

# DEVELOPMENT OF COARSE-GRAINED MOLECULAR MODELS OF AQUEOUS POLYACRYLAMIDE



**A Thesis**

*submitted towards partial fulfillment of*

*BS-MS dual degree program*

*in*

**CHEMISTRY**

*by*

**PALLAVI BANERJEE**

**(20111019)**

*Under the guidance of*

**DR. NITISH NAIR**

**SHELL TECHNOLOGY CENTRE, BANGALORE**

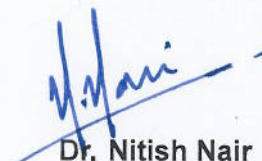
**DEPARTMENT OF CHEMISTRY  
INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE**

---

## CERTIFICATE

This is to certify that this dissertation entitled **Development of Coarse-grained Molecular Models of Aqueous Polyacrylamide** towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents the research carried out by **Pallavi Banerjee**, at Shell Technology Centre, Bangalore under the supervision of Dr. Nitish Nair, Computational Researcher, Computational Centre of Expertise during the academic year 2015-2016.

28th March, 2016



Dr. Nitish Nair

## DECLARATION

I hereby declare that the matter embodied in the report entitled **Development of Coarse-grained Molecular Models of Aqueous Polyacrylamide** are the results of the investigations carried out by me at the Computational Centre of Expertise (CCOE), Shell Technology Centre, Bangalore under the supervision of Dr. Nitish Nair and the same has not been submitted elsewhere for any other degree.

28th March, 2016



**Pallavi Banerjee**

---

# ACKNOWLEDGEMENT

This thesis owes its existence to a bunch of people who deserve special mention for the direct/indirect inputs toward making this work a success. Foremost, I would like to thank my project supervisor, Dr. Nitish Nair (Shell) for the unremitting guidance and encouragement throughout the tenure of the project. The many fruitful discussions with him rooted in me the essence of rational inquiry and logical thinking, not to mention the art behind effective presentation and documentation. I thank him for being enormously patient with me and for never saying no. My heartfelt thanks also goes to Dr. Sudip Roy (Shell) for being a mentor and having helped me every step of the way with his encyclopedia of expertise and knowledge. My gratitude to the entire team of Computational Centre of Expertise (CCOE) at Shell Technology Centre Bangalore for the intellectually stimulating as well as convivial environment of work. I also thank my TAC advisor, Dr. Arnab Mukherjee (IISER Pune) for his insightful suggestions, and also for the invaluable training he has imparted to me in my years at IISER. I would like to thank Reman Kumar to have helped me out of tricky situations in this undertaking, and to the all those who have taken the effort to answer my queries on the GROMACS and VOTCA forums.

I am indebted to Shell for equipping me with immense computational power and resources that led to an accelerated development in the work, and to IISER Pune to allow me to conduct my Master's thesis at Shell.

A big thanks to my friends from IISER interning at Shell for all the fun-filled, almost never-ending discussions over tea and lunch that rendered a more lively atmosphere to the workplace. And lastly, a shout-out to my family for being my emotional pillar throughout this journey.

**Pallavi Banerjee**

---

# ABSTRACT

Viscosity-enhancing polymers such as Polyacrylamide (PAM) are gaining traction as candidates for Enhanced Oil Recovery (EOR). This popularity encourages the computational study of PAM to understand its behavior under various solution conditions. On account of the involvement of large length and time scales in the description of its properties, coarse-graining the atomistic models has become a major subject of interest.

This thesis is focused on developing mesoscale models of PAM in an aqueous environment, from an atomistic base. We have followed three approaches towards achieving a coarse-grained model: MARTINI, Iterative Boltzmann Inversion (IBI), and a hybrid scheme of integrating IBI and MARTINI. The objective was to evaluate the reliability of these approaches in representing the structure and thermodynamics of the target system.

We have reproduced the global structural properties (radius of gyration,  $R_G$ , and end-to-end distance,  $R_{ee}$ ) to a reasonable extent with the CG system developed within the MARTINI framework, although the local structure could not be precisely captured. The results with IBI and IBI+MARTINI of bonded distributions, and radial distribution functions (RDFs) show an absolute replication of the local structure of a single chain of PAM in water. The viscosity results show that the all-IBI method fails to mimic the dynamical property, whereas IBI+MARTINI was successful in mimicking the trend exhibited by the atomistic system. We also show that the derived potentials fail to reproduce the structure beyond 4 wt% of concentration of solution, entailing the need for re-parameterization of potentials for higher concentrations.

# CONTENTS

<b>Acknowledgement</b> . . . . .	iii
<b>Abstract</b> . . . . .	iv
List of Figures . . . . .	vii
List of Tables . . . . .	viii
<b>1. Introduction</b> . . . . .	1
1.1 Computational Modeling . . . . .	2
1.2 Outline . . . . .	3
<b>2. Theory</b> . . . . .	4
2.1 Molecular Dynamics . . . . .	4
2.1.1 Force Field . . . . .	4
2.1.2 The MD Algorithm . . . . .	6
2.2 Coarse Graining . . . . .	8
2.2.1 MARTINI force field . . . . .	8
2.2.2 Iterative Boltzmann Inversion (IBI)[1][2] . . . . .	10
<b>3. Methodology</b> . . . . .	12
3.1 Atomistic PAM melt . . . . .	12
3.2 Atomistic PAM in water . . . . .	13
3.3 MARTINI coarse-graining . . . . .	13
3.3.1 Mapping scheme . . . . .	13
3.3.2 CG potentials . . . . .	13
3.3.3 CG simulation details . . . . .	14
3.4 Coarse-graining by IBI . . . . .	15
3.4.1 Coarse-graining procedure . . . . .	15
3.4.2 CG simulation details . . . . .	16
3.5 Combining IBI and MARTINI . . . . .	17
3.5.1 Coarse-graining procedure . . . . .	17
3.5.2 CG simulation details . . . . .	17
3.6 Viscosity calculations . . . . .	17

---

<b>4. Results and Discussion</b> . . . . .	19
4.1 Validation of atomistic force field . . . . .	19
4.2 MARTINI . . . . .	20
4.3 Iterative Boltzmann Inversion . . . . .	23
4.4 Combining IBI and MARTINI . . . . .	25
4.5 Multi-chains . . . . .	28
4.6 Viscosity enhancement . . . . .	30
<b>5. Conclusions and Future Outlook</b> . . . . .	32
<b>Appendices</b> . . . . .	34
A Glass Transition Temperature . . . . .	35
B Radius of gyration . . . . .	35
C End-to-end distance . . . . .	36
D Radial Distribution Function (RDF) . . . . .	36
E Transverse Current Autocorrelation Function (TCAF) . . . . .	37
<b>Bibliography</b> . . . . .	41

# LIST OF FIGURES

1.1	Hierarchy of simulation approaches[3]	3
2.1	Molecular Dynamics algorithm	5
2.2	Schematic: The leap-frog principle [4]	8
3.1	Mapping scheme	14
3.2	Derivation of bonded parameters from CG runs	15
4.1	Density vs N	19
4.2	Glass Transition Temperature	20
4.3	MARTINI bonded distributions	20
4.4	Time series of $R_g$	21
4.5	Time series of $R_{ee}$	21
4.6	$R_g$ vs N	22
4.7	$\log R_g$ vs $\log N$ (atomistic)	22
4.8	$\log R_g$ vs $\log N$ (CG)	22
4.9	IBI bonded distributions	23
4.10	IBI non-bonded distributions	24
4.11	Comparison of $R_g$ (IBI)	25
4.12	Comparison of $R_{ee}$ (IBI)	25
4.13	IBI+MARTINI bonded distributions	26
4.14	IBI+MARTINI non-bonded distributions	27
4.15	Comparison of $R_g$ (IBI+MARTINI)	27
4.16	Comparison of $R_{ee}$ (IBI+MARTINI)	27
4.17	Multichains: Comparison of $R_g$	29
4.18	Multichains: Comparison of $R_{ee}$	29
4.19	Viscosity vs wave vector	30
4.20	Viscosities with varying models	31
A.1	Schematic: Glass transition temperature	35
B.1	Schematic: Radius of gyration	35
C.1	Schematic: End-to-end distance	36
D.2	Schematic: Radial Distribution Function	37



# LIST OF TABLES

3.1	Bonded parameters for MARTINI coarse-grained simulations . . . . .	14
4.1	Comparison of structural properties ( $R_g$ and $R_{ee}$ ) between atomistic and CG simulations . . . . .	21
4.2	Comparison of structural properties ( $R_g$ and $R_{ee}$ ) between atomistic and CG simulations . . . . .	25
4.3	Comparison of structural properties ( $R_g$ and $R_{ee}$ ) between atomistic and CG simulations . . . . .	27

# 1. INTRODUCTION

In the world of growing energy needs, oil continues to hold a leading position in the energy sector. The world would almost sputter to a standstill without oil. Owing to the increasing demands of technology and power, the future of crude oil production has become a growing concern. This energy crisis is attributed to a dearth of oil resources and difficulties associated with locating fecund oil fields. Traditional techniques of oil recovery allow only about 20-40% of the oil to be extracted from underground reservoirs, leaving a majority of it untapped. Enhanced oil recovery (EOR) techniques provide possibilities to access the stranded amount, adding another 20% to the recovered quantity [5]. Enhanced oil recovery can be divided into three categories:

- Thermal Recovery : Heat is imparted to the reservoir in the form of steam to reduce the viscosity of oil.
- Chemical Injection : Chemicals such as polymers or surfactants mixed with water are introduced into the reservoirs for efficient water flooding.
- Gas Injection : Injection of gases such as nitrogen, carbon dioxide, natural gas, so that they dissolve in the oil and reduce its viscosity, thus enhancing its flow.

Polymer flooding is one of the most employed methods in the area of chemical Enhanced Oil Recovery (cEOR). It involves injecting long-chain, hydrophilic polymers in order to render the water more viscous and improve volumetric sweep efficiency of the water-flood. One such polymer widely used in cEOR is Polyacrylamide (PAM). Polyacrylamide is a water-soluble acrylic resin, produced by free radical polymerization of acrylamide. Aqueous solutions of this long-chain, flexible polymer exhibit enhanced viscous behavior when compared with plain water [6]. On account of its viscoelastic properties, PAM has grabbed the attention of both experimentalists and computational scientists alike. There exists a bountiful store of literature on the experimental studies of PAM, but only a limited amount of computational investigation of the polymer has been conducted [6][7][8].

## 1.1 Computational Modeling

Experimental studies are subject to several limitations, such as impracticality of achieving extreme environmental conditions, a great many trial-and-errors leading up to wastage of resources, limited control over experimental variables. Experiments fail to provide answers to how exactly a process occurs. Computational modeling and simulation studies enable researchers to determine properties that are experimentally inaccessible, and also to predict outcomes via a sweep of parameter space. Computer simulations help build a bridge between theory and experiment, allowing a better insight into the problem at hand.

There exists a hierarchy in the simulation techniques, and active research seeks to combine models of different scales popularly known as multiscale modeling. Fig. 1.1 makes a comparison among the many computational approaches of investigating chemical/biological systems of interest. The finest details of a system can be studied by the first-principles approach of quantum mechanics. This method promises chemical accuracy but is limited to very small systems. Next on the ladder comes classical molecular dynamics (MD) which can be construed as a coarser definition of quantum mechanics. Similarly, coarse-grained MD can replace classical atomistic MD by integrating over certain degrees of freedom in the latter. There are two ways of traversing from one level to another:

- bottom-up : Details at a finer scale are used for parameterization at a coarser scale
- top-down : Properties at a larger scale are the ingredients for designing a more detailed model.

Atomistic studies on polymers suffer from the drawback of long relaxation times for high-molecular-weight chains. Coarse-graining comes to the rescue in such cases. Coarse-grained simulations involve eliminating unnecessary degrees of freedom to probe length and time ranges of systems beyond the reach and span of atomistic models. MARTINI is one such coarse-graining approach that has been designed to represent the partitioning free energy between the differently polar phases of a large library of chemicals [9]. Yet another coarse-graining scheme is Iterative Boltzmann Inversion (IBI) which targets the structural properties of the system [1]. With coarser systems, spanning greater length and time scales becomes feasible, which provides the benefit of reduced computational time and effort. Coarse-graining has its own pitfalls though. Since many degrees of freedom are lost in the process of coarse-graining, not

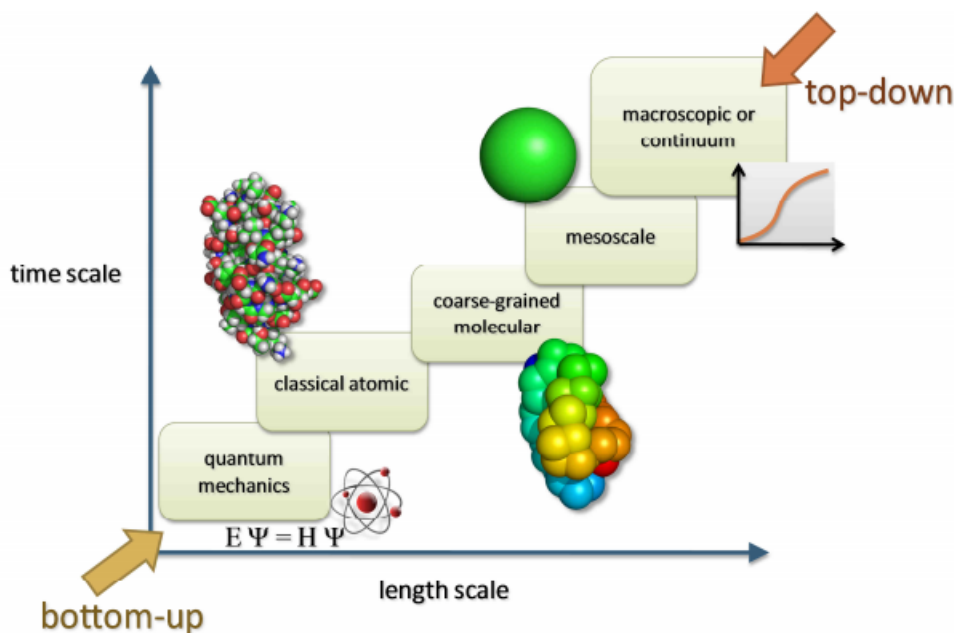


Fig. 1.1: Hierarchy of simulation approaches[3]

many properties of interest can be targeted at the same time. While aiming to achieve computational efficiency and enhanced sampling, the intricate details of the system are compromised upon.

## 1.2 Outline

The main agenda of this thesis is the construction of a methodology that would unify both the structure and thermodynamics of the system of interest. Besides, it is a study of the structural and rheological properties of aqueous polyacrylamide.

Chapter 2 discusses the theory behind the methodology adopted, beginning with the basics of Molecular Dynamics (MD), the basic tool that has been employed throughout the work. It also talks about the approach to building a mesoscale model from an all-atom foundation. In chapter 3, the details of the simulations carried out have been delineated. Chapter 4 presents the results of the investigation along with an elucidation of the results. Chapter 5 lists the conclusions.

## 2. THEORY

This chapter reviews the theory behind the methodology adopted in the work. It begins with a theoretical background on Molecular Dynamics, the simulation technique ubiquitously employed throughout this study. Further on, the procedure of coarse-graining starting from an atomistic base has been discussed. Two types of coarse-graining approaches have been implemented in this work. One of them being the MARTINI [9] approach and the other is Iterative Boltzmann Inversion [1].

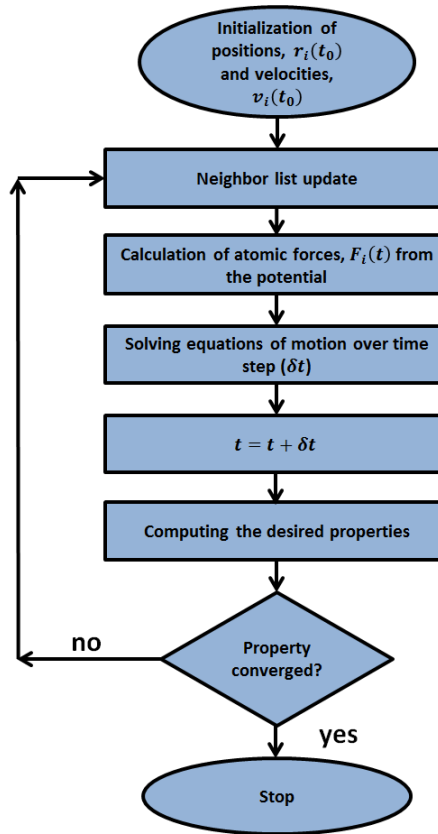
### 2.1 Molecular Dynamics

Particle simulations can be split into two: stochastic and deterministic. Monte-Carlo(MC) works on stochastic algorithms, whereas Molecular Dynamics(MD) follows a deterministic approach. Molecular Dynamics is a technique formulated on the principles of classical and statistical mechanics [10][11][12]. Information is generated at the microscopic level that includes positions and velocities of atoms. The fundamental principle is to apply Newton's equations motion on N interacting particles. Microscopic motion of particles are translated to macroscopic properties by the aid of statistical mechanics. The trajectory of the coordinates is generated by integration of the classical equations of motion. The process of solving these equations is repeated until the system properties do not change with time, in other words, the system reaches its equilibrium. The basic flow of the MD algorithm is shown in Fig. 2.1

#### 2.1.1 Force Field

The communication between the atoms constituting a classical system is governed by the system's Hamiltonian, more popularly known as a force field. A functional form of the degrees of freedom that defines the inter-atomic potential of the system is known as a force-field. The potential function can be divided into two parts:

- Bonded potential: Includes bond vibrations, angle-bending vibrations, proper and improper dihedrals.
- Non-bonded potential: Van der waal's forces are computed by Lennard-Jones or Buckingham potential forms and electrostatic interactions are defined by Coulomb



**Fig. 2.1:** Molecular Dynamics algorithm

potential. The computation of non-bonded interactions is dependent on the neighbor list.

The function is expressed as the sum of the constituent energy terms.

$$E = E_{\text{bond}} + E_{\text{angle}} + E_{\text{tor}} + E_{\text{imp}} + E_{\text{vdw}} + E_{\text{el}} \quad (2.1)$$

The force field package used in this study was GROMOS-53A6 [13]. The following is an overview of the potential forms that make up the system's energy. The summation over all  $N$  bonds results in the total bond energy.

$$E_{\text{bond}}(r) = \sum_{n=1}^{N_b} \frac{1}{4} K_r (r_n^2 - r_0^2)^2 \quad (2.2)$$

where  $K_r$  is the force constant and  $r_0$  the bond length at equilibrium. The angle potential is defined by the following relation:

$$E_{\text{angle}}(\theta) = \sum_{n=1}^{N_\theta} \frac{1}{2} K_{\theta_n} (\cos\theta_n - \cos\theta_0)^2 \quad (2.3)$$

where  $K_\theta$  is the force constant and  $\theta_0$  is the angle at equilibrium. The torsion potential energy is obtained by:

$$E_{\text{tor}} = \sum_{n=1}^{N_\phi} K_{\phi_n} (1 + \cos(\delta_n) \cos(m_n \phi_n)) \quad (2.4)$$

where  $\delta_n$  is the phase shift going from 0 to  $\pi$ ,  $m_n$  is the angle's multiplicity  $\phi_n$ . The potential energy associated with restricting atoms to a particular arrangement, like placing them all in a plane, or a tetrahedral arrangement etc., is the improper dihedral energy which follows the form, summing over all  $N_\xi$  dihedral interaction points:

$$E_{\text{imp}} = \sum_{n=1}^{N_\xi} \frac{1}{2} K_{\xi_n} (\xi_n - \xi_{0n})^2 \quad (2.5)$$

where  $K_\xi$  is the improper dihedral force constant and  $\xi_0$  the equilibrium improper dihedral angle. The non-bonded interactions are computed by the Lennard-Jones 12/6 interaction function.

$$E_{\text{vdw}} = \sum_{i < j} \left( \frac{C12}{r_{ij}^{12}} - \frac{C6_{ij}}{r_{ij}^6} \right) \quad (2.6)$$

The parameters  $C12_{ij}$  and  $C6_{ij}$  are decided by the atom types and the nature of interaction, by the following combination rules:

$$C12_{ij} = \sqrt{C12_{ii} C12_{jj}} \quad C6_{ij} = \sqrt{C6_{ii} C6_{jj}} \quad (2.7)$$

where  $C12_{ii}$ ,  $C6_{ii}$ ,  $C12_{jj}$ ,  $C6_{jj}$  are self-interaction terms,  $C12_{ij}$ ,  $C6_{ij}$  are cross-interaction terms, and  $r_{ij}$  the distance between atoms  $i$  and  $j$ . Electrostatic interactions are defined by the Coulomb potential function acting over all the non-bonded pairs of atoms.

$$E_{\text{el}} = \sum_{i < j} \frac{q_i q_j}{4\pi\epsilon r_{ij}} \quad (2.8)$$

where  $q_i$ ,  $q_j$  are the partial charges on atoms  $i$  and  $j$  respectively,  $\epsilon$  is the dielectric permittivity of the medium,  $r_{ij}$  is the distance between atoms  $i$  and  $j$ .

### 2.1.2 The MD Algorithm

The atoms constituting the system are allotted coordinates and velocities at the start of the run. A neighbor list is generated to compute the non-bonded forces. The forces are only counted for the pairs for which the distance between is less than the provided cut-off radius. Having defined the potential energy, the calculation of force on every particle is performed, which takes the longest computational time in the entire process. The

force on any atom is obtained by differentiating the potential with respect to position.

$$F_i = -\frac{\delta V}{\delta r_i} \quad (2.9)$$

The positions of the atoms are updated by numerically solving Newton's equations of motion.

$$a_i = \frac{d^2 r_i}{dt^2} = \frac{F_i}{m_i} \quad (2.10)$$

The Verlet algorithm is commonly used to generate the trajectory of individual atoms [14]. A Taylor series expansion on the updated position is performed, both forward and backward in time:

$$r(t + dt) = r(t) + v(t)dt + \frac{1}{2}a(t)dt^2 + \frac{1}{6}b(t)dt^3 + \mathcal{O}(dt^4) \quad (2.11)$$

$$r(t - dt) = r(t) - v(t)dt + \frac{1}{2}a(t)dt^2 - \frac{1}{6}b(t)dt^3 + \mathcal{O}(dt^4) \quad (2.12)$$

Adding the two equations:

$$r(t + dt) = 2r(t) - r(t - dt) + a(t)dt^2 + \mathcal{O}(dt^4) \quad (2.13)$$

The acceleration,  $a(t)$ , is force divided by the mass of the concerned particle.

$$a(t) = -\frac{1}{m}dU(r(t)) \quad (2.14)$$

The leap-frog algorithm [15] a variant of the Verlet algorithm. The positions  $r$  at time  $t$  and velocities  $v$  at time  $t - \frac{1}{2}$  are used to update both positions and velocities.

$$r(t + dt) = r(t) + v(t + \frac{dt}{2}) \quad (2.15)$$

$$v(t + \frac{dt}{2}) = v(t - \frac{dt}{2}) + a(t)dt \quad (2.16)$$

The velocities at time  $t$  are:

$$v(t) = \frac{1}{2}[v(t - \frac{1}{2}dt) + v(t + \frac{1}{2}dt)] \quad (2.17)$$



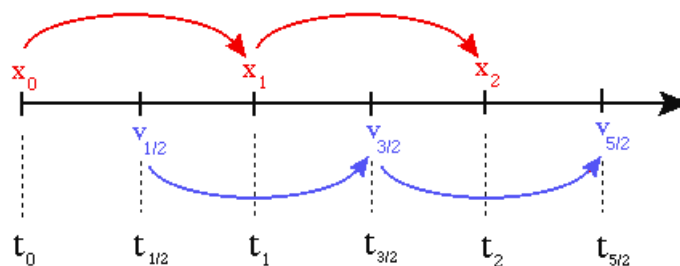


Fig. 2.2: Schematic: The leap-frog principle [4]

## 2.2 Coarse Graining

Atomistic simulations suffer from limitations in length and time scales, owing to which processes that take much longer than nanoseconds cannot be studied, for e.g. chain relaxation of polymers, protein folding, or lipid self-assembly. Coarse-graining is a possible remedy to this problem. It provides the benefit of much fewer degrees of freedom and results in less-detailed, smoother potentials. Groups of atoms are clustered into beads based on the chemical nature of the cluster. This philosophy allows the retention of the specific chemistry of the system while allowing faster computation. As a result of the removal of the finer details of the system, it becomes feasible to use a larger time-step which increases the computational efficiency by orders of magnitude. Several coarse-graining approaches have been devised, each depending on the properties of interest: MARTINI [9], Iterative Boltzmann Inversion [1], force matching [16], relative entropy minimization [17], to name a few. This study deals with MARTINI and IBI. Apart from these two well-known methodologies, we have implemented a novel scheme of integrating the MARTINI and IBI approaches. The following sections provide details of the three approaches of coarse-graining employed in the study.

### 2.2.1 MARTINI force field

MARTINI [9] is a coarse-grained force field designed specifically for MD simulations on biomolecules. The assignment of parameters has been made so as to represent phases of different polarity. At present, MARTINI force-fields for proteins [18], lipids [19], DNA [20], polymers [21], nanoparticles [22] are available.

The construction of the MARTINI model is based on mapping four heavy atom neighbors in the molecule to one coarse-grained bead. A three-to-one mapping scheme has also been introduced to allow higher resolution in cases of ring molecules. These are the S-type beads. Different particle types have been defined depending on the chemical property of the bead, with the aim of mimicking the chemical identity of the all-atom base. Four types of beads have been parameterized: polar(P), non-polar(N), apo-

lar(C), and charged(Q) [23]. Further categorization provides the following sub-types: d = hydrogen donor, a = hydrogen acceptor, da = both, 0 = none, and a polarity index that runs from 1 to 5 in increasing magnitude.

The non-bonded interactions are defined by the Lennard-Jones(LJ) 12-6 potential energy function.

$$U_{LJ}(r) = 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r} \right)^{12} - \left( \frac{\sigma_{ij}}{r} \right)^6 \right] \quad (2.18)$$

where  $\sigma_{ij}$  is the shortest extent of approach between i and j and  $\epsilon_{ij}$  is the well depth of the potential [9]. Every pair of particles has been assigned a  $\sigma_{ij}$  and  $\epsilon_{ij}$  depending on the nature of the interacting particle types ij. The assignment of the LJ parameters to the particles has been done on the basis of experimental evidence, particularly the free energy of hydration, the free energy of vaporization, and the partitioning free energies between phases [23]. The Coulomb function defines the electrostatics of the system [9].

$$U_{el}(r) = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r} \quad (2.19)$$

where the relative dielectric constant,  $\epsilon_r = 15$  to provide explicit screening,  $q_i, q_j$  are the charges on the particles. The electrostatic potential is relevant solely for the systems containing Q type (charged) beads, as the rest of the beads carry no partial charges on them. Bonded parameters are derived from atomistic bonded distributions, as will be described in the next chapter. A weak harmonic potential function  $V_{bond}(r)$  is made to describe bonded interactions of the coarse-grained system with the derived values of the force constant  $K_{bond}$  and equilibrium bond length  $r_{bond}$ .

$$V_{bond}(r) = \frac{1}{2} K_{bond} (r - r_{bond})^2 \quad (2.20)$$

Similarly, angle vibrations are described a weak harmonic potential.

$$V_{angle}(\theta) = \frac{1}{2} K_{angle} [\cos(\theta) - \cos(\theta_0)]^2 \quad (2.21)$$

where  $K_{angle}$  is the derived force constant and  $\theta_0$  is the angle at equilibrium. If needed, proper dihedrals are also incorporated into the model, especially in cases where a secondary structure representation is targeted, like in cases of proteins. The MARTINI philosophy has already met many applications, yet there are shortcomings to it. One of the drawbacks of MARTINI is that it does not mimic the atomistic structure accurately. Iterative Boltzmann Inversion(IBM) wins on this ground. The next section describes the IBM method of coarse-graining in detail.

### 2.2.2 Iterative Boltzmann Inversion (IBI)[1][2]

IBI is a coarse-graining strategy that follows a systematic bottom-up approach to mimic the structure of the underlying atomistic foundation [1]. The mechanism is built upon an assumption that bonded and non-bonded interactions are independent of each other, and hence the total potential energy of the system,  $U$  is the sum of bonded  $U_b$  and non-bonded  $U_{nb}$  potential energies.

$$U = \sum U_b + \sum U_{nb} \quad (2.22)$$

Bonded potentials (bonds, angles, dihedrals) are obtained by a Boltzmann inversion of the bonded distributions of the reference system. This method has been designed to work on a canonical ensemble.  $P(q)$  is the Boltzmann distribution corresponding to the degree of freedom in question[2].

$$P(q) = Z^{-1} \exp[-\beta U(q)] \quad (2.23)$$

where  $Z = \int \exp[-\beta U(q)] dq$  is the canonical partition function,  $\beta = 1/k_B T$  [2].  $P(q)$  is obtained from the trajectory of a well-sampled all-atom foundation. It is further assumed that the bonded interactions are not correlated, which allows the factorization of the probability of states.

$$P(r, \theta, \phi) = P_r(r) P_\theta(\theta) P_\phi(\phi) \quad (2.24)$$

Histograms of the distributions should be rescaled to obtain a volume-normalized distribution function.

$$P_r(r) = \frac{H_r(r)}{4\pi r^2}, \quad P_\theta(\theta) = \frac{H_\theta(\theta)}{\sin\theta}, \quad P_\phi(\phi) = H_\phi(\phi) \quad (2.25)$$

A Boltzmann inversion on the distribution functions gives us the respective potentials, shown in (2.27).

$$U(r, \theta, \phi) = U_r(r) + U_\theta(\theta) + U_\phi(\phi) \quad (2.26)$$

$$U_q(q) = -k_B T \ln P_q(q) \quad (2.27)$$

This inverted potential is, in actuality, the free energy, which is approximated as the potential energy. In most cases, the inverted potential suffices to be the bonded potential energy function for the coarse-grained runs in attribution to the stiffness of the bonded interactions. However, this does not hold true for the non-bonded interactions.

Non-bonded potential functions are derived from an iterative refinement of the energy function obtained from a Boltzmann-inverted radial distribution function,  $g(r)$ , from the

atomistic reference. The directly inverted potential serves as a reasonable first guess for the iterative procedure. Such a refinement is performed for every non-bonded interaction of the system. The CG potential is optimized by the following relation:

$$U^{n+1} = U^n + k_B T \ln \left( \frac{g_i^{\text{CG}}(r)}{g^{\text{target}}(r)} \right) \quad (2.28)$$

The potentials are said to be converged at the step where  $g_i(r)$  matches  $g^{\text{target}}(r)$ . The extent of convergence is defined by the following error function:

$$f_{\text{merit}} = \int w(r) (g_i^{\text{CG}}(r) - g^{\text{target}}(r))^2 dr \quad (2.29)$$

where  $w(r) = \exp(-r)$ , a weighting function to avoid strong deviations at small distances [24]. The same iterative refining has to be operated on bonded interactions too if the direct Boltzmann inverted potential does not serve as the suitable potential energy function. For IBI on bonded interactions, the  $g(r)$  in equations (2.28) and (2.29) gets replaced by  $P(r)$ .

## 3. METHODOLOGY

Here we present a layout of the methodology adopted in the study. The flow of the work is as follows: (a) atomistic simulations of polyacrylamide melt, (b) atomistic simulations of aqueous polyacrylamide which acts as the reference upon which the coarse-grained model is constructed, (c) coarse-graining by MARTINI, (d) IBI, and (e) integration of IBI and MARTINI. All the MD simulations have been performed using the GROMACS 5.0.5 [25] suite of programs.

### 3.1 Atomistic PAM melt

The polymer chain was constructed in an atactic manner using the polymer builder associated with the software CULGI. Copies of the polymer were packed into a simulation box to create a homogeneous melt system. The purpose of performing melt simulations of the polymer is to obtain the chain length that would adequately represent the characteristics of PAM used in experiments. A united-atom force field, GROMOS-53A6 [13], was used to design the atomistic model with charges borrowed from the paper by Wang, *et al* [26]. The integrator used was leap-frog [15] with 2 fs as the time step. The steepest descent algorithm was used to energy-minimize the system [27]. The simulations were set up at temperature, 300 K, and pressure, 1 bar, with coupling constants of 0.2 ps and 1.5 ps respectively, using the Berendsen weak-coupling scheme [28]. Reaction-field method [29] defined the electrostatics with a cut-off of 1.2 nm and dielectric constant of 3.5. Lennard-Jones interactions were truncated at 1.2 nm, applying dispersion corrections for potential and pressure. With a cut-off radius of 1.2 nm, the neighbor list was updated at every 10th step. The Linear Constraint Solver (LINCS) algorithm [30] applied bond constraints.

Single chains were relaxed in vacuum before packing them into a simulation box. NPT equilibration of PAM chains with 5-80 monomers were carried out individually at  $T = 300\text{K}$ . The density of the melt in each case was noted. The minimum chain length that represented the experimental density of PAM was picked for further tests. To confirm the validity of the force-field and the chosen chain length, the glass transition temperature,  $T_g$ , of the polymeric melt was obtained (refer Sec. A of the Appendices for a detailed description). This process required four cycles of simulated annealing until

the system's potential energy at 300 K reached a minimum. The system was heated from 300 K to 800 K, cooled down to 250 K with leaps of 50 K, and the potential energy was extracted in each cycle. On reaching the minimum potential energy, extended runs of 10 ns were carried out at every temperature range and the corresponding average density was gathered at each temperature. The resulting density-temperature profile gives us the  $T_g$ .

## 3.2 Atomistic PAM in water

The force-field for the polymer used was the same as that of the melt, while SPC-E [31] was used for water. A single relaxed chain of the polymer in a cubic box of edge 5.63 nm, with 5698 SPC-E [31] water molecules comprised the system. Steepest descent [27] was performed to energy-minimize the structure, followed by an NVT equilibration for 1 ns to relax the solvent around the polymer. An NPT equilibration for 5 ns and further an NPT production run for 300 ns were carried out. The simulation details were same as that of the melt except for the inclusion of Particle Mesh Ewald (PME) electrostatics [32] in the aqueous case with a real space cut-off of 1.2 nm and PME order of 4. The temperature was maintained at 300 K with the Nose-Hoover [33][34] thermostat and pressure at 1 bar with the Parrinello-Rahman [35] barostat with coupling constants of 0.4 ps and 1.5 ps respectively.

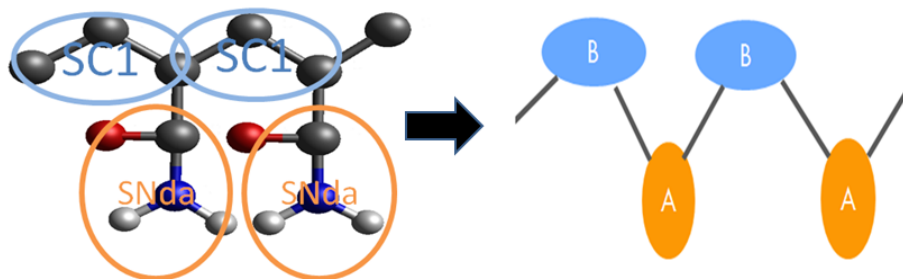
## 3.3 MARTINI coarse-graining

### 3.3.1 Mapping scheme

The atomistic polymer chain was mapped to a chain consisting of two types of beads, A and B. The mapping scheme has been derived from Wang, et al [26]. Bead A represents the pendant amide and bead B represents the backbone carbon chain, both following a mapping of 3 heavy atoms to 1 bead. The mapping scheme is shown in Fig.3.1. The center of mass of the cluster of atoms comprising a bead is its coordinate. MARTINI water is incorporated into the coarse-grained system, wherein each water bead represents a clump of 4 molecules of atomistic water of bead type P4.

### 3.3.2 CG potentials

The bonded potentials of the CG system were built to match the target atomistic potentials. These include bond AB, angle ABA, and angle BAB. The reference bonded distributions were obtained from mapping the atomistic trajectory to a coarse-grained



**Fig. 3.1:** Mapping scheme with MARTINI bead definitions (refer Sec.2.2.1 for the definition of bead types)

one with the aforementioned mapping definition. A Boltzmann inversion, as shown in eqn. (3.1) was performed over the target distributions to obtain the potentials over which harmonic fits were made, which then serve as input potentials for the CG simulations.

$$U(x) = -k_B T \ln P(x) \quad (3.1)$$

where  $U(x)$  is the inverted potential,  $T$  is the temperature, and  $P(x)$  the probability of atomistic bonded distribution. Harmonic fitting is performed with the following equation:

$$U(r) = \frac{1}{2} k (x - x_0)^2 \quad (3.2)$$

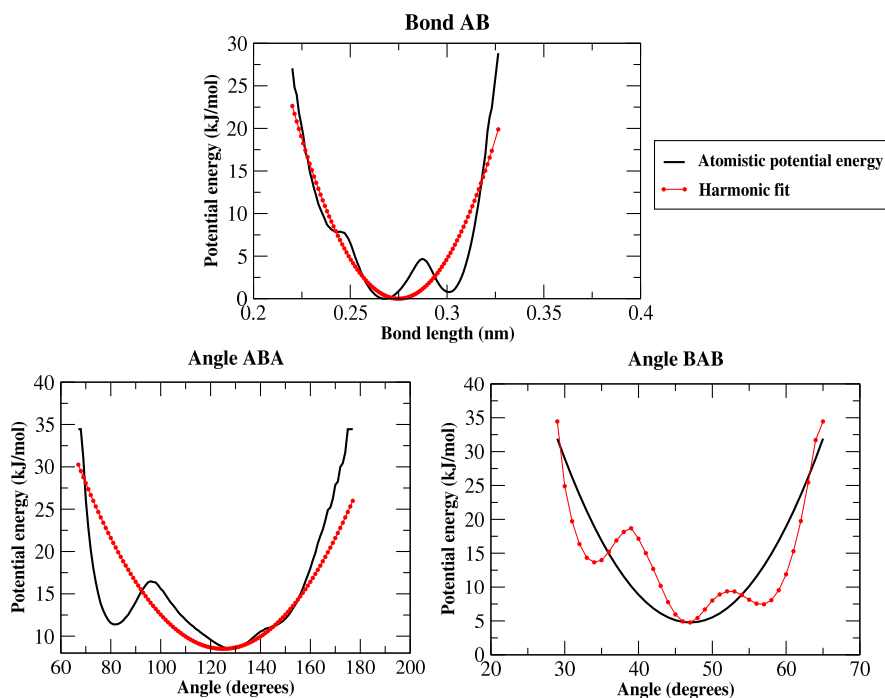
where  $x$  is either a bond or angle coordinate. Fig. 3.2 shows individual bonded potential energies obtained from the atomistic run and the harmonic fit performed resulting in the input parameters for the CG runs as shown in Tab. 3.1.

	Equilibrium value	Force constant
Bond AB	0.275nm	15003.8kJmol <sup>-1</sup> nm <sup>-2</sup>
Angle ABA	125°	42.45kJmol <sup>-1</sup> rad <sup>-2</sup>
Angle BAB	47°	550kJmol <sup>-1</sup> rad <sup>-2</sup>

**Tab. 3.1:** Bonded parameters for MARTINI coarse-grained simulations

### 3.3.3 CG simulation details

The three derived, bonded potentials were incorporated into the CG force-field. The LJ parameters of the existing MARTINI bead-types, namely,  $\sigma$  and  $\epsilon$ , imparted form to the non-bonded interactions. NPT simulations of the polymer of the same chain length as the atomistic chain were run in MARTINI water. Electrostatic interactions were computed by the Reaction-field zero algorithm [29] assisted by a Coulomb modifier. LJ potential was gradually turned off to reach zero from 0.9 nm to 1.2 nm and electrostatic



**Fig. 3.2:** Derivation of bonded parameters from CG runs

interactions from 0.0 nm to 1.2 nm. At every 10th step, the neighbors were updated. The integrator used was leap-frog [15] with a time-step of 20 fs. For the initial NPT equilibration of 5 ns, the thermostat used was V-rescale [36] with temperature at 300 K and the barostat was Berendsen [28] with pressure at 1 bar with coupling constants of 1 ps and 4 ps respectively. NPT production run was done for 300 ns. Nose-Hoover [33][34] thermostat and Parrinello-Rahman [35] barostat with coupling constants of 4 ps and 12 ps maintained constant temperature and pressure.

## 3.4 Coarse-graining by IBI

### 3.4.1 Coarse-graining procedure

The mapping scheme followed for the polymer chain was the same as described in Sec. 3.3.1. Each water SPC-E water molecule is mapped to one superatom of water where the superatom's center is the center of mass of the three atoms constituting a water molecule. The coarse-graining kit VOTCA [37] has been employed in this work to perform the iterative procedure of coarse-graining. The potentials for the bonded interactions were the Boltzmann-inverse of the distributions obtained from the reference trajectory after the mapping [38]. In the mapped polymer there exist 4 bonded types: bond AB, angle ABA, angle BAB, and a dihedral angle BABA. The dihedral angle has been included to represent the structure more precisely in comparison with the con-



structured MARTINI force-field. The non-bonded interactions cover 6 types: A-A, B-B, A-B, A-SOL, B-SOL, SOL-SOL, where SOL is the water bead where one molecule of water was mapped to one bead. The Boltzmann-inverse of the target RDFs obtained from the reference trajectory after mapping serve as the guess potentials to begin the iterative procedure. The non-bonded potentials exclude 1-2, 1-3, and 1-4 interactions, owing to the inclusion of bond, angle and dihedral potential inputs.

Since the non-bonded interactions are highly correlated, they were optimized by updating them in a sequential manner. The imposed  $f_{\text{merit}}$  (refer eqn. (2.29)) value to quantify the convergence of potentials was 0.5. Before being fed as input potentials, the Boltzmann-inverted numerical potentials were smoothed by cubic splines. Each step of the iterative procedure was run for 10 ns to allow sufficient sampling. The IBI method focuses on reproducing the structure of the system, and hence, loses on the required thermodynamics. Since the potentials were optimized based on structure alone in the absence of a barostat, the pressure of the CG system deviated from the pressure of the atomistic reference. To combat this issue a linear pressure correction [39] was applied to the long-range section of each of the optimized CG non-bonded potentials. The correction term is [38]:

$$dU(r) = A \left( 1 - \frac{r}{r_c} \right) \quad (3.3)$$

where  $A = -(dP)0.1k_B T$ ,  $r_c$  is the cutoff for non-bonded potentials,  $dP$  is the difference in pressure from the atomistic value. In our procedure, pressure correction was applied to all the interactions at the same time after every complete cycle of potential update. This process was made to continue until both structure and pressure reached satisfactory convergence.

### 3.4.2 CG simulation details

The CG-MD simulations were run at constant NVT (300K) with GROMACS-5.1 [25]. The system consists of a cubic box with PBC invoked. The leap-frog was made to describe the Langevin dynamics, with a time step of 4 fs. The Langevin thermostat maintains the temperature at 300 K with a friction constant,  $\gamma = 0.5\tau^{-1}$ , where  $\tau = 0.2$  ps is the time constant. The cutoff for both the neighbor listing and van der Waals interactions was 1.2 nm with the neighbors being renewed every 10 time-steps. All potentials were fed in the tabulated form to perform extensive CG-MD simulations.

## 3.5 Combining IBI and MARTINI

### 3.5.1 Coarse-graining procedure

The mapping scheme for the polymer was the same as in Sec. 3.3.1. A single chain of the mapped polymer (PAM30) was solvated in 1424 MARTINI water beads, maintaining the same concentration as that of the reference system. To incorporate the effect of MARTINI water, the interactions: W-W, A-W and B-W were accounted for by the non-bonded potentials derived from the all-MARTINI system (refer Sec. 3.3). A 'W' bead is one MARTINI water bead. These potentials are obtained by converting the analytical potentials corresponding to each of the three non-bonded pairs to a numerical (tabulated) form. Since the bonded potentials obtained from a direct Boltzmann inversion failed to result in similar bonded distributions as the atomistic, the iterative refining of potentials was applied on both bonded, and non-bonded, non-water interactions: A-A, B-B, and A-B. Unlike the general optimization procedure of potentials in order of relative strengths:  $U_{\text{bond}} \rightarrow U_{\text{angle}} \rightarrow U_{\text{nonbonded}} \rightarrow U_{\text{dihedral}}$ , this work follows the order of updating only one interaction potential at a step in a cycle that goes over all of the four bonded and three non-bonded interactions.

### 3.5.2 CG simulation details

CG MD simulations were run with a time step of 4 fs in the NVT ensemble. Each iteration was 10 ns long. The integrator describing Langevin dynamics was leap-frog. The Langevin thermostat maintains the temperature at 300 K with a friction constant,  $\gamma = 0.5\tau^{-1}$ , where  $\tau = 0.2$  ps is the time constant. The cutoff for the neighbor list update was 1.2 nm. Following the methodology adopted for MARTINI (refer Sec. 3.3.3), non-bonded potentials were made to smoothly reach zero at the cutoff value, 1.2 nm, by applying the force-switch algorithm [40] between 0.9 nm and 1.2 nm. The implementation of the force-switch algorithm does the pressure-correction in the system and thereby, no additional pressure-correction needed to be done on every potential being updated in the IBI procedure.

## 3.6 Viscosity calculations

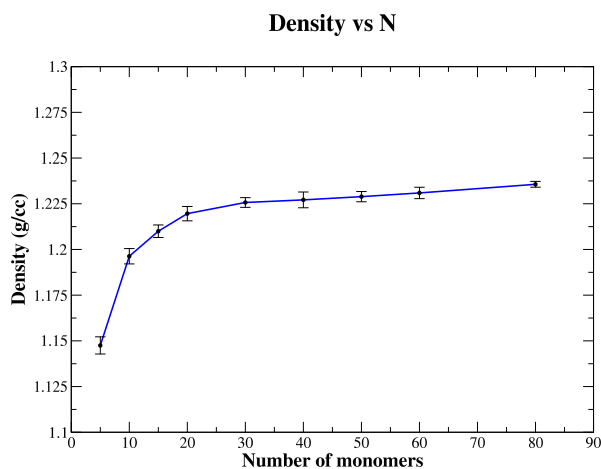
In our simulations, the rheological behavior of systems has been studied by calculating the bulk viscosity of aqueous polymer solutions with the Transverse Current Autocorrelation Function method (TCAF). (refer Sec. E of the Appendices for the background). We conducted 3 trials for each model: (a) all-atom (b) MARTINI (c) IBI (d) IBI+MARTINI. The polymer concentration in water varied from 0 wt% (0 chains) to 28.05 wt% (15

chains). The initial NPT equilibration lasted for 2 ns, followed by 5 ns of NPT production run at 300 K and 1 bar pressure. The trajectory was recorded every 10 fs for each system. An exponential fit is made over the decaying TCAF with the decay constant transformed into the shear viscosity of the system. [41].

## 4. RESULTS AND DISCUSSION

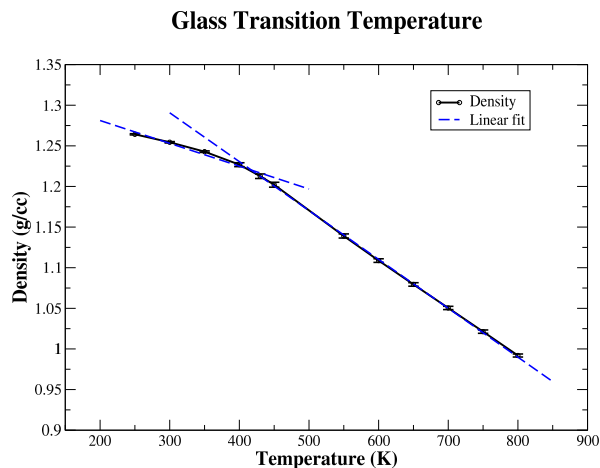
### 4.1 Validation of atomistic force field

The force-field chosen to carry out MD simulations for both polyacrylamide melt and aqueous polyacrylamide was a united-atom force-field, GROMOS-53A6 [13]. As described in Sec. 3.1, the density of the melt as a function of the chain length was noted. Fig. 4.1 shows that the density begins to saturate after  $N = 30$  with the value,  $1.226 \pm 0.003$  g/cc lying in the range of experimental densities reported for PAM: 1.189 g/cc [42], 1.21 g/cc [43], 1.302 g/cc [44]. We singled out  $N=30$  as the shortest polymer that reproduced the experimental density. Additionally, longer chains will require more intense computational effort. With PAM30 as the system, further validation of



**Fig. 4.1:** Density of polymeric melt vs Number of monomers,  $N$ , to obtain the chain length that mimics PAM from experiments

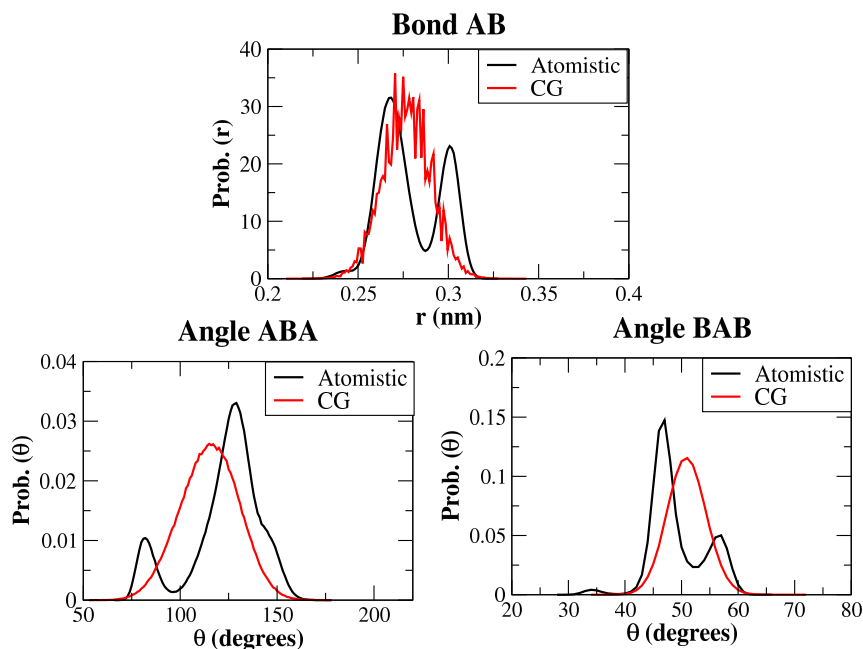
the force-field and the chain length was made by calculating the  $T_g$  of PAM30 melt as in Sec. 3.1. Once the minimum potential energy was reached, the densities were plotted against the corresponding temperatures. The point at which the slope of the density profile changes abruptly is where the polymeric melt transitions from a glassy state (more dense) to loose, rubbery state (less dense). As can be seen in Fig. 4.2, the point of intersection of the two states of the polymer occurs at temperature ( $T_g$ ) = 415.49 K. The  $T_g$  obtained from simulations is quite close (error  $\approx 5.14\%$ ) to the most cited experimental  $T_g$ , 438 K [45].



**Fig. 4.2:** Glass transition temperature obtained from simulated annealing on a polymeric melt of PAM30

## 4.2 MARTINI

We study the structural aspects of our MARTINI-based coarse-grained system of a single chain of PAM30 in water. The bonded distributions resulting from feeding harmonic fit parameters derived in Sec. 3.3 have been shown in Fig. 4.3. The CG

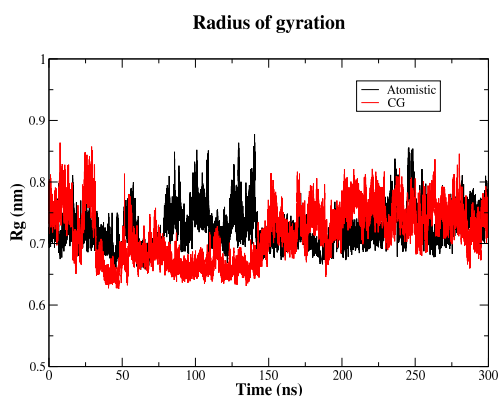


**Fig. 4.3:** Comparison of the bonded distributions between all-atom and CG systems

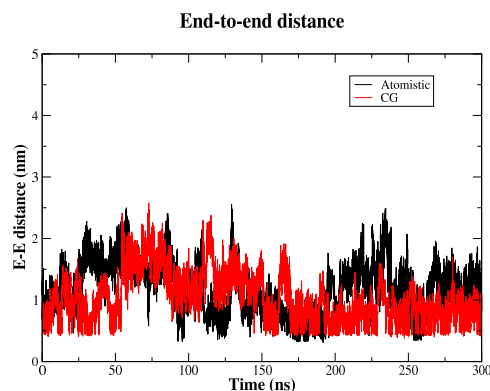
bonded distributions overlap with the underlying atomistic distributions. Although the distributions do not peak at the same value for both, the histograms span the same

range of values.

The structural properties of interest are the radius of gyration,  $R_g$ , and the end-to-end distance,  $R_{ee}$ , of the polymer chain. These properties have been described in Sec. B and Sec. C of Appendices. Figs. 4.4 and 4.5 compare the structural properties between the all-atom and coarse-grained systems. Table 4.1 places average values



**Fig. 4.4:** Time series of radius of gyration ( $R_g$ ) of a single chain of PAM30 in water



**Fig. 4.5:** Time series of end-to-end ( $R_{ee}$ ) of a single chain of PAM30 in water

of  $R_g$  and  $R_{ee}$  of the constructed MARTINI system against those of the all-atom model.

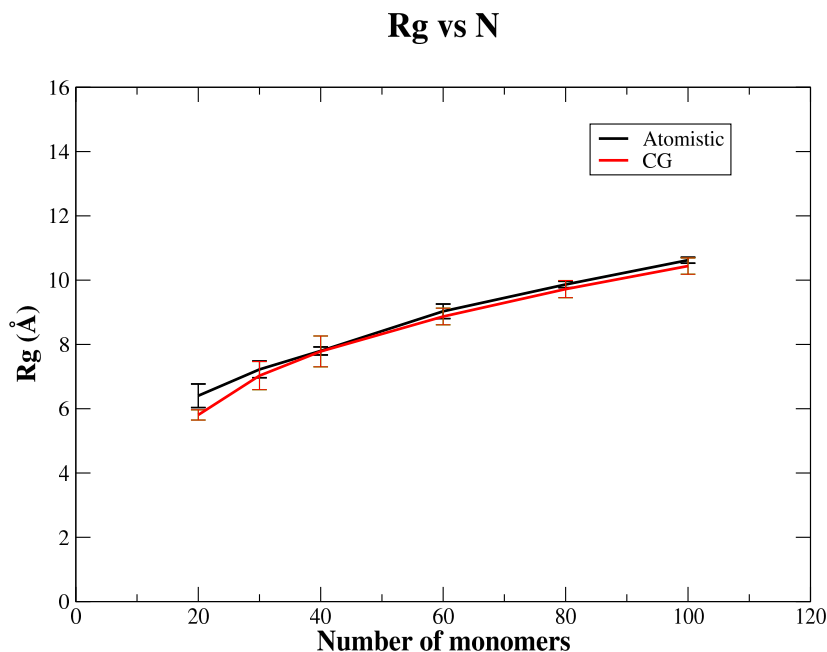
	$R_g$ (nm)	$R_{ee}$ (nm)
Atomistic	$0.721 \pm 0.028$	$1.180 \pm 0.421$
Coarse-grained	$0.715 \pm 0.045$	$1.029 \pm 0.399$

**Tab. 4.1:** Comparison of structural properties ( $R_g$  and  $R_{ee}$ ) between atomistic and CG simulations

In order to test how applicable our CG force-field was for chain lengths other than PAM30, we computed the radius of gyration for PAM chains in aqueous systems with  $N=20-100$  monomers. Fig.4.6 shows the scaling behavior of  $R_g$  vs  $N$  for atomistic and CG systems. The scaling relation is shown in eqn. (4.1), and (4.2).

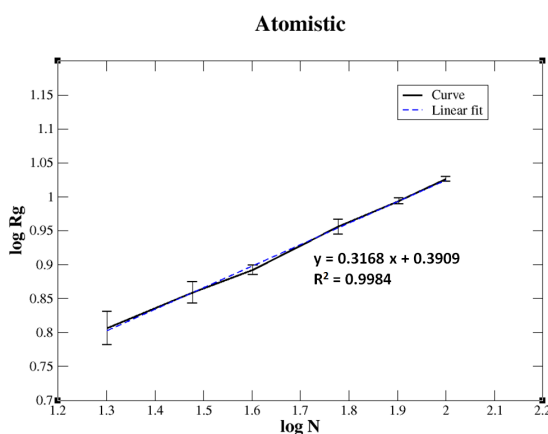
$$R_g \propto N^\lambda \quad (4.1)$$

$$\log R_g = \lambda \log N + C \quad (4.2)$$

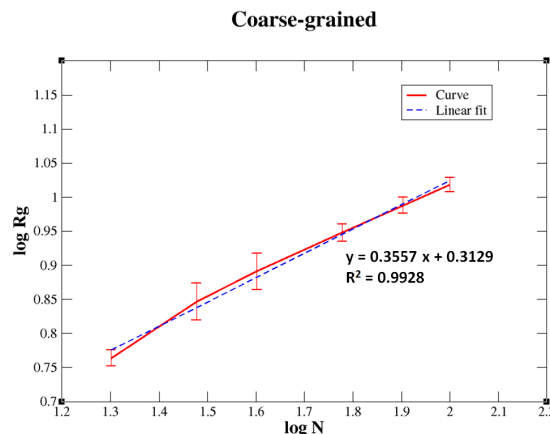


**Fig. 4.6:** Radius of gyration as a function of chain length for atomistic and CG systems

Log-log plots in Fig.4.7 and 4.8 show similar scaling trends for the atomistic and CG representations.



**Fig. 4.7:** Relation of  $\log R_g$  with  $\log N$  for atomistic system

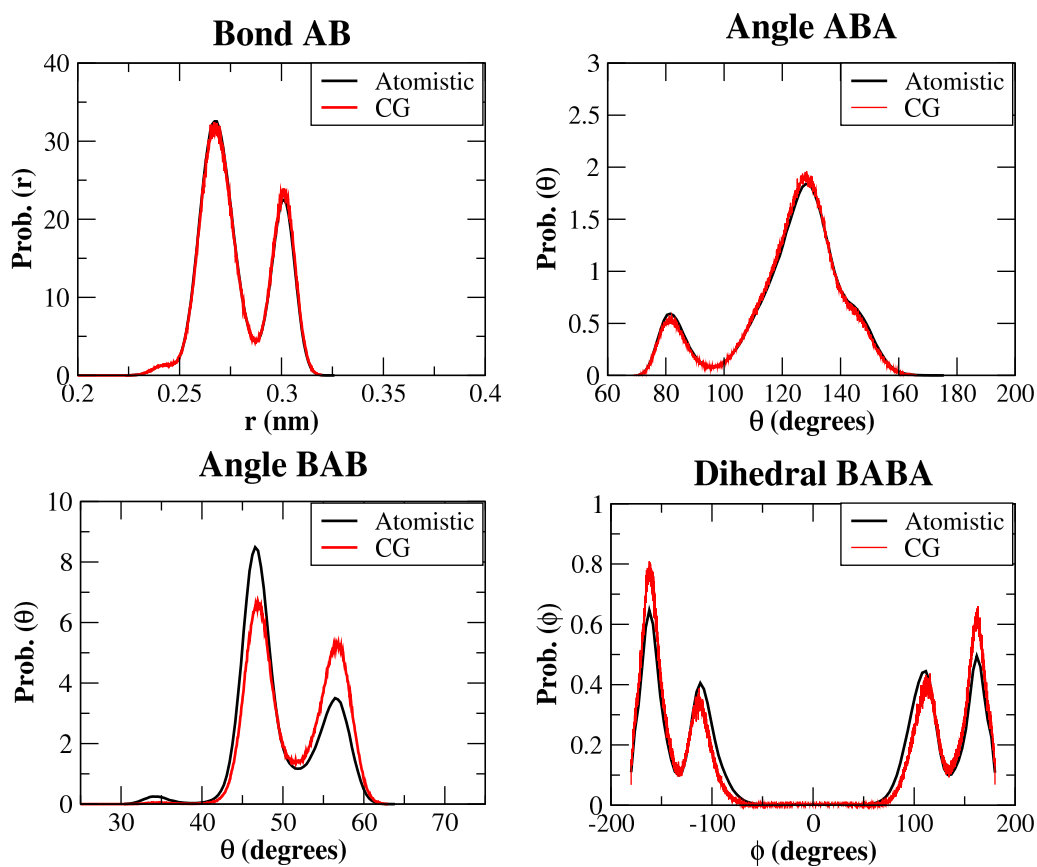


**Fig. 4.8:** Relation of  $\log R_g$  with  $\log N$  for coarse-grained system

The plots of Figs. 4.6, 4.7, 4.8 show that the exponent values,  $\lambda$ , from both atomistic, 0.3168, and coarse-grained simulations, 0.3557, are very close. This suggests that the CG system matches the atomistic reference quite well.

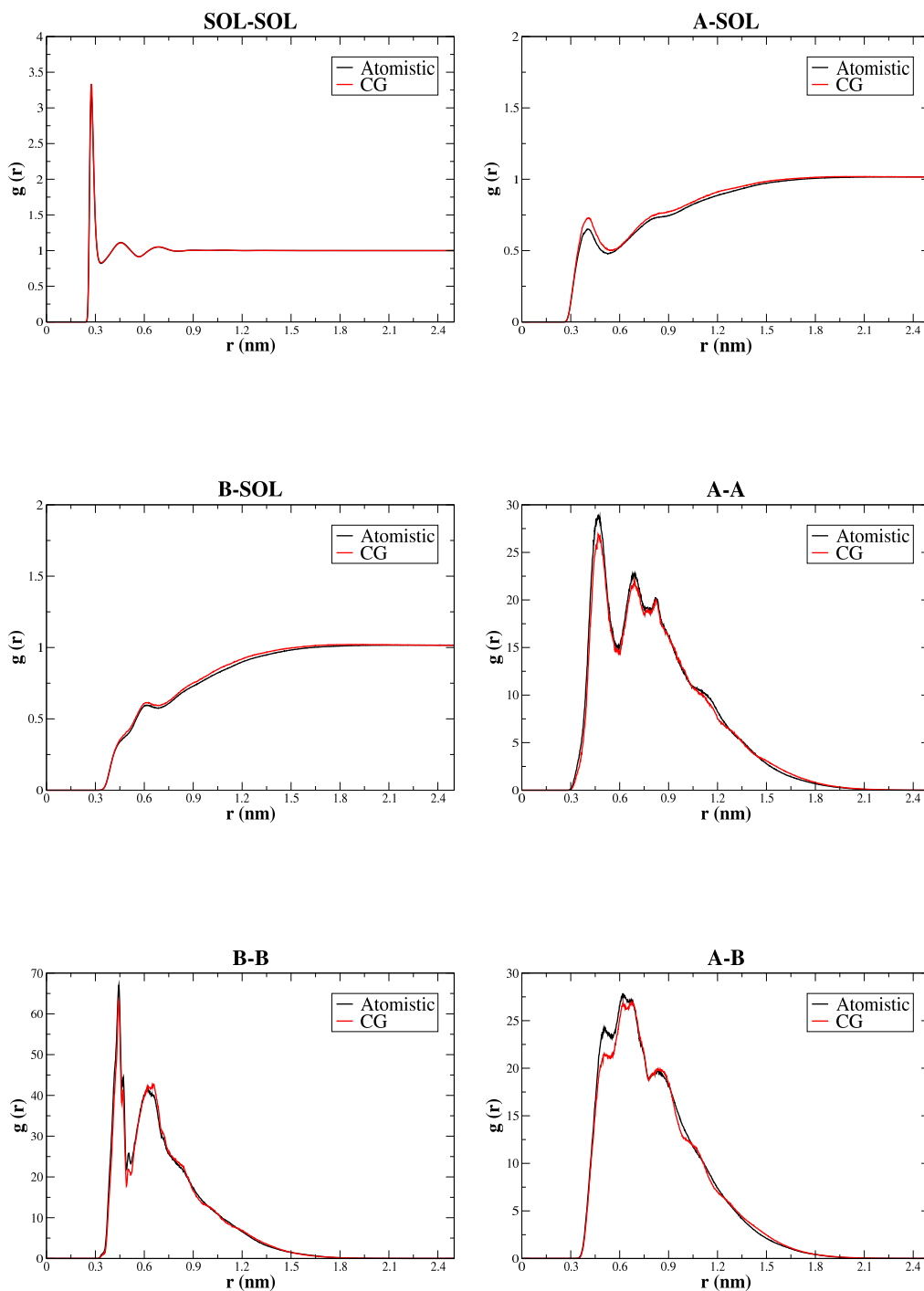
### 4.3 Iterative Boltzmann Inversion

The results presented in this section correspond to the procedure detailed in Sec. 3.4. Convergence was achieved in 74 steps of iteration. Fig. 4.9 shows the comparison of the bonded distributions and Fig. 4.10 of the non-bonded distributions between all-atom and CG models.



**Fig. 4.9:** Comparison of bonded distributions between atomistic and CG models



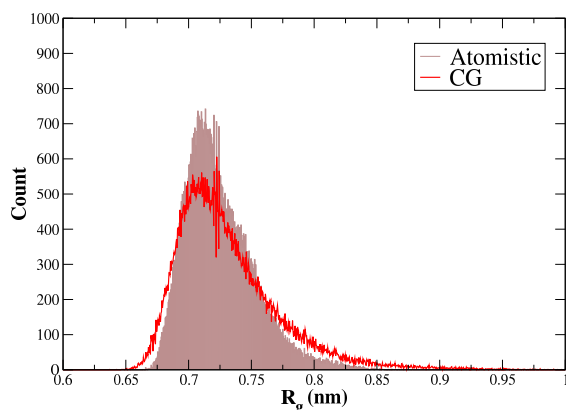


**Fig. 4.10:** Reproduction of the atomistic non-bonded interactions by the CG system.

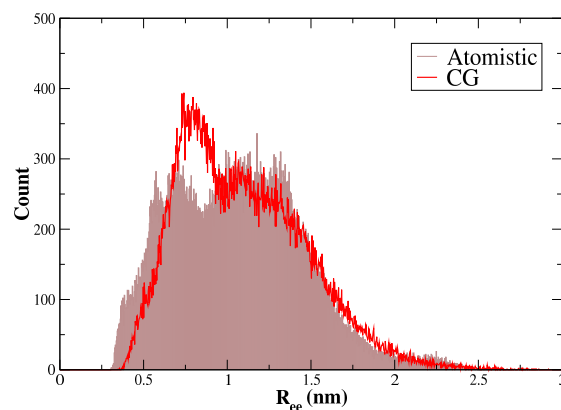
The local structural arrangement of the polymer system is best described by the bonded probability distributions and radial distribution functions (RDFs) (refer Sec. D for the description) between the centers of the pseudo-atoms (CG beads). The RDFs have been computed after excluding the three nearest neighbors. The above results clearly indicate that the intra-molecular and intermolecular interactions have been successfully

captured by our CG model. The pressure correction implemented according to eqn. (3.3) lowered the pressure from 2422.62 bar to 1.59 bar, a value close to the pressure of the all-atom system, i.e., 1.21 bar. A consistent pressure value makes sure that the thermodynamics of the system is not disturbed.

Static structural properties such as the  $R_g$  and  $R_{ee}$  describe the global structure of polymers. The following plots compare the structural properties of the all-atom and CG models.



**Fig. 4.11:** Histograms of radius of gyration ( $R_g$ ) of CG against the all-atom system



**Fig. 4.12:** Histograms of end-to-end distance ( $R_{ee}$ ) of CG against the all-atom system

	$R_g$ (nm)	$R_{ee}$ (nm)
Atomistic	$0.724 \pm 0.028$	$1.061 \pm 0.390$
Coarse-grained	$0.731 \pm 0.041$	$1.093 \pm 0.377$

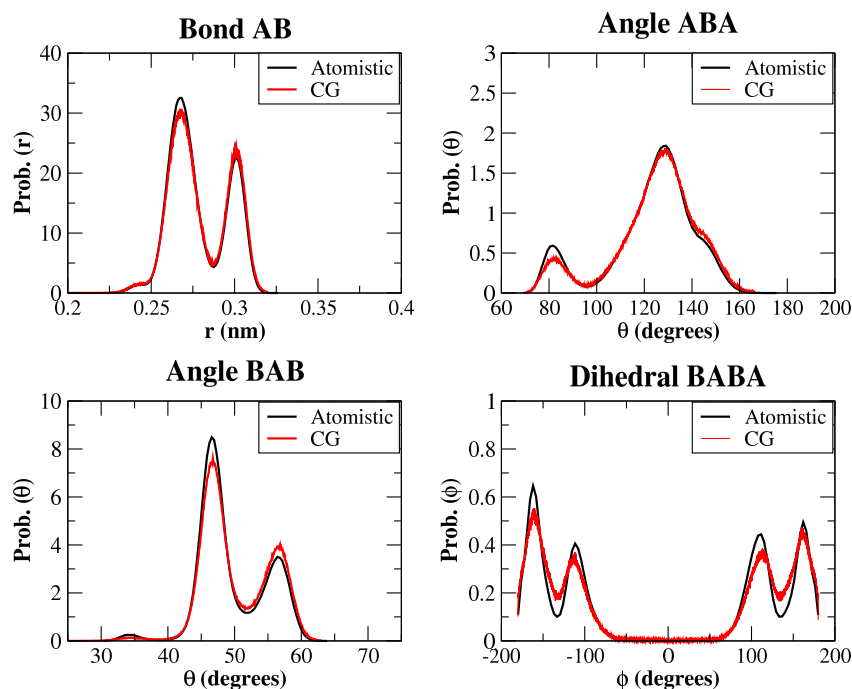
**Tab. 4.2:** Comparison of structural properties ( $R_g$  and  $R_{ee}$ ) between atomistic and CG simulations

The above histograms obtained from the CG model do not exactly match the atomistic in shape because of insufficient sampling. Nevertheless, the average values of the structural properties of the CG lie close to the all-atom system. Also, the peaks in the distributions of the systems lie around the same value. This suggests that our CG model is an adequate representation of the structure of the atomistic model.

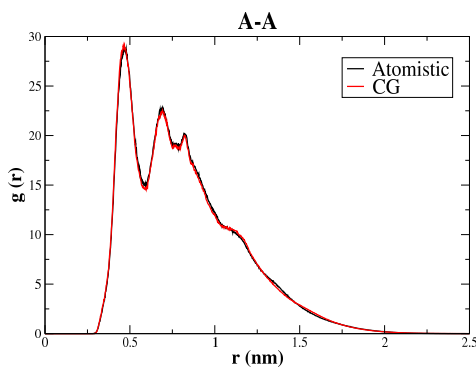
## 4.4 Combining IBI and MARTINI

This section reports the results of the coarse-graining methodology followed in Sec. 3.5. A good level of convergence was achieved in 60 steps of the iterative procedure

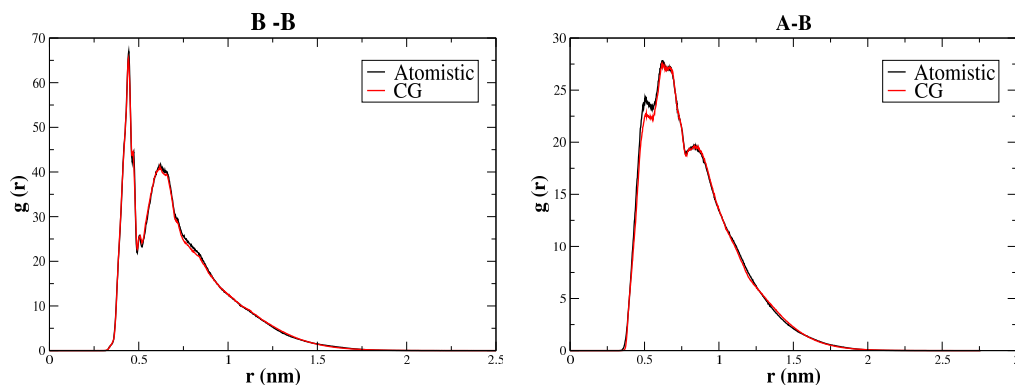
that sequentially iterated over the bonded and non-bonded interactions. A comparison of the bonded distributions is shown in Fig. 4.13 and of non-bonded in Fig. 4.14.



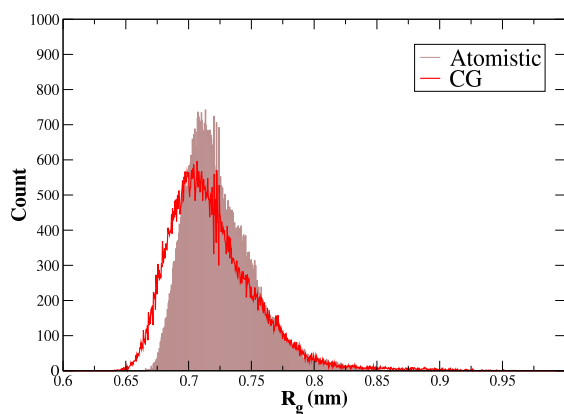
**Fig. 4.13:** Comparison of bonded distributions between atomistic and CG models



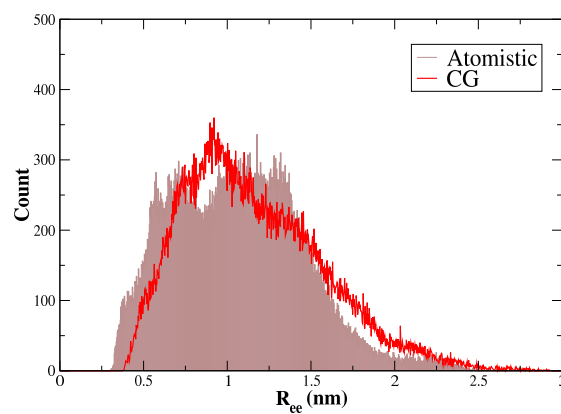
The above plots indicate that the local structure of the CG model perfectly matches the all-atom system at short and long distances. The nearest three neighbors have been excluded while computing the RDFs. Hence, both the intra- and inter-molecular arrangements have been replicated by our coarser model. To study the global structure of the CG polymer, we looked at the  $R_g$  and  $R_{ee}$  of the systems in question. The results are shown in Figs. 4.15 and 4.16.



**Fig. 4.14:** Reproduction of the atomistic RDFs by the CG model.



**Fig. 4.15:** Histograms of radius of gyration ( $R_g$ ) of CG against all-atom system



**Fig. 4.16:** Histograms of end-to-end distance ( $R_{ee}$ ) of CG against all-atom system

	$R_g$ (nm)	$R_{ee}$ (nm)
Atomistic	$0.724 \pm 0.028$	$1.061 \pm 0.390$
Coarse-grained	$0.719 \pm 0.036$	$1.156 \pm 0.417$

**Tab. 4.3:** Comparison of structural properties ( $R_g$  and  $R_{ee}$ ) between atomistic and CG simulations

The CG bonded distributions of this hybrid model are closer to the atomistic than those obtained by the model derived from IBI alone (refer Sec. 4.3). This is because the hybrid model optimizes potentials over bonded and non-bonded interactions whereas the other model only iterates over the non-bonded interactions. The CG distributions of  $R_g$  and  $R_{ee}$  show good overlap with the atomistic distributions. There is only a tiny shift in the  $R_g$  of the CG model to the left. This is because the distance between the center of mass of a peripheral bead and the chain's center is always less than the distance between the center of mass of a peripheral atom and chain's center. This difference

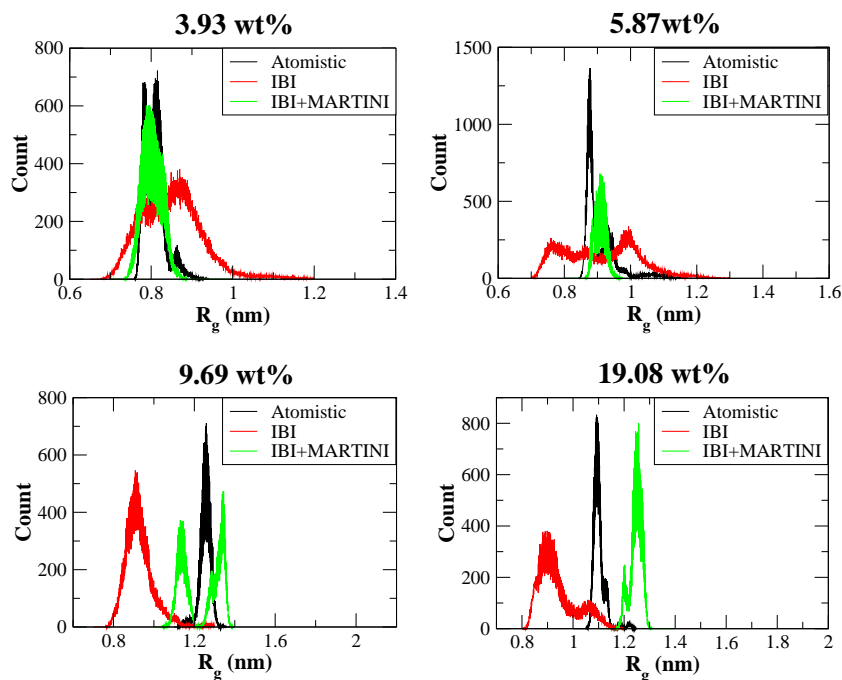
leads to only a slightly smaller  $R_g$  of the CG polymer than the atomistic chain. The slight shift to the right in  $R_{ee}$  of the CG polymer can also be explained along the same lines. The distance between the end atoms of the atomistic chain will always be smaller than the distance between the end super-atoms of the CG polymer.

From an investigation of the structural properties of our coarse-grained models, we infer that the MARTINI model does not fulfill the intricate details of the bonded distributions, i.e., intra-molecular arrangements are not exactly mimicked. However, the MARTINI model represents the global structure fairly well. On the other hand, both the model obtained from IBI and that obtained from IBI+MARTINI parameterization show an excellent match with the structure of the atomistic system. The IBI+MARTINI model offers an additional advantage of fewer optimisable non-bonded interactions compared to the pure IBI model. The greater the number of non-bonded potentials to optimize, the longer it takes to reach convergence because these potentials are cross-correlated. Moreover, the IBI+MARTINI model allows faster simulations owing to the inclusion of MARTINI water which offers a 4 times higher level of coarse-graining over the pure IBI system.

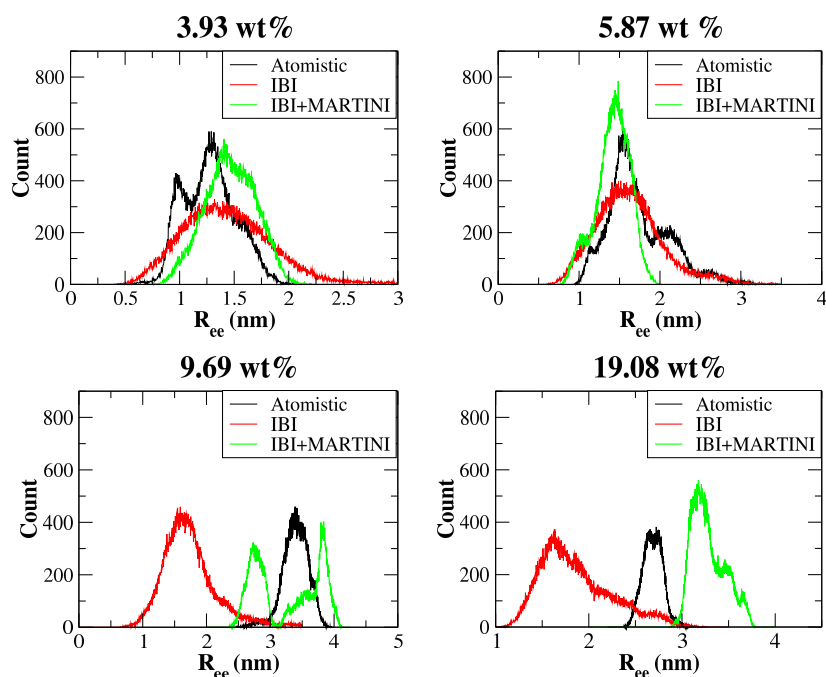
Having arrived at CG models that satisfy the structural properties of a single PAM chain in water, we decided to test the derived potentials on systems with higher concentrations of PAM.

## 4.5 Multi-chains

This section discusses the usefulness of the potentials derived from a single chain of PAM in water when applied to higher concentrations (more than one chain) of the aqueous polymeric system. We use the potentials from the IBI model and IBI+MARTINI for the test. Figs. 4.17 and 4.18 show the extent of overlap between the CG and all-atom probability distributions of  $R_g$  and  $R_{ee}$  respectively.



**Fig. 4.17:** Comparison of  $R_g$  between atomistic and CG models

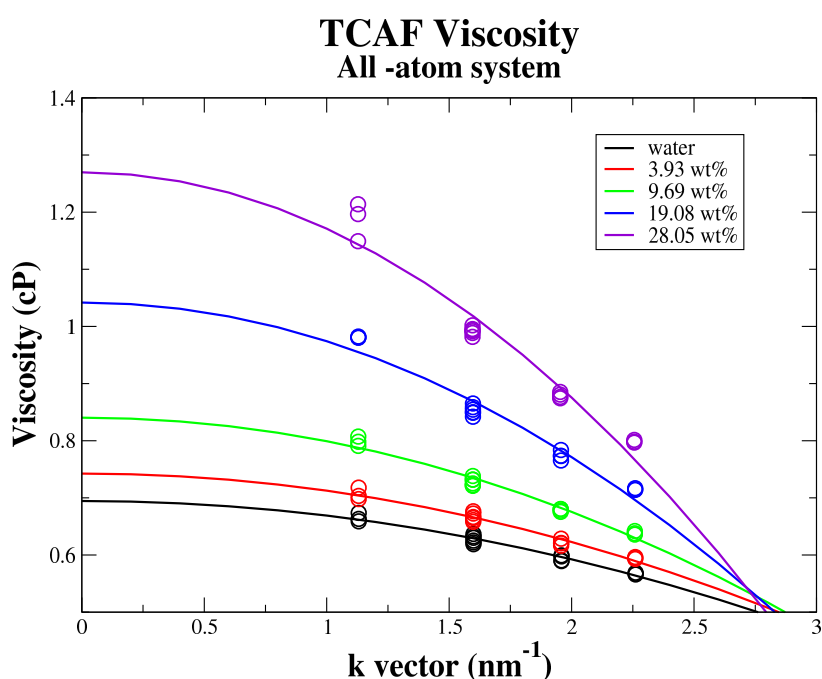


**Fig. 4.18:** Comparison of  $R_{ee}$  between atomistic and CG models

From the above results, we infer that single chain potentials work only qualitatively for the concentration of 3.93 weight%. As we move towards higher concentrations, we see that the CG distributions progressively worsen. These results suggest that a re-optimization of potentials would be needed to design coarse-grained systems in the higher concentration regimes.

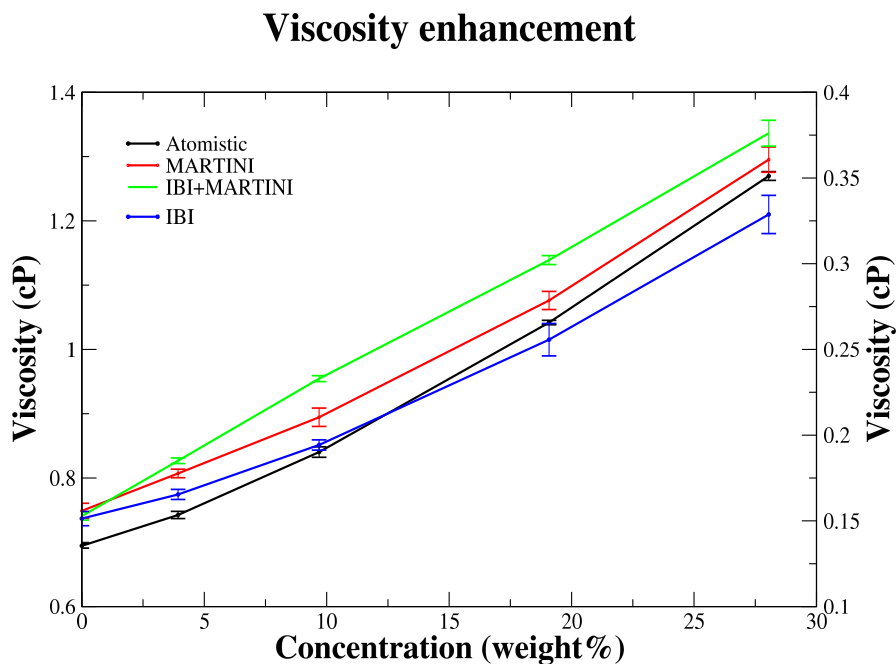
## 4.6 Viscosity enhancement

The results reported in this section correspond to the viscosity calculations made following the procedure in Sec. 3.6. Viscosity values have been reported for 4 models: atomistic, MARTINI, IBI, IBI+MARTINI, spanning concentrations from 0 to 28 weight% of the PAM solution. Following eqn.(E.10), individual plots of viscosity,  $\mu$ , as a function of wave vector,  $k$ , were generated. Fig.4.19 shows the TCAF viscosity plot for atomistic PAM in SPC-E water. The extrapolated value at  $k = 0$  of each TCAF curve gives us the bulk viscosity of the corresponding system. As expected, with increasing concentrations of the polymeric solution the viscosity of the system was enhanced.



**Fig. 4.19:** Viscosity wrt wave vector with increasing concentrations for the all-atom system of PAM in SPC-E water

The trend in viscosity enhancement as a function of increasing concentrations of the polymeric solution has been shown in Fig.4.20 for all the four model force-fields. Although the IBI model was able to represent the increasing trend of viscosity with



**Fig. 4.20:** Viscosity as a function of concentration for all the three CG models against the all-atom system. The curve corresponding to the IBI model (blue) follows the y axis on the right, while the rest three curves follow the left y axis.

concentration, the magnitude of the viscosity values are poorly represented. The viscosity obtained for water from this model was  $0.151 \pm 0.004$  cP, a very low value when compared to the viscosity of SPC-E water,  $0.695 \pm 0.003$  cP. This huge difference can be attributed to the disappearance of many degrees of freedom entailed by the coarse-graining process. The MARTINI water model has been parameterized to reproduce the thermodynamics of the atomistic system, and as a consequence its viscosity is in close agreement with the viscosity of SPC-E water. The experimental viscosity reported for water is 0.89 cP [46]. The CG MARTINI water model is closer to the experimental viscosity than the all-atom SPC-E water. Due to the upper shift in viscosity of water, the vertical offset in the trend against the atomistic curve is maintained throughout. Hence, it is justified that the IBI+MARTINI model wins over the pure IBI system in representing the dynamical behavior.



## 5. CONCLUSIONS AND FUTURE OUTLOOK

Molecular dynamics simulations have been conducted throughout this study for both the atomistic and the coarse-grained models. The objective of this work was to arrive at the most appropriate method of development of a coarse-grained force-field for aqueous polyacrylamide. With GROMOS-53A6 as the force-field and PAM of 30 monomers as the appropriate chain length, the density and the glass transition temperature of PAM melt were successfully reproduced. Three methods have been evaluated to design a coarse-grained system from a united-atom description of an infinitely dilute system of PAM in water. The MARTINI model sufficiently represented the global structure of the polymer chain but not the intricacies of the local structure. In order to accurately represent the structure, a CG system was modeled by the Iterative Boltzmann Inversion (IBI) scheme. The IBI method captured both the local and global structure of the polymer chain. Since the IBI model is constructed based on structure alone, the thermodynamics of the system is often distorted. To take care of this limitation, a novel mechanism of integrating the philosophies of IBI and MARTINI was devised. This hybrid model could successfully reproduce the local and global structure of the polymer. Owing to the incorporation of MARTINI water, this model was 4 times faster than the pure IBI model with the added benefit of fewer fit parameters. While the pure IBI and IBI+MARTINI models correctly predicted the trend in viscosity enhancement, only the latter were in close match with the absolute values of viscosity obtained from atomistic simulations. We have also shown that in terms of replication of structure, both the IBI and IBI+MARTINI models are not applicable beyond  $\approx 4$  wt% of concentration. Combining the results of static and dynamic properties, we conclude that the MARTINI model provides a loose representation of structure but the viscosity trend is adequately maintained; both the IBI and IBI+MARTINI models quantitatively capture the structure at infinite dilution, and qualitatively so for polymer concentrations up to 4 wt%. Only IBI+MARTINI replicates the values and trend in viscosities as a function of polymer concentration. This implies that to work with higher concentrations of the polymer, a re-optimization of potentials will be needed.

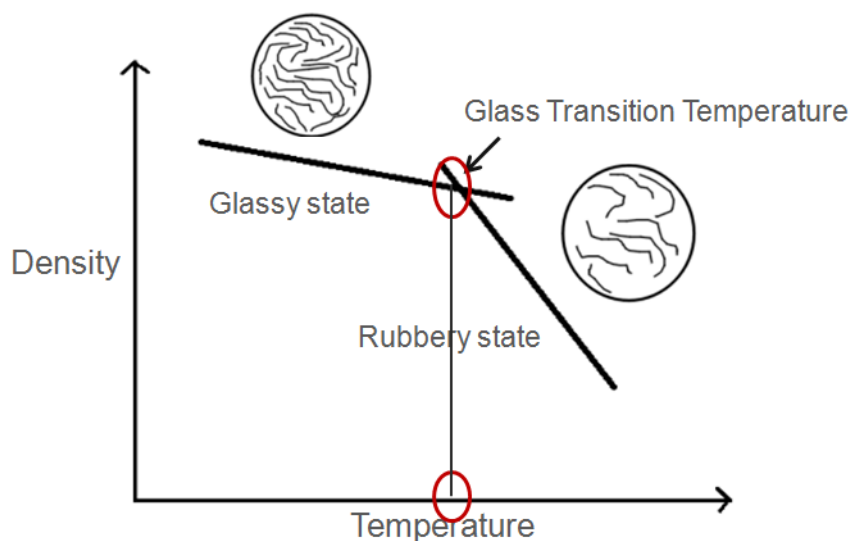
The future work involves devising schemes to generate potentials applicable to larger ranges of concentration. A greater challenge lies in extending our study to polyelectrolytes such as Hydrolyzed-PAM (HPAM), since these are more extensively employed

in oil fields than neutral polymers like PAM. Another plan is to develop a better atomistic forcefield for aqueous PAM, which will lead to the generation of more refined coarse-grained models as well.

# Appendices

## A Glass Transition Temperature

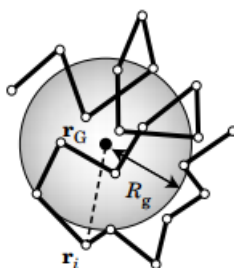
Glass transition temperature ( $T_g$ ) is the principle indicator in the evaluation of thermal properties of amorphous polymers. At temperatures below  $T_g$ , amorphous polymers are in the glassy/rigid state wherein the molecules in the polymer can only vibrate about their position. The chains begin to wiggle around due to the onset of their segmental motion at the  $T_g$ . This state (rubbery state) renders some softness and flexibility to the system. As depicted in the following schematic, the density profile against temperature changes considerably at the  $T_g$ . It is this concept that we have taken inspiration from to approve of the atomistic force field of the polymer.



**Fig. A.1:** A schematic description of the glass transition temperature ( $T_g$ )

## B Radius of gyration

The radius of gyration quantifies the size of the polymer chain. It provides a measure of how compact or extended the polymer chain is. This quantity holds great significance as it can be obtained experimentally too.



**Fig. B.1:** A schematic representation of radius of gyration,  $R_g$ , of a polymer chain [47]

$R_g$  is given by:

$$R_g^2 = \frac{1}{N+1} \sum_{i=0}^N \langle (\mathbf{r}_i - \mathbf{r}_G)^2 \rangle \quad (\text{B.1})$$

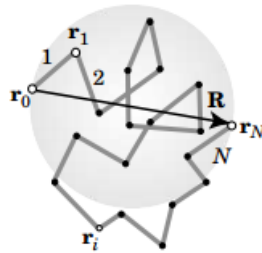
where  $R_g^2$  is the mean square of the distance between the beads and the center of mass of the chain,  $r_G$ , and  $N$  is the number of linkages.

## C End-to-end distance

Another way to measure the chain dimension is by calculating the end-to-end distance ( $R_{ee}$ ) of the chain. It is the distance between two terminal beads, given by:

$$R_{ee}^2 = \langle \mathbf{R}^2 \rangle = \langle (\mathbf{r}_N - \mathbf{r}_0)^2 \rangle \quad (\text{C.2})$$

where  $r_0$  is the position of the first and  $r_N$  of the last bead of the chain.



**Fig. C.1:** A schematic representation of the end-to-end distance,  $R_{ee}$ , of a polymer chain [47]

## D Radial Distribution Function (RDF)

The radial distribution function is a function that describes the radial arrangement of atoms around each other. It provides an estimate of the order/disorder in the system. As shown in Fig.D.2 the entire volume is spliced into shells of thickness  $\delta r$ . The probability of locating an atom in shell  $\delta r$  at a distance  $r$  of the reference atom is the RDF,  $g(r)$ . The mathematic definition:

$$g(r) = \frac{n(r)}{\rho 4\pi r^2 \delta r} \quad (\text{D.3})$$

where  $n(r)$  is the average number of atoms in a shell of thickness  $\delta r$  at  $r$ ,  $\rho$  is the bulk density.

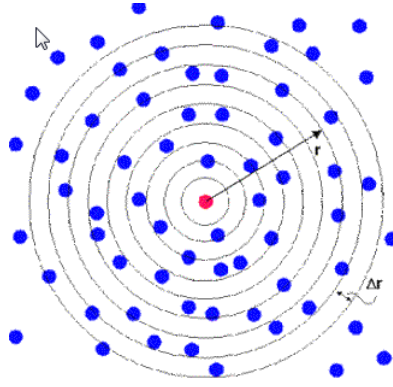


Fig. D.2: Discretization of space for the evaluation of the RDF [48]

## E Transverse Current Autocorrelation Function (TCAF)

At small length scales, random thermal motion of particles result in momentum gradients. Transport coefficients are obtained by an analysis of these gradients [41]. The microscopic transverse momentum field,  $u_{\perp}$  is given by:

$$[41]u_{\perp}(\mathbf{k}, t) = \sum_{j=1}^N \hat{\mathbf{k}}_{\perp} \cdot \mathbf{p}^j(t) \sin[\mathbf{k} \cdot \mathbf{r}^j(t)] \quad (\text{E.4})$$

$$[41]u_{\perp}(\mathbf{k}, t) = \sum_{j=1}^N \hat{\mathbf{k}}_{\perp} \cdot \mathbf{p}^j(t) \cos[\mathbf{k} \cdot \mathbf{r}^j(t)] \quad (\text{E.5})$$

where  $\mathbf{p}^j$  is the momentum of molecule  $j$ , and the  $\mathbf{r}^j$  is the center of mass of molecule  $j$ . The vector  $\hat{\mathbf{k}}_{\perp}$  is the unit vector normal to  $\mathbf{k}$  [41]. The  $\mathbf{k}$  values are: The

$$[41]\mathbf{k} = \frac{2\pi}{L} (n_1, n_2, n_3) \quad (\text{E.6})$$

where  $n_1, n_2, n_3$  represent crystallographic indices. The TCAF is thus obtained [41]:

$$C_{\perp}(\mathbf{k}_x, t) = \langle u_{\perp}(\mathbf{k}_x, t) u_{\perp}(\mathbf{k}_x, 0) \rangle \quad (\text{E.7})$$

With very small values of  $k$ , and large values of  $t$ ,  $C_{\perp}(\mathbf{k}, t)$  decays as:

$$C_{\perp}(\mathbf{k}_x, t) \sim e^{-\left(\frac{\mu k_x^2}{\rho}\right)t} \quad (\text{E.8})$$

$C_{\perp}(\mathbf{k}, t)$  is fitted to an exponential decay to extract the shear viscosity [41]. However, since at short times the behavior of  $C_{\perp}(\mathbf{k}, t)$  is not exponential, a relaxation-time ap-

proximation is made on the momentum transport [49] [50]. With this approximation,  $C_{\perp}(k, t)$  decays with the following function [49] [51]:

$$C_{\perp}(k, t) = \frac{1}{2} \left[ 1 - \frac{1}{\omega} \right] \exp \left[ -\frac{(1 + \omega)t}{2\tau} \right] + \frac{1}{2} \left[ 1 + \frac{1}{\omega} \right] \exp \left[ -\frac{(1 - \omega)t}{2\tau} \right] \quad (\text{E.9})$$

where  $\omega = \sqrt{1 - 4\tau(\mu/\rho)k^2}$ .

The values for  $\mu$  and  $\tau$  are extracted from fitting the simulation curves to eqn.(E.9). For very small values of  $k$ , eqn.(E.9) reduces to eqn.(E.8). and for large  $k$  values  $C_{\perp}(k, t)$  exhibits damped oscillations because  $\omega$  becomes an imaginary quantity.  $\mu$  is an even function of  $k$ , and to order  $k^2$ , the relation becomes [41]:

$$\mu(k) = \mu_{\infty} + ak^2 \quad (\text{E.10})$$

where  $\mu_{\infty}$  is the infinite limit of  $\mu$  obtained by extrapolating the relation to  $k \rightarrow 0$  limit. The value,  $\mu_{\infty}$ , is the bulk viscosity of the system.

# BIBLIOGRAPHY

- [1] C. Peter and K. Kremer, "Multiscale simulation of soft matter systems—from the atomistic to the coarse-grained level and back," *Soft Matter*, vol. 5, no. 22, pp. 4357–4366, 2009.
- [2] V. Rühle, *Morphology and charge transport in conjugated polymers*. PhD thesis, Mainz, Univ., Diss., 2010, 2010.
- [3] <http://www.engr.ucsb.edu>, "multiscale diagram; date accessed: March 4, 2016." <http://www.engr.ucsb.edu/~shell/che210d/Multiscale.pdf>.
- [4] <http://einstein.drexel.edu/>, "The leapfrog integrator; date accessed: March 3, 2016." [http://einstein.drexel.edu/courses/Comp\\_Phys/Integrators/leapfrog/](http://einstein.drexel.edu/courses/Comp_Phys/Integrators/leapfrog/).
- [5] A. Abidin, T. Puspasari, and W. Nugroho, "Polymers for enhanced oil recovery technology," *Procedia Chemistry*, vol. 4, pp. 11 – 16, 2012. The International Conference on Innovation in Polymer Science and Technology.
- [6] W.-M. Kulicke, R. Kniewske, and J. Klein, "Preparation, characterization, solution properties and rheological behaviour of polyacrylamide," *Progress in Polymer Science*, vol. 8, no. 4, pp. 373 – 468, 1982.
- [7] J. Francois, D. Sarazin, T. Schwartz, and G. Weill, "Polyacrylamide in water: molecular weight dependence of  $\langle r^2 \rangle$  and  $[\eta]$  and the problem of the excluded volume exponent," *Polymer*, vol. 20, no. 8, pp. 969–975, 1979.
- [8] G. Misra and S. Bhattacharya, "Determination of the molecular weight of polyacrylamide fractions by osmometry," *European Polymer Journal*, vol. 15, no. 2, pp. 125–128, 1979.
- [9] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. De Vries, "The martini force field: coarse grained model for biomolecular simulations," *The Journal of Physical Chemistry B*, vol. 111, no. 27, pp. 7812–7824, 2007.
- [10] B. J. Alder and T. Wainwright, "Studies in molecular dynamics. i. general method," *The Journal of Chemical Physics*, vol. 31, no. 2, pp. 459–466, 1959.
- [11] A. Rahman, "Correlations in the motion of atoms in liquid argon," *Physical Review*, vol. 136, no. 2A, p. A405, 1964.
- [12] M. Karplus and J. A. McCammon, "Molecular dynamics simulations of biomolecules," *Nature Structural & Molecular Biology*, vol. 9, no. 9, pp. 646–652, 2002.
- [13] C. Oostenbrink, A. Villa, A. E. Mark, and W. F. Van Gunsteren, "A biomolecular force field based on the free enthalpy of hydration and solvation: the gromos force-field parameter sets 53a5 and 53a6," *Journal of computational chemistry*, vol. 25, no. 13, pp. 1656–1676, 2004.
- [14] L. Verlet, "Computer" experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules," *Physical review*, vol. 159, no. 1, p. 98, 1967.



- [15] R. Hockney, S. Goel, and J. Eastwood, "Quiet high-resolution computer models of a plasma," *Journal of Computational Physics*, vol. 14, no. 2, pp. 148–158, 1974.
- [16] F. Ercolessi and J. B. Adams, "Interatomic potentials from first-principles calculations: the force-matching method," *EPL (Europhysics Letters)*, vol. 26, no. 8, p. 583, 1994.
- [17] A. Chaimovich and M. S. Shell, "Coarse-graining errors and numerical optimization using a relative entropy framework," *The Journal of chemical physics*, vol. 134, no. 9, p. 094112, 2011.
- [18] L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, and S.-J. Marrink, "The martini coarse-grained force field: extension to proteins," *Journal of chemical theory and computation*, vol. 4, no. 5, pp. 819–834, 2008.
- [19] S. V. Bennun, M. I. Hoopes, C. Xing, and R. Faller, "Coarse-grained modeling of lipids," *Chemistry and physics of lipids*, vol. 159, no. 2, pp. 59–66, 2009.
- [20] J. J. Uusitalo, H. I. Ingólfsson, P. Akhshi, D. P. Tieleman, and S. J. Marrink, "Martini coarse-grained force field: extension to dna," *Journal of chemical theory and computation*, vol. 11, no. 8, pp. 3932–3945, 2015.
- [21] H. Lee, A. H. de Vries, S.-J. Marrink, and R. W. Pastor, "A coarse-grained model for polyethylene oxide and polyethylene glycol: conformation and hydrodynamics," *The journal of physical chemistry B*, vol. 113, no. 40, pp. 13186–13194, 2009.
- [22] R. S. D'Rozario, C. L. Wee, E. J. Wallace, and M. S. Sansom, "The interaction of c60 and its derivatives with a lipid bilayer via molecular dynamics simulations," *Nanotechnology*, vol. 20, no. 11, p. 115102, 2009.
- [23] S. J. Marrink and D. P. Tieleman, "Perspective on the martini model," *Chemical Society Reviews*, vol. 42, no. 16, pp. 6801–6822, 2013.
- [24] H. Hu, Z. Lu, and W. Yang, "Fitting molecular electrostatic potentials from quantum mechanical calculations," *Journal of chemical theory and computation*, vol. 3, no. 3, pp. 1004–1013, 2007.
- [25] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, "Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," *SoftwareX*, vol. 1, pp. 19–25, 2015.
- [26] H. Wang, H. Zhang, C. Liu, and S. Yuan, "Coarse-grained molecular dynamics simulation of self-assembly of polyacrylamide and sodium dodecylsulfate in aqueous solution," *Journal of colloid and interface science*, vol. 386, no. 1, pp. 205–211, 2012.
- [27] W. H. Press, S. Teukolsky, W. Vetterling, and B. Flannery, "Numerical recipes in fortran (cambridge)," 1992.
- [28] H. J. Berendsen, J. v. Postma, W. F. van Gunsteren, A. DiNola, and J. Haak, "Molecular dynamics with coupling to an external bath," *The Journal of chemical physics*, vol. 81, no. 8, pp. 3684–3690, 1984.
- [29] L. Onsager, "Electric moments of molecules in liquids," *Journal of the American Chemical Society*, vol. 58, no. 8, pp. 1486–1493, 1936.
- [30] B. Hess, H. Bekker, H. J. Berendsen, J. G. Fraaije, *et al.*, "Lincs: a linear constraint solver for molecular simulations," *Journal of computational chemistry*, vol. 18, no. 12, pp. 1463–1472, 1997.
- [31] P. G. Kusalik, I. M. Svishchev, *et al.*, "The spatial structure in liquid water," *Science(Washington)*, vol. 265, no. 5176, pp. 1219–1221, 1994.

- [32] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, "A smooth particle mesh ewald method," *The Journal of chemical physics*, vol. 103, no. 19, pp. 8577–8593, 1995.
- [33] S. Nosé, "A molecular dynamics method for simulations in the canonical ensemble," *Molecular physics*, vol. 52, no. 2, pp. 255–268, 1984.
- [34] W. G. Hoover, "Canonical dynamics: equilibrium phase-space distributions," *Physical Review A*, vol. 31, no. 3, p. 1695, 1985.
- [35] M. Parrinello and A. Rahman, "Polymorphic transitions in single crystals: A new molecular dynamics method," *Journal of Applied physics*, vol. 52, no. 12, pp. 7182–7190, 1981.
- [36] G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," *The Journal of chemical physics*, vol. 126, no. 1, p. 014101, 2007.
- [37] V. Rühle, C. Junghans, A. Lukyanov, K. Kremer, and D. Andrienko, "Versatile object-oriented toolkit for coarse-graining applications," *Journal of chemical theory and computation*, vol. 5, no. 12, pp. 3211–3223, 2009.
- [38] B. Bayramoglu and R. Faller, "Modeling of polystyrene under confinement: Exploring the limits of iterative boltzmann inversion," *Macromolecules*, vol. 46, no. 19, pp. 7957–7976, 2013.
- [39] D. Reith, M. Pütz, and F. Müller-Plathe, "Deriving effective mesoscale potentials from atomistic simulations," *Journal of computational chemistry*, vol. 24, no. 13, pp. 1624–1636, 2003.
- [40] E. L. B. H. M.J. Abraham, D. van der Spoel and the GROMACS development team, "Gromacs user manual version 5.1-beta1."
- [41] B. J. Palmer, "Transverse-current autocorrelation-function calculations of the shear viscosity for molecular liquids," *Physical Review E*, vol. 49, no. 1, p. 359, 1994.
- [42] www.chemicalbook.com, "density of polyacrylamide; date accessed: March 9, 2016." [http://www.chemicalbook.com/ChemicalProductProperty\\_EN\\_CB7390058.htm](http://www.chemicalbook.com/ChemicalProductProperty_EN_CB7390058.htm).
- [43] <http://polymer.nims.go.jp>, "density of polyacrylamide; date accessed: March 9, 2016." [http://polymer.nims.go.jp/index\\_en.html](http://polymer.nims.go.jp/index_en.html).
- [44] <https://scientificpolymer.com>, "density of polyacrylamide; date accessed: March 9, 2016." <http://scientificpolymer.com/density-of-polymers-by-density/>.
- [45] <https://www.sigmaaldrich.com/>, "glass transition temperature of polyacrylamide; date accessed: March 9, 2016." [https://www3.nd.edu/~hgao/thermal\\_transitions\\_of\\_homopolymers.pdf](https://www3.nd.edu/~hgao/thermal_transitions_of_homopolymers.pdf).
- [46] R. C. Weast, "Crc handbook of chemistry and physics.," 1986.
- [47] I. Teraoka, *Frontmatter and Index*. Wiley Online Library, 2002.
- [48] <http://www.ccp5.ac.uk/>, "rdf diagram; date accessed: March 19, 2016." [http://www.ccp5.ac.uk/DL\\_POLY/Democritus/Theory/rdf.html](http://www.ccp5.ac.uk/DL_POLY/Democritus/Theory/rdf.html).
- [49] J. P. Boon and S. Yip, *Molecular hydrodynamics*. Courier Corporation, 1980.
- [50] D. Forster, "Hydrodynamic fluctuations, broken symmetry, and correlation functions," in *Reading, Mass., WA Benjamin, Inc.(Frontiers in Physics. Volume 47)*, 1975. 343 p., vol. 47, 1975.
- [51] R. Vogelsang and C. Hoheisel, "Thermal conductivity of a binary-liquid mixture studied by molecular dynamics with use of lennard-jones potentials," *Physical Review A*, vol. 35, no. 8, p. 3487, 1987.