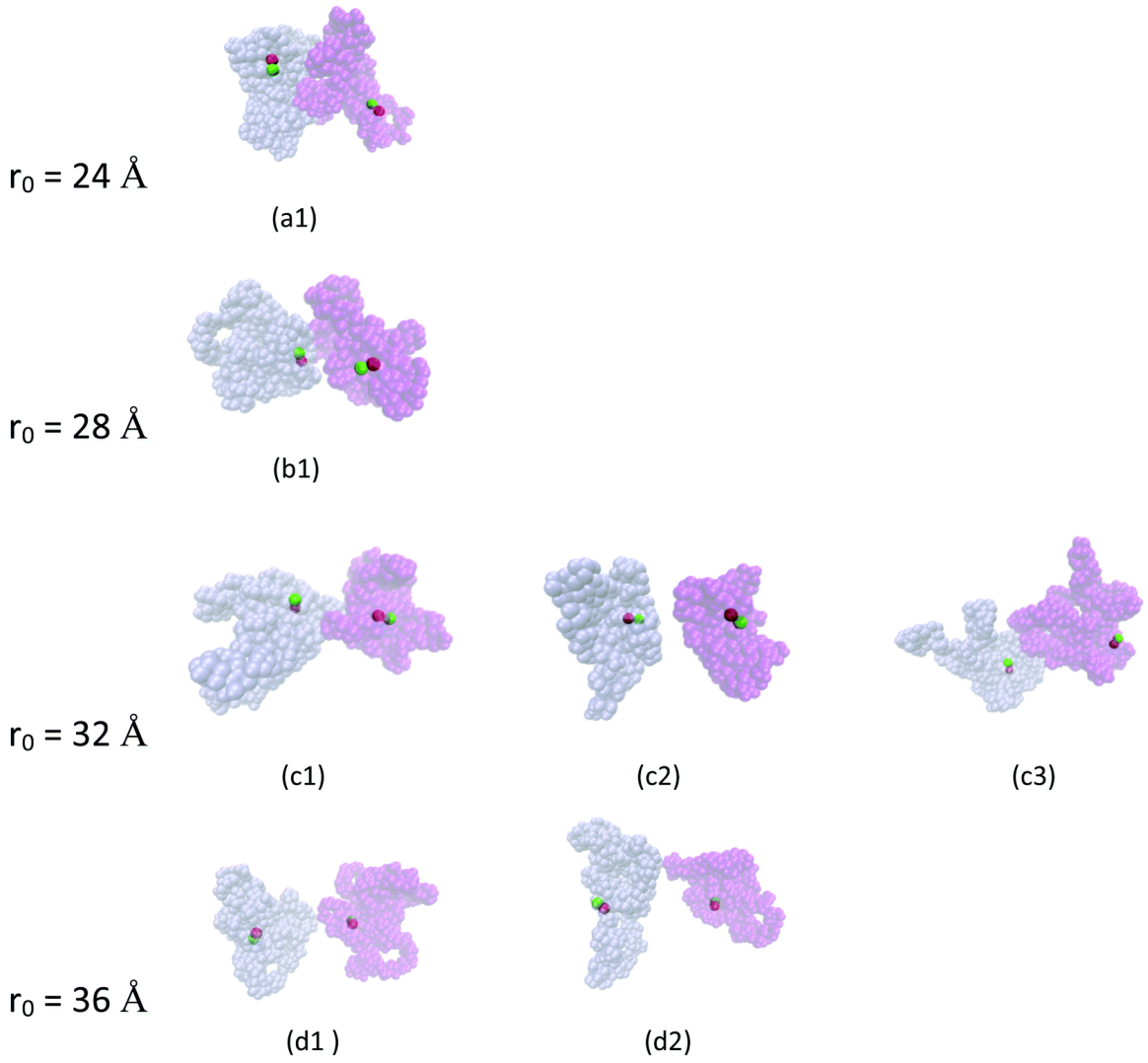

Figure 2.8. Snapshots corresponding to the peaks in the dipoledipole distribution (Figure 4.6) from different umbrella sampling windows (r0 = 24, 28, 32 and 36 Å) of the two c-PNMG monomers (DP=100) (one shown in gray and the other in magenta). In the case of r_0 = 32 Å, three distinct peaks are seen and hence three snapshots are shown, (c1) corresponding to the largest peak, (c2) to the second largest and (c3) to the third peak. Similarly for r_0 = 36 Å, two peaks are seen and (d1) corresponds to the largest peak and (d2) to the second largest peak. The positive and negative ends of each polymer are represented by a red and green sphere, respectively. The methanol solvent molecules are not shown for clarity.

2.4 Conclusions

Several computational methods have been used to study the aggregation of cyclic and linear polypeptoids in methanol, and compared with experimental results. Analyses of the trajectories from the simulations indicate that the dipoledipole interaction along with the attractive solvophobic effect (expulsion of the solvating molecules between the two cyclic polypeptoid chains) favor the dimerization whereas the effective repulsive interaction arising due to the solvation of the dipoles greatly reduces the attractive component in aggregation. The resulting cluster now consists of two dipoles that are exposed to the solvent allowing for further aggregation. However, the repulsive forces due to the solvation of the dipoles suggests that the aggregation is limited to small clusters as larger clusters would necessitate the burial, to some extent, of the dipoles in the interior. This is corroborated by the experimental data that shows the presence of monomeric to small oligomeric clusters.

2.5 References

[1] Pu Du, Ang Li, Xin Li, Yueheng Zhang, Changwoo Do, Lilin He, Steven W Rick, Vijay T John, Revati Kumar, and Donghui Zhang, "Aggregation of cyclic polypeptoids bearing zwitterionic end-groups with attractive dipole–dipole and solvophobic interactions: A study by small-angle neutron scattering and molecular dynamics simulation", *Physical Chemistry Chemical Physics* **19**(22), pp. 14388–14400 (2017).

[2] Francesco Sciortino, Piero Tartaglia, and Emanuela Zaccarelli, "One-dimensional cluster growth and branching gels in colloidal systems with short-range depletion attraction and screened electrostatic repulsion", *The Journal of Physical Chemistry B* **109**(46), pp. 21942–21953 (2005).

[3] P Douglas Godfrin, Néstor E Valadez-Pérez, Ramon Castaneda-Priego, Norman J Wagner, and Yun Liu, "Generalized phase behavior of cluster formation in colloidal dispersions with competing interactions", *Soft matter* **10**(28), pp. 5061–5071 (2014).

[4] Anna Stradner, Helen Sedgwick, Frédéric Cardinaux, Wilson CK Poon, Stefan U Egelhaaf, and Peter Schurtenberger, "Equilibrium cluster formation in concentrated protein solutions and colloids", *Nature* **432**(7016), pp. 492 (2004).

[5] B-Y Ha and AJ Liu, "Kinetics of bundle growth in dna condensation", *EPL (Europhysics Letters)* **46**(5), pp. 624 (1999).

[6] H Sedgwick, K Kroy, A Salonen, MB Robertson, SU Egelhaaf, and WCK Poon, "Non-equilibrium behavior of sticky colloidal particles: beads, clusters and gels", *The European Physical Journal E* **16**(1), pp. 77–80 (2005).

[7] Clare J Dibble, Michael Kogan, and Michael J Solomon, "Structure and dynamics of colloidal depletion gels: Coincidence of transitions and heterogeneity", *Physical Review E* **74**(4), pp. 041403 (2006).

[8] Andrew I Campbell, Valerie J Anderson, Jeroen S van Duijneveldt, and Paul Bartlett, "Dynamical arrest in attractive colloids: The effect of long-range repulsion", *Physical review letters* **94**(20), pp. 208301 (2005).

[9] Guangnan Meng, Natalie Arkus, Michael P Brenner, and Vinothan N Manoharan, "The free-energy landscape of clusters of attractive hard spheres", *Science* **327**(5965), pp. 560–563 (2010).

[10] Liang Hong, Angelo Cacciuto, Erik Luijten, and Steve Granick, "Clusters of charged janus spheres", *Nano letters* **6**(11), pp. 2510–2514 (2006).

[11] Francesco Sciortino, Stefano Mossa, Emanuela Zaccarelli, and Piero Tartaglia, "Equilibrium cluster phases and low-density arrested disordered states: the role of short-range attraction and long-range repulsion", *Physical review letters* **93**(5), pp. 055701 (2004).

[12] Peter J Lu, Jacinta C Conrad, Hans M Wyss, Andrew B Schofield, and David A Weitz, "Fluids of clusters in attractive colloids", *Physical Review Letters* **96**(2), pp. 028306 (2006).

[13] Bianca M Mladek, Dieter Gottwald, Gerhard Kahl, Martin Neumann, and Christos N Likos, "Formation of polymorphic cluster phases for a class of models of purely repulsive soft spheres", *Physical review letters* **96**(4), pp. 045701 (2006).

[14] Peter Schurtenberger, Richard A Chamberlin, George M Thurston, John A Thomson, and George B Benedek, "Observation of critical phenomena in a protein-water solution", *Physical review letters* **63**(19), pp. 2064 (1989).

[15] AA Louis, PG Bolhuis, JP Hansen, and EJ Meijer, "Can polymer coils be modeled as soft colloids?", *Physical review letters* **85**(12), pp. 2522 (2000).

[16] Martin Konieczny, Christos N Likos, and Hartmut Löwen, "Soft effective interactions between weakly charged polyelectrolyte chains", *The Journal of chemical physics* **121**(10), pp. 4913–4924 (2004).

[17] CN Likos, M Schmidt, H Löwen, M Ballauff, D Pötschke, and P Lindner, "Soft interaction between dissolved flexible dendrimers: theory and experiment", *Macromolecules* **34**(9), pp. 2914–2920 (2001).

[18] A Jusufi, M Watzlawek, and H Löwen, "Effective interaction between star polymers", *Macromolecules* **32**(13), pp. 4470–4473 (1999).

[19] CN Likos, H Löwen, M Watzlawek, B Abbas, O Jucknischke, J Allgaier, and D Richter, "Star polymers viewed as ultrasoft colloidal particles", *Physical review letters* **80**(20), pp. 4450 (1998).

[20] CN Likos, S Rosenfeldt, N Dingenouts, M Ballauff, P Lindner, N Werner, and F Vögtle, "Gaussian effective interaction between flexible dendrimers of fourth generation: A theoretical and experimental study", *The Journal of chemical physics* **117**(4), pp. 1869–1877 (2002).

[21] Marco Bernabei, Petra Bacova, Angel J Moreno, Arturo Narros, and Christos N Likos, "Fluids of semiflexible ring polymers: effective potentials and clustering", *Soft Matter* **9**(4), pp. 1287–1300 (2013).

[22] Mohammed Zakaria Slimani, Petra Bacova, Marco Bernabei, Arturo Narros, Christos N Likos, and Angel J Moreno, "Cluster glasses of semiflexible ring polymers", *ACS macro letters* **3**(7), pp. 611–616 (2014).

[23] Boualem Hammouda, "The mystery of clustering in macromolecular media", *Polymer* **50**(22), pp. 5293–5297 (2009).

[24] Boualem Hammouda, Derek L Ho, and Steve Kline, "Insight into clustering in poly (ethylene oxide) solutions", *Macromolecules* **37**(18), pp. 6932–6937 (2004).

[25] B-Y Ha and Andrea J Liu, "Charge oscillations and many-body effects in bundles of like-charged rods", *Physical Review E* **58**(5), pp. 6281 (1998).

[26] Mark L Henle and Philip A Pincus, "Equilibrium bundle size of rodlike polyelectrolytes with counterion-induced attractive interactions", *Physical Review E* **71**(6), pp. 060801 (2005).

[27] P-ÉA Melchy and MH Eikerling, "Physical theory of ionomer aggregation in water", *Physical Review E* **89**(3), pp. 032603 (2014).

[28] Anuj Shukla, Efstratios Mylonas, Emanuela Di Cola, Stephanie Finet, Peter Timmins, Theyencheri Narayanan, and Dmitri I Svergun, "Absence of equilibrium cluster phase in concentrated lysozyme solutions", *Proceedings of the National Academy of Sciences* **105**(13), pp. 5075–5080 (2008).

[29] IO Götze, HM Harreis, and CN Likos, "Tunable effective interactions between dendritic macromolecules", *The Journal of chemical physics* **120**(16), pp. 7761–7771 (2004).

[30] Andrew J Archer and Nigel B Wilding, "Phase behavior of a fluid with competing attractive and repulsive interactions", *Physical Review E* **76**(3), pp. 031501 (2007).

[31] Donghui Zhang, Samuel H Lahasky, Li Guo, Chang-Uk Lee, and Monika Lavan, "Polypeptoid materials: current status and future perspectives", *Macromolecules* **45**(15), pp. 5833–5841 (2012).

[32] Niklas Gangloff, Juliane Ulbricht, Thomas Lorson, Helmut Schlaad, and Robert Luxenhofer, "Peptoids and polypeptoids at the frontier of supra-and macromolecular engineering", *Chemical reviews* **116**(4), pp. 1753–1802 (2015).

[33] Hannah K Murnen, Adrianne M Rosales, Andrey V Dobrynin, Ronald N Zuckermann, and Rachel A Segalman, "Persistence length of polyelectrolytes with precisely located charges", *Soft Matter* **9**(1), pp. 90–98 (2013).

[34] Adrianne M Rosales, Hannah K Murnen, Steven R Kline, Ronald N Zuckermann, and Rachel A Segalman, "Determination of the persistence length of helical and non-helical polypeptoids in solution", *Soft Matter* **8**(13), pp. 3673–3680 (2012).

[35] Susan M Miller, Reyna J Simon, Simon Ng, Ronald N Zuckermann, Janice M Kerr, and Walter H Moos, "Proteolytic studies of homologous peptide and n-substituted glycine peptoid oligomers", *Bioorganic & medicinal chemistry letters* **4**(22), pp. 2657–2662 (1994).

[36] Li Guo, Jinhai Li, Zachary Brown, Kushal Ghale, and Donghui Zhang, "Synthesis and characterization of cyclic and linear helical poly (α-peptoid) s by n-heterocyclic carbene-mediated ring-opening polymerizations of n-substituted n-carboxyanhydrides", *Peptide Science* **96**(5), pp. 596–603 (2011).

[37] Philippe Armand, Kent Kirshenbaum, Richard A Goldsmith, Shauna Farr-Jones, Annelise E Barron, Kiet TV Truong, Ken A Dill, Dale F Mierke, Fred E Cohen, Ronald N Zuckermann, et al., "Nmr determination of the major solution conformation of a peptoid pentamer with chiral side chains", *Proceedings of the National Academy of Sciences* **95**(8), pp. 4309–4314 (1998).

[38] Kent Kirshenbaum, Annelise E Barron, Richard A Goldsmith, Philippe Armand, Erin K Bradley, Kiet TV Truong, Ken A Dill, Fred E Cohen, and Ronald N Zuckermann, "Sequence-specific polypeptoids: a diverse family of heteropolymers with stable secondary structure", *Proceedings of the National Academy of Sciences* **95**(8), pp. 4303–4308 (1998).

[39] Chang-Uk Lee, Ang Li, Kushal Ghale, and Donghui Zhang, "Crystallization and melting behaviors of cyclic and linear polypeptoids with alkyl side chains", *Macromolecules* **46**(20), pp. 8213–8223 (2013).

[40] Adrianne M Rosales, Hannah K Murnen, Ronald N Zuckermann, and Rachel A Segalman, "Control of crystallization and melting behavior in sequence specific polypeptoids", *Macromolecules* **43**(13), pp. 5627–5636 (2010).

[41] Corinna Fetsch and Robert Luxenhofer, "Thermal properties of aliphatic polypeptoids", *Polymers* **5**(1), pp. 112–127 (2013).

[42] Curt M Breneman and Kenneth B Wiberg, "Determining atom-centered monopoles from molecular electrostatic potentials. the need for high sampling density in formamide conformational analysis", *Journal of Computational Chemistry* **11**(3), pp. 361–373 (1990).

[43] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, . Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, "Gaussian09 Revision D.01", Gaussian Inc. Wallingford CT 2009.

[44] Ang Li, Lu Lu, Xin Li, LiLin He, Changwoo Do, Jayne C Garno, and Donghui Zhang, "Amidine-mediated zwitterionic ring-opening polymerization of n-alkyl n-carboxyanhydride: mechanism, kinetics, and architecture elucidation", *Macromolecules* **49**(4), pp. 1163–1171 (2016).

[45] William L Jorgensen and Julian Tirado-Rives, "The opls [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin", *Journal of the American Chemical Society* **110**(6), pp. 1657–1666 (1988).

[46] William L Jorgensen, David S Maxwell, and Julian Tirado-Rives, "Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids", *Journal of the American Chemical Society* **118**(45), pp. 11225–11236 (1996).

[47] Sung Hyun Park and Igal Szleifer, "Structural and dynamical characteristics of peptoid oligomers with achiral aliphatic side chains studied by molecular dynamics simulation", *The Journal of Physical Chemistry B* **115**(37), pp. 10967–10975 (2011).

[48] Leandro Martínez, Ricardo Andrade, Ernesto G Birgin, and José Mario Martínez, "Packmol: a package for building initial configurations for molecular dynamics simulations", *Journal of computational chemistry* **30**(13), pp. 2157–2164 (2009).

[49] Benoît Roux, "The calculation of the potential of mean force using computer simulations", *Computer physics communications* **91**(1-3), pp. 275–282 (1995).

[50] Marc Souaille and Benoıt Roux, "Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations", *Computer physics communications* **135**(1), pp. 40–57 (2001).

[51] Steve Plimpton, "Fast parallel algorithms for short-range molecular dynamics", *Journal of computational physics* **117**(1), pp. 1–19 (1995).

[52] Yi Liu, Vyacheslav S Bryantsev, Mamadou S Diallo, and William A Goddard Iii, "Pamam dendrimers undergo ph responsive conformational changes without swelling", *Journal of the*

American Chemical Society **131**(8), pp. 2798–2799 (2009).

[53] J Anthony Semlyen, *Cyclic polymers*, Springer (2000).

CHAPTER 3
MOLECULAR DYNAMICS STUDIES OF SELF-ASSEMBLED SEQUENCE-DEFINED SINGLY CHARGED POLYPEPTOIDS

3.1 Introduction

Polypeptoids, or poly N-substituted glycines, are a class of peptidomimetic polymers that are highly tunable, and hence an ideal model system to study self-assembly as a function of chemical groups in soft matter systems.[1,2] Polypeptoids are promising bioinspired materials for diverse applications ranging from drug carriers[3], antifouling coatings[4], antimicrobial agents[5], and other therapeutics[6] due to the hundreds of chemically diverse side chains.[7–11]

From a synthetic point of view, polypeptoids are particularly appealing since the chemical sequence can be controlled resulting in a model system that allows for great tunability of non-covalent secondary interactions such as electrostatics, van der Waals interactions, hydrogen boding, as well as hydrophobicity and hydrophilicity. For example, Kudirka and coworkers designed a type of peptoid block copolymer which can self-assembled into ultrathin, two-dimensional highly ordered nanosheets.[12] Their results demonstrated that intermolecular electrostatic interactions are the key factors in the nanosheet formation.[12] In addition, Sanii *et. al.* found that the hydrophobic air-water interface plays an important role in forming peptoid nanosheets. The peptoid monolayer collapses into a stable free, floating bilayer that can be stabilized by the air-water interface.[13] Moreover, unlike polypeptides, the peptoid α-carbon become achiral, and polypeptoids can not form secondary structures that are stabilized by hydrogen bonding of the backbone, such as α-helices and β-sheets, because of lacking hydrogen atoms on nitrogen groups.

The lack of stereoisomers and the synthetic tunability of these peptoids make them an ideal platform to study the effect of secondary/non-covalent interactions of the side groups on self-aggregation. These non-covalent interactions can arise due to the hydrophobic side groups, charged side groups, and hydrophilic neutral groups, which in turn can modulate the size and shape of the micelles formed in aqueous solutions of these polypeptoids. Sternhagen *et al.*[14] have carried out experimental studies on a series of 25-mer polypeptoids with five consecutive decyl hydrophobic monomers followed by 20 hydrophilic monomers, one of which is charged. They

found the position of the charged group has a significant effect on the aggregation number and micellar size. In this work, we focus on the molecular origins of this behavior as well as insight into miclellar structual properties and solvent effects, both of which are not directly obtainable by experiments.

Investigations of structural and dynamics properties of peptoid systems have been carried out using a number of theoretical and computational methods including quantum-mechanical (QM) modeling, molecular dynamics (MD) simulation, Monte Carlo (MC) simulations, and coarse-grained (CG) modeling which lack atomistic level details.[15–20] Atomistic MD simulations can, in principle, provide a complete description of the molecular structure of these systems, provided the simulation are long enough to sample the large conformational space of these systems.[21,22] Often coupling these simulations with enhanced sampling algorithms, such as parallel tempering,[23] or replica exchange molecular dynamics (REMD),[24] umbrella sampling,[25] and metadynamics[26,27] can greatly improve the sampling.[28–30] However, these simulations rely heavily on the quality of the underlying force fields which must accurately represent the peptoid potential energy landscape. In this work, we used the second generation of the general AMBER force field (GAFF2) to study the micellar systems mentioned earlier. We first demonstrate that GAFF2 parameters successfully reproduce both *cis* and *trans* conformations of a simple sarcosine dipeptoid in the gas phase, although it originally parameterized for generic organic molecules.[31] Atomistic MD simulations of self-assembly of sequence-defined singly charged polypeptoids in water were performed to explore the structural properties of the micelles formed and to shed the light on the non-covalent interactions that dominate micellar size and shape.

3.2 Methods

3.2.1 Force Field Validation

As mentioned in the introduction, the quality of a simulation study depends on the accuracy of the force field and hence validation of a force field is essential. Mukherjee *et al.*[20] reported that GAFF could successfully predict a range of experimental structures for peptoid systems in implicit solvent but worked poorly in explicit solvent. To improve the transferability of GAFF,

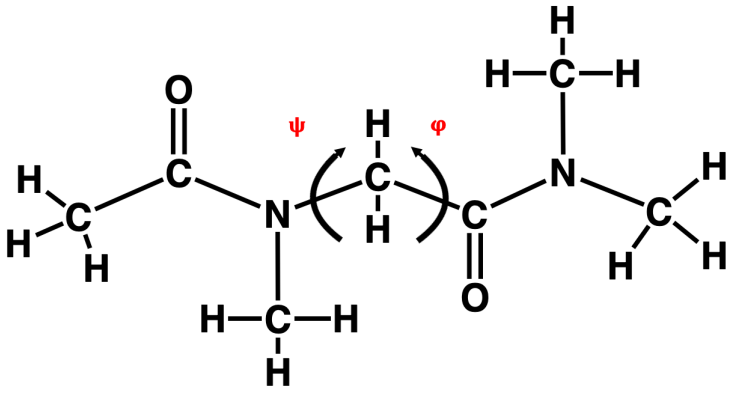

Figure 3.1. chemical structure of sarcosine dipeptoid.

Wang et al. developed the GAFF2 parameter sets by studying more model molecules for force field parameterizations and optimizing bonded force field parameters with high-quality quantum mechanical (QM) data. A simple peptoid analog with methyl side chain, sarcosine dipeptoid (shown in Figure 3.1) was chosen to be as the representative compound, and its potential energy profile as a function of the two dihedral angle ψ and ϕ (see Figure 3.1) in the gas phase was calculated using the GAFF2 force field and validated on the results from electronic structure scans. The partial charges of sarcosine for classical simulations were derived from restrained electrostatic potential (RESP) method[32] at the Hartree-Fock level of theory with the 6-31G* basis set using the Antechamber program. AMBER 16[33] was used to carried out classical simulations. For electronic structure calculations of the potential energy profile of the dihedral angles, the energies and optimized geometries were obtained at the B3LYP/6-31G* level of theory. All electronic structure calculations were performed with Gaussian 09.[34] These scans covered the entire 360 range of ψ, ϕ dihedral angles. We chose to use increments of 20 for the scans of the backbone dihedrals to capture the minima accurately.

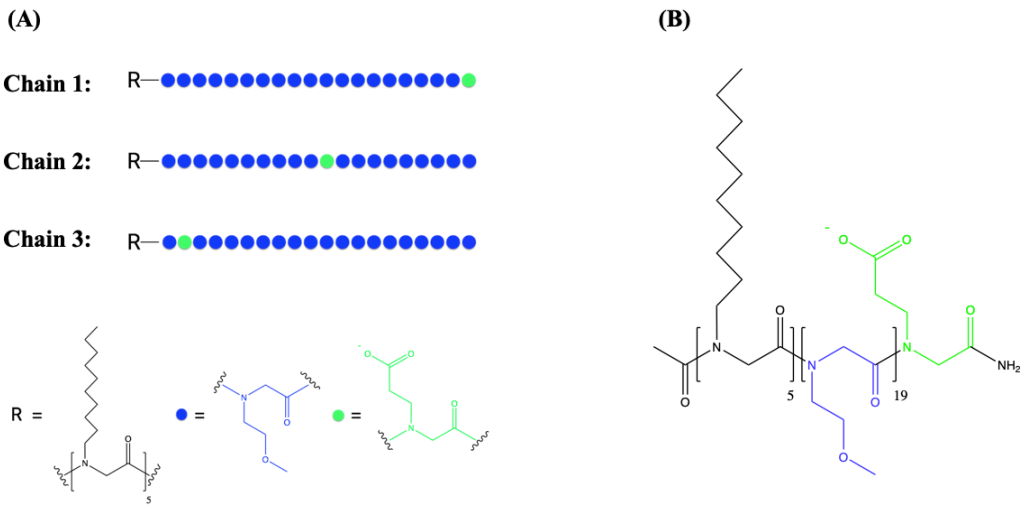

Figure 3.2. (A) A schematic of polypeptoid sequences and (B) the chemical structure representation of ionic peptoid block copolymer (chain 1). Note that R is for hydrophobic part, blue sphere for neutral monomer, and green sphere for charged monomer.

3.2.2 Molecular Dynamics Simulations

The sequences of amphiphilic ionic peptoid block copolymers were taken from the earlier study by Sternhagen *et al.*[14] and are shown in Figure 4.1. The MD simulations were carried out with the AMBER 16 software package[33] with the second generation of general AMBER force field (GAFF2) for all polymers.[31] Topologies and parameters were generated by LEAP and Antechamber modules in AMBER 16.[33,35] The initial coordinates of all three polymers were generated with OpenBabel[36] with corresponding SMILES strings. In addition, hydrophobic, neutral, and charged monomers (shown in Figure 4.1B) were built in order to calculate the partial charges. All monomers were optimized at the B3LYP/6-31G* level of theory, and restrained electrostatic potential (RESP) charges[32] were calculated at the level theory of HF/6-31G*. Geometry optimizations and charge calculations were performed by use of Gaussian 09 software package.[34] Each system composed of different polypeptoid surfactants immersed in a cubic box with pre-equilibrated TIP3P water[37] molecules. It worth noting that polypeptoid surfactants were put randomly in the box and the

distance between each pair of surfactants was at least 20 Å. In addition, Na+ counterions were added into the solvated system in order to neutralize the negative charge on the side chain of the peptoid. Table 4.1 lists the details of components in each system that was studied in this work. The number of polypeptoids, in each case was set to the aggregation number obtained by Sternhagen *et al.*[14] from neutron scattering experiments.

Table 3.1. An overview of MD simulations.

Chain #	# Surfactant	# Water	Concentration (w%)
1	27	339709	1.46
2	22	385127	1.05
3	13	295784	0.81

Each solvated system was minimized with steepest descent algorithm for 1000 steps and followed by 1000 steps of conjugate gradient minimization. In order to adequately sample configuration space, we carried out four simulations of the same system with different starting points that were generated as follows. The initial minimized structure for each case was heated gradually from 0 K to 300 K for trajectory 1 , 350 K for trajectory 2, 400 K for trajectory 3 and 450 K for trajectory 4 in 1 ns of MD, and followed by 10 ns run to cool back down to 300 K in the Canonical (NVT) ensemble, respectively. Each configuration was then followed by a 50 ns equilibration run and another 50 ns production run in the Isothermal-isobaric (NPT) ensemble with the temperature at 300K with an average pressure of 1 atm. The simulations were carried out using a time step of 2 fs for integration and 8 cut-off for nonbonded interactions, along with the particle mesh Ewald[38] for long-range interactions. The SHAKE algorithm[39] was used to constrained bonds involving hydrogen. The temperature was controlled using Langevin dynamics[40] with a collision frequency of 2.0 ps^{-1}, and the pressure was maintained by Isotropic position scaling[41] with a relaxation time of 2 ps.

3.3 Results and Discussion

3.3.1 Force Field Validation

2D Ramachandran potential energy surface (PES) plots were calculated using GAFF2 parameters and compared them to those generated using B3LYP/6-31G* level of theory (Figure 3.3). For the

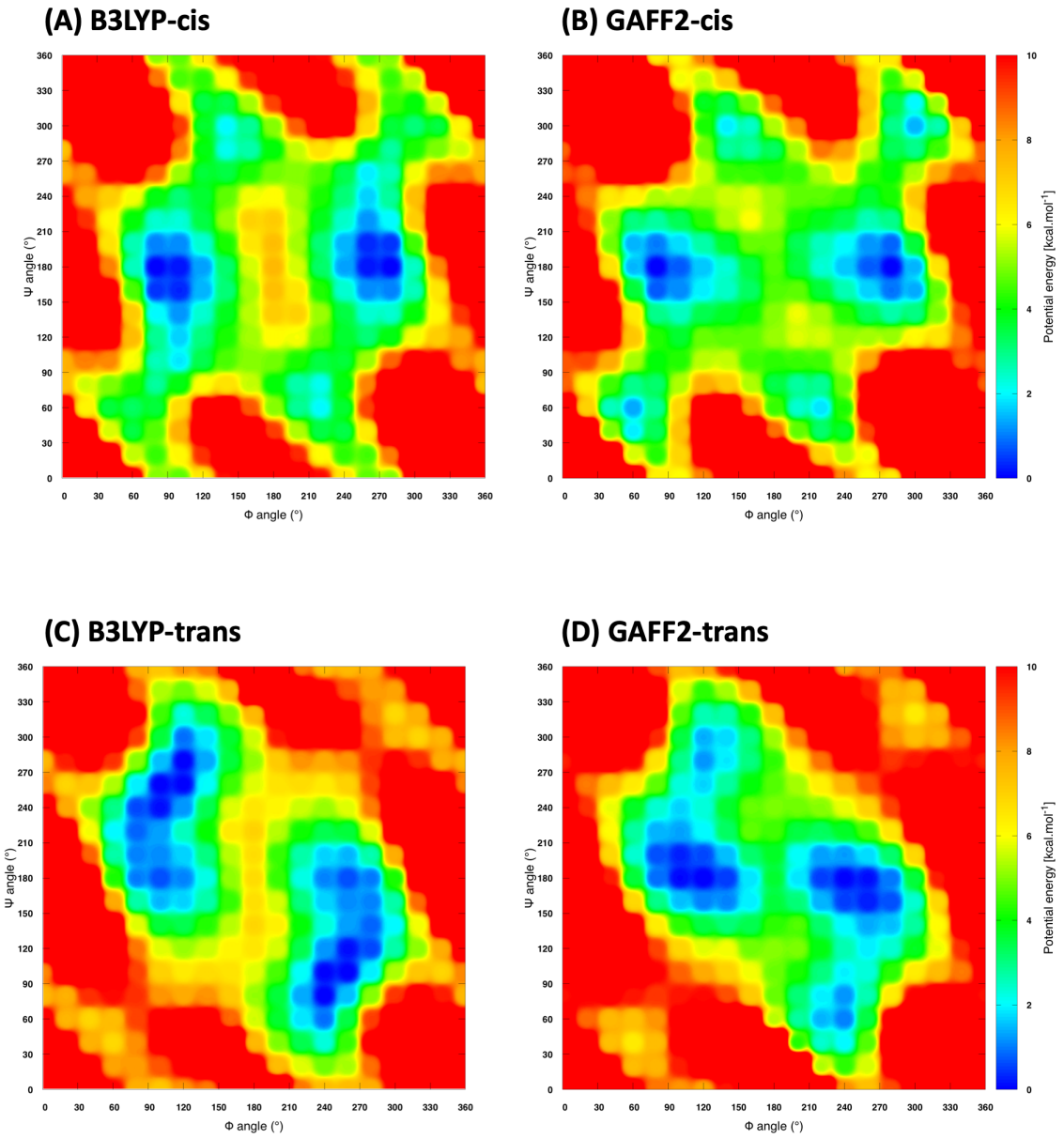

Figure 3.3. Potential energy (kcal/mol) surfaces of (ψ, ϕ) angles (degrees) for *cis* and *trans* sarcosine in vacuum using (B, D) GAFF2 in AMBER 16 and (A, C) B3LYP in Gaussian 09.

cis conformation, the B3LYP PES has two global minima centered at (90, 180) and (270, 180), while the GAFF2 PES similarly has two global minima center at (90, 180) and (280, 180). For the *trans* conformation, untuned GAFF2 predict two different global minima. However, the energy difference is less than 2 kcal/mol between these minima. All these PES plots are center-symmetric because of the achirality of the peptoid backbone. It is evident that GAFF2 correctly reproduce the main features of the QM calculations for both *cis* and *trans* peptoid conformations.

3.3.2 Molecular Dynamics Studies of Micelles

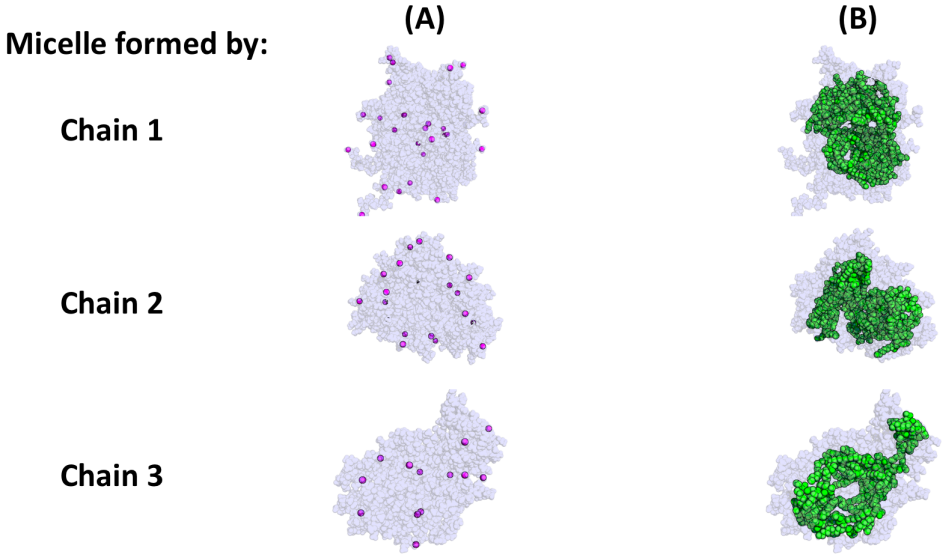

Figure 3.4. Illustration of the micelles of each studied surfactant in production simulation runs. (A)Solid pink balls represent the COO^- groups on the ionic monomers and transparent purple balls for other parts of micelles. (B)Solid green balls represent the hydrophobic parts of the polypeptoid chains and transparent purple balls for other parts of micelles. For clarity, counter ions and water molecules have been omitted.

As mentioned above, simulations in the study were started with molecular configurations with uniformly dispersed polypeptoid chains that are far from the equilibrium aggregated structure. However, during the equilibration, they all form micelles that remain stable throughout the entire simulation at constant temperature and pressure with an aggregation number that did not change. Each simulated micelle did not change markedly or convert to a very different structure. The analyses were carried out on the last 50 ns of the simulation, namely the production run, for each

of the four trajectories of each system and the average taken over all four trajectories (hence on 200 ns of data for each case).

Typical structural snapshots of micelles formed by chain 1 to 3 from production runs are shown in Figure 3.4. Ionic monomers with COO^- groups on the surfactant peptoid chains are found more likely to stay in the outer region, near the surface of the micelles, and hence exposed to the water.

Micelle size. As mentioned in the previous chapter on cyclic polypeptoids, neutron scattering experiments can be used to determine the radius of gyration (R_g) of the micelle, which in turn in an indicator of the size and compactness of the micelle. The radius of gyration from the small angle neutron scattering (SANS) experiments is determined from the micellar form factor P(Q). The form factors obtained from the SANS data and calculated from the MD trajectories are shown in Figure 3.5A. For comparison, form factors of micelles and water molecules that are within 5 $\AA$ of micelles were also calculated from the simulation (shown in Figure 3.5B). The Guinier region (the region of the curve for which Q values are $\leq 0.1\AA^{-1}$) of the scattering profiles calculated from simulations extends more towards high Q than experimental ones suggesting that the overall size of the simulated micelles are slightly smaller than their experiment counterparts. The R_g calculated from MD simulation trajectories and determined from Guinier analysis of the SANS data for all micelle systems was shown in Figure 4.2. For the last 50 ns of the simulations, the average R_g of the micelles formed by chain 1 to 3 are 3.12 ± 0.08 nm, 2.92 ± 0.05 nm, and 2.63 ± 0.05 nm, respectively. The SANS data gives larger R_g values for all three micelles. However, it should be noted that the SANS experiments were conducted on the peptoid on the peptoid micellar solutions at a concentration of 0.5 wt %, while our simulations were carried out at a higher concentration, since it enabled better sampling due to the comparatively smaller sizes required. Secondly, the SANS P(Q) can include solvating water molecules and clearly the form factors from simulations which include solvating waters (Figure 3.5B) show better agreement with the SANS data. Nonetheless, the R_g of the pure micelle (no water included) from simulations do compare reasonably well with the experimental data. Both the experimental and MD derived

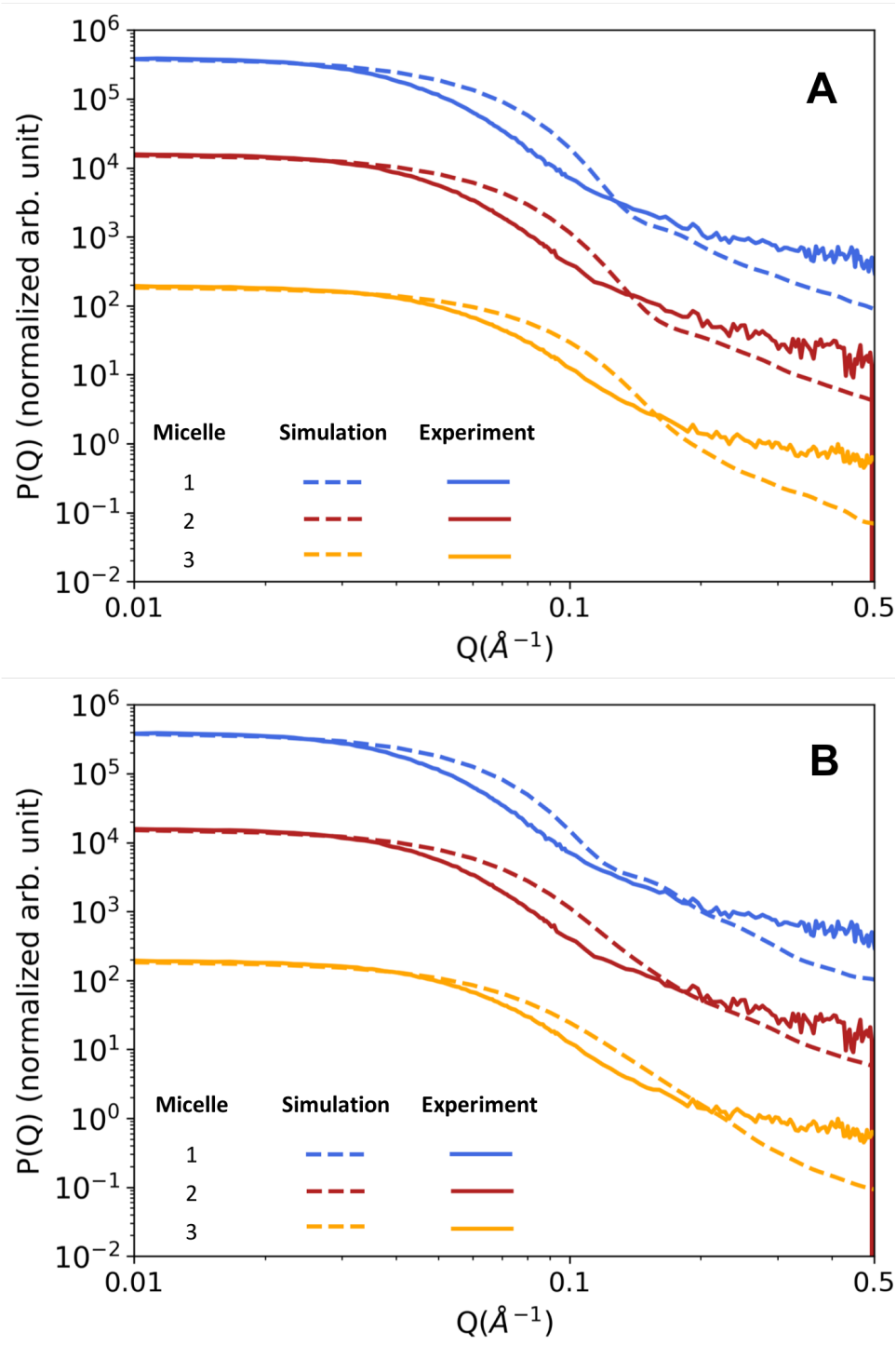

Figure 3.5. Form factors of the micelles formed by chain 1,2, and 3 from simulation in comparison with the experimental results obtained from small angle neutron scattering (SANS). Simulation P(Q) values calculated for (A) micelles only and (B) micelles and water molecules that are within 5 $\AA$ of micelles.

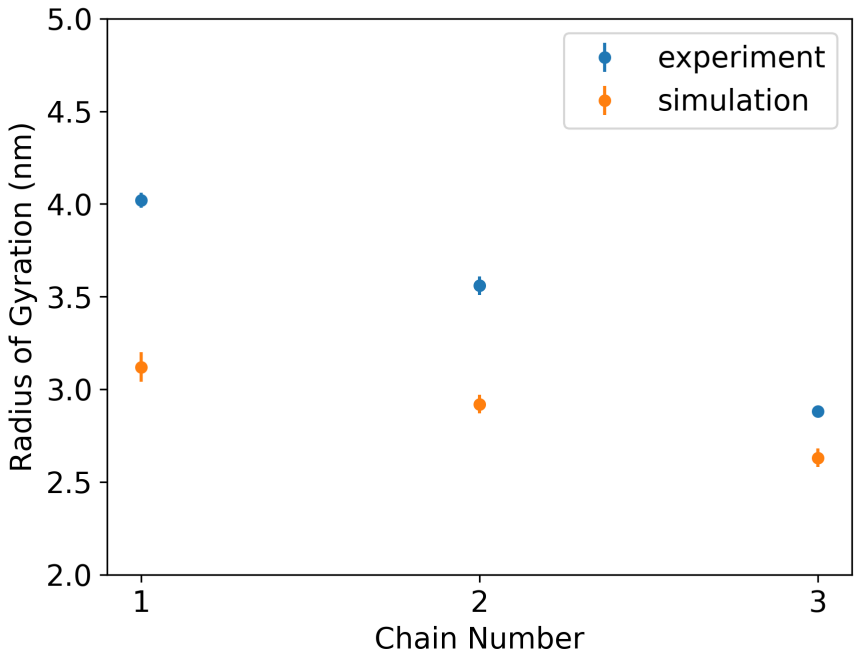

Figure 3.6. Radius of gyration (R_g) of the micelles formed by chain 1 to 3.

R_g values show a an apparent decrease in micelle size as the ionic monomer gets closer to the hydrophobic tail. All micelle systems showed a small fluctuation in the radius of gyration. It indicates that micelles are stable over the course of the simulation. While R_g can provide the quantitative information about the overall size of a micelle, it cannot give a detailed molecular picture of the internal structure of the micelle-water complex.

Micelle Shape. In order to determine the structural features of self-assembled micelles, the gyration tensor[42] of the micelles was calculated. The gyration tensor is defined as:

$$S = \frac{1}{N}\begin{bmatrix} \sum_i (x_i - x_{cm})^2 & \sum_i (x_i - x_{cm})(y_i - y_{cm}) & \sum_i (x_i - x_{cm})(z_i - z_{cm}) \\ \sum_i (x_i - x_{cm})(y_i - y_{cm}) & \sum_i (y_i - y_{cm})^2 & \sum_i (y_i - y_{cm})(z_i - z_{cm}) \\ \sum_i (x_i - x_{cm})(z_i - z_{cm}) & \sum_i (y_i - y_{cm})(z_i - z_{cm}) & \sum_i (z_i - z_{cm})^2 \end{bmatrix} \quad (3\text{–}1)$$

where the summation is calculated over N atoms of the micelle. The eigenvalues ($\lambda_1 > \lambda_2 > \lambda_3$) of the gyration tensor give the squared lengths of the principal axes of gyration, namely the radius

of gyration,

$$\lambda_1 + \lambda_2 + \lambda_3 = R_g^2 \tag{3–2}$$

The ratios of the eigenvalues indicate the deviation from a sphere-like shape of the micelle. In Figure 3.7, the ratios of $< \lambda_1 > / < \lambda_3 >$ and $< \lambda_2 > / < \lambda_3 >$ between eigenvalues of gyration tensor of the micelles form by chain 1 to 3 are shown. If the value of these ratios is close to 1, the shape of the micelle is spherical like. If the value of this ratio is large, the micelle is elongated. According to the previous studies in the literature, self-avoiding walk polymer models have averaged value ratios of $< \lambda_1 >:< \lambda_2 >:< \lambda_3 >= 14 : 2.98 : 1$.[43] Clearly, the micelle formed by chain 1 is more spherical than other two. In addition, the shape of micelles can be roughly quantified and confirmed by inspecting them in the simulation trajectories and visualizing the snapshots. From the ratio $< \lambda_2 > / < \lambda_3 >$ for the micelle formed by chain 3 in Figure 3.7, one clearly sees a large deviation from a sphere-like shape.

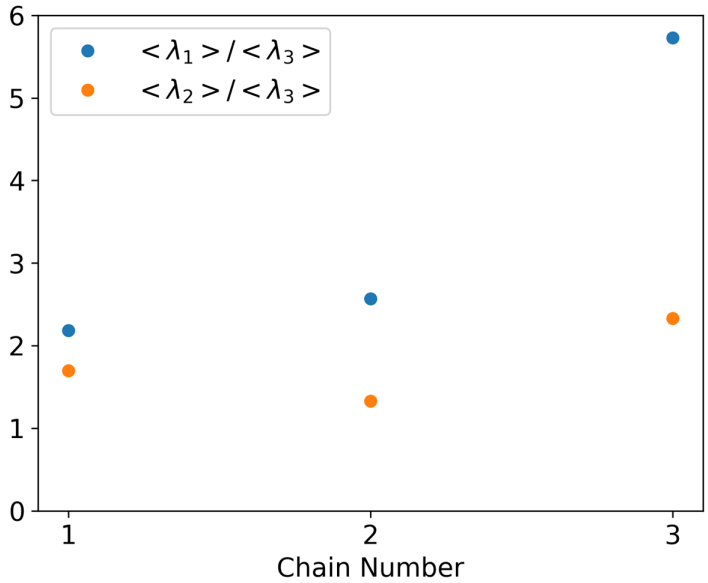

Figure 3.7. The ratios of the eigenvalues of gyration tensor.

Another way to characterize the shape of a micelle is the inertia tensor.[44–47] It is defined by:

$$I = \begin{bmatrix} \sum_i m_i(y_i^2 + z_i^2) & \sum_i -m_i(x_i y_i) & \sum_i -m_i(x_i z_i) \\ \sum_i -m_i(x_i y_i) & \sum_i m_i(x_i^2 + z_i^2) & \sum_i -m_i(x_i z_i) \\ \sum_i -m_i(x_i z_i) & \sum_i -m_i(y_i z_i) & \sum_i m_i(x_i^2 + y_i^2) \end{bmatrix} \tag{3–3}$$

where m is the mass of the individual atom, and the coordinates (x, y, z) are relative with the center of mass of the chain being the origin of the coordinate system. The key difference of moment of inertia tensor compared to the gyration tensor is that particle positions are weighted by mass in the inertia tensor, whereas only the particle positions are considered in the gyration tensor. The shape of the micelles can be analyzed using asphericity[48], S, as defined by:

$$S = \frac{3}{2} \frac{\sum_i^3 (\lambda_i - \overline{\lambda})^2}{(\sum_i^3 \lambda_i)^2} \tag{3–4}$$

For perfect spheres, the value of S is zero. The asphericity value of ionic peptoid micelles

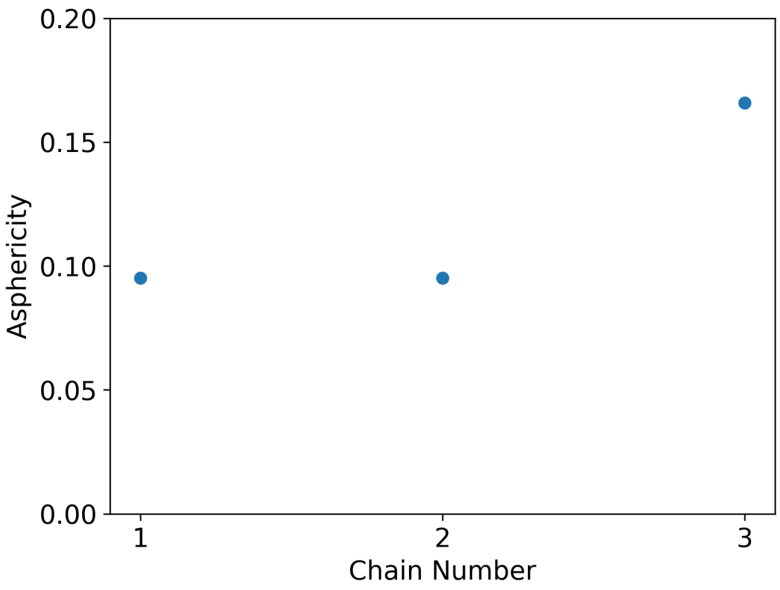

Figure 3.8. Asphericity of the micelles formed by chain 1 to 3.

increases from 0.095 to 0.165 as the ionic group on the backbone of the polypeptoid gets closer to the hydrophobic tail. The large micelle with the ionic group furthest from the hydrophobic parts shows spherical symmetry.

Solvent accessible surface area. Another quantify that can be used to study structural properties of the micelles is solvent accessible surface area (SASA).[49–51]. A brief description of the determination of the SASA is presented here. In general, a probe sphere molecule with a radius of 1.4 Å (mimicking the water molecule) is allowed to roll along the van der Waals surface of the micelle, and the summation of the contact area gives the total solvent accessible surface area. The distributions of total SASA and SASA per chain of the micelles are shown in Figure 3.9. The average total SASA values of the micelles formed by chain 1 to 3 are 58937 ± 1304,

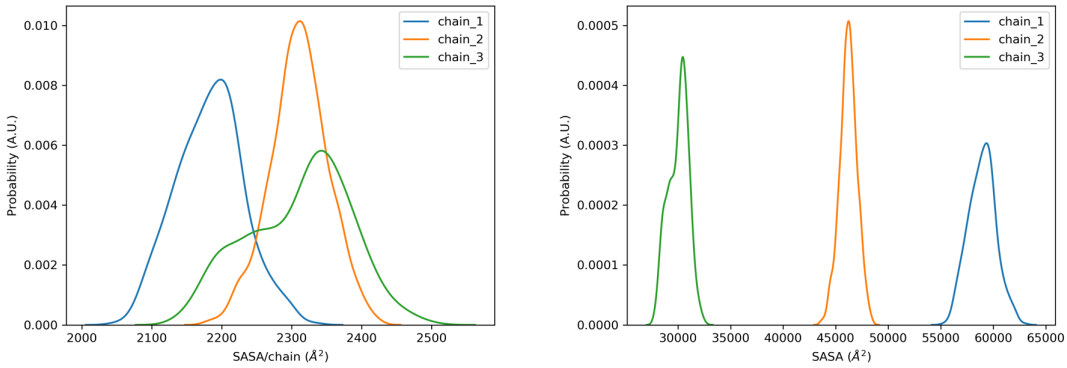

Figure 3.9. Solvent accessible surface area of the micelles.

46210 ± 825, and $30037 \pm 954 Å^2$, respectively. Due to the roughness of the micelle's surface, the average SASA values of each micelle is significantly larger than the SASA value of a perfect sphere with the same size. Unsurprisingly, the total SASA value is highest for the largest micelle 1 and the smallest for the micelle 3. Assuming spherical symmetry, the ratio of the average total SASA for the micelle formed by chain 1 to that formed by chain 2 should be approximately equal to the ratio of their R_g^2 (around 1.2:1 which is indeed the case). This indicates that although there is surface roughness, there is overall average spherical symmetry for these micelles. However, the total SASA ratio for micelle 1 to micelle 3 is almost 2 while the ratio of their R_g^2 is around $1.4:1$.

This clearly reiterates the non spherical nature of the micelle formed by chain 3. The SASA value per chain showed overlapping distributions. However, the micelle formed by chain 1 had a smaller average SASA value per chain as compared to micelle 2 and 3. The distribution for micelle 3 was the broadest with a shoulder at lower values. Typically, one would assume that the hydrophobic tail does not contribute to the SASA. The hydrophilic part can form a diffuse shell around the hydrophobic region and on average can be spherical but with a wide distribution. In the case of chain 3, the SASA has two regions. Our simulation trajectories clearly show that although the charged groups in chain 3 are close to the hydrophobic groups, they are nonetheless exposed to the solvent (charge-dipole interactions). The large SASA per chain with two distinct regions coupled with the smaller aggregation number indicate that the hydrophobic core is not perfectly screened by the hydrophilic section and is less compact. Combining this with the comparatively aspherical nature of the micelle suggests that the micellar shape is elongated. Clearly, the strong interactions between the surfactants and the water molecules mediate the shape of the micelles.

Scattering length density profiles. In order to further investigate the structural characteristics of the micelles formed by chain 1 to 3. We compared the excess neutron scattering length density profiles calculated from molecular dynamics simulations with the ones obtained from contrast variation SANS experiments. The excess neutron scattering length density profile,[52,53] $\Delta\rho$, which gives the difference of the scattering length density between the micelle and the solvent, takes the following expression:

$$\Delta\rho(r) = \rho_{micelle}(r) - \rho_{solvent} \tag{3–5}$$

where $\rho_{micelle}$ is the average scattering length density of the micelle, including surfactants and solvent molecules, $\rho_{solvent}$ is the average scattering length density of the solvent. The normalized neutron scattering length density profiles of the micelles which calculated from MD simulations and obtained from SANS experiments, are given in Figure 3.10. A similar trend has been observed for both simulation and experiment. It should be mentioned that D_2O, instead of H_2O, was used as solvent to calculate $\Delta\rho$, in order to compare with SANS experiments directly. It is clear to

see that $\Delta\rho$ exhibit their highest values in the molecular core region for all the micelles. The overall higher value for the normalized $\Delta\rho$ for micelle 3 suggests that more water molecules have penetrated into the micelle and a less compact core, compared to micelles 1 and 2. These results are consistent with the results mentioned above. The water penetration can be confirmed by calculating the water density profile, $H(r)$. The water density profiles of the micelles obtained from MD simulations are shown in Figure 3.11. When the distance r from the center of mass of the micelles) is between 10 to 20 Å, there are more penetrated water molecules in the micelles 3, while fewer waters are in the core region of the micelles 1 and 2. Moreover, the computed $H(r)$ shows good quantitative agreement with the computed $\Delta\rho$ in spatial region ranging from 10 $\AA$ to the micelle's surface along the radial direction of the micelle. $H(r)$ and $\Delta\rho$ values can be neglected at small r since large statistical fluctuations occur when the distance r is small.

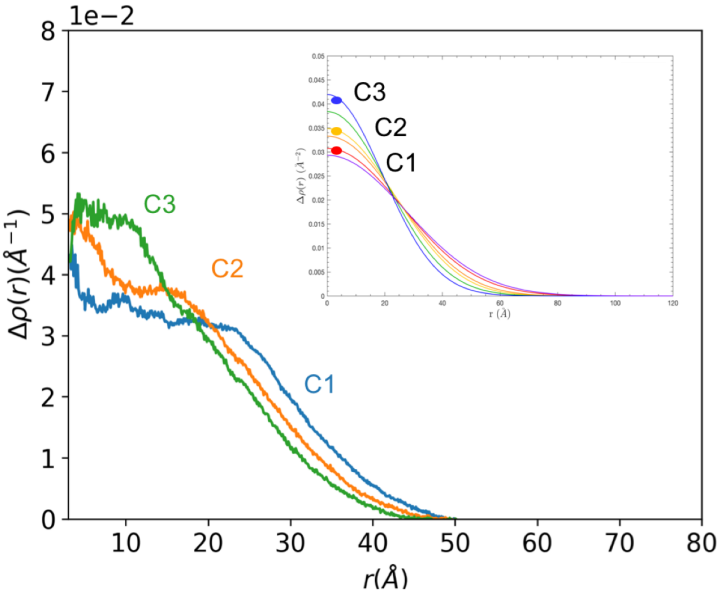

Figure 3.10. The neutron scattering length density profiles $\Delta\rho(r)$ of the micelles. The main figure shows $\Delta\rho(r)$ values calculated from molecular dynamics simulations. The experimental $\Delta\rho(r)$ obtained from contrast variation small angle neutron scattering is giving in the inset.

Ionic monomer solvation. The solvation of the ionic monomer can be understood by investigating the radial distribution functions (RDFs) between COO^- groups and waters near the micelle-water

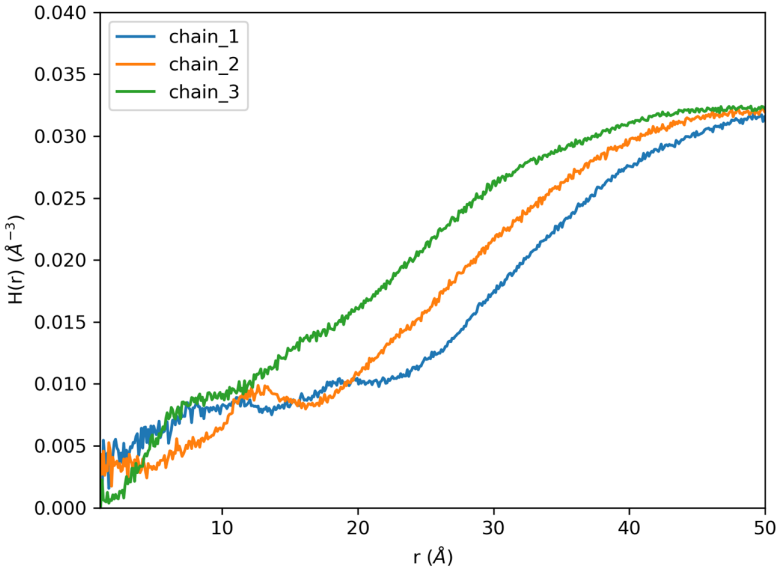

Figure 3.11. The water profiles $H(r)$ of the micelles.

interface. We calculated the coordination numbers of water in the first solvent shell by integrating the RDFs to the first minima of the RDF, namely 4.2 $\AA$ (Table 3.2). For the micelle formed by chain 1, there are more water molecules that are solvated around its ionic monomers. As

Table 3.2. Waters in the first solvent shell of COO^- groups. Note: waters are counted if distance between C(COO^{-1}) and O (H_2O) $\leq$ 4.2 Å.

Micelle formed by	n(water)
Chain 1	7.5
Chain 2	4.3
Chain 3	4.0

shown in Figure 3.4, ionic monomers are for the most part exposed to the solvent even when they are very close to the hydrophobic region. Overall the shape and structure of the micelles are governed by a number of non-covalent forces. The aggregation takes place due to the hydrophobic collapse of the hydrophobic regions of the polymers, as expected. The aggregation number itself is determined by the repulsion between the charged groups. The strong charge-dipole interactions between charge groups and waters force the ionic monomers to the micellar periphery and allows for greater penetration of the water molecules as well as a more elongated shape for the micelle formed by chain 3.

3.4 Conclusions

In this work, the all-atom simulations provide a detailed molecular level understanding of the structural aspect of micelles formed by sequence-defined peptoid block copolymers containing a single ionic monomer. Our results revealed the importance of solvent (dipole) - ionic group (charge) interactions as well as hydrophobic effects on the stability and structure of self-assembled charged polypeptoids as a function of placement of the charged group on the backbone. In future work, we will focus on studying more complex peptoids with additional side chains and multiple charges in different solvents.

3.5 References

[1] Jing Sun and Ronald N Zuckermann, "Peptoid polymers: a highly designable bioinspired material", *ACS nano* **7**(6), pp. 4715–4732 (2013).

[2] Donghui Zhang, Samuel H Lahasky, Li Guo, Chang-Uk Lee, and Monika Lavan, "Polypeptoid materials: current status and future perspectives", *Macromolecules* **45**(15), pp. 5833–5841 (2012).

[3] Abigail S Knight, Effie Y Zhou, Matthew B Francis, and Ronald N Zuckermann, "Sequence programmable peptoid polymers for diverse materials applications", *Advanced Materials* **27**(38), pp. 5665–5691 (2015).

[4] Jing Sun and Zhibo Li, "Peptoid applications in biomedicine and nanotechnology", In *Peptide Applications in Biomedicine, Biotechnology and Bioengineering* , pp. 183–213. Elsevier (2018).

[5] Nathaniel P Chongsiriwatana, James A Patch, Ann M Czyzewski, Michelle T Dohm, Andrey Ivankin, David Gidalevitz, Ronald N Zuckermann, and Annelise E Barron, "Peptoids that mimic the structure, function, and mechanism of helical antimicrobial peptides", *Proceedings of the National Academy of Sciences* **105**(8), pp. 2794–2799 (2008).

[6] Ronald N Zuckermann, "Peptoid origins", *Peptide Science* **96**(5), pp. 545–555 (2011).

[7] Kent Kirshenbaum, Annelise E Barron, Richard A Goldsmith, Philippe Armand, Erin K Bradley, Kiet TV Truong, Ken A Dill, Fred E Cohen, and Ronald N Zuckermann, "Sequence-specific polypeptoids: a diverse family of heteropolymers with stable secondary structure", *Proceedings of the National Academy of Sciences* **95**(8), pp. 4303–4308 (1998).

[8] Philippe Armand, Kent Kirshenbaum, Richard A Goldsmith, Shauna Farr-Jones, Annelise E Barron, Kiet TV Truong, Ken A Dill, Dale F Mierke, Fred E Cohen, Ronald N Zuckermann, et al., "Nmr determination of the major solution conformation of a peptoid pentamer with chiral side chains", *Proceedings of the National Academy of Sciences* **95**(8), pp. 4309–4314

(1998).

[9] Cindy W Wu, Kent Kirshenbaum, Tracy J Sanborn, James A Patch, Kai Huang, Ken A Dill, Ronald N Zuckermann, and Annelise E Barron, "Structural and spectroscopic studies of peptoid oligomers with α-chiral aliphatic side chains", *Journal of the American Chemical Society* **125**(44), pp. 13525–13530 (2003).

[10] Cindy W Wu, Tracy J Sanborn, Kai Huang, Ronald N Zuckermann, and Annelise E Barron, "Peptoid oligomers with α-chiral, aromatic side chains: sequence requirements for the formation of stable peptoid helices", *Journal of the American Chemical Society* **123**(28), pp. 6778–6784 (2001).

[11] Li Guo, Jinhai Li, Zachary Brown, Kushal Ghale, and Donghui Zhang, "Synthesis and characterization of cyclic and linear helical poly (α-peptoid) s by n-heterocyclic carbene-mediated ring-opening polymerizations of n-substituted n-carboxyanhydrides", *Peptide Science* **96**(5), pp. 596–603 (2011).

[12] Romas Kudirka, Helen Tran, Babak Sanii, Ki Tae Nam, Philip H Choi, Neeraja Venkateswaran, Ritchie Chen, Stephen Whitelam, and Ronald N Zuckermann, "Folding of a single-chain, information-rich polypeptoid sequence into a highly ordered nanosheet", *Peptide Science* **96**(5), pp. 586–595 (2011).

[13] Babak Sanii, Romas Kudirka, Andrew Cho, Neeraja Venkateswaran, Gloria K Olivier, Alexander M Olson, Helen Tran, R Marika Harada, Li Tan, and Ronald N Zuckermann, "Shaken, not stirred: collapsing a peptoid monolayer to produce free-floating, stable nanosheets", *Journal of the American Chemical Society* **133**(51), pp. 20808–20815 (2011).

[14] Garrett L Sternhagen, Sudipta Gupta, Yueheng Zhang, Vijay John, Gerald J Schneider, and Donghui Zhang, "Solution self-assemblies of sequence-defined ionic peptoid block copolymers", *Journal of the American Chemical Society* **140**(11), pp. 4100–4109 (2018).

[15] Philippe Armand, Kent Kirshenbaum, Alexis Falicov, Roland L Dunbrack Jr, Ken A Dill, Ronald N Zuckermann, and Fred E Cohen, "Chiral n-substituted glycines can form stable helical conformations", *Folding and Design* **2**(6), pp. 369–375 (1997).

[16] Kerstin Moehle and Hans-Jörg Hofmann, "Peptides and peptoidsa quantum chemical structure comparison", *Biopolymers* **38**(6), pp. 781–790 (1996).

[17] Glenn L Butterfoss, P Douglas Renfrew, Brian Kuhlman, Kent Kirshenbaum, and Richard Bonneau, "A preliminary survey of the peptoid folding landscape", *Journal of the American Chemical Society* **131**(46), pp. 16798–16807 (2009).

[18] Dina T Mirijanian, Ranjan V Mannige, Ronald N Zuckermann, and Stephen Whitelam, "Development and use of an atomistic charmm-based forcefield for peptoid simulation", *Journal of computational chemistry* **35**(5), pp. 360–370 (2014).

[19] P Douglas Renfrew, Timothy W Craven, Glenn L Butterfoss, Kent Kirshenbaum, and Richard Bonneau, "A rotamer library to enable modeling and design of peptoid foldamers", *Journal*

of the American Chemical Society **136**(24), pp. 8772–8782 (2014).

[20] Sudipto Mukherjee, Guangfeng Zhou, Chris Michel, and Vincent A Voelz, "Insights into peptoid helix folding cooperativity from an improved backbone potential", *The Journal of Physical Chemistry B* **119**(50), pp. 15407–15417 (2015).

[21] Scott A Hollingsworth and Ron O Dror, "Molecular dynamics simulation for all", *Neuron* **99**(6), pp. 1129–1143 (2018).

[22] Martin Karplus and J Andrew McCammon, "Molecular dynamics simulations of biomolecules", *Nature Structural & Molecular Biology* **9**(9), pp. 646 (2002).

[23] Robert H Swendsen and Jian-Sheng Wang, "Replica monte carlo simulation of spin-glasses", *Physical review letters* **57**(21), pp. 2607 (1986).

[24] Yuji Sugita and Yuko Okamoto, "Replica-exchange molecular dynamics method for protein folding", *Chemical physics letters* **314**(1-2), pp. 141–151 (1999).

[25] Glenn M Torrie and John P Valleau, "Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling", *Journal of Computational Physics* **23**(2), pp. 187–199 (1977).

[26] Alessandro Laio and Michele Parrinello, "Escaping free-energy minima", *Proceedings of the National Academy of Sciences* **99**(20), pp. 12562–12566 (2002).

[27] Omar Valsson, Pratyush Tiwary, and Michele Parrinello, "Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint", *Annual review of physical chemistry* **67**, pp. 159–184 (2016).

[28] Tatiana Maximova, Ryan Moffatt, Buyong Ma, Ruth Nussinov, and Amarda Shehu, "Principles and overview of sampling methods for modeling macromolecular structure and dynamics", *PLoS computational biology* **12**(4), pp. e1004619 (2016).

[29] Yi Isaac Yang, Qiang Shao, Jun Zhang, Lijiang Yang, and Yi Qin Gao, "Enhanced sampling in molecular dynamics", *The Journal of chemical physics* **151**(7), pp. 070902 (2019).

[30] Ayori Mitsutake, Yoshiharu Mori, and Yuko Okamoto, "Enhanced sampling algorithms", In *Biomolecular Simulations* , pp. 153–195. Springer (2013).

[31] Junmei Wang, Romain M Wolf, James W Caldwell, Peter A Kollman, and David A Case, "Development and testing of a general amber force field", *Journal of computational chemistry* **25**(9), pp. 1157–1174 (2004).

[32] Christopher I Bayly, Piotr Cieplak, Wendy Cornell, and Peter A Kollman, "A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the resp model", *The Journal of Physical Chemistry* **97**(40), pp. 10269–10280 (1993).

[33] D.S. Cerutti T.E. Cheatham III T.A. Darden R.E. Duke T.J. Giese H. Gohlke A.W. Goetz N. Homeyer S. Izadi P. Janowski J. Kaus A. Kovalenko T.S. Lee S. LeGrand P. Li C. Lin T.

Luchko R. Luo B. Madej D. Mermelstein K.M. Merz G. Monard H. Nguyen H.T. Nguyen I. Omelyan A. Onufriev D.R. Roe A. Roitberg C. Sagui C.L. Simmerling W.M. Botello-Smith J. Swails R.C. Walker J. Wang R.M. Wolf X. Wu L. Xiao D.A. Case, R.M. Betz and P.A. Kollman, "Amber 2016", University of California, San Francisco.

[34] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, . Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, "Gaussian09 Revision D.01", Gaussian Inc. Wallingford CT 2009.

[35] Junmei Wang, Wei Wang, Peter A Kollman, and David A Case, "Automatic atom type and bond type perception in molecular mechanical calculations", *Journal of molecular graphics and modelling* **25**(2), pp. 247–260 (2006).

[36] Noel M O'Boyle, Michael Banck, Craig A James, Chris Morley, Tim Vandermeersch, and Geoffrey R Hutchison, "Open babel: An open chemical toolbox", *Journal of cheminformatics* **3**(1), pp. 33 (2011).

[37] William L Jorgensen, Jayaraman Chandrasekhar, Jeffry D Madura, Roger W Impey, and Michael L Klein, "Comparison of simple potential functions for simulating liquid water", *The Journal of chemical physics* **79**(2), pp. 926–935 (1983).

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

[39] Jean-Paul Ryckaert, Giovanni Ciccotti, and Herman JC Berendsen, "Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes", *Journal of computational physics* **23**(3), pp. 327–341 (1977).

[40] Xiongwu Wu and Bernard R Brooks, "Self-guided langevin dynamics simulation method", *Chemical Physics Letters* **381**(3-4), pp. 512–518 (2003).

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

[42] Handan Arkın and Wolfhard Janke, "Gyration tensor based analysis of the shapes of polymer chains in an attractive spherical cage", *The Journal of chemical physics* **138**(5), pp. 054904

(2013).

[43] Manfred Bohn and Dieter W Heermann, "Diffusion-driven looping provides a consistent framework for chromatin organization", *PloS one* **5**(8), pp. e12218 (2010).

[44] Y Bruce Yu and Weizhen Wang, "Determinant of the inertial tensor and rotational entropy of random polymers", *The Journal of Physical Chemistry B* **103**(36), pp. 7676–7680 (1999).

[45] Joel W Cannon, Joseph A Aronovitz, and Paul Goldbart, "Equilibrium distribution of shapes for linear and star macromolecules", *Journal de Physique I* **1**(5), pp. 629–645 (1991).

[46] O Jagodzinski, E Eisenriegler, and K Kremer, "Universal shape properties of open and closed polymer chains: renormalization group analysis and monte carlo experiments", *Journal de Physique I* **2**(12), pp. 2243–2279 (1992).

[47] JD Honeycutt and D Thirumalai, "Static properties of polymer chains in porous media", *The Journal of Chemical Physics* **90**(8), pp. 4542–4559 (1989).

[48] Ruxandra I Dima and D Thirumalai, "Asymmetry in the shapes of folded and denatured states of proteins", *The Journal of Physical Chemistry B* **108**(21), pp. 6564–6570 (2004).

[49] Byungkook Lee and Frederic M Richards, "The interpretation of protein structures: estimation of static accessibility", *Journal of molecular biology* **55**(3), pp. 379–IN4 (1971).

[50] A Shrake and JA Rupley, "Environment and exposure to solvent of protein atoms. lysozyme and insulin", *Journal of molecular biology* **79**(2), pp. 351–371 (1973).

[51] Elizabeth Durham, Brent Dorr, Nils Woetzel, René Staritzbichler, and Jens Meiler, "Solvent accessible surface area approximations for rapid and accurate protein structure prediction", *Journal of molecular modeling* **15**(9), pp. 1093–1108 (2009).

[52] Xin Li, Kunlun Hong, Yun Liu, Chwen-Yang Shew, Emily Liu, Kenneth W Herwig, Gregory S Smith, Junpeng Zhao, Guangzhao Zhang, Stergios Pispas, et al., "Water distributions in polystyrene-block-poly [styrene-g-poly (ethylene oxide)] block grafted copolymer system in aqueous solutions revealed by contrast variation small angle neutron scattering study", *The Journal of chemical physics* **133**(14), pp. 144912 (2010).

[53] Bin Wu, Boutheïna Kerkeni, Takeshi Egami, Changwoo Do, Yun Liu, Yongmei Wang, Lionel Porcar, Kunlun Hong, Sean C Smith, Emily L Liu, et al., "Structured water in polyelectrolyte dendrimers: Understanding small angle neutron scattering results through atomistic simulation", *The Journal of chemical physics* **136**(14), pp. 144901 (2012).

CHAPTER 4
TOWARDS A COARSE-GRAINED MODEL OF THE PEPTOID BACKBONE: THE CASE OF N,N-DIMETHYLACETAMIDE

4.1 Introduction

Self-assembly of block copolymers in solution is relevant to a number of applications including biomedical uses.[2,3] Understanding the effect of secondary/non-covalent interactions on self-assembly can be challenging given the complexity in macromolecular systems. Polypeptoids, a class of highly tunable biomimetic analogues of peptides, are an ideal prototypical model system to study self-assembly such as micelle formation in aqueous environments.[4–6] In addition, polypeptoids themselves have many potential applicaitons including in the field of drug delivery.[7,8] A number of block copolypeptoids have been synthesized and their aggregation behavior in liquid solution have been studied by experiments and computational invertigations in recent years.[9–13] Atomistic (AA) molecular dynamics simulations have also been used to study solvation and self-assembly of peptoid systems. For example, the atomic resolution structure of bilayer peptoid nonosheets was determined by molecular simulation with using CHARMM based atomistic force field developed by Mirijanian *et al.*[14] Prakash *et al.* used molecular simulations to compare the backbone flexibility of a model peptoid with a corresponding peptide in water and at surfaces.[15] The OPLS[16,17] based force field has previously been used to model peptoid systems, including the studies of structural and dynamical characterisitics of peptoid oligomers and mechanism of aggregation of cyclic polypeptoids in methanol solution.[18] However, the computational expensse limits simulation times and system size to a few hundred nanoseconds and nanometers, respectively. Although the use of enhanced sampling algorithms can aid the sampling of configuration space, the dependece of these algorithm to a few select variables such as temperature, geometric collective variables, can often pose a challenge to sampling rare events from atomistic simulations.[19–21]

Reproduced from P. Du, S. W. Rick and R. Kumar, *Phys. Chem. Chem. Phys.*, 2018, **20**, 23386-23396 (https://www.doi.org/10.1039/C8CP03283A)[1] with permission from the PCCP Owner Societies.

Coarse-grained (CG) models, on the other hand, are computationally cheaper alternatives wherein atomistic level details are removed while retaining the relevant important physics, thus making it easier to study more complicated systems, including self-assembly in soft matter.[22–24] By reducing some degrees of freedom and only presenting the most important ones, CG modeling can span scales efficiently.[25–27]

Coarse-graining methodologies,[28,29] especially in the area of biological applications,[30] have become increasingly important in recent times starting from the seminal work of Warshel and Levitt.[31] As already mentioned this typically involves a reduction int he number of degrees of freedom resulting in a lower resolution model that is significantly more efficient but nonetheless incorporates the relevant physics of the problem under study. At the very least, light atoms such as hydrogen are folded into the nearest heavy atoms, the so-called united atom approach. Further coarse-graining combines groups of heavy atoms to form pseudo-atoms, as is the case in the MARTINI force field.[32,33] The parameterization of the force field can follow a number of different paths and can use experimental data or reference higher resolution simulation data or a combination of both. The method adopted here is a so-called "physics-based force field" wherein one retains the functional forms used in typical all-atom simulations such as bonded and non-bonded interactions. The higher resolution all-atom simulations provide data in the training set that can be used to parameterize the model using iterative Boltzmann,[34] force-matching,[35–37] or Newton inversion methods,[38] to name just a few.

To develop a coarse-grained model of complex peptoid systems, a bottom-up approach is being used for parameterizing CG models to accurately reproduce structural properties and is validated on thermodynamic properties of AA models. The typical polypeptoids used in micelle formation consist of a backbone with hydrophobic hydrocarbons, neutral ether based hydrophilic and charged carboxylate based hydrophilic side chains. The underlying coarse-graining philosophy is to develop CG models of each chemical group. The CG models of the hydrocarbon and ether chains have already been developed.[39,40]

In this study, a CG model of a simple amide system, N,N-dimethylacetamide (DMA) is developed as it contains the repeating motif of the peptoid backbone. Hence it presents an important step towards establishing CG models for peptoid systems. A key focus of this work is to reproduce not just the structure of liquid DMA but also the solvation structure of DMA in water. This is particularly important in order to develop peptoid CG models that can accurately describe the solvation properties of the polymer in aqueous solutions and hence enable accurate investigations of self-assembly in an aqueous environment. This is in contrast to a complex anisotropic implicit solvent CG model of peptoids developed by Haxton *et al.* to study two dimensional peptoid based nanosheets.[41]

It has been recognized that water plays an essential role in determining the structure and properties of peptoid systems.[15,42] A CG model of water developed Molinero and co-workers, based only on short range Stillinger-Weber (SW) potential,[43] is used in this work to describe the interactions between DMA and water. The OPLS-UA[16] force field is adopted to describe the intramolecular interactions between DMA molecules with the elimination of hydrogen sties. The focus of this study is to develop intermolecular parameters for DMA-DMA and DMA-water interactions using a "reduced range" approach that allows the retention of molecular features but the comparatively short range interactions, involving three-body forms in some cases, leads to computational efficiency over conventional all-atom models.[39]

In this study, we develop a CG model of DMA, laying the groundwork for developing CG models of polypeptoids. This CG model is parameterized on AA simulations, using a functional form that is a hybridized approach involving the OPLS-UA force field and the SW "reduced range" potential form. The parameterization scheme is based on the method used by Kumar and Skinner[44] to develop an all-atom model of liquid water with three-body interactions.[44] It is optimized to reproduce structural and thermodynamic properties of DMA in liquid solution, specifically features of a number of relevant radial distribution functions of pure DMA and dilute DMA aqueous solution and is validated on liquid/vapor properties, data from all-atom simulations of concentrated aqueous solutions of DMA, as well as simulation of a small peptoid oligomer in

water. The reference data is from simulations using the all-atom OPLS-AA force field. In addition, the computational efficiency of the CG model is discussed and compared to the higher resolution AA models.

The paper is organized in the following manner. The computational methods of the CG model parameterization scheme and the simulation set up are described in Section 2. The results are presented and discussed in Section 3 and the conclusions are outlined in Section 4.

4.2 Methods

The functional form of the non-bonded interactions used in this study is based on the SW[43,45] formalism that includes the short range two- and three-body terms. The potential energy of the system is given by

$$E = \sum_{i}\sum_{j>i}\phi_2(r_{ij} + \sum_{i}\sum_{j\neq i}\sum_{k>j}\phi_3(r_{ij}, r_{ik}, \theta_{ijk}) \tag{4–1}$$

Here the two-body term $\phi_2(r_{ij})$ is given by

$$\phi_2(r_{ij}) = A\epsilon[B(\frac{\sigma}{r_{ij}})^p - (\frac{\sigma}{r_{ij}})^q]exp(\frac{\sigma}{r_{ij} - a\sigma}) \tag{4–2}$$

and the three-body term $\phi_3(r_{ij}, r_{ik}, \theta_{ijk})$ by

$$\phi_3(r_{ij}, r_{ik}, \theta_{ijk}) = \lambda\epsilon[cos\theta_{ijk} - cos\theta_0]^2 exp(\frac{\gamma\sigma}{r_{ij} - a\sigma}exp(\frac{\gamma\sigma}{r_{ik} - a\sigma}) \tag{4–3}$$

Both terms are taken to be zero at distances greater than $a\sigma$. In this work, as with the original formalism, p is set to 4 and q to 0. The constants A and B are taken to be the same as in the original formalism with $A = 7.049556277$, $B = 0.60222455$. The two parameters in the two-body term, namely the energy scaling variable ϵ and the particle size σ, are part of the parameterization set. The three-body term is characterized by a scaling factor λ and an angular term θ, both of which are included in the parameterization procedure. The angle θ_0 takes into account the preferred orientational ordering of the solvation shell around the central ith atom and in the case of the neat water model, mW, developed by Molinero and co-workers is $109.47°$.[45] The term λ scales the three-body angular interaction. Both the two- and three-body terms are relatively short range, with the range controlled by the parameter a (since the potential goes to

zero at $a\sigma$) and to a lesser extent by γ, and the two are set to the values used in the mW water model, namely $a = 1.8$ and $\gamma = 1.2$.

The parameterization involves optimizing the set (ϵ,σ,λ,θ_0) for the different non-bonded interactions based on data from all-atom simulations. The details of the all-atom simulations and the parameterization scheme are presented below.

4.2.1 All-atom Reference Simulations

The reference all-atom simulations were performed using the TIP3P[46] water model along with the OPLS-AA force field for the DMA molecules. Two sets of all-atom reference simulations were performed, the first of a dilute solution of DMA in water (0.2 mol/L) and the second of pure liquid DMA. The dilute aqueous simulation of DMA was carried out in a cubic box, with an initial box length of 20 Å, consisting of a single DMA and 524 TIP3P water molecules. The liquid DMA simulations consisted of 512 DMAs in a cubic box of initial length 40 Å. A third set of simulations of a more concentrated aqeuous solution of DMA (0.5 mol/L) were carried out with the AA model. The system consisted of 20 DMA molecules along with 1700 water molecules in a box of length 40 Å. In addition to the above set of simulations, the DMA liquid-vapor system was also simulated. The system consisted of the same number of DMA molecules as above but with a box size of 40 Å* 40 Å* 80 Å, with the liquid-vapor interface perpendicular to the z-axis. Finally, simulations of a simple polypeptoid solution were also performed. The system contained two simple ten-mer polypeptoid chains (See Figure 4.1b) and 1700 water molecules in a 40 Åcubic box.

In each case, except as noted, the simulations were carried out within the isobaric isothermal (NPT) ensemble at 300 K and 1 atm with a 1 fs time step, using the LAMMPS MD package.[47] Each system was equilibrated for 10 ns and followed by a 10 ns production run. The temperature was controlled with a Nose-Hoover thermostat[48] and barostat.[49] Long range Coulomb interactions were calculated using the Ewald summation, specifically the Particle-Particle-Particle-Mesh (PPPM) method with the desired relative error in forces set to 0.0001.[50] The Lennard-Jones cutoff distance

was set to 10 Å. For the calculations of the diffusion constant of pure liquid DMA, a 5 ns production run in the NVE assemble was carried out.

Free energy of solvation of DMA in water was carried out using finite difference thermodynamic integration (FDTI),[51] using a soft core potential to avoid singularities.[52–54] The free energy was calculated using a 10 ns NPT run with 20 integration points, equally spaced between 0 and 1, with a time step of 1 fs using the GROMACS[55] software.

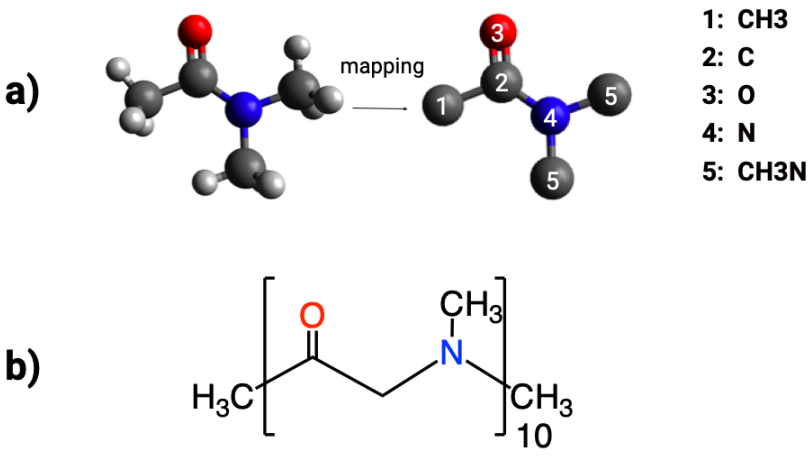

Figure 4.1. (a) Mapping of DMA molecule from the AA to the CG model. The white balls represent H, the gray ones C, the blue ones N, and the red ones O. There are 5 total atom types in the CG DMA molecule, which are 1: CH_3, 2: C, 3: O, 4: N, 5: CH_3N. (b) Schematic of the simple peptoid chain.

4.2.2 Coarse-grained Model Development

Figure 4.1 shows atom types and the mapping of the DMA molecule from AA to the CG representation. The interactions between DMA and water are represented by SW potentials. Intramolecular interactions are taken from the force field, OPLS-UA. Initial guesses for the parameters of each type of atom in DMA were selected from the monatomic model of water (mW) and methane, which were previously developed by Molinero and co-workers.[56]

The CG simulations of dilute DMA with mW waters, liquid DMA, concentrated aqueous DMA, polypeptoid solution and the DMA liquid-vapor cases were performed in LAMMPS with a 5 fs time step under the same simulation conditions as the reference AA simulations. The intramolecular terms were the same as the united atom OPLS-UA model.[16] The suitability of the 5fs time step was tested using a 5 ns NVE simulation that showed virtually no energy drift, despite large fluctuations in the energy. The free energy of solvation was carried out using the LAMMPS suite of programs. The free energy calculations were done using FDTI, with a soft core potential to avoid singularities. The free energy was calculated using 100 integration points, equally spaced between 0 and 1, and were run for 5 million steps, with a time step of 5 fs. The subroutines to carry out TI calculations in LAMMPS were taken from the work of Gyawali *et al.*[39,57] For the CG peptoid solution simulations, the parameters developed for DMA were used, with the parameters for the CH_2 group in the peptoid set to the values obtained for the CH_3N methyl group in DMA.

4.2.3 Parameterization of the DMA CG Model

The development of the CG model is primarily focused on reproducing the solvation structure of water around the DMA molecules. The energy scales ϵ and λ, particle size σ, and three-body angle θ_0 are the parameters that govern the solvation behavior. One can obtain estimates of θ_0 from either the angular distribution in the first solvation shell from AA simulations or from experimental information such as the tetrahedral geometry of water, leaving just the particle size and energy scale for parameterization. The overall parameterization process is outlined in the flowchart shown in Figure 4.2. An iterative parameterization procedure, described below, is adopted wherein one starts with zeroth order model that is progressively optimized to reproduce properties in the training set, in this case details of important redial distribution functions, from the all atom reference simulation data.

As mentioned in the introduction, the parameterization scheme is based on the one adopted by Kumar and Skinner in the development of an improved all-atom water model and is similar in spirit to the Newton inversion method. In order to illustrate the method, consider a system that

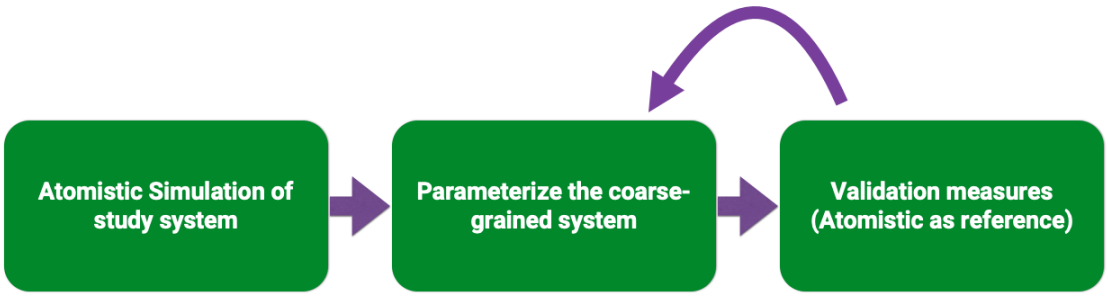

Figure 4.2. A flowchart of CG model parameterization scheme.

consists of just one type of particle that interacts via the SW potential form. In that case, the total intermolecular interaction energy can be expressed as a sum of two-body energies E_{12} and three-body energies E_{123}. Thus, one can write:

$$E_{total} = \omega_1 E_{ij} + \omega_2 E_{ijk} \tag{4–4}$$

where ω_1 and ω_2 are scaling parameters and are initially set to 1, the so-called zeroth order model with a particular set of parameters (ϵ,σ,λ,θ_0). Based on the scheme by Kumar *et al.*, ω_i can be determined in a systematic manner.[44] Since, for this simple case, there are two parameters that need to be determined, two observables/properties (P_1 and P_2) are chosen to be part of the training set. Each observable/property will in general be a function of both ω_1 and ω_2. Hence the total derivative of each P_j is given by

$$dP_j = \sum_i \frac{\partial P_j}{\partial \omega_i} d\omega_i \tag{4–5}$$

The derivatives $\partial P_j / \partial P_j$ are calculated numerically using a series of 5 simulations with parameters (ω_1,ω_2), ($\omega_1 + \delta\omega_1$,$\omega_2$),($\omega_1 - \delta\omega_1$,$\omega_2$), ($\omega_1$,$\omega_2 + \delta\omega_2$), ($\omega_1$,$\omega_2 - \delta\omega_2$). Based on these numerical

derivatives and the difference, ΔP_j, between the value of each P_j from the simulation at (ω_1,ω_2) and reference/experimental value, the next generation values of (ω_1,ω_2) can be determined by solving the following set of equations for $\Delta\omega_i$:

$$\Delta P_j = \sum_i \frac{\partial P_j}{\partial \omega_i} \Delta\omega_i \tag{4–6}$$

The new values of (ω_1,ω_2) now form the first-generation model and the process is repeated until convergence is reached. In the above example, only one type of particle was considered and hence only two scaling parameters, one scaling the three-body and the other the two-body interactions, are needed. The method can be extended to systems with more than one type of particle (as is the case for the DMA systems under study) leading to many two- and three- body scaling parameters, with the corresponding number of properties in the training set. The training set can be from higher level simulation data, from experiments or from a combination of both and can be extended to more complicated systems with more than one type of particle. Therefore, for a given set of σ values, the energy scale factors for the two- and three-body terms can be determined.

In the current development of the CG model, data from the reference all-atom OPLS-AA simulations on liquid DMA and the dilute solution of DMA in water are used. Specifically, the position of the first maximum of relevant radial distribution functions(O atom of water with the different DMA atoms in the dilute aqueous solution as well as intermolecular DMA atoms in the liquid DMA solution) as well as the associated coordination number around each atom type are the properties in the training set.

The initial σ and ϵ and values for O and C atoms were taken from the work of Molinero *et al.* while the initial N atom parameters set was taken from the OPLS-UA force field. The Lorentz-Bertholet mixing rules were used to obtain the remaining σ and ϵ values. The three-body angle θ was set to 109.47° for all the atom types, based on previous work on water and methane. The three-body energy scaling parameter λ was initially set to the value for the mW water model. The AA simulations did not show any preferred orientation of the solvation environment of the

N, and since the N atom is buried in DMA molecule one does not expect that the three-body terms were scaled and the scaling parameters were determined using the iterative algorithm outlined above. First, the DMA-DMA parameters were determined using the iterative algorithm on simulations with pure liquid DMA. Ten interactions were scaled, namely five two-body terms (C-C, O-O, N-N, CH_3-CH_3 and CH_3N-CH_3N) and five three-body terms (O-O-O, CH_3N-O-CH_3N, O-CH_3N-CH_3N, O-CH_3-CH_3, and CH_3-O-CH_3). For the three-body terms only those involving O with distinct sites were chosen. For the DMA-water interactions, the dilute DMA systems were simulated. In this case, five two-body terms (C-O_w, CH_3N-O_w, C-O_w, CH_3-O_w and O-O_w) and five three-body terms (O_w-O-O_w, O-O_w-O_w, O_w-CH_3N-O_w, CH_3N-O_w-O_w and O_w-C-O_w) were considered in the parameterization scheme. O_w refers to the mW CG water. In each case a 10x10 matrix was determined, solved to obtain the new set of parameters, and the procedure carried out iteratively until convergence was reached. For each case, convergence was achieved in less than ten iterations. After the first generation model was developed, a series of new starting points were developed by changing the various σ values. The iterative procedure was applied to each set. The final model was one for which the root mean square error was the least(less than 1).

It should be noted that for the two-body interactions that were not varied, the Lorentz-Bertholet mixing rules were used with the scaling factor of the scaled relevant two-body interactions absorbed into the corresponding ϵ values. The mixing rules are thus implicitly included in the paramerization procedure. Care should be taken when using mixing rules since they can lead to nonphysical results. However, the bottom-up approach with increasing complexity that is being adopted to develop the force field for real peptoid systems, starting from the alkane chains (hydrophobic side chains by Gyawali *et al.*),[39] to the peptoid backbone (current work) to more complex side chains (future work) to some extent mitigates this risk.

4.3 Results and Discussion

At the end of the parameterization procedure, the only three-body terms that were found to be non-zero were the water interactions with the carbonyl oxygen involving the $O_wO_{DMA}O_w$, and the $O_wO_wO_{DMA}$ angle, where O_w refers to the mW water and O_{DMA} to the carbonyl O atom of

the DMA. Since the N atom is buried inside the molecule, there is no special preferred orientation of the water molecules or other DMA molecules with the N atom and hence the three-body terms involving the N atoms turn out to be essentially zero as expected. The C atoms, similar to the case of methane and other aliphatic chains in previous studies, also did not result in non-zero three-body terms. The values of each of the parameters are listed in Table 4.1.

Table 4.1. Parameters of (a)two- and (b) three- body interactions for the CG DMA together with previously published parameters for water.[45], The scaling terms obtained from the parameterization step are folded into the ϵ and λ parameters.

(a)

Type i	Type j	ϵ(kcal/mol)	σ(Å)
mW	mW	6.189	2.3925
mW	CH_3	0.40	3.7
mW	C	0.08	4.8
mW	O	3.20	2.2
mW	N	0.015	4.4
mW	CH_3N	0.40	3.7
CH_3	CH_3	0.1	4.6
CH_3	C	0.09	4.9
CH_3	O	0.16	4.2
CH_3	N	0.1	5.05
CH_3	CH_3N	0.1	4.6
C	C	0.09	5.2
C	O	0.1	4.5
C	N	0.09	5.35
C	CH_3N	0.09	4.9
O	O	0.6	4.8
O	N	0.115	4.55
O	CH_3N	0.16	4.2
N	N	0.1	5.5
N	CH_3N	0.1	5.05
CH_3N	mW	0.1	4.6

(b)

Type i	Type j	Type k	λ	$cos\theta$
mW	mW	mW	23.15	-0.333333333
O	mW	mW	10.00	-0.333333333
mW	O	mW	10.00	-0.333333333

The O_{DMA}-O_w, C_{DMA}-O_w, $CH_{3\text{-}DMA}$-O_w, N_{DMA}-O_w and the CH_3N_{DMA}-O_w radial distribution functions are shown in Figure 4.3 and the associated coordination numbers in Table 4.2. Note

that these atom types are based on the representations in Figure 1. The CG models does indeed give a fair representation of the solvation environment of the DMA in water and reproduces reasonably well both the positions of the first peak and the corresponding peak height. However, the CG model radial distribution functions are slightly shifted to smaller distances compared to the AA model, suggesting that the CG DMA molecules, on average, are more hydrophilic than the AA model. The liquid DMA radial distribution functions (C-C, O-O, N-N, CH_3N-CH_3N and CH_3-CH_3) are shown in Figure 4.4 and the coordination numbers are tabulated in Table 4.3. Once again, the representation of the intermolecular DMA-DMA structure is reproduced relatively well by the CG model when compared to the AA case. Specifically, the first peak of the O-O and CH_3N-CH_3N show really good agreement, while for the N atom that is buried inside, the peak position is reproduced whereas the peak height is not. The same holds true for the C-C radial distribution functions. This is not necessarily surprising since these groups are further apart from each other. The CH_3-CH_3 radial distribution function is not very strongly structured and hence is difficult to reproduce very accurately using this method. The CG model seems to suggest a stronger attraction of the CH_3-CH_3) as compared to the AA model, since the radial distribution function is shifted to a smaller distance for the CG case as compared to the AA. Nonetheless, the dominant intermolecular interactions are fairly well reproduced by the CG model both for pure DMA and dilute DMA in water.

Table 4.2. The position (R) of the first maximum in the corresponding radial distribution function and the coordination number (C) for dilute DMA system.

	R(Å)		C	
	AA	CG	AA	CG
O_{DMA}-O_w	2.75	2.75	2.8	4.3
C_{DMA}-O_w	3.75	3.65	4.5	6.2
CH_{3DMA}-O_w	3.75	3.95	19.2	19.9
N_{DMA}-O_w	4.75	4.55	31.5	30.5
CH_3N_{DMA}-O_w	3.85	4.05	14.5	21.2

The radial distribution functions were part of the training set of the model and hence for validation different properties are calculated and compared to both experiment and the AA model in the following section.

Table 4.3. The position (R) of the first maximum in the corresponding radial distribution function and the coordination number (C) for liquid DMA system.

	R(Å)		C	
	AA	CG	AA	CG
O-O	5.65	5.65	12.5	12.3
C-C	5.65	5.75	13.5	13.6
N-N	6.05	5.85	13.1	13.7
CH_3N-CH_3N	2.45	2.45	1.0	1.0
CH_3-CH_3	4.15	5.75	6.3	10.6

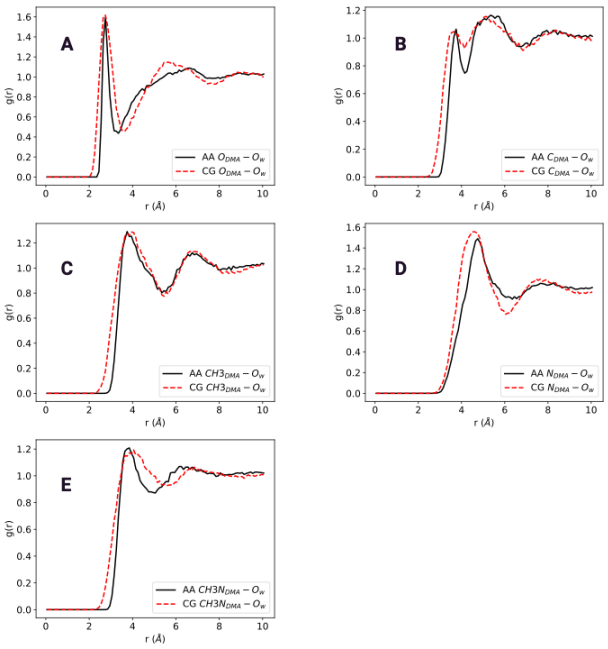

Figure 4.3. The liquid DMA radial distribution function at 300 K for AA and CG models. For the AA case, CH_3 refers to the carbon atom of the methyl group bonded to the carbonyl group, CH_3N refers to the carbon atom of the methyl group bonded to the N atom, C refers to the carbon atom of the carbonyl group and O to the oxygen atom of the carbonyl group.

Liquid DMA properties

A number of properties of pure liquid DMA were calculated and tabulated in Table 4. The self-diffusion coefficient of DMA was calculated from a 5 ns NVE run using the Einstein equation:[58]

$$D = \lim_{t\to\infty} \frac{1}{6t} < |r(t) - r(0)|^2 > \tag{4–7}$$

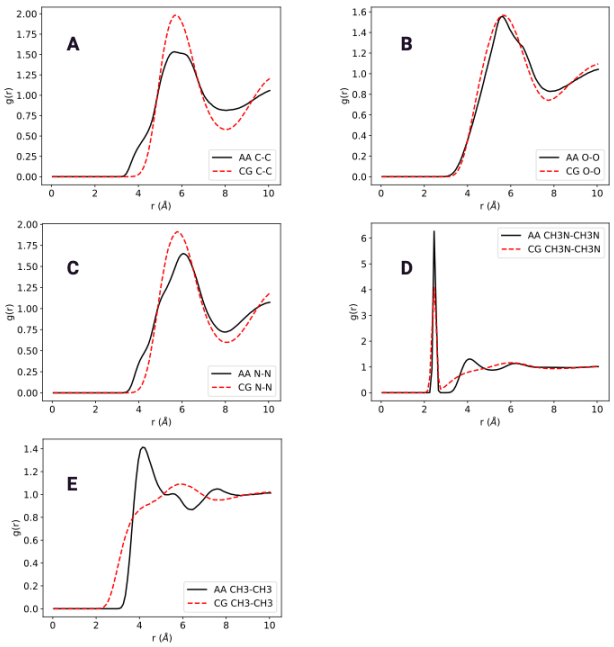

Figure 4.4. The waterDMA radial distribution functions for dilute DMA solution at 300 K for AA and CG models. For the AA, CH_3 refers to the C atom of the methyl group bonded to the carbonyl group, CH_3N refers to the carbon atom of methyl group bonded to the nitrogen atom, C refers to the carbon atom of the carbonyl group and O to the oxygen atom of the carbonyl group, and O_w refers to the water oxygen atom.

where $r(t)$ is the position of the center of mass of DMA at time t. The CG diffusion constant is a little more than twice the value of the AA, which is not surprising since the dynamics of CG models are faster than the corresponding all-atom systems.

The isothermal compressibility of liquid DMA at 300 K around density of $\rho_0 = 937 kg/m^3$ was found from:[59]

$$K_T = -\frac{1}{V}(\frac{\partial V}{\partial p}) \approx (\frac{ln(\frac{\rho_2}{\rho_1})}{< p_2 > - < p_1 >})_T \tag{4–8}$$

where ρ_2 is $1.01\rho_0$ and ρ_1 is $0.99\rho_0$. p_2 and p_1 were the pressures computed from a 10 ns NVT simulation. Surprisingly the CG model reproduces the experimental value more accurately than the AA model to which it was parameterized. The AA model shows a lower compressibility as compared to the CG model. This could be related to the CH_3-CH_3 radial distribution for CG being shifted to smaller distances as compared to the AA model.

The enthalpy of vaporization was computed from:[58]

$$\Delta H_{vap} = < E_{pot}(g) > - < E_{pot}(l) > + k_b T \tag{4–9}$$

where $E_{pot}(g)$ is the gas-phase potential energy and $E_{pot}(l)$ is the liquid phase potential energy. The CG model gives a higher value for the enthalpy of vaporization as compared to the AA model and the experimental value, but is within around 10% of the experimental value.

The next set of studies focuses on aqueous solutions of DMA as well as the simplest polypeptoid that can be formed with DMA as the repeating unit.

Dilute aqueous DMA solution

The solvation free energy of DMA from the CG simulations is very close to the value of the AA model, although both are around 1 kcal/mol higher than the experimental value. The fidelity of the solvation free energy of the CG model to the AA model along with the agreement between the radial distribution function data between the G and AA models does suggest that the CG model reproduces the solvation environment around the DMA molecule as seen in the AA simulations.

Concentrated DMA aqueous solution and simple peptoid oligomer solution

The various radial distribution functions of the concentrated DMA-water system and the simple peptoid systems are shown in Figure 4.5 and Figure 4.6 respectively. Despite not being part of the training set, the CG model gives a good representation of the solvation environment of the DMA and the simple peptoid in the aqueous system, although once again the CG model is slightly more hydrophilic than the corresponding AA model.

It is clear from both the agreement with the training data as well as the concentrated DMA simulations that the CG model provides a reasonable description of the DMA in bulk conditions. The next set of studies determines its suitability for interfacial regions. The air-liquid interface is an example for an extreme case of an asymmetrical environment and thus is an excellent system to test the applicability of the CG force field.

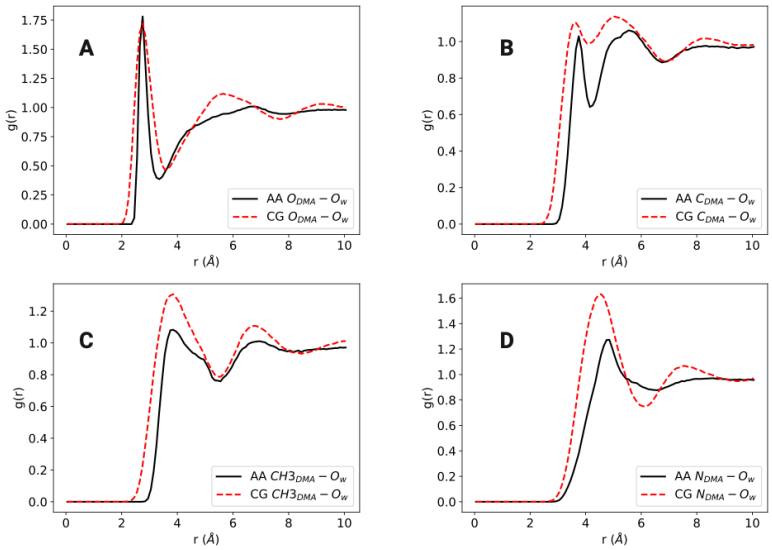

Figure 4.5. The radial distribution functions for the concentrated aqueous DMA solution at 300 K for AA and CG models. For the AA case, CH_3 refers to the carbon atom of the methyl group bonded to the carbonyl group, CH_3N refers to the carbon atom of the methyl group bonded to the nitrogen atom, C refers to the carbon of the carbonyl group and O to the oxygen atom of the carbonyl group, and O_w refers to the water oxygen atom.

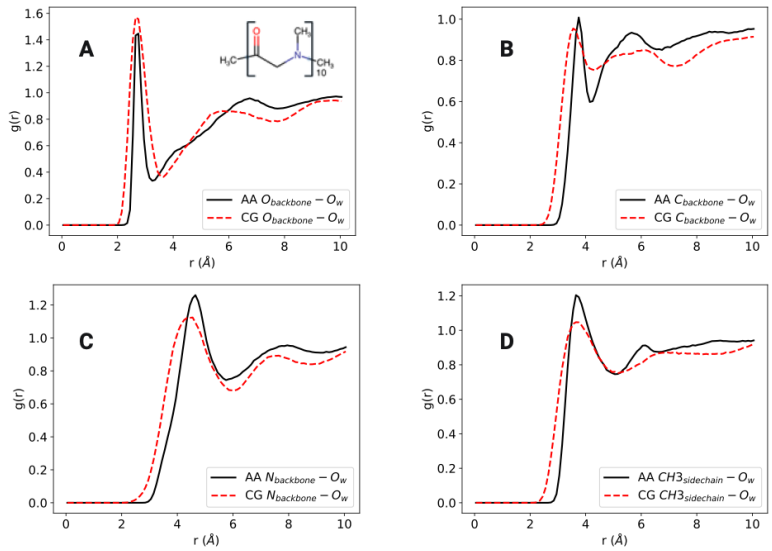

Figure 4.6. The radial distribution functions for the dilute polypeptoid solution at 300 K for AA and CG models. $C_{backbone}$ and $O_{backbone}$ refer to the carbon and oxygen atom of the carbonyl group, $N_{backbone}$ refers to the nitrogen atom on the backbone, and $CH3_{sidechain}$ refers to the carbon atom of methyl group bonded to the nitrogen atom.

Surface tension of liquid DMA

The surface tension of the liquid DMA liquid-air interface was also calculated from the liquid DMA-air slab simulations at 300 K. In order to determine the surface tension γ the following relationship between γ and the pressure tensors, evaluated from the simulation, was used:[58]

$$\gamma = \frac{Lz}{2}[< p_N > - < p_T >] \tag{4–10}$$

where L_z is the box length of the direction perpendicular to the interface and $< p_N >$ and $< p_T >$ are pressure tensors perpendicular and tangential to the liquid-air interface, respectively.[59] The surface tension of CG model is compared to the value from the AA model along with the experimental data[60] in Table 4.4. The CG model shows a surprisingly better agreement with experiment than the AA model.

Table 4.4. Comparison of DMA models and experiment

Properties	AA DMA(300 K)	CG DMA (300 K)	Experiment
Surface tension γ (mN/m)	32 ± 1	39 ± 1	37.13(293.15 K)[60]
Diffusion coefficient D $(10^{-5} cm^2/S)$	0.8	1.8	
Isothermal compressibility K_T (Gpa^{-1})	-0.53 ± 0.01	-6.1 ± 0.01	0.63 (298.15 K)[61]
Solvation free energy ΔG_{solv} $(kcal/mol)$	-7.5 ± 0.2	-7.7 ± 0.1	-8.5 (298.15 K)[62]
Enthalpy of vaporization ΔH_{vap} $(kcal/mol)$	12.40 ± 0.01	13.4 ± 0.2	12.01 (298.15 K)[61]

Liquid and vapor density of liquid DMA at 300 K

In order to determine the density of the two phases of pure DMA at 300 K, NVT simulations of the liquid DMA-air interface were carried out. The density profile of DMA (using the center of mass of each DMA), $\rho(z)$, is plotted in Figure 7 as a function of distance, z, from the center mass of the liquid slab in the direction perpendicular to the interface. While both the AA and the CG models show density oscillations near the interfacial region, the oscillations are much more pronounced for the CG model. The stronger CH_3-CH_3 interactions of the CG model as compared to the AA model could result in this increased structuring at the interface. This can also result

in differences in the orientation of the DMA molecules in the interfacial region. The distribution of the angle made by the CH_3-C vector of DMA molecules (with Z-coordinate of the center mass between 19 and 21 Å, the beginning of the interfacial region as seen in Figure 4.7) with the normal to the xy plane (air-liquid interface is along the z-direction, perpendicular to the xy plane) is shown in Figure 4.8. Although both the AA and CG models show a broad distribution, the AA has a broad peak in the 80-90° region whereas the CG shifted to around 120°.

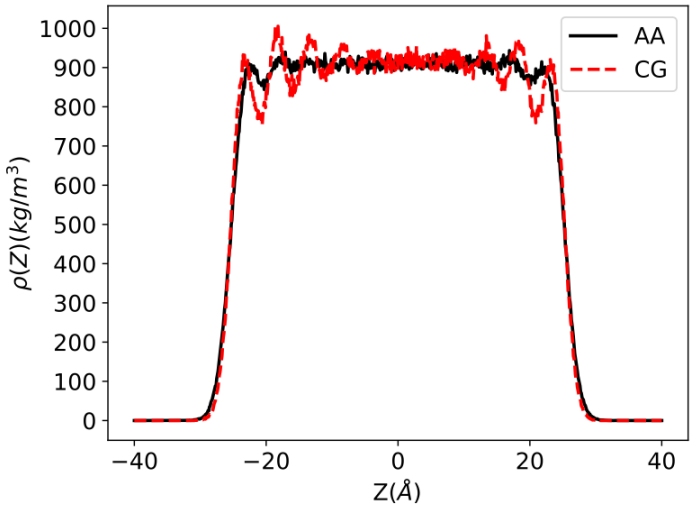

Figure 4.7. Density profile of DMA liquidvapor interface at 300 K for both the AA and CG simulations. Z represents the distance in the Z direction from the center of mass of the liquid and is perpendicular to the liquidvapor interface.

The density profiles in Figure 4.7 are fit to the following function,

$$\rho(z) = \frac{\rho_l + \rho_v}{2} - \frac{\rho_l - \rho_v}{2} tanh \frac{z - z_0}{d} \tag{4–11}$$

where ρ_l and ρ_v are the densities of the liquid and vapor phase, respectively, d is the width of the vapor-liquid interface and z_0 is the position of the Gibbs dividing surface of the interface.[63,64] The densities of the two phases are thereby determined from this fit and tabulated in Table 4.5 for both the CG model and the AA model along with the experimental value for the liquid phase. The CG model liquid density is in good agreement with the AA model, despite the differences

in interfacial orientation, and the liquid density compares reasonably well with the experimental value, although both CG and AA densities are higher than the experimental value by around 3%.

Table 4.5. The liquid density (ρ_L) and vapor density (ρ_V) for DMA at 300 K together with experiment data for the liquid at 298.15 K

AA		CG		Experiment
$\rho_L(kg/m^3)$	$\rho_v(kg/m^3)$	$\rho_L(kg/m^3)$	$\rho_V(kg/m^3)$	$\rho_L(kg/m^3)$
907.33	0.109	905.21	0.82	937[65]

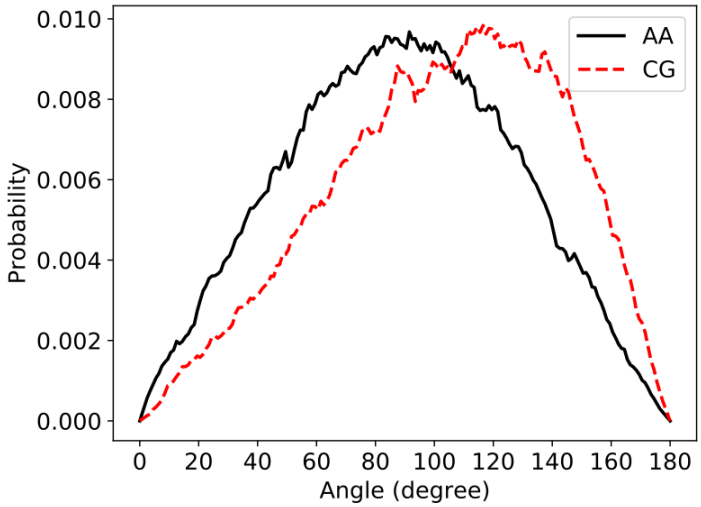

Figure 4.8. Distribution of the angle made by the CH_3C vector of DMA molecules (with the z-coordinate of the center of mass between 19 and 21 Å) with the normal to the xy plane. (Note the interface is along the z-axis.)

The good agreement between the CG and AA models for the surface tension and the liquid density, despite these parameters not being in the training set, does indicate that the reduced range formalism based CG model can reproduce the results for the structure and thermodynamics of the bulk and the interfacial region obtained from the higher level AA simulations. This could be in part due to a fortuitous cancellation of errors but is probably also due to the attention to an accurate description of the solvation structure that is the underlying philosophy in the development of this CG model. Previous works with SW CG models have also found good agreement of water/vapor, liquid alkane/water, liquid alkane/vapor, and liquid ether/vapor surface tensions.[39,40,45,46]

Finally, benchmark simulations on liquid DMA were performed to test the performance of the CG model against the AA. Table 4.6 shows our CG DMA model is at least 27 times faster than the AA DMA model. The scaling behavior clearly improves with increase in the system size.

Table 4.6. Computational efficiency for three different systems including 2560, 12800 and 64000 DMAs using 20, 40, 60, 80 and 100 CPU cores. All benchmark simulations were performed on the QB2 computational nodes of Louisiana Optical Network Infrastructure (LONI), each with 2.8 GHz E5- 2680v2 Xeon processors and 64 GB memory using the LAMMPS software.

	2560 DMAs			12800 DMAs			64000 DMAs		
CPUs (cores)	AA (ns/day)	CG (ns/day)	AA/CG ratio	AA (ns/day)	CG (ns/day)	AA/CG ratio	AA (ns/day)	CG (ns/day)	AA/CG ratio
20	4.44	118.38	27	0.89	25.80	29	0.17	5.72	34
40	8.14	218.61	27	1.71	47.39	28	0.33	10.88	33
60	10.80	304.60	28	2.44	69.26	28	0.50	14.89	30
80	11.94	361.60	30	3.16	89.54	28	0.67	19.58	29
100	15.05	437.72	29	4.04	109.30	27	0.82	24.20	30

Future work will concentrate on developing compatible CG models for different side chains of the peptoid system, namely hydrophilic neutral groups such as ethers as well as charged groups such as carboxylate.

4.4 Conclusions

A new CG model of N,N-dimethylacetamide, a representative of the peptoid backbone, has been developed. The underlying philosophy of the CG model development is to reproduce the solvation structure in the liquid of a more complex all atom model, OPLS-AA in this case, using reduced range functional forms that include three-body terms. Without electrostatics or any long range interactions, the CG DMA model is at least 27 times faster than AA model for the pure liquid system. The DMA molecules were represented as 6 heavy atoms interacting through short range potentials. The CG DMA model gives a good representation of the liquid DMA system, aqueous DMA solution as well as the liquid DMA-vapor interface. The pure bulk DMA solvation structure as well as the dilute DMA aqueous solution solvation structure are part of the model training set. The CG model was validated by a comparison to results from AA simulations of the surface tension, enthalpy of vaporization, isothermal compressibility and liquid density of pure

DMA, the solvation free energy of DMA in water as well as the solvation structure of both a concentrated aqueous solution of DMA as well as a simple peptoid oligomer aqueous solutions.

4.5 References

[1] Pu Du, Steven W Rick, and Revati Kumar, "Towards a coarse-grained model of the peptoid backbone: the case of n, n-dimethylacetamide", *Physical Chemistry Chemical Physics* **20**(36), pp. 23386–23396 (2018).

[2] Hongbo Feng, Xinyi Lu, Weiyu Wang, Nam-Goo Kang, and Jimmy Mays, "Block copolymers: Synthesis, self-assembly, and applications", *Polymers* **9**(10), pp. 494 (2017).

[3] Suzana Pereira Nunes, "Block copolymer membranes for aqueous solution applications", *Macromolecules* **49**(8), pp. 2905–2916 (2016).

[4] Jing Sun and Ronald N Zuckermann, "Peptoid polymers: a highly designable bioinspired material", *ACS nano* **7**(6), pp. 4715–4732 (2013).

[5] Corinna Fetsch, Arlett Grossmann, Lisa Holz, Jonas F Nawroth, and Robert Luxenhofer, "Polypeptoids from n-substituted glycine n-carboxyanhydrides: hydrophilic, hydrophobic, and amphiphilic polymers with poisson distribution", *Macromolecules* **44**(17), pp. 6746–6758 (2011).

[6] Adrianne M Rosales, Rachel A Segalman, and Ronald N Zuckermann, "Polypeptoids: a model system to study the effect of monomer sequence on polymer properties and self-assembly", *Soft Matter* **9**(35), pp. 8400–8414 (2013).

[7] Donghui Zhang, Samuel H Lahasky, Li Guo, Chang-Uk Lee, and Monika Lavan, "Polypeptoid materials: current status and future perspectives", *Macromolecules* **45**(15), pp. 5833–5841 (2012).

[8] Niklas Gangloff, Juliane Ulbricht, Thomas Lorson, Helmut Schlaad, and Robert Luxenhofer, "Peptoids and polypeptoids at the frontier of supra-and macromolecular engineering", *Chemical reviews* **116**(4), pp. 1753–1802 (2016).

[9] Hidenori Otsuka, Yukio Nagasaki, and Kazunori Kataoka, "Self-assembly of block copolymers", *Materials Today* **4**(3), pp. 30–36 (2001).

[10] Corinna Fetsch, Jens Gaitzsch, Lea Messager, Giuseppe Battaglia, and Robert Luxenhofer, "Self-assembly of amphiphilic block copolypeptoids–micelles, worms and polymersomes", *Scientific reports* **6**, pp. 33491 (2016).

[11] Garrett L Sternhagen, Sudipta Gupta, Yueheng Zhang, Vijay John, Gerald J Schneider, and Donghui Zhang, "Solution self-assemblies of sequence-defined ionic peptoid block copolymers", *Journal of the American Chemical Society* **140**(11), pp. 4100–4109 (2018).

[12] Corinna Fetsch and Robert Luxenhofer, "Highly defined multiblock copolypeptoids: Pushing the limits of living nucleophilic ring-opening polymerization", *Macromolecular rapid communications* **33**(19), pp. 1708–1713 (2012).

[13] Adrianne M Rosales, Hannah K Murnen, Steven R Kline, Ronald N Zuckermann, and Rachel A Segalman, "Determination of the persistence length of helical and non-helical polypeptoids in solution", *Soft Matter* **8**(13), pp. 3673–3680 (2012).

[14] Dina T Mirijanian, Ranjan V Mannige, Ronald N Zuckermann, and Stephen Whitelam, "Development and use of an atomistic charmm-based forcefield for peptoid simulation", *Journal of computational chemistry* **35**(5), pp. 360–370 (2014).

[15] Arushi Prakash, Marcel D Baer, Christopher J Mundy, and Jim Pfaendtner, "Peptoid backbone flexibilility dictates its interaction with water and surfaces: a molecular dynamics investigation", *Biomacromolecules* **19**(3), pp. 1006–1015 (2018).

[16] William L Jorgensen and Julian Tirado-Rives, "The opls [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin", *Journal of the American Chemical Society* **110**(6), pp. 1657–1666 (1988).

[17] William L Jorgensen, David S Maxwell, and Julian Tirado-Rives, "Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids", *Journal of the American Chemical Society* **118**(45), pp. 11225–11236 (1996).

[18] Pu Du, Ang Li, Xin Li, Yueheng Zhang, Changwoo Do, Lilin He, Steven W Rick, Vijay T John, Revati Kumar, and Donghui Zhang, "Aggregation of cyclic polypeptoids bearing zwitterionic end-groups with attractive dipole–dipole and solvophobic interactions: a study by small-angle neutron scattering and molecular dynamics simulation", *Physical Chemistry Chemical Physics* **19**(22), pp. 14388–14400 (2017).

[19] Alessandro Barducci, Jim Pfaendtner, and Massimiliano Bonomi, "Tackling sampling challenges in biomolecular simulations", In *Molecular modeling of proteins* , pp. 151–171. Springer (2015).

[20] Luca Monticelli and Emppu Salonen, *Biomolecular simulations: methods and protocols* volume 924, Springer (2013).

[21] Tatiana Maximova, Ryan Moffatt, Buyong Ma, Ruth Nussinov, and Amarda Shehu, "Principles and overview of sampling methods for modeling macromolecular structure and dynamics", *PLoS computational biology* **12**(4), pp. e1004619 (2016).

[22] Marissa G Saunders and Gregory A Voth, "Coarse-graining methods for computational biology", *Annual review of biophysics* **42**, pp. 73–93 (2013).

[23] J McCarty, IY Lyubimov, and MG Guenza, "Multiscale modeling of coarse-grained macromolecular liquids", *The Journal of Physical Chemistry B* **113**(35), pp. 11876–11886 (2009).

[24] Carlos F Lopez, Steve O Nielsen, Preston B Moore, John C Shelley, and Michael L Klein, "Self-assembly of a phospholipid langmuir monolayer using coarse-grained molecular dynamics simulations", *Journal of Physics: Condensed Matter* **14**(40), pp. 9431 (2002).

[25] Sebastian Kmiecik, Dominik Gront, Michal Kolinski, Lukasz Wieteska, Aleksandra Elzbieta Dawid, and Andrzej Kolinski, "Coarse-grained protein models and their applications", *Chemical reviews* **116**(14), pp. 7898–7936 (2016).

[26] Sophia P Hirakis, Britton Warren Boras, Lane W Votapka, Robert Dean Malmstrom, Andrew D McCulloch, and Rommie E Amaro, "Bridging scales through multiscale modeling: a case study on protein kinase a", *Frontiers in physiology* **6**, pp. 250 (2015).

[27] Shina CL Kamerlin, Spyridon Vicatos, Anatoly Dryga, and Arieh Warshel, "Coarse-grained (multiscale) simulations in studies of biophysical and chemical systems", *Annual review of physical chemistry* **62**, pp. 41–64 (2011).

[28] M Guenza, "Thermodynamic consistency and other challenges in coarse-graining models", *The European Physical Journal Special Topics* **224**(12), pp. 2177–2191 (2015).

[29] Sergei Izvekov, Jessica MJ Swanson, and Gregory A Voth, "Coarse-graining in interaction space: A systematic approach for replacing long-range electrostatics with short-range potentials", *The Journal of Physical Chemistry B* **112**(15), pp. 4711–4724 (2008).

[30] Valentina Tozzini, "Multiscale modeling of proteins", *Accounts of chemical research* **43**(2), pp. 220–230 (2009).

[31] Michael Levitt and Arieh Warshel, "Computer simulation of protein folding", *Nature* **253**(5494), pp. 694 (1975).

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

[33] Siewert J Marrink and D Peter Tieleman, "Perspective on the martini model", *Chemical Society Reviews* **42**(16), pp. 6801–6822 (2013).

[34] W Schommers, "Pair potentials in disordered many-particle systems: A study for liquid gallium", *Physical Review A* **28**(6), pp. 3599 (1983).

[35] Furio Ercolessi and James B Adams, "Interatomic potentials from first-principles calculations: the force-matching method", *EPL (Europhysics Letters)* **26**(8), pp. 583 (1994).

[36] Sergei Izvekov and Gregory A Voth, "A multiscale coarse-graining method for biomolecular systems", *The Journal of Physical Chemistry B* **109**(7), pp. 2469–2473 (2005).

[37] Sergei Izvekov, Michele Parrinello, Christian J Burnham, and Gregory A Voth, "Effective force fields for condensed phase systems from ab initio molecular dynamics simulation: A new method for force-matching", *The Journal of chemical physics* **120**(23), pp. 10896–10913 (2004).

[38] Alexander Lyubartsev, Alexander Mirzoev, LiJun Chen, and Aatto Laaksonen, "Systematic coarse-graining of molecular models by the newton inversion method", *Faraday discussions* **144**, pp. 43–56 (2010).

[39] Gaurav Gyawali, Samuel Sternfield, Revati Kumar, and Steven W Rick, "Coarse-grained models of aqueous and pure liquid alkanes", *Journal of chemical theory and computation* **13**(8), pp. 3846–3853 (2017).

[40] Bryan Raubenolt, Gaurav Gyawali, Wenwen Tang, Katy S Wong, and Steven W Rick, "Coarse-grained simulations of aqueous thermoresponsive polyethers.", *Polymers* **10**(5) (2018).

[41] Thomas K Haxton, Ronald N Zuckermann, and Stephen Whitelam, "Implicit-solvent coarse-grained simulation with a fluctuating interface reveals a molecular mechanism for peptoid monolayer buckling", *Journal of chemical theory and computation* **12**(1), pp. 345–352 (2015).

[42] Melissa L Hebert, Dhaval S Shah, Phillip Blake, J Phillip Turner, and Shannon L Servoss, "Tunable peptoid microspheres: effects of side chain chemistry and sequence", *Organic & biomolecular chemistry* **11**(27), pp. 4459–4464 (2013).

[43] Frank H Stillinger and Thomas A Weber, "Computer simulation of local order in condensed phases of silicon", *Physical review B* **31**(8), pp. 5262 (1985).

[44] R Kumar and JL Skinner, "Water simulation model with explicit three-molecule interactions", *The Journal of Physical Chemistry B* **112**(28), pp. 8311–8318 (2008).

[45] Valeria Molinero and Emily B Moore, "Water modeled as an intermediate element between carbon and silicon", *The Journal of Physical Chemistry B* **113**(13), pp. 4008–4016 (2008).

[46] William L Jorgensen, "Quantum and statistical mechanical studies of liquids. 10. transferable intermolecular potential functions for water, alcohols, and ethers. application to liquid water", *Journal of the American Chemical Society* **103**(2), pp. 335–340 (1981).

[47] Steve Plimpton, "Fast parallel algorithms for short-range molecular dynamics", *Journal of computational physics* **117**(1), pp. 1–19 (1995).

[48] Denis J Evans and Brad Lee Holian, "The nose–hoover thermostat", *The Journal of chemical physics* **83**(8), pp. 4069–4074 (1985).

[49] Glenn J Martyna, Douglas J Tobias, and Michael L Klein, "Constant pressure molecular dynamics algorithms", *The Journal of Chemical Physics* **101**(5), pp. 4177–4189 (1994).

[50] Brock A Luty and Wilfred F van Gunsteren, "Calculating electrostatic interactions using the particle- particle particle- mesh method with nonperiodic long-range interactions", *The Journal of Physical Chemistry* **100**(7), pp. 2581–2587 (1996).

[51] M Mezei, "The finite difference thermodynamic integration, tested on calculating the hydration free energy difference between acetone and dimethylamine in water", *The Journal*

of chemical physics **86**(12), pp. 7084–7088 (1987).

[52] Pavel V Klimovich, Michael R Shirts, and David L Mobley, "Guidelines for the analysis of free energy calculations", *Journal of computer-aided molecular design* **29**(5), pp. 397–411 (2015).

[53] Thomas Steinbrecher, InSuk Joung, and David A Case, "Soft-core potentials in thermodynamic integration: Comparing one-and two-step transformations", *Journal of computational chemistry* **32**(15), pp. 3253–3263 (2011).

[54] Thomas C Beutler, Alan E Mark, René C van Schaik, Paul R Gerber, and Wilfred F Van Gunsteren, "Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations", *Chemical physics letters* **222**(6), pp. 529–539 (1994).

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

[56] Liam C Jacobson and Valeria Molinero, "A methane- water model for coarse-grained simulations of solutions and clathrate hydrates", *The Journal of Physical Chemistry B* **114**(21), pp. 7302–7311 (2010).

[57] Gyawali, Gaurav and Rick, Steven W, "A git repository for the free energy", https://github.com/ggyawali/pair_sw_soft/tree/master (2017), [Online; accessed 10-May-2019].

[58] Carl Caleman, Paul J van Maaren, Minyan Hong, Jochen S Hub, Luciano T Costa, and David van der Spoel, "Force field benchmark of organic liquids: density, enthalpy of vaporization, heat capacities, surface tension, isothermal compressibility, volumetric expansion coefficient, and dielectric constant", *Journal of chemical theory and computation* **8**(1), pp. 61–74 (2011).

[59] Kazi A Motakabbir and M Berkowitz, "Isothermal compressibility of spc/e water", *Journal of Physical Chemistry* **94**(21), pp. 8359–8362 (1990).

[60] RK Shukla, Atul Kumar, Urvashi Srivastava, Naveen Awasthi, and JD Pandey, "Critical evaluation of surface tension of binary liquid mixtures from associated and nonassociated processes at various temperatures: an experimental and theoretical study", *Canadian Journal of Physics* **91**(3), pp. 211–220 (2012).

[61] Yitzhak Marcus, *The properties of solvents*, Wiley (1998).

[62] Devleena Shivakumar, Joshua Williams, Yujie Wu, Wolfgang Damm, John Shelley, and Woody Sherman, "Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the opls force field", *Journal of chemical theory and computation* **6**(5), pp. 1509–1519 (2010).

[63] José Alejandre, Dominic J Tildesley, and Gustavo A Chapela, "Molecular dynamics simulation of the orthobaric densities and surface tension of water", *The Journal of chemical*

physics **102**(11), pp. 4574–4583 (1995).

[64] Ryuji Sakamaki, Amadeu K Sum, Tetsu Narumi, and Kenji Yasuoka, "Molecular dynamics simulations of vapor/liquid coexistence using the nonpolarizable water models", *The Journal of chemical physics* **134**(12), pp. 124708 (2011).

[65] H Iloukhani and K Khanlarzadeh, "Densities, viscosities, and refractive indices for binary and ternary mixtures of n, n-dimethylacetamide (1)+ 2-methylbutan-2-ol (2)+ ethyl acetate (3) at 298.15 k for the liquid region and at ambient pressure", *Journal of Chemical & Engineering Data* **51**(4), pp. 1226–1231 (2006).

CHAPTER 5
CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

In this dissertation, the primary focus is to use molecular dynamics simulations to gain insight into the molecular mechanisms that govern self-assembly of polypeptoids in solution. As described in Chapter 2, atomistic simulations were performed to investigate the aggregation behavior of cyclic polypeptoids bearing zwitterionic end-groups in a low dielectric solvent, namely methanol, wherein small cluster formation was observed experimentally, ranging from a single polymer chain to small oligomers. Simulation results reveal that it is due to attractive dipole-dipole interactions for the initial approach of the cyclic polypeptoids. However, the attractive solvophobic effect starts to dominate while the effective repulsive interaction due to solvation of the dipoles drastically reduces the attractive dipole-dipole components.

In chapter 3, atomistic molecular dynamics simulations were carried out to study the self-assembly of sequence-defined singly charged polypeptoids in water solutions. Each of these polymers consists of five hydrophobic groups, 19 hydrophilic neutral groups, and one ionic monomer at the different position of the backbone. Simulation results reveal that water-polymer sidechain interactions play a critical role in the self-assembly of these polypeptoid systems. Charge-dipole secondary interactions (ionic side chain-water) dictate micelle size and shape, water penetration, and other micellar properties. Hence, micelle properties can be modulated by tuning these secondary interactions.

Although atomistic simulations can provide a complete description of these systems, the computational expense limits simulation time and length scales. Coarse-grained (CG) models, on the other hand, are computationally cheaper alternatives wherein atomistic level details are removed while retaining the relevant essential physics, thus making it easier to study more complicated systems, including self-assembly in soft matter. In chapter 4, A CG model of N, N-dimethyl acetamide (DMA), for further studying polypeptoid micellar interactions has been developed. The model was parameterized on atomistic simulations, using a hybridized

approach involving the OPLS-UA force field and the Stillinger-Weber potential. The model gives a good representation of the DMA systems and reproduces several thermodynamics properties. Furthermore, it showed good agreement with data obtained from all-atom simulations of a small simple peptoid oligopeptoid chain in water. Without long-ranged interactions and the absence of interaction sites on hydrogen atoms, the CG DMA model is at least 27 times faster than the higher resolution all-atom model.

5.2 Future Work

In chapter 3, the polypeptoids that were studied were singly charged, and the solvent was water. Based on the preliminary simulation results I have, the behavior of these peptoid polymers in a low dielectric constant solvent (methanol) are drastically different, as confirmed with experimental data. In future work, comprehensive studies of these peptoids (differing in the number of charged groups, the size of the hydrophobic region etc.) in solutions with different solvents will be carried out. I expect this will make a significant contribution to understanding the role of solvent polarity in the self-assembly of peptoid systems as well as the role of charges on micellar shape. Future efforts will also be focused on investigating the effect of polymer concentration, the addition of salts, and charge fraction on the micellar structure.

The CG model of DMA presented in chapter 3 represents a first step toward the systematic parameterization of a variety of peptoid sequences. In future work, we will focus on parameterizing several different peptoid side chains, including the side chains mentioned in chapter 2 and chapter 3. Once force field parameters for the CG model of peptoid are accurately obtained, aggregation or micelles formation by both charged and neutral polypeptoids will be studied. Simulations will include explicit pure water solvation as well as a range of added sodium salts at various temperatures. Molecular structural properties of aggregates or micelles will be analyzed for each condition.

APPENDIX A
COPYRIGHT PERMISSION I

The following is proof of copyright permission for the reuse of two papers that were published originally by the Royal Society of Chemistry.

Re-use permission requests

Material published by the Royal Society of Chemistry and other publishers is subject to all applicable copyright, database protection, and other rights. Therefore, for any publication, whether printed or electronic, permission must be obtained to use material for which the author(s) does not already own the copyright. This material may be, for example, a figure, diagram, table, photo or some other image.

Author reusing their own work published by the Royal Society of Chemistry

You do not need to request permission to reuse your own figures, diagrams, etc, that were originally published in a Royal Society of Chemistry publication. However, permission should be requested for use of the whole article or chapter except if reusing it in a thesis. If you are including an article or book chapter published by us in your thesis please ensure that your co-authors are aware of this.

Reuse of material that was published originally by the Royal Society of Chemistry must be accompanied by the appropriate acknowledgement of the publication. The form of the acknowledgement is dependent on the journal in which it was published originally, as detailed in 'Acknowledgements'.

The content is a screenshot of Licences, copyright & permissions page on the Royal Society of Chemistry's website (https://www.rsc.org/journals-books-databases/journal-authors-reviewers/licences-copyright-permissions/#deposition-sharing, accessed 28 August 2018).

APPENDIX B
COPYRIGHT PERMISSION II

The following is proof of copyright permission for the reused figure in this dissertation:

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Sep 27, 2019

This Agreement between Pu Du ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4677161495399
License date	Sep 27, 2019
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Advanced Materials
Licensed Content Title	Sequence Programmable Peptoid Polymers for Diverse Materials Applications
Licensed Content Author	Ronald N. Zuckermann, Matthew B. Francis, Effie Y. Zhou, et al
Licensed Content Date	Apr 8, 2015
Licensed Content Volume	27
Licensed Content Issue	38
Licensed Content Pages	27
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	3
Original Wiley figure/table number(s)	Figure 3
Will you be translating?	No

The content is a screenshot of permission details from RightsLink's website (https://copyright.com).

LIST OF REFERENCES

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

Adcock, S. A. and McCammon, J. A. (2006). Molecular dynamics: survey of methods for simulating the activity of proteins. *Chemical reviews*, 106(5):1589–1615.

Alejandre, J., Tildesley, D. J., and Chapela, G. A. (1995). Molecular dynamics simulation of the orthobaric densities and surface tension of water. *The Journal of chemical physics*, 102(11):4574–4583.

Allen, M. P. and Tildesley, D. J. (2012). *Computer simulation in chemical physics*, volume 397. Springer Science & Business Media.

Aprà, E., Bylaska, E. J., Dean, D. J., Fortunelli, A., Gao, F., Krstić, P. S., Wells, J. C., and Windus, T. L. (2003). Nwchem for materials science. *Computational Materials Science*, 28(2):209–221.

Archer, A. J. and Wilding, N. B. (2007). Phase behavior of a fluid with competing attractive and repulsive interactions. *Physical Review E*, 76(3):031501.

Arkın, H. and Janke, W. (2013). Gyration tensor based analysis of the shapes of polymer chains in an attractive spherical cage. *The Journal of chemical physics*, 138(5):054904.

Armand, P., Kirshenbaum, K., Falicov, A., Dunbrack Jr, R. L., Dill, K. A., Zuckermann, R. N., and Cohen, F. E. (1997). Chiral n-substituted glycines can form stable helical conformations. *Folding and Design*, 2(6):369–375.

Armand, P., Kirshenbaum, K., Goldsmith, R. A., Farr-Jones, S., Barron, A. E., Truong, K. T., Dill, K. A., Mierke, D. F., Cohen, F. E., Zuckermann, R. N., et al. (1998). Nmr determination of the major solution conformation of a peptoid pentamer with chiral side chains. *Proceedings of the National Academy of Sciences*, 95(8):4309–4314.

Aronovitz, J. and Nelson, D. (1986). Universal features of polymer shapes. *Journal de physique*, 47(9):1445–1456.

Avsar, S. Y., Kyropoulou, M., Di Leone, S., Schoenenberger, C.-A., Meier, W. P., and Palivan, C. G. (2018). Biomolecules turn self-assembling amphiphilic block co-polymer platforms into biomimetic interfaces. *Frontiers in chemistry*, 6.

Barducci, A., Pfaendtner, J., and Bonomi, M. (2015). Tackling sampling challenges in biomolecular simulations. In *Molecular modeling of proteins*, pages 151–171. Springer.

Barrat, J.-L., Baschnagel, J., and Lyulin, A. (2010). Molecular dynamics simulations of glassy polymers. *Soft Matter*, 6(15):3430–3446.

Barron, A. E. and Zuckerman, R. N. (1999). Bioinspired polymeric materials: in-between proteins and plastics. *Current opinion in chemical biology*, 3(6):681–687.

Bayly, C. I., Cieplak, P., Cornell, W., and Kollman, P. A. (1993). A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the resp model. *The Journal of Physical Chemistry*, 97(40):10269–10280.

Berendsen, H. J., Postma, J. v., van Gunsteren, W. F., DiNola, A., and Haak, J. (1984). Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81(8):3684–3690.

Bernabei, M., Bacova, P., Moreno, A. J., Narros, A., and Likos, C. N. (2013). Fluids of semiflexible ring polymers: effective potentials and clustering. *Soft Matter*, 9(4):1287–1300.

Best, R. B., Zhu, X., Shim, J., Lopes, P. E., Mittal, J., Feig, M., and MacKerell Jr, A. D. (2012). Optimization of the additive charmm all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain $\chi 1$ and $\chi 2$ dihedral angles. *Journal of chemical theory and computation*, 8(9):3257–3273.

Beutler, T. C., Mark, A. E., van Schaik, R. C., Gerber, P. R., and Van Gunsteren, W. F. (1994). Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chemical physics letters*, 222(6):529–539.

Bohn, M. and Heermann, D. W. (2010). Diffusion-driven looping provides a consistent framework for chromatin organization. *PloS one*, 5(8):e12218.

Breneman, C. M. and Wiberg, K. B. (1990). Determining atom-centered monopoles from molecular electrostatic potentials. the need for high sampling density in formamide conformational analysis. *Journal of Computational Chemistry*, 11(3):361–373.

Brooks, B. R., Brooks III, C. L., Mackerell Jr, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., et al. (2009). Charmm: the biomolecular simulation program. *Journal of computational chemistry*, 30(10):1545–1614.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. a., and Karplus, M. (1983). Charmm: a program for macromolecular energy, minimization, and dynamics calculations. *Journal of computational chemistry*, 4(2):187–217.

Butterfoss, G. L., Renfrew, P. D., Kuhlman, B., Kirshenbaum, K., and Bonneau, R. (2009). A preliminary survey of the peptoid folding landscape. *Journal of the American Chemical Society*, 131(46):16798–16807.

Butterfoss, G. L., Yoo, B., Jaworski, J. N., Chorny, I., Dill, K. A., Zuckermann, R. N., Bonneau, R., Kirshenbaum, K., and Voelz, V. A. (2012). De novo structure prediction and experimental characterization of folded peptoid oligomers. *Proceedings of the National Academy of Sciences*, 109(36):14320–14325.

Caleman, C., van Maaren, P. J., Hong, M., Hub, J. S., Costa, L. T., and van der Spoel, D. (2011). Force field benchmark of organic liquids: density, enthalpy of vaporization, heat capacities, surface tension, isothermal compressibility, volumetric expansion coefficient, and dielectric constant. *Journal of chemical theory and computation*, 8(1):61–74.

Campbell, A. I., Anderson, V. J., van Duijneveldt, J. S., and Bartlett, P. (2005). Dynamical arrest in attractive colloids: The effect of long-range repulsion. *Physical review letters*, 94(20):208301.

Cannon, J. W., Aronovitz, J. A., and Goldbart, P. (1991). Equilibrium distribution of shapes for linear and star macromolecules. *Journal de Physique I*, 1(5):629–645.

Car, R. and Parrinello, M. (1985). Unified approach for molecular dynamics and density-functional theory. *Physical review letters*, 55(22):2471.

Cavalli, A., Ferrara, P., and Caflisch, A. (2002). Weak temperature dependence of the free energy surface and folding pathways of structured peptides. *Proteins: Structure, Function, and Bioinformatics*, 47(3):305–314.

Chongsiriwatana, N. P., Patch, J. A., Czyzewski, A. M., Dohm, M. T., Ivankin, A., Gidalevitz, D., Zuckermann, R. N., and Barron, A. E. (2008). Peptoids that mimic the structure, function, and mechanism of helical antimicrobial peptides. *Proceedings of the National Academy of Sciences*, 105(8):2794–2799.

Cohen, A. J., Mori-Sánchez, P., and Yang, W. (2011). Challenges for density functional theory. *Chemical reviews*, 112(1):289–320.

Cramer, C. J. and Truhlar, D. G. (2009). Density functional theory for transition metals and transition metal chemistry. *Physical Chemistry Chemical Physics*, 11(46):10757–10816.

Crapster, J. A., Guzei, I. A., and Blackwell, H. E. (2013). A peptoid ribbon secondary structure. *Angewandte Chemie International Edition*, 52(19):5079–5084.

D.A. Case, R.M. Betz, D. C. T. C. I. T. D. R. D. T. G. H. G. A. G. N. H. S. I. P. J. J. K. A. K. T. L. S. L. P. L. C. L. T. L. R. L. B. M. D. M. K. M. G. M. H. N. H. N. I. O. A. O. D. R. A. R. C. S. C. S. W. B.-S. J. S. R. W. J. W. R. W. X. W. L. X. and Kollman, P. Amber 2016. University of California, San Francisco.

Darden, T., York, D., and Pedersen, L. (1993). Particle mesh ewald: An n log (n) method for ewald sums in large systems. *The Journal of chemical physics*, 98(12):10089–10092.

Daura, X., Mark, A. E., and Van Gunsteren, W. F. (1998). Parametrization of aliphatic chn united atoms of gromos96 force field. *Journal of computational chemistry*, 19(5):535–547.

Dibble, C. J., Kogan, M., and Solomon, M. J. (2006). Structure and dynamics of colloidal depletion gels: Coincidence of transitions and heterogeneity. *Physical Review E*, 74(4):041403.

Dima, R. I. and Thirumalai, D. (2004). Asymmetry in the shapes of folded and denatured states of proteins. *The Journal of Physical Chemistry B*, 108(21):6564–6570.

Dobson, C. M., Šali, A., and Karplus, M. (1998). Protein folding: a perspective from theory and experiment. *Angewandte Chemie International Edition*, 37(7):868–893.

Dong, Y., Li, Q., and Martini, A. (2013). Molecular dynamics simulation of atomic friction: A review and guide. *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, 31(3):030801.

Du, P., Li, A., Li, X., Zhang, Y., Do, C., He, L., Rick, S. W., John, V. T., Kumar, R., and Zhang, D. (2017). Aggregation of cyclic polypeptoids bearing zwitterionic end-groups with attractive dipole–dipole and solvophobic interactions: A study by small-angle neutron scattering and molecular dynamics simulation. *Physical Chemistry Chemical Physics*, 19(22):14388–14400.

Du, P., Rick, S. W., and Kumar, R. (2018). Towards a coarse-grained model of the peptoid backbone: the case of n, n-dimethylacetamide. *Physical Chemistry Chemical Physics*, 20(36):23386–23396.

Durham, E., Dorr, B., Woetzel, N., Staritzbichler, R., and Meiler, J. (2009). Solvent accessible surface area approximations for rapid and accurate protein structure prediction. *Journal of molecular modeling*, 15(9):1093–1108.

Eastman, P., Friedrichs, M. S., Chodera, J. D., Radmer, R. J., Bruns, C. M., Ku, J. P., Beauchamp, K. A., Lane, T. J., Wang, L.-P., Shukla, D., et al. (2012). Openmm 4: a reusable, extensible, hardware independent library for high performance molecular simulation. *Journal of chemical theory and computation*, 9(1):461–469.

Ercolessi, F. and Adams, J. B. (1994). Interatomic potentials from first-principles calculations: the force-matching method. *EPL (Europhysics Letters)*, 26(8):583.

Evans, D. J. and Holian, B. L. (1985). The nose–hoover thermostat. *The Journal of chemical physics*, 83(8):4069–4074.

Feng, H., Lu, X., Wang, W., Kang, N.-G., and Mays, J. W. (2017). Block copolymers: Synthesis, self-assembly, and applications. *Polymers*, 9(10):494.

Fetsch, C., Gaitzsch, J., Messager, L., Battaglia, G., and Luxenhofer, R. (2016). Self-assembly of amphiphilic block copolypeptoids–micelles, worms and polymersomes. *Scientific reports*, 6:33491.

Fetsch, C., Grossmann, A., Holz, L., Nawroth, J. F., and Luxenhofer, R. (2011). Polypeptoids from n-substituted glycine n-carboxyanhydrides: hydrophilic, hydrophobic, and amphiphilic polymers with poisson distribution. *Macromolecules*, 44(17):6746–6758.

Fetsch, C. and Luxenhofer, R. (2012). Highly defined multiblock copolypeptoids: Pushing the limits of living nucleophilic ring-opening polymerization. *Macromolecular rapid communications*, 33(19):1708–1713.

Fetsch, C. and Luxenhofer, R. (2013). Thermal properties of aliphatic polypeptoids. *Polymers*, 5(1):112–127.

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, Jr., J. A., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, ., Foresman, J. B., Ortiz, J. V., Cioslowski, J., and Fox, D. J. Gaussian09 Revision D.01. Gaussian Inc. Wallingford CT 2009.

Gangloff, N., Ulbricht, J., Lorson, T., Schlaad, H., and Luxenhofer, R. (2015). Peptoids and polypeptoids at the frontier of supra-and macromolecular engineering. *Chemical reviews*, 116(4):1753–1802.

Glinka, C., Barker, J., Hammouda, B., Krueger, S., Moyer, J., and Orts, W. (1998). The 30 m small-angle neutron scattering instruments at the national institute of standards and technology. *Journal of applied crystallography*, 31(3):430–445.

Godfrin, P. D., Valadez-Pérez, N. E., Castaneda-Priego, R., Wagner, N. J., and Liu, Y. (2014). Generalized phase behavior of cluster formation in colloidal dispersions with competing interactions. *Soft matter*, 10(28):5061–5071.

Gorske, B. C., Mumford, E. M., Gerrity, C. G., and Ko, I. (2017). A peptoid square helix via synergistic control of backbone dihedral angles. *Journal of the American Chemical Society*, 139(24):8070–8073.

Götze, I., Harreis, H., and Likos, C. (2004). Tunable effective interactions between dendritic macromolecules. *The Journal of chemical physics*, 120(16):7761–7771.

Guenza, M. (2015). Thermodynamic consistency and other challenges in coarse-graining models. *The European Physical Journal Special Topics*, 224(12):2177–2191.

Guo, L., Lahasky, S. H., Ghale, K., and Zhang, D. (2012). N-heterocyclic carbene-mediated zwitterionic polymerization of n-substituted n-carboxyanhydrides toward poly (α-peptoid) s: kinetic, mechanism, and architectural control. *Journal of the American Chemical Society*, 134(22):9163–9171.

Guo, L., Li, J., Brown, Z., Ghale, K., and Zhang, D. (2011). Synthesis and characterization of cyclic and linear helical poly (α-peptoid) s by n-heterocyclic carbene-mediated ring-opening polymerizations of n-substituted n-carboxyanhydrides. *Peptide Science*, 96(5):596–603.

Guo, L. and Zhang, D. (2009). Cyclic poly (α-peptoid) s and their block copolymers from n-heterocyclic carbene-mediated ring-opening polymerizations of n-substituted n-carboxylanhydrides. *Journal of the American Chemical Society*, 131(50):18072–18074.

Gyawali, G., Sternfield, S., Kumar, R., and Rick, S. W. (2017). Coarse-grained models of aqueous and pure liquid alkanes. *Journal of chemical theory and computation*, 13(8):3846–3853.

Gyawali, Gaurav and Rick, Steven W (2017). A git repository for the free energy. https://github.com/ggyawali/pair_sw_soft/tree/master. [Online; accessed 10-May-2019].

Ha, B.-Y. and Liu, A. (1999). Kinetics of bundle growth in dna condensation. *EPL (Europhysics Letters)*, 46(5):624.

Ha, B.-Y. and Liu, A. J. (1998). Charge oscillations and many-body effects in bundles of like-charged rods. *Physical Review E*, 58(5):6281.

Hagen, S. J., Hofrichter, J., and Eaton, W. A. (1995). Protein reaction kinetics in a room-temperature glass. *Science*, 269(5226):959–962.

Hammouda, B. (1993). Sans from homogeneous polymer mixtures: a unified overview. In *Polymer Characteristics*, pages 87–133. Springer.

Hammouda, B. (2009). The mystery of clustering in macromolecular media. *Polymer*, 50(22):5293–5297.

Hammouda, B., Ho, D. L., and Kline, S. (2004). Insight into clustering in poly (ethylene oxide) solutions. *Macromolecules*, 37(18):6932–6937.

Harrison, J. A., Schall, J. D., Knippenberg, M. T., Gao, G., and Mikulski, P. T. (2008). Elucidating atomic-scale friction using molecular dynamics and specialized analysis techniques. *Journal of Physics: Condensed Matter*, 20(35):354009.

Haxton, T. K., Zuckermann, R. N., and Whitelam, S. (2015). Implicit-solvent coarse-grained simulation with a fluctuating interface reveals a molecular mechanism for peptoid monolayer buckling. *Journal of chemical theory and computation*, 12(1):345–352.

Hebert, M. L., Shah, D. S., Blake, P., Turner, J. P., and Servoss, S. L. (2013). Tunable peptoid microspheres: effects of side chain chemistry and sequence. *Organic & biomolecular chemistry*, 11(27):4459–4464.

Henle, M. L. and Pincus, P. A. (2005). Equilibrium bundle size of rodlike polyelectrolytes with counterion-induced attractive interactions. *Physical Review E*, 71(6):060801.

Hirakis, S. P., Boras, B. W., Votapka, L. W., Malmstrom, R. D., McCulloch, A. D., and Amaro, R. E. (2015). Bridging scales through multiscale modeling: a case study on protein kinase a. *Frontiers in physiology*, 6:250.

Hockney, R. W. (1970). The potential calculation and some applications. *Methods Comput. Phys.*, 9:136.

Hollingsworth, S. A. and Dror, R. O. (2018). Molecular dynamics simulation for all. *Neuron*, 99(6):1129–1143.

Honeycutt, J. and Thirumalai, D. (1989). Static properties of polymer chains in porous media. *The Journal of Chemical Physics*, 90(8):4542–4559.

Hong, L., Cacciuto, A., Luijten, E., and Granick, S. (2006). Clusters of charged janus spheres. *Nano letters*, 6(11):2510–2514.

Hrubỳ, M., Filippov, S. K., and Štěpánek, P. (2016). Biomedical application of block copolymers. *Macromolecular Self-assembly*, pages 231–250.

Huang, K., Wu, C. W., Sanborn, T. J., Patch, J. A., Kirshenbaum, K., Zuckermann, R. N., Barron, A. E., and Radhakrishnan, I. (2006). A threaded loop conformation adopted by a family of peptoid nonamers. *Journal of the American Chemical Society*, 128(5):1733–1738.

Iloukhani, H. and Khanlarzadeh, K. (2006). Densities, viscosities, and refractive indices for binary and ternary mixtures of n, n-dimethylacetamide (1)+ 2-methylbutan-2-ol (2)+ ethyl acetate (3) at 298.15 k for the liquid region and at ambient pressure. *Journal of Chemical & Engineering Data*, 51(4):1226–1231.

Izvekov, S., Parrinello, M., Burnham, C. J., and Voth, G. A. (2004). Effective force fields for condensed phase systems from ab initio molecular dynamics simulation: A new method for force-matching. *The Journal of chemical physics*, 120(23):10896–10913.

Izvekov, S., Swanson, J. M., and Voth, G. A. (2008). Coarse-graining in interaction space: A systematic approach for replacing long-range electrostatics with short-range potentials. *The Journal of Physical Chemistry B*, 112(15):4711–4724.

Izvekov, S. and Voth, G. A. (2005). A multiscale coarse-graining method for biomolecular systems. *The Journal of Physical Chemistry B*, 109(7):2469–2473.

Jacobson, L. C. and Molinero, V. (2010). A methane- water model for coarse-grained simulations of solutions and clathrate hydrates. *The Journal of Physical Chemistry B*, 114(21):7302–7311.

Jagodzinski, O., Eisenriegler, E., and Kremer, K. (1992). Universal shape properties of open and closed polymer chains: renormalization group analysis and monte carlo experiments. *Journal de Physique I*, 2(12):2243–2279.

Jang, S. S., Çağin, T., and Goddard III, W. A. (2003). Effect of cyclic chain architecture on properties of dilute solutions of polyethylene from molecular dynamics simulations. *The Journal of chemical physics*, 119(3):1843–1854.

Jensen, F. (2017). *Introduction to computational chemistry*. John wiley & sons.

Jorgensen, W. L. (1981). Quantum and statistical mechanical studies of liquids. 10. transferable intermolecular potential functions for water, alcohols, and ethers. application to liquid water. *Journal of the American Chemical Society*, 103(2):335–340.

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics*, 79(2):926–935.

Jorgensen, W. L., Madura, J. D., and Swenson, C. J. (1984). Optimized intermolecular potential functions for liquid hydrocarbons. *Journal of the American Chemical Society*, 106(22):6638–6646.

Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J. (1996a). Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118(45):11225–11236.

Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J. (1996b). Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118(45):11225–11236.

Jorgensen, W. L. and Tirado-Rives, J. (1988). The opls [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, 110(6):1657–1666.

Jusufi, A., Watzlawek, M., and Löwen, H. (1999). Effective interaction between star polymers. *Macromolecules*, 32(13):4470–4473.

Kakizawa, Y. and Kataoka, K. (2002). Block copolymer self-assembly into monodispersive nanoparticles with hybrid core of antisense dna and calcium phosphate. *Langmuir*, 18(12):4539–4543.

Kamerlin, S. C., Vicatos, S., Dryga, A., and Warshel, A. (2011). Coarse-grained (multiscale) simulations in studies of biophysical and chemical systems. *Annual review of physical chemistry*, 62:41–64.

Karplus, M. and McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nature Structural & Molecular Biology*, 9(9):646.

Kirshenbaum, K., Barron, A. E., Goldsmith, R. A., Armand, P., Bradley, E. K., Truong, K. T., Dill, K. A., Cohen, F. E., and Zuckermann, R. N. (1998). Sequence-specific polypeptoids: a diverse family of heteropolymers with stable secondary structure. *Proceedings of the National Academy of Sciences*, 95(8):4303–4308.

Kirshenbaum, K., Zuckermann, R. N., and Dill, K. A. (1999). Designing polymers that mimic biomolecules. *Current opinion in structural biology*, 9(4):530–535.

Klimovich, P. V., Shirts, M. R., and Mobley, D. L. (2015). Guidelines for the analysis of free energy calculations. *Journal of computer-aided molecular design*, 29(5):397–411.

Kline, S. R. (2006). Reduction and analysis of sans and usans data using igor pro. *Journal of applied crystallography*, 39(6):895–900.

Kmiecik, S., Gront, D., Kolinski, M., Wieteska, L., Dawid, A. E., and Kolinski, A. (2016). Coarse-grained protein models and their applications. *Chemical reviews*, 116(14):7898–7936.

Knight, A. S., Zhou, E. Y., Francis, M. B., and Zuckermann, R. N. (2015). Sequence programmable peptoid polymers for diverse materials applications. *Advanced Materials*,

27(38):5665–5691.

Konieczny, M., Likos, C. N., and Löwen, H. (2004). Soft effective interactions between weakly charged polyelectrolyte chains. *The Journal of chemical physics*, 121(10):4913–4924.

Kudirka, R., Tran, H., Sanii, B., Nam, K. T., Choi, P. H., Venkateswaran, N., Chen, R., Whitelam, S., and Zuckermann, R. N. (2011). Folding of a single-chain, information-rich polypeptoid sequence into a highly ordered nanosheet. *Peptide Science*, 96(5):586–595.

Kumar, R. and Skinner, J. (2008). Water simulation model with explicit three-molecule interactions. *The Journal of Physical Chemistry B*, 112(28):8311–8318.

Lahasky, S. H., Serem, W. K., Guo, L., Garno, J. C., and Zhang, D. (2011). Synthesis and characterization of cyclic brush-like polymers by n-heterocyclic carbene-mediated zwitterionic polymerization of n-propargyl n-carboxyanhydride and the grafting-to approach. *Macromolecules*, 44(23):9063–9074.

Laio, A. and Parrinello, M. (2002). Escaping free-energy minima. *Proceedings of the National Academy of Sciences*, 99(20):12562–12566.

Lee, B. and Richards, F. M. (1971). The interpretation of protein structures: estimation of static accessibility. *Journal of molecular biology*, 55(3):379–IN4.

Lee, B.-C., Connolly, M. D., and Zuckermann, R. N. (2009). Bioinspired polymers for nanoscience research. *ChemInform*, 40(32):i.

Lee, C.-U., Li, A., Ghale, K., and Zhang, D. (2013). Crystallization and melting behaviors of cyclic and linear polypeptoids with alkyl side chains. *Macromolecules*, 46(20):8213–8223.

Levitt, M. and Warshel, A. (1975). Computer simulation of protein folding. *Nature*, 253(5494):694.

Li, A., Lu, L., Li, X., He, L., Do, C., Garno, J. C., and Zhang, D. (2016). Amidine-mediated zwitterionic ring-opening polymerization of n-alkyl n-carboxyanhydride: mechanism, kinetics, and architecture elucidation. *Macromolecules*, 49(4):1163–1171.

Li, X., Hong, K., Liu, Y., Shew, C.-Y., Liu, E., Herwig, K. W., Smith, G. S., Zhao, J., Zhang, G., Pispas, S., et al. (2010). Water distributions in polystyrene-block-poly [styrene-g-poly (ethylene oxide)] block grafted copolymer system in aqueous solutions revealed by contrast variation small angle neutron scattering study. *The Journal of chemical physics*, 133(14):144912.

Likos, C., Löwen, H., Watzlawek, M., Abbas, B., Jucknischke, O., Allgaier, J., and Richter, D. (1998). Star polymers viewed as ultrasoft colloidal particles. *Physical review letters*, 80(20):4450.

Likos, C., Rosenfeldt, S., Dingenouts, N., Ballauff, M., Lindner, P., Werner, N., and Vögtle, F. (2002). Gaussian effective interaction between flexible dendrimers of fourth generation: A theoretical and experimental study. *The Journal of chemical physics*, 117(4):1869–1877.

Likos, C., Schmidt, M., Löwen, H., Ballauff, M., Pötschke, D., and Lindner, P. (2001). Soft interaction between dissolved flexible dendrimers: theory and experiment. *Macromolecules*, 34(9):2914–2920.

Liu, Y., Bryantsev, V. S., Diallo, M. S., and Goddard Iii, W. A. (2009). Pamam dendrimers undergo ph responsive conformational changes without swelling. *Journal of the American Chemical Society*, 131(8):2798–2799.

Lopez, C. F., Nielsen, S. O., Moore, P. B., Shelley, J. C., and Klein, M. L. (2002). Self-assembly of a phospholipid langmuir monolayer using coarse-grained molecular dynamics simulations. *Journal of Physics: Condensed Matter*, 14(40):9431.

Louis, A., Bolhuis, P., Hansen, J., and Meijer, E. (2000). Can polymer coils be modeled as soft colloids? *Physical review letters*, 85(12):2522.

Lu, P. J., Conrad, J. C., Wyss, H. M., Schofield, A. B., and Weitz, D. A. (2006). Fluids of clusters in attractive colloids. *Physical Review Letters*, 96(2):028306.

Luty, B. A. and van Gunsteren, W. F. (1996). Calculating electrostatic interactions using the particle- particle particle- mesh method with nonperiodic long-range interactions. *The Journal of Physical Chemistry*, 100(7):2581–2587.

Luxenhofer, R., Fetsch, C., and Grossmann, A. (2013). Polypeptoids: A perfect match for molecular definition and macromolecular engineering? *Journal of Polymer Science Part A: Polymer Chemistry*, 51(13):2731–2752.

Lyubartsev, A., Mirzoev, A., Chen, L., and Laaksonen, A. (2010). Systematic coarse-graining of molecular models by the newton inversion method. *Faraday discussions*, 144:43–56.

Mackerell Jr, A. D., Feig, M., and Brooks III, C. L. (2004). Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of computational chemistry*, 25(11):1400–1415.

Mannige, R. V., Haxton, T. K., Proulx, C., Robertson, E. J., Battigelli, A., Butterfoss, G. L., Zuckermann, R. N., and Whitelam, S. (2015). Peptoid nanosheets exhibit a new secondary-structure motif. *Nature*, 526(7573):415.

Marcus, Y. (1998). *The properties of solvents*. Wiley.

Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P., and De Vries, A. H. (2007). The martini force field: coarse grained model for biomolecular simulations. *The journal of physical chemistry B*, 111(27):7812–7824.

Marrink, S. J. and Tieleman, D. P. (2013). Perspective on the martini model. *Chemical Society Reviews*, 42(16):6801–6822.

Martínez, L., Andrade, R., Birgin, E. G., and Martínez, J. M. (2009). Packmol: a package for building initial configurations for molecular dynamics simulations. *Journal of computational*

chemistry, 30(13):2157–2164.

Martyna, G. J., Tobias, D. J., and Klein, M. L. (1994). Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics*, 101(5):4177–4189.

Maximova, T., Moffatt, R., Ma, B., Nussinov, R., and Shehu, A. (2016). Principles and overview of sampling methods for modeling macromolecular structure and dynamics. *PLoS computational biology*, 12(4):e1004619.

McCarty, J., Lyubimov, I., and Guenza, M. (2009). Multiscale modeling of coarse-grained macromolecular liquids. *The Journal of Physical Chemistry B*, 113(35):11876–11886.

Melchy, P.-É. and Eikerling, M. (2014). Physical theory of ionomer aggregation in water. *Physical Review E*, 89(3):032603.

Meng, G., Arkus, N., Brenner, M. P., and Manoharan, V. N. (2010). The free-energy landscape of clusters of attractive hard spheres. *Science*, 327(5965):560–563.

Mezei, M. (1987). The finite difference thermodynamic integration, tested on calculating the hydration free energy difference between acetone and dimethylamine in water. *The Journal of chemical physics*, 86(12):7084–7088.

Miller, S. M., Simon, R. J., Ng, S., Zuckermann, R. N., Kerr, J. M., and Moos, W. H. (1994). Proteolytic studies of homologous peptide and n-substituted glycine peptoid oligomers. *Bioorganic & medicinal chemistry letters*, 4(22):2657–2662.

Mirijanian, D. T., Mannige, R. V., Zuckermann, R. N., and Whitelam, S. (2014). Development and use of an atomistic charmm-based forcefield for peptoid simulation. *Journal of computational chemistry*, 35(5):360–370.

Mitsutake, A., Mori, Y., and Okamoto, Y. (2013). Enhanced sampling algorithms. In *Biomolecular Simulations*, pages 153–195. Springer.

Mladek, B. M., Gottwald, D., Kahl, G., Neumann, M., and Likos, C. N. (2006). Formation of polymorphic cluster phases for a class of models of purely repulsive soft spheres. *Physical review letters*, 96(4):045701.

Moehle, K. and Hofmann, H.-J. (1996). Peptides and peptoidsa quantum chemical structure comparison. *Biopolymers*, 38(6):781–790.

Möhle, K. and Hofmann, H.-J. (1996). Peptides and peptoids-a systematic structure comparison. In *Molecular modeling annual*, volume 2, pages 307–311. Springer.

Molinero, V. and Moore, E. B. (2008). Water modeled as an intermediate element between carbon and silicon. *The Journal of Physical Chemistry B*, 113(13):4008–4016.

Monticelli, L. and Salonen, E. (2013). *Biomolecular simulations: methods and protocols*, volume 924. Springer.

Motakabbir, K. A. and Berkowitz, M. (1990). Isothermal compressibility of spc/e water. *Journal of Physical Chemistry*, 94(21):8359–8362.

Mukherjee, S., Zhou, G., Michel, C., and Voelz, V. A. (2015). Insights into peptoid helix folding cooperativity from an improved backbone potential. *The Journal of Physical Chemistry B*, 119(50):15407–15417.

Murnen, H. K., Rosales, A. M., Dobrynin, A. V., Zuckermann, R. N., and Segalman, R. A. (2013). Persistence length of polyelectrolytes with precisely located charges. *Soft Matter*, 9(1):90–98.

Murphy, J. E., Uno, T., Hamer, J. D., Cohen, F. E., Dwarki, V., and Zuckermann, R. N. (1998). A combinatorial approach to the discovery of efficient cationic peptoid reagents for gene delivery. *Proceedings of the National Academy of Sciences*, 95(4):1517–1522.

Nunes, S. P. (2016). Block copolymer membranes for aqueous solution applications. *Macromolecules*, 49(8):2905–2916.

O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., and Hutchison, G. R. (2011). Open babel: An open chemical toolbox. *Journal of cheminformatics*, 3(1):33.

Otsuka, H., Nagasaki, Y., and Kataoka, K. (2001). Self-assembly of block copolymers. *Materials Today*, 4(3):30–36.

Park, S. H. and Szleifer, I. (2011). Structural and dynamical characteristics of peptoid oligomers with achiral aliphatic side chains studied by molecular dynamics simulation. *The Journal of Physical Chemistry B*, 115(37):10967–10975.

Parr, R. G. (1980). Density functional theory of atoms and molecules. In *Horizons of Quantum Chemistry*, pages 5–15. Springer.

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

Plimpton, S. (1995). Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics*, 117(1):1–19.

Ponder, J. W. and Richards, F. M. (1987). An efficient newton-like method for molecular mechanics energy minimization of large molecules. *Journal of Computational Chemistry*, 8(7):1016–1024.

Prakash, A., Baer, M. D., Mundy, C. J., and Pfaendtner, J. (2018). Peptoid backbone flexibilility dictates its interaction with water and surfaces: a molecular dynamics investigation. *Biomacromolecules*, 19(3):1006–1015.

Raubenolt, B., Gyawali, G., Tang, W., Wong, K. S., and Rick, S. W. (2018). Coarse-grained simulations of aqueous thermoresponsive polyethers. *Polymers*, 10(5).

Renfrew, P. D., Craven, T. W., Butterfoss, G. L., Kirshenbaum, K., and Bonneau, R. (2014). A rotamer library to enable modeling and design of peptoid foldamers. *Journal of the American Chemical Society*, 136(24):8772–8782.

Rosales, A. M., Murnen, H. K., Kline, S. R., Zuckermann, R. N., and Segalman, R. A. (2012). Determination of the persistence length of helical and non-helical polypeptoids in solution. *Soft Matter*, 8(13):3673–3680.

Rosales, A. M., Murnen, H. K., Zuckermann, R. N., and Segalman, R. A. (2010). Control of crystallization and melting behavior in sequence specific polypeptoids. *Macromolecules*, 43(13):5627–5636.

Rosales, A. M., Segalman, R. A., and Zuckermann, R. N. (2013). Polypeptoids: a model system to study the effect of monomer sequence on polymer properties and self-assembly. *Soft Matter*, 9(35):8400–8414.

Roux, B. (1995). The calculation of the potential of mean force using computer simulations. *Computer physics communications*, 91(1-3):275–282.

Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of computational physics*, 23(3):327–341.

Sakamaki, R., Sum, A. K., Narumi, T., and Yasuoka, K. (2011). Molecular dynamics simulations of vapor/liquid coexistence using the nonpolarizable water models. *The Journal of chemical physics*, 134(12):124708.

Samaddar, P., Deep, A., and Kim, K.-H. (2018). An engineering insight into block copolymer self-assembly: Contemporary application from biomedical research to nanotechnology. *Chemical Engineering Journal*, 342:71–89.

Sanii, B., Kudirka, R., Cho, A., Venkateswaran, N., Olivier, G. K., Olson, A. M., Tran, H., Harada, R. M., Tan, L., and Zuckermann, R. N. (2011). Shaken, not stirred: collapsing a peptoid monolayer to produce free-floating, stable nanosheets. *Journal of the American Chemical Society*, 133(51):20808–20815.

Saunders, M. G. and Voth, G. A. (2013). Coarse-graining methods for computational biology. *Annual review of biophysics*, 42:73–93.

Schommers, W. (1983). Pair potentials in disordered many-particle systems: A study for liquid gallium. *Physical Review A*, 28(6):3599.

Schurtenberger, P., Chamberlin, R. A., Thurston, G. M., Thomson, J. A., and Benedek, G. B. (1989). Observation of critical phenomena in a protein-water solution. *Physical review letters*, 63(19):2064.

Sciortino, F., Mossa, S., Zaccarelli, E., and Tartaglia, P. (2004). Equilibrium cluster phases and low-density arrested disordered states: the role of short-range attraction and long-range

repulsion. *Physical review letters*, 93(5):055701.

Sciortino, F., Tartaglia, P., and Zaccarelli, E. (2005). One-dimensional cluster growth and branching gels in colloidal systems with short-range depletion attraction and screened electrostatic repulsion. *The Journal of Physical Chemistry B*, 109(46):21942–21953.

Sedgwick, H., Kroy, K., Salonen, A., Robertson, M., Egelhaaf, S., and Poon, W. (2005). Non-equilibrium behavior of sticky colloidal particles: beads, clusters and gels. *The European Physical Journal E*, 16(1):77–80.

Semlyen, J. A. (2000). *Cyclic polymers*. Springer.

Shivakumar, D., Williams, J., Wu, Y., Damm, W., Shelley, J., and Sherman, W. (2010). Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the opls force field. *Journal of chemical theory and computation*, 6(5):1509–1519.

Shrake, A. and Rupley, J. (1973). Environment and exposure to solvent of protein atoms. lysozyme and insulin. *Journal of molecular biology*, 79(2):351–371.

Shukla, A., Mylonas, E., Di Cola, E., Finet, S., Timmins, P., Narayanan, T., and Svergun, D. I. (2008). Absence of equilibrium cluster phase in concentrated lysozyme solutions. *Proceedings of the National Academy of Sciences*, 105(13):5075–5080.

Shukla, R., Kumar, A., Srivastava, U., Awasthi, N., and Pandey, J. (2012). Critical evaluation of surface tension of binary liquid mixtures from associated and nonassociated processes at various temperatures: an experimental and theoretical study. *Canadian Journal of Physics*, 91(3):211–220.

Simon, R. J., Kania, R. S., Zuckermann, R. N., Huebner, V. D., Jewell, D. A., Banville, S., Ng, S., Wang, L., Rosenberg, S., and Marlowe, C. K. (1992). Peptoids: a modular approach to drug discovery. *Proceedings of the National Academy of Sciences*, 89(20):9367–9371.

Slimani, M. Z., Bacova, P., Bernabei, M., Narros, A., Likos, C. N., and Moreno, A. J. (2014). Cluster glasses of semiflexible ring polymers. *ACS macro letters*, 3(7):611–616.

Souaille, M. and Roux, B. (2001). Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Computer physics communications*, 135(1):40–57.

Steinbrecher, T., Joung, I., and Case, D. A. (2011). Soft-core potentials in thermodynamic integration: Comparing one-and two-step transformations. *Journal of computational chemistry*, 32(15):3253–3263.

Sternhagen, G. L., Gupta, S., Zhang, Y., John, V., Schneider, G. J., and Zhang, D. (2018). Solution self-assemblies of sequence-defined ionic peptoid block copolymers. *Journal of the American Chemical Society*, 140(11):4100–4109.

Stillinger, F. H. and Weber, T. A. (1985). Computer simulation of local order in condensed phases of silicon. *Physical review B*, 31(8):5262.

Stradner, A., Sedgwick, H., Cardinaux, F., Poon, W. C., Egelhaaf, S. U., and Schurtenberger, P. (2004). Equilibrium cluster formation in concentrated protein solutions and colloids. *Nature*, 432(7016):492.

Sugita, Y. and Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. *Chemical physics letters*, 314(1-2):141–151.

Sun, J. and Li, Z. (2018). Peptoid applications in biomedicine and nanotechnology. In *Peptide Applications in Biomedicine, Biotechnology and Bioengineering*, pages 183–213. Elsevier.

Sun, J. and Zuckermann, R. N. (2013). Peptoid polymers: a highly designable bioinspired material. *ACS nano*, 7(6):4715–4732.

Swendsen, R. H. and Wang, J.-S. (1986). Replica monte carlo simulation of spin-glasses. *Physical review letters*, 57(21):2607.

Swope, W. C., Andersen, H. C., Berens, P. H., and Wilson, K. R. (1982). A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of Chemical Physics*, 76(1):637–649.

Torrie, G. M. and Valleau, J. P. (1977). Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics*, 23(2):187–199.

Tozzini, V. (2009). Multiscale modeling of proteins. *Accounts of chemical research*, 43(2):220–230.

Valsson, O., Tiwary, P., and Parrinello, M. (2016). Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint. *Annual review of physical chemistry*, 67:159–184.

Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., et al. (2010). Charmm general force field: A force field for drug-like molecules compatible with the charmm all-atom additive biological force fields. *Journal of computational chemistry*, 31(4):671–690.

Verlet, L. (1967). Computer" experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Physical review*, 159(1):98.

Verlet, L. (1968). Computer" experiments" on classical fluids. ii. equilibrium correlation functions. *Physical Review*, 165(1):201.

Vitkup, D., Ringe, D., Petsko, G. A., and Karplus, M. (2000). Solvent mobility and the protein'glass' transition. *Nature Structural & Molecular Biology*, 7(1):34.

Wang, D., Tong, G., Dong, R., Zhou, Y., Shen, J., and Zhu, X. (2014). Self-assembly of supramolecularly engineered polymers and their biomedical applications. *Chemical Communications*, 50(81):11994–12017.

Wang, J., Wang, W., Kollman, P. A., and Case, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of molecular graphics and modelling*, 25(2):247–260.

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004). Development and testing of a general amber force field. *Journal of computational chemistry*, 25(9):1157–1174.

Weiner, P. K. and Kollman, P. A. (1981). Amber: Assisted model building with energy refinement. a general program for modeling molecules and their interactions. *Journal of Computational Chemistry*, 2(3):287–303.

Weiser, L. J. and Santiso, E. E. (2017). Molecular modeling studies of peptoid polymers. *AIMS MATERIALS SCIENCE*, 4(5):1029–1051.

Weiser, L. J. and Santiso, E. E. (2019). A cgenff-based force field for simulations of peptoids with both cis and trans peptide bonds. *Journal of computational chemistry*.

Wignall, G. D., Littrell, K. C., Heller, W. T., Melnichenko, Y. B., Bailey, K. M., Lynn, G. W., Myles, D. A., Urban, V. S., Buchanan, M. V., Selby, D. L., et al. (2012). The 40 m general purpose small-angle neutron scattering instrument at oak ridge national laboratory. *Journal of Applied Crystallography*, 45(5):990–998.

Wu, B., Kerkeni, B., Egami, T., Do, C., Liu, Y., Wang, Y., Porcar, L., Hong, K., Smith, S. C., Liu, E. L., et al. (2012). Structured water in polyelectrolyte dendrimers: Understanding small angle neutron scattering results through atomistic simulation. *The Journal of chemical physics*, 136(14):144901.

Wu, C. W., Kirshenbaum, K., Sanborn, T. J., Patch, J. A., Huang, K., Dill, K. A., Zuckermann, R. N., and Barron, A. E. (2003). Structural and spectroscopic studies of peptoid oligomers with α-chiral aliphatic side chains. *Journal of the American Chemical Society*, 125(44):13525–13530.

Wu, C. W., Sanborn, T. J., Huang, K., Zuckermann, R. N., and Barron, A. E. (2001a). Peptoid oligomers with α-chiral, aromatic side chains: sequence requirements for the formation of stable peptoid helices. *Journal of the American Chemical Society*, 123(28):6778–6784.

Wu, C. W., Sanborn, T. J., Zuckermann, R. N., and Barron, A. E. (2001b). Peptoid oligomers with α-chiral, aromatic side chains: effects of chain length on secondary structure. *Journal of the American Chemical Society*, 123(13):2958–2963.

Wu, X. and Brooks, B. R. (2003). Self-guided langevin dynamics simulation method. *Chemical Physics Letters*, 381(3-4):512–518.

Yang, Y. I., Shao, Q., Zhang, J., Yang, L., and Gao, Y. Q. (2019). Enhanced sampling in molecular dynamics. *The Journal of chemical physics*, 151(7):070902.

Yu, Y. B. and Wang, W. (1999). Determinant of the inertial tensor and rotational entropy of random polymers. *The Journal of Physical Chemistry B*, 103(36):7676–7680.

Zhang, D., Lahasky, S. H., Guo, L., Lee, C.-U., and Lavan, M. (2012). Polypeptoid materials: current status and future perspectives. *Macromolecules*, 45(15):5833–5841.

Zhou, Y. and Karplus, M. (1999). Interpreting the folding kinetics of helical proteins. *Nature*, 401(6751):400.

Zuckermann, R. N. (1993). The chemical synthesis of peptidomimetic libraries: Current opinion in structural biology 1993, 3: 580–584. *Current Opinion in Structural Biology*, 3(4):580–584.

Zuckermann, R. N. (2011). Peptoid origins. *Peptide Science*, 96(5):545–555.

Zuckermann, R. N., Kerr, J. M., Kent, S. B., and Moos, W. H. (1992). Efficient method for the preparation of peptoids [oligo (n-substituted glycines)] by submonomer solid-phase synthesis. *Journal of the American Chemical Society*, 114(26):10646–10647.

Zuckermann, R. N. and Kodadek, T. (2009). Peptoids as potential therapeutics. *Curr. Opin. Mol. Ther*, 11(3):299–307.

VITA

Pu Du was born in Guangan, China. In 2014, Pu graduated from Nicholls State University, with a bachelor's degree in Chemistry. In the same year, he began his Ph.D. study in chemistry under the supervision of Dr. Revati Kumar at Louisiana State University. In 2017, Pu started to pursue a Master's degree in computer science at Louisiana State University.

Pu likes movies, Sichuan food, college football, and cryptocurrencies.