

SCUOLA NORMALE SUPERIORE DI PISA  
CLASSE DI SCIENZE

DECEMBER 2009

---

Modeling and application of a  
coarse-grained force field for  
surfactants and polymeric  
membranes.

---

TESI DI PERFEZIONAMENTO  
IN CHIMICA

*Author:*  
Paolo ELVATI

*Supervisor:*  
Prof. Dr. Michele PARRINELLO

In light of knowledge attained, the happy achievement seems almost a matter of course, and any intelligent student can grasp it without too much trouble. But the years of anxious searching in the dark, with their intense longing, their alterations of confidence and exhaustion and the final emergence into the light – only those who have experienced it can understand it.

Albert Einstein

The mind is its own place, and in itself  
Can make a heav'n of hell, a hell of heav'n.

John Milton

---

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Fuel cells . . . . .	2
1.1.1	Direct methanol fuel cell . . . . .	3
1.2	Modeling . . . . .	5
1.2.1	Microscopic simulations . . . . .	6
1.3	Thesis structure . . . . .	8
<b>2</b>	<b>Force field parametrization</b>	<b>11</b>
2.1	Introduction . . . . .	11
2.1.1	Simulation details . . . . .	13
2.2	Bonded interactions . . . . .	13
2.3	Iterative Boltzmann Inversion . . . . .	17
2.3.1	Theory . . . . .	18
2.3.2	Results . . . . .	22
2.4	Fit of macroscopic properties . . . . .	28
2.4.1	Pure liquids . . . . .	30
2.4.2	Non-miscible species . . . . .	30
2.5	Charged species . . . . .	34
2.5.1	Charge-charge interactions . . . . .	34
2.5.2	Free energy perturbation . . . . .	35
2.6	Conclusions . . . . .	42
<b>3</b>	<b>Force field testing</b>	<b>43</b>
3.1	Polyethylene chains . . . . .	44
3.2	Solvents . . . . .	46
3.2.1	Temperature dependence . . . . .	47

## CONTENTS

---

3.3	Methanol and water mixtures . . . . .	49
3.4	Efficiency considerations . . . . .	53
3.5	Conclusions . . . . .	55
<b>4</b>	<b>FF Applications</b>	<b>57</b>
4.1	Surfactants . . . . .	57
4.1.1	Sodium dodecylbenzene sulfonate . . . . .	58
4.1.2	Micelles . . . . .	60
4.1.3	Conclusions . . . . .	65
4.2	SNOVA Membrane . . . . .	67
4.2.1	System setup . . . . .	68
4.2.2	Results . . . . .	70
4.2.3	Conclusions . . . . .	73
<b>5</b>	<b>Thesis summary</b>	<b>75</b>
<b>A</b>	<b>Bond and bend fitting tables</b>	<b>79</b>
<b>B</b>	<b>Force field summary tables</b>	<b>81</b>
	<b>Abbreviations and symbols</b>	<b>83</b>
	<b>Acknowledgments</b>	<b>85</b>
	<b>List of figures</b>	<b>87</b>
	<b>List of tables</b>	<b>89</b>
	<b>Bibliography</b>	<b>91</b>

---

# Chapter 1

## Introduction

Since ancient times the availability of energy, and the corresponding possibility of producing various forms of work, has been a crucial factor for the development and the wellness of society. Almost all human activity, from food production, to transportation and communication relies on the existence of some sort of energy source. Before the industrial revolution the vast majority of energy derived from human and animal activity. However in the last few centuries more and more power is produced by using other sources (*e.g.*, carbon, oil, uranium, wind, sun), which has had the consequence of greatly reducing the human labor and greatly increasing the quantity of energy at our disposal. On the other hand, the dependence of states and economies on new resources and raw materials was, and still is, the cause of new friction among them. Moreover the large exploitation of natural resources has great effect on the world climate (*e.g.*, greenhouse effect) and pollution, raising increased concerns in public opinion.

Trends in energy consumption also show that the availability of energy will be a central problem in the decades to come. In 2003 the world's energy consumption stood at  $4.43 \cdot 10^{20}$  J and is expected to double in a few years, largely due to the rising energy demands of fast developing countries, such as China and India. In addition, the combustion of fossil fuel (oil, coal and natural gas), which currently accounts for 85% of the global energy supply, produces an enormous amount of carbon dioxide, one of the major greenhouse gases responsible for global warming, as well as other polluting combustion byproducts.

Contrary to animal work, which inherits the efficiency of biological processes refined in millions years of evolution, other power sources are far from optimal and have much room for improvement in all the aspects related to production, transformation and storage of energy. It is therefore of no surprise that these topics are so central in research, due to

their political, social and technological repercussions.

In this general picture, the availability of a device that can store power in a very efficient way and deliver it with high efficiency, with a very limited impact on the environment, is of extreme interest. Fuel cells (FC) offer many of these advantages and are therefore the most likely candidate for energy devices of the next few decades.

## 1.1 Fuel cells

FCs comprise a wide variety of systems, with greatly different characteristics, properties and applications. Despite the huge overall differences, FCs can be described as galvanic cells that produce electricity by electrochemical conversion of fuels (*e.g.*, hydrogen, methanol) in heat and steam. The energy is stored as a chemical bond and is converted with extremely high efficiency (from 40% up to 90% if co-generation is used), because no mechanical work is involved. This allows theoretical thermodynamic efficiencies that can be greater than one, as well as longer lifetime of the devices. FCs also have the advantage of a very good scalability (minimal efficiency loss when increasing the size of the device), while having negligible noise and low chemical emissions.

A few examples of the most common kind of FCs, with some properties are summarized in Table 1.1.

Schematically, the basic components of a fuel cell are the anode, the anodic catalyst layer, an electrolyte, the cathodic catalyst layer and the cathode, all stacked and sealed to prevent direct leakage of gases and liquids between the anodic and cathodic compartments.

**Table 1.1:** Data for different types of fuel cells. The carrier column lists the main mobile ion present in the electrolyte. Data taken from [1, 2, 3, 4].

FC	Carrier	$T_{work}$ (°C)	Electrolyte	Power (MW)
Alkaline	$\text{OH}^-$	50-200	Alkaline solution	0.01-0.1
Direct methanol	$\text{H}^+$	20-80	Polymeric membrane	up to 1
Molten carbonate	$\text{CO}_3^{2-}$	600-700	Alkaline carbonate	10 - 100
Phosphoric acid	$\text{H}^+$	150-220	Phosphoric acid	up to 10
Proton exchange membrane	$\text{H}^+$	30-100	Polymeric membrane	0.5
Solid oxide	$\text{O}^{2-}$	700-1000	Ceramic	100

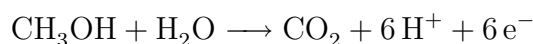
The anode and cathode consist of porous gas diffusion layers, usually made of highly electron and very low proton conductivity materials, such as thin layers of porous graphite. This layer ensures a homogeneous distribution of the reactants and a fast removal of the products, while supporting the electrodes. One of the most common catalysts is platinum for low temperature fuel cells and nickel for high temperature fuel cell, albeit to increase the kinetic efficiency or reduce poisoning of the electrodes, various alloys (*e.g.*, Pt-Ru, Pt-Mo, Ni oxides) are usually employed. Finally an electrolyte made of a material that provides high proton and extremely low electron conductivity is placed between the two electrodes.

The invention of fuel cells dates back from 1839 by Sir William Robert Grove who developed the *Grove cell* [5] and few years later the first hydrogen fuel cell [6]. The principle however was discovered by Christian Friedrich Schönbein [7] a professor at the University of Basel, who was in mail contact with Sir Grove, making FCs one of the oldest electrical conversion technologies known. However, due to a lack of driving need the FCs were not subject to extended investigation until the Gemini and Apollo space programs. Despite the huge improvements over the last thirty years, still many of the mechanisms in FCs are poorly understood and different limitations still need to be overcome to make these kind of devices widely applicable.

### 1.1.1 Direct methanol fuel cell

Considering only the low-temperature region (30 - 100°C), proton exchange membrane FC (PEMFC) and direct methanol FC (DMFC) are the most attractive power sources for a wide range of applications. In DMFCs and PEMFCs, sometimes called solid polymer electrolyte FCs, the electrolyte is a proton exchange membrane, like Nafion, a sulfonated tetrafluoroethylene based fluoropolymer, discovered in the late 1960s [8].

The key difference between DMFCs and PEMFCs is the fuel: in the DMFC the power is produced by direct conversion of liquid methanol, instead of hydrogen, without any reforming step. The overall mechanism for the methanol oxidation is:



where the oxidation of methanol proceeds by several deprotonation steps, which produce different stable and unstable intermediates [9].

Methanol is an attractive fuel option because it can be produced by renewable biomass resources and has a high energy density, since it is a liquid at standard operating conditions.

Moreover, the higher safety and easier handling and storage of methanol with respect to  $H_2$  make DMFCs more suitable for common small mobile devices, such as cellphones and laptops.

However the usage of methanol introduces a new series of problems [10]. On one hand, the catalytic activity for methanol oxidation, compared to hydrogen oxidation, is not very efficient. This is worsened by the fact that choice of possible catalysts is much narrower compared to PEMFCs: in fact only Pt and Pt-based alloys can be used. Moreover carbon monoxide, an intermediate product of methanol oxidation, can poison Pt by blocking the metallic surface and therefore hindering any further reaction. On the other hand the membranes used in DMFCs were developed for PEMFCs, thus optimized to maximize the proton conduction but neglecting methanol blocking. The permeation of methanol from the anode compartment to the cathode catalyst through the membrane, *i.e.*, methanol crossover, leads to mixed electrochemical potential (methanol oxidation interferes with the oxygen reduction at the cathode), which causes a decrease in the overall cell performance. Methanol crossover, which happens mainly due to diffusion and electro-osmotic drag, depends on a number of factors such as the anode performance, the concentration of methanol in the fuel feed and the working temperature. The most important, however, is the membrane methanol permeability and thickness. Considering the fact that almost two thirds of the voltage loss in a DMFCs is due to methanol crossover, it is clear why most of the research is focused on the improvement of the membrane performance.

However, trying to understand the physical and chemical processes that occur in the electrolyte is intricate. Experimental approaches are hampered by the structural and thermal anisotropy of the membranes, and the results depend on many other variables (*e.g.*, catalytic layer contact, reagent diffusion) giving contradictory outcomes [11]. Another problem arises from the difficulty of obtaining truly reproducible measurements, due to the fact that many of the properties of the membrane (and of the polymer in general) strongly depend on the thermal and mechanical history of the sample and, as shown recently [12], often the standardization procedures are not effective. The experimental difficulties and the poor understanding of the exact details of many underlying mechanisms (*e.g.*, ionic and parasitic transport), force one to use only empirical models, thereby making the effects of any membrane modification difficult, if not impossible to predict.

These are the reasons why even the most studied FC membrane, *i.e.*, Nafion, about which more than 33000 papers have been published to date, continues to elude a consistent description. Specifically a detailed understanding of the morphology of the membrane,



especially when hydrated, and of the transport mechanisms is still incomplete, despite the huge research effort.

## 1.2 Modeling

Since DMFC's membranes have a low degree of spatial order, which makes it difficult to deduce unambiguous information from experiments alone, many computational modeling studies have been used.

There are different global methodologies for modeling fuel cells. The easiest classification is into macroscopic and microscopic models. The microscopic models seek to model transport on an individual pore level, whereas the macroscopic ones are continuum models and average over this level. The microscopic models provide more realistic conditions, but they require much more knowledge of the micro-structure and are much more expensive in terms of computation time.

Macroscopic models are therefore more common for fuel cells and although some microscopic details should be incorporated into them, in such approaches, the exact micro-structural details of the modeled domain are neglected. Instead, the domain is treated as a randomly arranged porous structure that can be described by a small number of parameters such as porosity and surface area per unit volume. Transport properties within the domain are averaged over the electrode volume, so that averaging is performed over a region that is small compared to the size of the electrode but large compared to the micro-structure.

Starting from the 70s, several macroscopic fuel cell models were produced [13], and for membranes they can be grouped as diffusive models [14, 15] (where the membrane is modeled as a single phase), hydraulic models [16, 17] (two-phase membrane) and various combinations of the two (hydraulic-diffusive models [18, 19] and combination models [20, 21]).

However, the success of these approaches is largely limited to Nafion and its various chemical homologues (*e.g.*, Flemion, Aciplex and Haflon), and even in the best cases they can only relate some measurable macroscopic properties to other macroscopic properties and still fail in relating the molecular scale with the final properties of the membrane. To bridge this gap and to connect the the chemical and topological structure of the polymer to the properties of the membranes, a less phenomenological approach should be used and microscopic details must be included in the model.

### 1.2.1 Microscopic simulations

In recent years, impressive advances have been made in molecular modeling and numerical simulations. Some of these can be ascribed to the increase in computational power, but of fundamental importance was the development of a wide variety of methods, each of them optimized to analyze problems in a restricted region of the time and length scales. Computer simulations, such as those using Monte Carlo (MC) or molecular dynamics (MD) methods, are important tools and are routinely used to study molecular systems in condensed phase or solution. This is because these methods can provide numerical solutions to many problems of statistical mechanics by calculating ensemble averages for experimental observables.

Polymers, with their large variety of important length and time scales, pose a formidable challenge for computational simulations. From the spatial point of view the relevant scales range from bond length, on the order of  $\text{\AA}$ , to at least the length of the chain, which can easily reach the order of the micrometers. Moreover, polymers also suffer from the detailed treatment of the fast modes, because time propagation of the system is limited by the frequency of the stiffest mode. Consequently, slow modes and structural reorganization (*e.g.*, end-to-end distance, reptation movements, phase transitions), which happen in much longer time scales, can hardly be equilibrated or correctly sampled.

#### Coarse-graining

Polymers are by no means an exception: soft materials, biomolecular systems and also many process that occur in liquids, occur over length and time scales that are beyond the current capabilities of atomic-level simulations. Therefore different approaches continue to be developed or improved to overcome these limitations.

A common strategy employed by many of these methods is to cluster groups of atoms into new coarse-grained (CG) sites. The reduction in the total number of degrees of freedom of the system as well as a choice of computationally efficient interaction between the CG sites, allows for a significant improvement in the accessible spatial and temporal scales. As such the idea of the CG approach is to simplify the representation of the molecules by reducing the molecular-scale information while retaining as much of the key physical information about the system as possible.

There are numerous different CG models that vary enormously depending on the choice made during modeling of the target system. CG models can be either on-lattice, which are faster (*e.g.*, lattice Boltzmann methods [22]) or off-lattice, which are more versatile

and realistic [23, 24]. Also the decisions pertaining the nature of the dynamics, which can be conservative or dissipative, generated different models, like the dissipative particle dynamics (DPD) [25, 26, 27, 28] where, in general, [29, 30] energy conservation is sacrificed to efficiently simulate suspension of particles of millions or billions of atoms. The methods differ widely also on the level of coarse-graining, ranging from almost all-atom (AA) representations, as in the united atom model, to colloidal modeling (*e.g.*, the hybrid molecular dynamics of Malevanets and Kapral [31]), passing through the one-bead CG models for proteins, which use a single interaction center to represent a whole amino acid [32]. Recently, also mixed or multiple resolution models were used, such as the multi-scale methods used by Goddard [33]. Finally, even for the same level of coarse-graining and same dynamics, many different methods arise from the choice of the functional form and the parametrization procedure of the potential, such as the force matching method of Izvekov [34], the simplex optimization techniques of Reith [35] or the MARTINI force field [36].

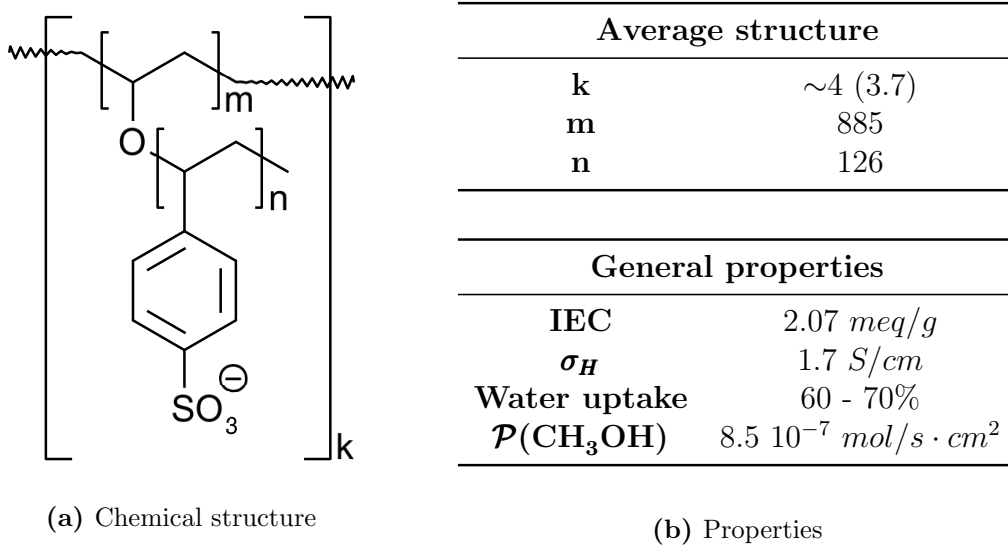
Although CG models may provide a highly efficient computational tool for rapidly investigating the relevant configuration space, the conclusions drawn with such methods can be misleading unless the CG model is consistent with a higher-resolution representation. The challenge in the development of the CG model is therefore the parametrization of the interaction force field, which effectively bridges a realistic AA model with the computationally efficient low-resolution CG description.

## 1.3 Thesis structure

This thesis is about the parametrization of a CG force field and its applications to a novel polymeric membrane, called SNOVA. This is a very promising electrolyte for DMFCs, due to its high proton conductivity and low methanol permeation. The chemical structure of the polymer and some of its characteristics are presented in Figure 1.1.

While very promising as an electrolyte, there is no information regarding the morphology of the hydrated polymer and the transport mechanisms occurring in the membrane. Therefore any future improvement as well as macroscopic modeling is seriously hampered by the lack of understanding of the behavior of the electrolyte under working conditions.

As noted earlier, the AA simulations of such systems are computationally too expensive. Thus, the aim of this work is to build a CG model of the polymer, where only the essential parts of the macromolecule are included. The interaction potentials between the resulting CG beads are then chosen to accurately represent the underlying system.



**Figure 1.1:** Some properties and average chemical formula of structure of the chain composing the SNOVA membrane. The methanol permeability of the membrane  $\mathcal{P}$  was obtained at 30 °C, using a 6 M methanol solution as anodic fuel [37].

With this task in mind, in the second chapter the parametrization techniques and the result of the modeling procedure will be presented. In particular, *a)* the choice of the CG grouping, *b)* the fitting procedure for the bonded potentials and *c)* the different techniques

used to determine the non-bonded potentials (*i.e.*, iterative Boltzmann inversion, fitting to thermodynamic properties and electrostatic interaction modeling).

As previously discussed, there is still no universal solution to the CG problem. Therefore, no matter how careful the planning and the modeling of the CG force fields are, there is no guarantee that they will reproduce the properties of interest. Hence in the third chapter the obtained force field is tested by considering increasingly complex systems. Firstly the behavior of polyethylene chains in vacuum is considered to test the bonded potentials. Secondly macroscopic properties of simple CG homogeneous solutions are compared with the values obtained from experiments and AA simulations, with a particular focus on temperature dependence. Finally the non-ideal behavior of the methanol-water solutions in the CG representation with respect to the experiments is discussed.

After these validations, the force field can be applied to more interesting applications, and therefore in the fourth chapter the applications to real systems of interest are presented. Firstly, the aggregation properties of a very common surfactant, sodium dodecylbenzene sulfonate (SDBS) are shown, and secondly, preliminary results on the topology of the SNOVA membrane are presented.

Following this, a detailed summary of the obtained results and the appendices, which contain many tables listing exhaustive results used in the force field parametrization, are presented.



---

# Chapter 2

## Force field parametrization

### 2.1 Introduction

As mentioned previously (see Section 1.2.1) the CG procedure consists in building a low-resolution description of the system under study by representing groups of many atoms with single beads. This dimensional reduction is reflected on the properties of the simulated system, and limits the information that can be extracted. Therefore the choice of the subset of the degrees of freedom that are integrated out and the parametrization of the interactions in the new representation are of fundamental importance. There is no unique or preferred way to determine the mapping between an AA to a CG representation, however, previous work [38] shows that taking into account the chemical structure and the functionality of the molecules is a sensible strategy to capture the relevant properties of the system. In this work the target macromolecule is a sodium polyethylene-*g*-polystyrene sulfonate membrane (see Figure 1.1) swollen by a methanol-water mixture. The PE backbone of the polymer was coarse-grained by lumping 3 carbons and their hydrogen together into one bead to have a relevant computational gain but also preserve some structural information. On the other hand, for the grafted chains more chemical diversity was retained by using three beads to model the different functional groups of the PS repeating units: one bead for the sulfonic group, one for the benzene ring and one for the aliphatic part. Finally, both the solvent and counterion molecules were treated as one CG molecule each (see Table 2.1).

Akin to an atomistic force field, the CG units interact through a CG force field consisting of bonded and non-bonded interactions. Here, the potential acting on each bead was

**Table 2.1:** Mapping between the chemical fragments in atomistic representation and the coarse-grained units.

CG unit	AA fragment
B	$-\text{C}_6\text{H}_4-$
C2	$-\text{CH}_2-\text{CH}_2-$
C3	$-\text{CH}_2-\text{CH}_2-\text{CH}_2-$
M	$\text{CH}_3\text{OH}$
Na	$\text{Na}^+$
S	$-\text{SO}_3^-$
W	$\text{H}_2\text{O}$

described as the sum of the following terms:

$$\begin{aligned}
 U_{tot} &= U_{bonded} + U_{non-bonded} \\
 &= (U_{bond} + U_{angle}) + (U_{vdw} + U_{el})
 \end{aligned}$$

where  $U_{bond}$  is the potential describing the bond stretching,  $U_{angle}$  represents the angle bending energy,  $U_{vdw}$  accounts for van der Waals interactions and  $U_{el}$  for the electrostatic. Contrary to AA force fields, the term related to the torsional energy was neglected because in the CG resolution the characteristic lengths become large enough to make torsional contribution quite small, and therefore more appropriately included in the non-bonded interactions.

All these contributions to the potential energy are correlated to some extent, and as such the order of the parametrization is relevant [39]. The most direct solution is the simultaneous optimization of all the terms composing the potential energy. While possible [35, 40, 41], this will greatly reduce the transferability of the obtained CG force field. An alternative solution to minimize the inevitable correlation introduced by the FF parametrization is to sequentially obtain the terms in the order of relative strength [35]:

$$U_{bond} \longrightarrow U_{angle} \longrightarrow U_{non-bonded}$$

In the rest of this chapter, the fitting procedures used to obtain the CG force field parameters is presented. After some general operative details concerning the simulations, the method used for the bonded interactions is described, followed by the approaches used for non-bonded potentials. The latter can be subdivided in three groups: fit pair corre-



lation functions against reference radial distribution functions, *i.e.*, iterative Boltzmann inversion (Section 2.3), fitting against experimental density and surface/interface tension data (Section 2.4), and methods to model charged species (Section 2.5). The list of all the obtained parameters can be found in the Appendix B.

### 2.1.1 Simulation details

All the AA simulations of this chapter, were carried out with the NAMD code [42] using the AMBER force fields (GAFF [43] and FF03 [44]). The r-RESPA [45] algorithm was used to speed up the simulations by evaluating the long-ranged electrostatic interactions every 2 fs while all other interactions were computed every fs; the particle mesh Ewald method [46] was used to evaluate the electrostatic interactions. Constant temperature was maintained at 325 K, which is a typical working condition of the membrane in the fuel cell, by using a Langevin thermostat with a frequency constant of 1 ps<sup>-1</sup> while constant pressure was obtained by using the Langevin piston method [47, 48].

The coarse-grained molecular dynamics simulations were carried out using the CM<sup>3</sup>D code [49]. Constant temperature simulations were performed at 325 K using a Nose-Hoover chain [50, 51] of length three with a frequency constant of 0.5 ps<sup>-1</sup>. All the potentials were truncated at a cutoff radius of 1.8 nm and tail corrections to the pressure were added when appropriate. A multiple time step scheme (r-RESPA) was also used to improve the computational efficiency: long range forces were evaluated every 10 fs while all other interactions were evaluated every 2 fs. It should be stressed that a conservative choice of the CG time step was chosen; however at the price of a slight increasing energy drift, the time step can be pushed to 30 fs for long range interactions and to 2.5 fs for the short one.

## 2.2 Bonded interactions

The bonded interactions between connected sites are described by harmonic bond stretching and harmonic bending potentials:

$$U_{bond}(r) = \frac{1}{2}k_b(r - r_0)^2 \tag{2.1}$$

$$U_{angle}(\theta) = \frac{1}{2}k_a(\theta - \theta_0)^2 \tag{2.2}$$

in which  $r_0$  and  $\theta_0$  are parameters that set the equilibrium distance and angle respectively, and  $k_b$  and  $k_a$  are the corresponding force constants. To obtain the reference distributions of structural observables AA MD simulations were performed. To avoid the inclusion of terms coming from AA intermolecular forces in the CG bonded potentials, the distance,  $P(r)$ , and angle,  $P(\theta)$ , distributions were obtained from runs of single molecules in vacuum (100 ns with a long range forces cutoff of 1.8 nm). Details about the alkane used to model the potentials can be found in Tables B.1 and B.2.

The relation between the structural distributions and the effective potential depends on the ensemble. In the canonical ensemble the probability distributions to have a given distance  $r'$  or angle  $\theta'$  are:

$$P(r') = \frac{1}{N} \iiint dr d\theta d\phi r^2 \sin(\theta) \delta(r - r') \exp(-\beta U_{bond}(r)) \quad (2.3)$$

$$P(\theta') = \frac{1}{N} \iiint dr d\theta d\phi r^2 \sin(\theta) \delta(\theta - \theta') \exp(-\beta U_{angle}(\theta)) \quad (2.4)$$

where  $\delta$  is the Dirac's delta and  $N$  the normalization factor.

The fitting procedure of the AA distribution can be carried out in many different ways, because equations (2.3) and (2.4) can be rearranged differently, *e.g.*, by fitting  $P(r)$  with  $c_0 r^2 e^{-\beta U_{bond}(r)}$  or  $P(r)/r^2$  with  $c_0 e^{-\beta U_{bond}(r)}$ . In the limiting case of a perfect fit all the procedures give equivalent results, but in general each procedure weights the points of the distributions in different ways. An *a priori* choice of the best method can be made if the error function for each distribution is known. However, this is often unfeasible, therefore in this work the method that gives parameters that best reproduce the original AA distribution was chosen. This *a posteriori* approach involves some extra simulations because the result of each fitting has to be tested, but the procedure can be done for just one molecule, so the overall cost is negligible.

The case of the angular potential for C2 beads is now presented, in order to show how much the fitting procedure influences the results. After obtaining the target distribution  $P(\theta)$  from an AA simulation of one hexane molecule, three procedures were used to analyze it:

1. Fit  $P(\theta)$  with  $y(x) = a_2 \cdot \sin(x) \cdot \exp(-a_1 \cdot (x - a_0)^2)$
2. Fit  $P(\theta)/\sin(\theta)$  with  $y(x) = a_2 \cdot \exp(-a_1 \cdot (x - a_0)^2)$
3. Fit  $P(\cos(\theta))$  with  $y(x) = a_2 \cdot \exp(-a_1 \cdot (\arccos(x) - a_0)^2)$

where  $a_0$ ,  $a_1$ ,  $a_2$  are the free parameters that will determine the potential. The three methods are not equivalent because each of them has different residuals and as such the minimization of  $\chi^2$  gives different results. Another reason for the difference is the numerical error, which is not the same for all the functions and that can become extremely big near singularity points, like  $\pi$ . As shown in Table 2.2, different results are obtained depending on the method and, to some extent, on the size of the bins.

**Table 2.2:** Results of different fitting methods of the C2–C2–C2 distribution obtained from an AA simulation of hexane in vacuum. The details for each procedure are written in the text.  $\theta_0$  is expressed in degrees,  $k_a$  in K/rad<sup>2</sup> and the unit for the distribution bin depends on the method: rad for the first two, dimensionless for the last.

Bin size	Method 1		Method 2		Method 3	
	$\theta_0$	$k_a$	$\theta_0$	$k_a$	$\theta_0$	$k_a$
0.02	180	3470	177	102000	180	12800
0.01	180	3480	177	80000	180	24600
0.005	180	3480	180	35300	180	27500
0.001	180	3480	180	32700	180	28600
0.0001	180	3480	180	33600	180	28700

To choose the best fitting procedure among these three, the same number of CG hexane simulations were done using the set of parameters obtained with each method. All the resulting angle distributions were then again analyzed with all the methods; the results are reported in Table 2.3.

**Table 2.3:** Consistency test of the CG parameters obtained with different procedures. Each line of the table corresponds to a different  $P(\theta)$  obtained from a CG simulation done using the parameter listed in the first column.  $\theta_0$  is expressed in degrees and  $k_a$  in K/rad<sup>2</sup>.

Input		Method 1		Method 2		Method 3	
$\theta_0$	$k_a$	$\theta_0$	$k_a$	$\theta_0$	$k_a$	$\theta_0$	$k_a$
180	3500	179	3600	178	3800	178	3800
180	33600	176	$6.5 \cdot 10^8$	178	$6.5 \cdot 10^8$	180	$2.2 \cdot 10^5$
180	28700	177	$2.6 \cdot 10^8$	178	$6.5 \cdot 10^8$	177	$3.5 \cdot 10^6$

From the results it is clear that the first method is the only one which is working properly

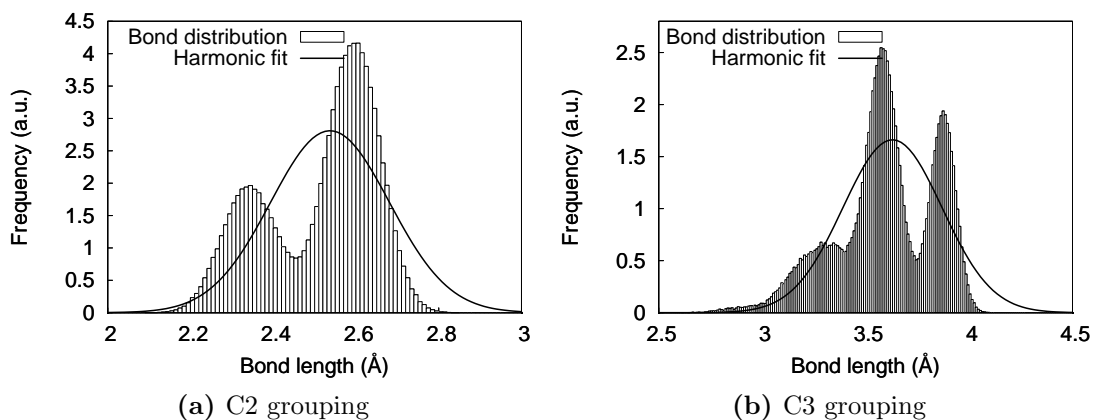
in this case. Moreover, by choosing this method there is no need for further refinement because the potential is self consistent: the small discrepancy between the input and the obtained  $k_a$  is within the error.

The most likely reason for these results is that the numerical convergence is ill conditioned, and the last two methods start too far from the solution. In fact when starting from a better value also method two and three give almost the same solution, as shown in the first line of Table 2.3.

Similar tests were carried for the distance distribution, although the differences were not so marked. Following the results of these tests the parameters for the bonded interactions were generated by least square fitting of

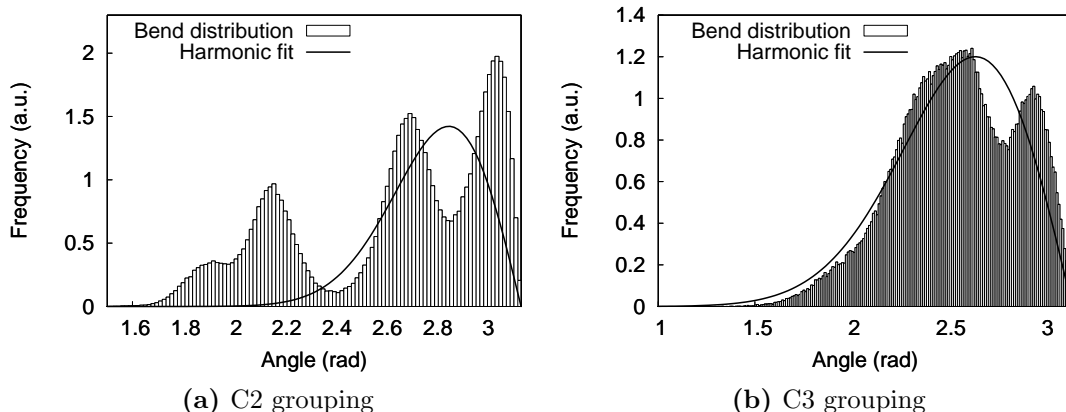
- $P(r)$  with  $y(x) = a_2 \cdot r^2 \cdot \exp(-a_1 \cdot (x - a_0)^2)$
- $P(\theta)$  with  $y(x) = a_2 \cdot \sin(x) \cdot \exp(-a_1 \cdot (x - a_0)^2)$

The parameters of all the fittings are collected the Appendix A. As an example typical AA distributions and their fit are presented in Figures 2.1 and 2.2.



**Figure 2.1:** Dodecane bond distribution and its fit for different CG grouping.

From these plots it is evident that the harmonic potential can give a good approximation yet still fail to completely reproduce all the features present in the AA distributions. The biggest deviation is represented by the C2 angle potential (Figure 2.2a) where the configurations with an angle smaller than 2.3 rad (about  $130^\circ$ ) are not reproduced in the CG model. Despite these limitations however, this model is still able to capture important features of the polyethylene chain, as shown in the next chapter (Section 3.1).



**Figure 2.2:** Dodecane angle distribution and its fit for different CG grouping.

It is interesting to note that, contrary to common practice, in this force field there is no distinction between terminal and central alkane beads, *e.g.*, between  $\text{CH}_3\text{-CH}_2\text{-}$  and  $\text{-CH}_2\text{-CH}_2\text{-}$ . The reason is that the analysis of the bond and bend distribution do not show any systematic discrepancy between the end and the rest of the chain. This is likely to be related to the simple functional form of the interaction, although the minor differences present in the AA distributions are likely to be caught with a more detailed model, *e.g.*, tabulated potentials.

## 2.3 Iterative Boltzmann Inversion

The choice of the model for non-bonded interaction is very important in the parametrization of systems that shows a complex phase behavior. Therefore, in the effort to correctly reproduce the key properties of the swollen membrane, different approaches were used to account for non-bonded interactions. It has been already stressed that in the process of building a CG force field, a crucial point is to incorporate into the CG model as much of the neglected atomistic and electronic degrees of freedom as possible. On the other hand, classical simulations cannot fully reproduce the physical properties of a system, due to the simplification in the treatment of multi-body interactions and due to the dimensional reduction [52]. However, as long as only a limited set of properties is of interest, good approximation is obtained by using effective site-site pair-potentials, which average out many-body interactions. It is therefore crucial to find a balance between these two opposing factors. This is accomplished by carefully selecting the properties that should be correctly

reproduced in the new representation.

A widely used option is to reproduce experimental or computed values of macroscopic properties that are considered relevant for the case under study. In general however, this approach is limited by the fact that many details about the local structure are lost and that only simple analytical functions can be used to model the interactions, so that a limited set of parameters has to be optimized. While in general a simple analytic potential can give a good representation of the underlying interaction, it is unable to fully capture many details of the original system. Therefore in the present work many interactions were modeled by using an iterative Boltzmann inversion (IBI) [53] method, which has been widely proven to work extremely well [39, 54].

### 2.3.1 Theory

In the IBI method an effective two-body potential is obtained by iteratively modifying a tentative potential until it is able to reproduce a target radial distribution function (rdf),  $g_{tar}$ . While in principle experimentally obtained rdfs should be used, the structural distributions of interest are often not available or hard and expensive to measure with sufficient accuracy. Therefore it is a common practice to obtain them from *ab-initio* or classical simulations.

Once  $g_{tar}$  and the initial potential ( $u_0$ ) are chosen a series of simulations are carried out, where after the potential is modified in a stepwise fashion by adding a correction term  $\Delta u_i$ . The most common expression for the iterative correction is

$$u_{i+1} = u_i + \Delta u_i = u_i + k_B T \ln \left( \frac{g_i}{g_{tar}} \right) \quad (2.5)$$

It is important to note that this approach is simply a convenient numerical algorithm [55] useful to find the unique [56] potential associated with  $g_{tar}$ , which is also a stationary point of Equation 2.5. Moreover, the Henderson theorem, which uniquely links a rdf with a two-body potential, is not valid when modeling bonded beads, because the spherical symmetry of the potential is broken by the orientational correlation introduced by bonded interactions. Therefore in this work when dealing with such cases, the effective potential was considered reliable if obtained from two different initial guesses, normally the potential of mean force (PMF) and a LJ function (see below for details).

The main limitation of this method is often identified as the number of iterations needed to reach the final potential, *i.e.*, convergence speed, which is not easily predicted

or controlled. Also the numerical stability of the method is non-trivial due to potentially diverging quantities (*e.g.*, the logarithm of the ratio of rdfs at short range) and the statistical and numerical errors. A systematic solution of these problems was never published, although many different partial efforts can be found in the literature.

### Starting potential

The most common choice for  $u_0$  is the PMF of the system, which is simply related to the target rdf,  $u_0 = -k_B T \ln(g_{tar})$ . It should be noted that this potential is not a stationary solution of equation (2.5), except for infinitely diluted systems, because  $-k_B T \ln(g)$  is effectively a free energy and not a potential.

However, many other choices are possible. As an alternative a simple analytical function (*e.g.*, LJ-like) which resembles the shape and the position of the first well of the PMF, was used here. This choice does not suffer from numerical noise that derives from the statistical error accumulated when computing the rdfs, and can converge faster due to the absence of density related multiple peaks present in the PMF.

### Convergence

In an attempt to accelerate the convergence of the IBI method and obtain smooth pressure-corrected potentials, the original procedure introduced by Soper [53] was slightly modified by summing at each iteration an additional term. The new term is built starting from the consideration that at each iteration the exact and unknown interaction potential,  $u(r)$ , can be rewritten as the sum of three terms: the potential of the  $i$ -th iteration,  $u_i(r)$ , the potential correction computed after that iteration,  $\Delta u_i(r)$ , and an extra term,  $\epsilon_i(r)$ , which contains all the other contributions. Therefore Equation (2.5) becomes:

$$u_{i+1} = u_i + \Delta u_i + \epsilon_i \tag{2.6}$$

The extra term can be added as long as it becomes negligible for  $g_i \rightarrow g_{tar}$ , as in that limit Equation (2.6) still can have a stationary solution.

To evaluate  $\epsilon_i(r)$  an extra constraint was added by forcing the CG system to have the experimental density. In the IBI procedure the correct density can be imposed in different ways, *e.g.*, using alternating steps of potential and pressure correction [57] or changed *a posteriori* [58] by taking advantage of the fact that Equation (2.5) can be satisfied by different potential within the numerical and statistical error of the exact potential. Here

instead, at each step a pressure  $P$  of 1 atm was imposed to the system. The goal of this additional constraint is to remove the post-processing step and also to reduce the number of iterations needed by improving the convergence of the attractive part of the potential.

In a classical simulations  $P$  can be written as the sum of two terms  $P = P_{ideal} + P(u(r))$ , *i.e.*, the pressure of an ideal gas ( $P_{ideal} = \rho k_B T$ ) plus the pressure due to two body interactions:

$$P(u(r)) = -\frac{2}{3}\pi\rho^2 \int_0^\infty r^3 g(r) \frac{d}{dr} u(r) dr \quad (2.7)$$

where  $\rho$  is the density of the system. By substituting Equation (2.6), Equation (2.7) becomes

$$P(u) \simeq P(u_{i+1}) = P_{ideal} + P(u_i) + P(\Delta u_i) + P(\epsilon_i) \quad (2.8)$$

Therefore, if the system is constrained to a certain pressure, it is possible to calculate  $P(\epsilon_i)$  at each iteration and from this the extra correction to the potential,  $\epsilon_i$ . In order to do so,  $\epsilon_i(r)$  has to be defined. There is a lot of freedom in this choice, and the limitations arise mainly from the integrability of the function in  $[0, \infty)$ , but supplementary properties are desirable, such as

- $\epsilon_i(r)$  is at least  $C^2$  in  $[0, r_c]$
- $\epsilon_i(r)$  and at least its first and second derivative, are zero at  $r_c$ . This avoids any discontinuities at the cutoff and also justifies the absence of contribution to tail corrections.
- $\epsilon_i(r)$  has a simple multiplicative dependence on at least one parameter to simplify the evaluation of  $\epsilon_i$  from  $P(\epsilon_i)$ .

Although different choices are possible [59]  $\epsilon_i$  was defined as:

$$\epsilon(r) = \begin{cases} a(x-c) \frac{1}{r_m-c} \exp\left[\frac{r_c-r_m}{r_m-c}\right] \exp\left[\frac{(r_m-r_c)^2}{(r_m-c)(x-r_c)}\right] & \text{if } r \leq r_c \\ 0 & \text{if } r > r_c \end{cases} \quad (2.9)$$

This function resembles a LJ function, but its extremal point ( $\epsilon'(r_m) = 0$ ) and node ( $\epsilon(c) = 0$ ) are independent, allowing more flexibility; it is  $C^\infty$  and is zero at  $r_c$  with all its derivatives. By applying equation (2.7) and (2.9) in equation (2.8),  $\epsilon_i(r)$  is obtained from

$$a = \frac{P_{target} - P_i - P(\Delta u)}{P(\tilde{\epsilon})} \quad (2.10)$$



where  $P_i$  ( $= P_{ideal} + P(u_i(r))$ ) is the average equilibrium pressure of the system at the  $i$ -th iteration and  $\epsilon = a \tilde{\epsilon}$ . It is worth nothing that  $g(r)$  is approximated with  $g_i(r)$  for the purpose of computing both the pressure corrections  $P(\Delta u)$  and  $P(\tilde{\epsilon})$ . However, besides the fact that the pressure of the next iteration is not exactly the target value, this approximation has no consequence on the procedure, because the difference becomes smaller with the convergence of the potential.

To evaluate the effectiveness of any modification to the standard IBI procedure, a measure of the convergence is needed, *i.e.*, a way to determine the “distance” to the exact solution. To evaluate this distance, the difference between the actual and the target rdf, is considered:

$$\chi_s^2 = \frac{1}{n} \sum_{j=1}^n (g_i(r_j) - g_{tar}(r_j))^2 \quad (2.11)$$

where  $i$  denotes the iteration, and  $n$  denotes the number of bins of the rdf.  $\chi_s^2$  however is not always a good measure of the convergence because in the calculation of the rdf the accuracy of each bin can vary considerably. Therefore, to ensure that regions of high uncertainty are not over-influencing the convergence, a second parameter is defined, which weights each contribution by the respective variance ( $\sigma^2$ ):

$$\chi^2 = \frac{1}{n} \sum_{j=1}^n \frac{(g_i(r_j) - g_{tar}(r_j))^2}{\sigma_i^2(r_j)} \quad (2.12)$$

### Numerical and statistical errors

As already noted by Soper [53], the numerical noise can be problematic both at short range, when the rdf approaches zero, and near  $r_c$ . The reason for short range error is twofold. First, the statistical uncertainty of the rdf is of the order of  $1/\sqrt{4\pi\rho N r^2 \Delta r}$ , where  $N$  is the number of the particles in the simulation and  $\Delta r$  is the bin width. Second, the potential correction depends on  $\ln(g_i/g_{tar})$ , which is a numerically ill conditioned quantity (it oscillates rapidly if the two rdfs do not go to zero at exactly the same speed). On the other hand, at long range the statistical and numerical error has heavy impact on the pressure correction term, which depends on  $r^3 du(r)/dr$ , making the procedure extremely sensitive to numerical noise of  $u(r)$ .

Hence it is clear that steps have to be taken to avoid that the error grows in amplitude with each iteration of the algorithm, eventually making the estimated potential useless. The treatment of the error, however, is not simple, because any procedure is equivalent to

introduction in Equation (2.5) of a weighting factor for  $\Delta u_i$  that depends both on  $r$  and  $i$ , potentially hindering the convergence. Particular care should be taken in the modification of the potential at short distance, as long range convergence can only be achieved after the repulsive part has already converged. On the other hand, the short range interaction changes fast and has bigger errors, making this part of the curve more sensitive to simple smoothing techniques such as running averages, which should therefore be avoided.

After trying different approaches, three were extensively tested in this work:

**Smoothing with Akima cubic splines [60].** With this technique  $\Delta u$  is interpolated with a piecewise curve built from polynomials of degree three or less. It differs with respect to the standard cubic spline by the conditions imposed at the data points; the result is that it guarantees a better stability of the curve near the outliers, removing artificial oscillation typical of cubic splines in these conditions.

**Smoothing with Bézier cubic splines. [61]** This technique consists in a splining algorithm that uses Bézier curves, *i.e.*, linear combination of Bernstein basis polynomials. With respect to the Akima cubic splines, this procedure gives much smoother curves at the price of losing more details, as the curve is not forced to pass through the data points.

**Windowing.** The correction terms are multiplied by two Fermi-like functions ( $1/(1 + e^{c_0(x-c_1)})$ ), so that it smoothly goes to zero at the border of the interval of interest. This technique was used in conjunction with the previous ones and is particularly important to avoid discontinuities at the cutoff.

## 2.3.2 Results

In principle both the pressure correction and the smoothing procedures can seriously limit the converging speed to the exact solution. Therefore before applying the procedure to compute the CG interactions needed for the force field, some tests were conducted to check and optimize the algorithm.

### Methodology optimization

Preliminary tests were carried on benzene and hexane (C3 grouping) boxes. The parameters of some of the systems used to test the methods are listed in Table 2.4.

The target rdfs were obtained from AA MD NVT simulations of 4 ns at the experimental density. Each iteration involved 1 ns long simulation of the equivalent CG system in a canonical ensemble. The values of the average pressure,  $\chi_s^2$  and  $\chi^2$  after 30 iterations are reported in Table 2.5, together with the total energy drift associated with the actual potential. This was measured by doing a 1 ns CG microcanonical simulation and by doing a linear fit of the total energy. For comparison a system with analytical LJ potential for the intermolecular interaction was also simulated.

**Table 2.4:** Details on the different systems used to test the effect of different corrections and error treatments on the convergence of the IBI method. In the first three columns information about the species, the total number of molecules and the size of the simulation box are given. After that is listed if the pressure correction was used and the shape of the starting potential: LJ or mean force potential (MFP). The smoothing column lists the techniques used to reduce numerical and statistical error of  $\Delta u_i$ : ACS refers to Akima cubic spline fitting, BS to Bézier spline fitting, W to the windowing procedure. *blj* and *hlj* are not actual IBI system but reference NVE runs done with analytical LJ potential to set a reference energy drift.

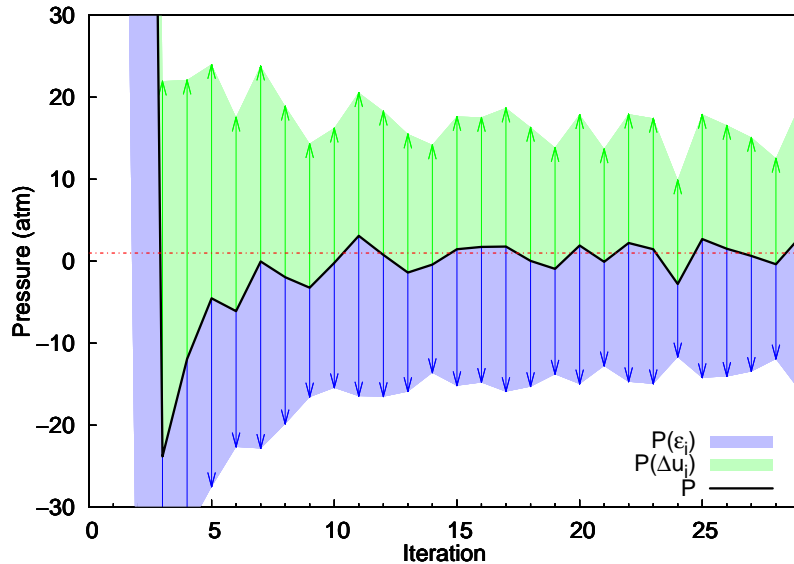
Label	Species	Molecules	L(Å)	$P_{corr}$	$\Delta u_0$	Smoothing
b1	B	1293	58.26	no	MFP	ACS
b2	B	1293	58.26	yes	MFP	W + ACS
b3	B	1293	58.26	yes	LJ	W + ACS
b4	B	1293	58.26	no	MFP	BS
b5	B	1293	58.26	no	MFP	W + BS
b6	B	1293	58.26	yes	MFP	W + BS
b7	B	1293	58.26	no	LJ	W + BS
<i>blj</i>	B	1293	58.26	-	-	-
h1	C3-C3	1169	64.24	no	MFP	W + BS
h2	C3-C3	1169	64.24	yes	MFP	W + BS
h3	C3-C3	1169	64.24	no	LJ	W + BS
h4	C3-C3	1169	64.24	yes	LJ	W + BS
h5	C3-C3	1169	64.24	no	MFP	W + ACS
<i>hlj</i>	C3-C3	1169	64.24	-	-	-

From the table the first evident result is that adding the pressure correction term does not impact the convergence speed, while on the other side it guarantees the pressure to be correct.

To better understand this behavior, the evolution of the pressure was analyzed (Figure 2.3), showing that the pressure converges after a few iterations, despite the fact that

**Table 2.5:** Pressure, convergence and energy drift for different iterative corrections and error treatments for the IBI procedure. The values reported refer to the potential and radial distribution function after 30 iterations. Details on the systems are reported in Table 2.4

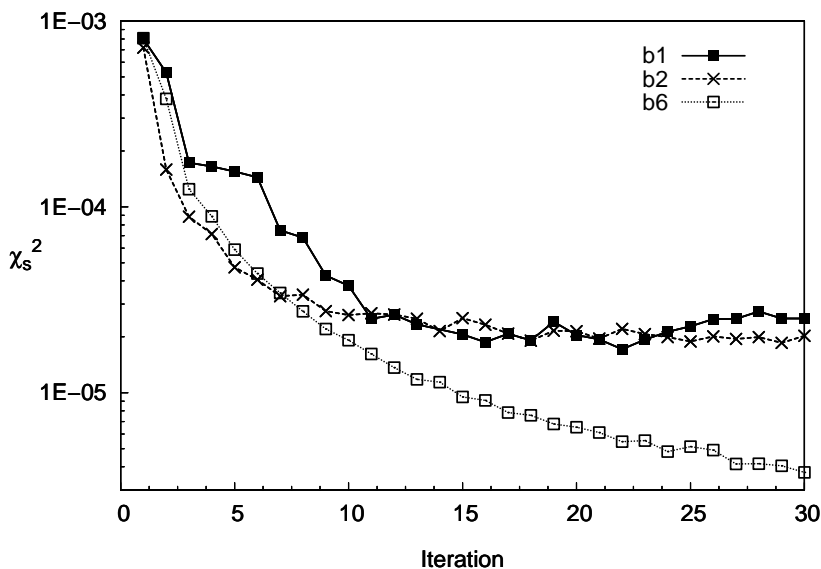
Simulation	P (atm)	$\chi_s^2(10^{-5})$	$\chi^2(10^{-3})$	Drift (K/atom ns)
b1	1300	1.43	1.66	$15.0 \pm 0.2$
b2	16	1.40	3.60	$57.7 \pm 0.2$
b3	6	0.22	1.60	$23.7 \pm 0.2$
b4	1300	0.35	0.33	$4.6 \pm 0.2$
b5	1460	0.36	0.51	$2.7 \pm 0.2$
b6	1	0.37	0.51	$5.7 \pm 0.2$
b7	500	6.53	2.84	$0.08 \pm 0.2$
<i>blj</i>	-	-	-	$10^{-4} \pm 0.5 \cdot 10^{-4}$
h1	1550	0.04	0.2	$1.6 \pm 0.1$
h2	3	0.06	0.2	$3.6 \pm 0.1$
h3	800	0.15	3.6	$0.2 \pm 0.1$
h4	1	0.13	2.1	$-0.3 \pm 0.1$
h5	820	0.06	0.4	$10.0 \pm 0.1$
<i>hlj</i>	-	-	-	$10^{-3} \pm 0.2 \cdot 10^{-3}$



**Figure 2.3:** Evolution of the average pressure of a box of CG benzene hexane (system h2 in Table 2.4) with the iterations. The arrows show the direction and the magnitude of the pressure modifications due to the different correction terms, while the red line corresponds to 1 atm.

the system was started far from equilibrium (about 1200 atm). After 10 iterations the pressure oscillates around the target value and the pressure correction changes the potential less than 1 K at each iteration, which is at least one order of magnitude less than the correction due to  $\Delta u_i$ . Therefore it is not surprising that after 30 iterations the convergence is completely dominated by the choice of the smoothing functions of  $\Delta u_i$  and not by the pressure correction. Furthermore the fact that after a few iterations  $\epsilon_i(r)$  becomes of the same order of magnitude of the error, allows casting the algorithm again in the form of equation 2.5. Hence, considering the advantages and the negligible computational cost of the pressure correction term, it was concluded to include it in the parametrization of the CG force field.

The influence on convergence of the other parameters, like the starting potential or the smoothing techniques, is not trivial. The tests show that Bézier splining gives smaller drifts and, in general, a better convergence (see also Figure 2.4).



**Figure 2.4:** Plot of  $\chi_s^2$  for different optimization methods of the IBI applied to B~B intramolecular interaction. The details of the systems are explained in Table 2.4.

On the other hand the choice of  $u_0$  gives different results depending on the smoothing techniques. In particular, Bézier splining does not work very well when starting from LJ potentials, resulting in an order of magnitude bigger values of  $\chi_s^2$  and  $\chi^2$  after 30 iterations, even though smoother functions are obtained, which reflects in a smaller drift. It should be noted that the exact values for the drift need to be taken only as an indication of trend

because they depend on many other parameters (*e.g.*, the time step), and on details of the practical implementations of MD algorithms as well.

### Force field modeling

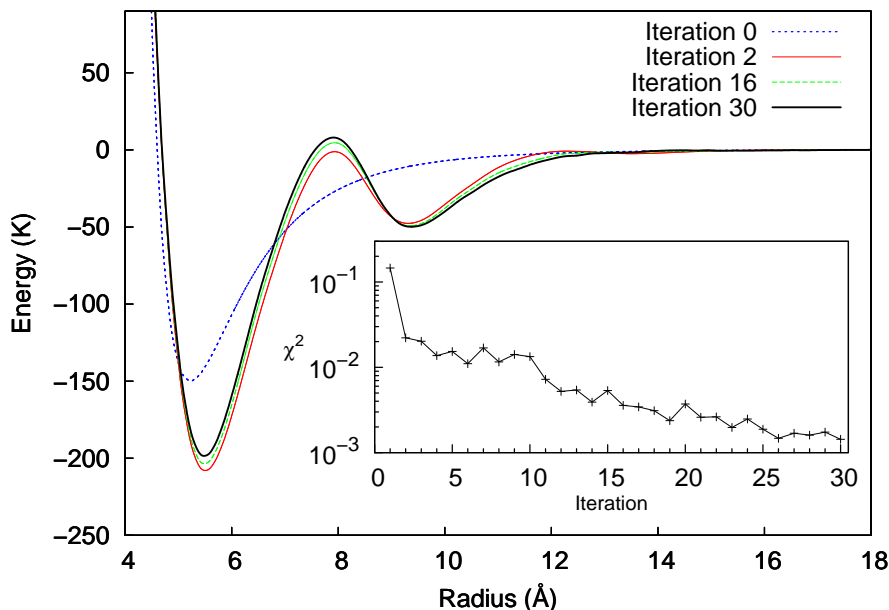
The results collected in the previous section reveal that a careful treatment of the numerical problems and the introduction of an extra pressure correction term, leads to a procedure which gives good potential with a few dozens iterations. Therefore for the purpose of modeling the force field interactions, the correction at each iteration was smoothed with Bézier curves and windowing techniques and the additional pressure correction term was used.

The general computational details are the same mentioned in the previous section: the target rdfs were obtained through AA NVT MD simulations 4 to 6 ns long, while the CG system was simulated on the same ensemble for 1 ns at each iteration. Further details on the systems can be found in Table 2.6.

**Table 2.6:** System used to model non-bonded interactions with the IBI method. The  $\chi_s^2$ ,  $\chi^2$  and total energy drift refer to the final potential. The drift is expressed in K/(atom ns).

Interaction	Species	Molecules	L(Å)	$\chi_s^2(10^{-6})$	$\chi^2(10^{-4})$	Drift
B~B	B	1293	58.26	1.0	3.2	0.024
B~C2	B-C2	1197	63.08	1.1	2.6	-0.312
B~C3	B	209	40.0	7.2	16.0	0.171
	C3	141				
B~M	B	209	40.0	21.7	9.8	1.79
	M	455				
C2~C2	C2-C2-C2	1169	64.24	0.22	2.3	-0.35
C2~C3	C2-C2-C2	584	64.22	2.2	2.4	0.007
	C3-C3	584				
C3~C3	C3-C3	1169	64.24	2.1	14.3	0.130

As shown in this table, IBI method was not only used for pure systems, but also for miscible solvents. In this case the mixtures were built by mixing the same volume of the two species considered at the experimental density at 50 °C.



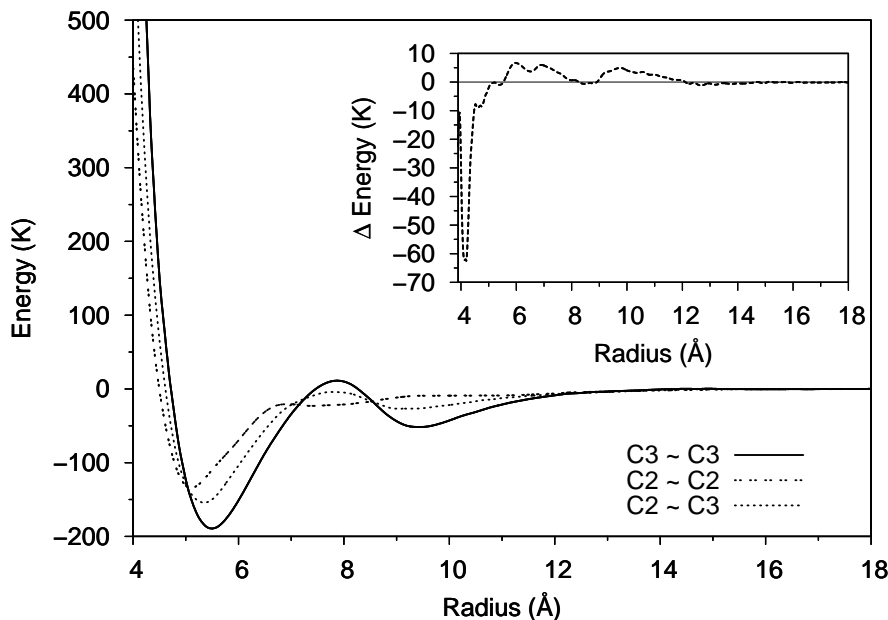
**Figure 2.5:** C3~C3 intermolecular potential for CG hexane for different iterations. In the inset the trend of the  $\chi^2$  corresponding to each iteration is plotted.

The main problem when dealing with mixture is that a certain degree of inhomogeneity is always present. Even in perfectly miscible solvents, such as water and methanol, at the atomic level some degree of clustering between molecules of the same specie can be observed. However, as long as these clusters are small and evenly distributed in the solvent, the rdf is still meaningful and the IBI method converges to a potential, which on average reproduces the same level of aggregation. Even so the applicability of IBI method to such types of systems is not obvious, and although the convergence for benzene-methanol and benzene-hexane mixture was good, for other systems which form larger clusters (*e.g.*, methanol-hexane) other methods had to be used (see Section 2.4).

As an example the convergence of C3~C3 potential is reported in Figure 2.5. From the analysis of the potential plots is clear that already after a few iterations the potential is near to the final function. However the  $\chi^2$  plot suggests that a few more iterations are needed to achieve a better agreement.

The final converged potentials used in this thesis are shown in Figures 2.6 and 2.7. In particular is extremely interesting to note that the C2~C3 interaction is very similar to the arithmetic average of the potential for C2~C2 and C3~C3, as shown in the inset of Figure 2.6. The small relative difference suggests that, at least in a few cases, mixing

rules for the IBI potentials are possible. However it is not known at this time if this is a coincidence or if there are deeper reasons for this.



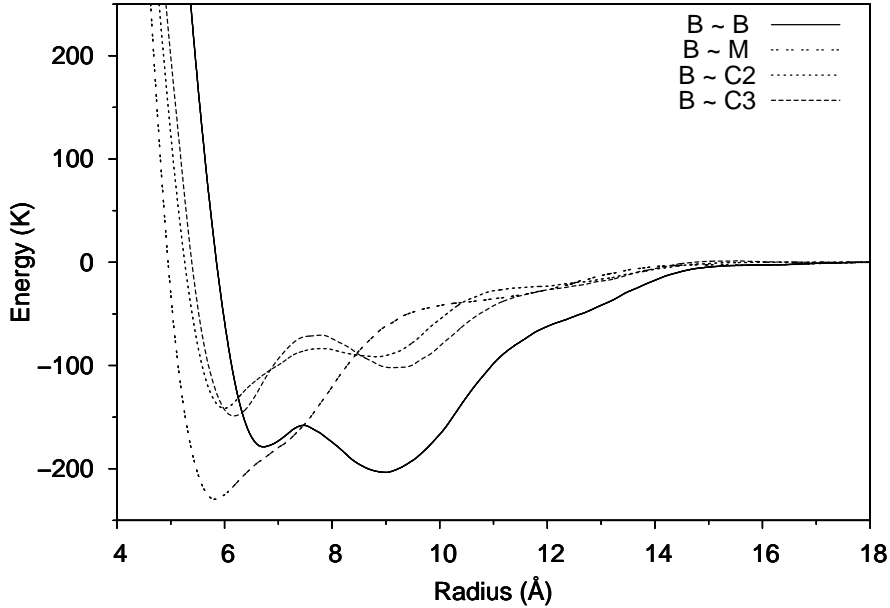
**Figure 2.6:** Effective potentials for the interaction between CG beads composing alkane molecules. In the inset the difference between the C2~C3 potential and the arithmetic average of C3~C3 and C2~C2 is plotted.

## 2.4 Fit of macroscopic properties

While the application of the IBI method gives very good results and it is useful to keep some information of local atomistic configurations in the CG representation, its applicability relies on the possibility of computing the rdf for the species of interest. It is therefore problematic to model segregated or anisotropic systems in this way, *e.g.*, like liquid hexane-water mixtures which phase separate. Also, the lack [62] of a systematic way to combine the tabulated potentials obtained through IBI can be a serious limitation. These type of rules are very useful when studying a system in different mixture of solvents, because they remove the need of recomputing solvent interactions for each concentration.

Therefore when dealing with such cases, following the example of Nielsen *et al.* [63] and Shinoda *et al.* [64], the density and surface or interfacial tension data (either from experiments or from AA simulations) were used as target properties to model the CG





**Figure 2.7:** Effective potentials that model the CG interaction of benzene beads with other CG beads.

interactions. The functional form of the effective pair-potential used in these fitting was:

$$U_{LJ}(r) = c \cdot \varepsilon \left[ \left( \frac{\sigma}{r} \right)^n - \left( \frac{\sigma}{r} \right)^m \right] \quad \text{with } c = \frac{n}{n-m} \cdot \left( \frac{n}{m} \right)^{\frac{m}{n-m}} \quad (2.13)$$

where  $\sigma$  and  $\varepsilon$  are defined such that  $U(\sigma) = 0$  and  $U_{min} = -\varepsilon$ . For the parameters  $n$  and  $m$ , the pairs 9-6 (LJ9-6), 12-4 (LJ12-4) and 12-6 (LJ12-6) were used. The choice of both density and surface/interface tension is of key importance in correctly reproducing aggregation and morphology of polyelectrolytes and surfactants. Moreover these properties form a complementary target in the optimization of the LJ parameters, since the density mostly depends on  $\sigma$ , which determines the size of the particles, while the interfacial tension is more dependent on  $\varepsilon$ , which sets the potential energy scale.

The density ( $\rho$ ) is straightforwardly computed from the isothermal-isobaric ensemble (NPT) simulations. The isothermal compressibility ( $\beta_T$ ) is obtained from the volume fluctuations of the box simulated in the NPT ensemble according to the following formula [65]:

$$\beta_T = \frac{\langle \delta V^2 \rangle}{k_B T V} \quad (2.14)$$

where  $k_B$  is the Boltzmann constant,  $T$  is the temperature,  $V$  the average volume and  $\langle \delta V^2 \rangle$  is ensemble average of the mean square deviation of the volume fluctuations. Surface and interfacial tension ( $\gamma$ ) are computed from canonical simulations of two-dimensional slabs of liquid oriented along the xy plane using [66]:

$$\gamma = \frac{L_z}{2} \left\{ P_{zz} - \frac{P_{xx} + P_{yy}}{2} \right\} \quad (2.15)$$

where  $L_z$  is the size of the periodic box along the z direction and  $P_{ii}$  (with  $i = x, y, z$ ) are the diagonal terms of the stress tensor. To obtain the interfacial tension in the periodic box two slabs of different liquids were placed in contact, while for surface tension the slab was simulated in vacuum.

The details of the fitting slightly differ for pure solvents, namely water and methanol, and non-miscible species as explained in the next paragraphs.

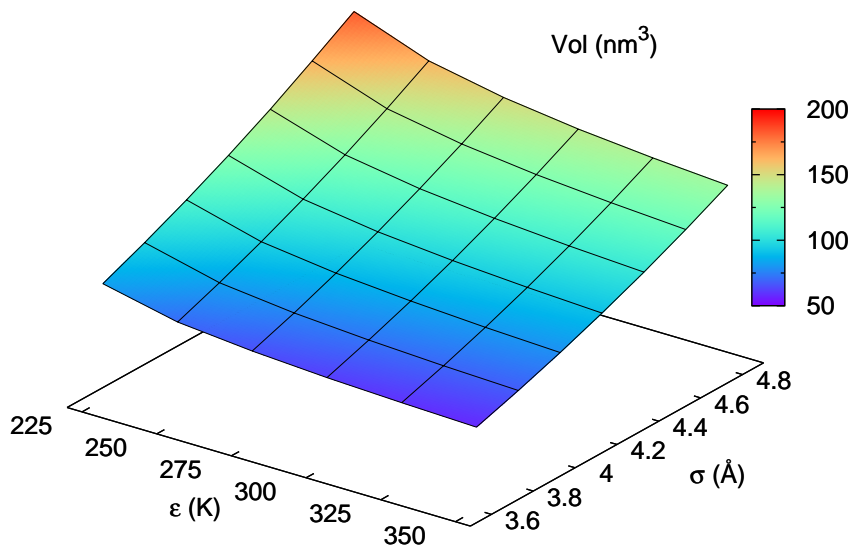
### 2.4.1 Pure liquids

The simultaneous optimization of two (or more) correlated parameters can be very complex and time consuming. Hence to simplify the search of the parameters first  $\rho$ ,  $\beta_T$  and  $\gamma$  were computed on a wide range of different values of  $\varepsilon$  and  $\sigma$ . In order to compute  $\rho$  and  $\beta_T$ , a box containing 1000 LJ particles with periodic boundary conditions was run for 4 ns in an NPT ensemble. For  $\gamma$ , slabs generated by replicating z-elongated box containing 2000 units, were simulated for 10 ns in a canonical ensemble.

From these surfaces (see Figures 2.8, 2.9 and 2.10) the isolines corresponding to the experimental values of each thermodynamic properties were interpolated, and the point of intersection between the isolines corresponding to the target  $\rho$  and  $\gamma$  was taken as starting point for further refinement. One advantage of this procedure is that once the surfaces corresponding to a macroscopic properties are built, they can be used for any solvent and for any choice of target properties.

### 2.4.2 Non-miscible species

The methods outlined above cannot be readily applied to obtain the interaction parameters of non-miscible species. As mention before, the IBI method has to be excluded because the rdfs cannot be computed, while thermodynamic fitting, as applied to pure water and methanol suffers from a scarcity of experimental data at the condition of interest and of

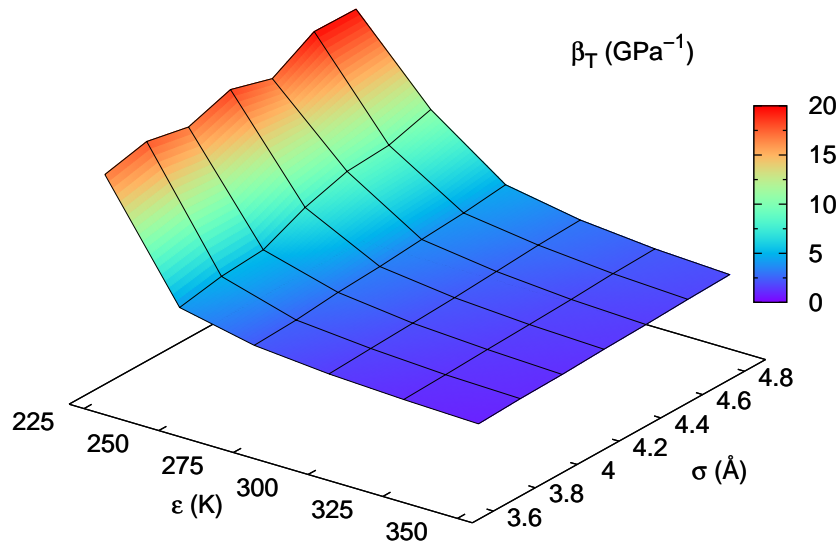


**Figure 2.8:** Average volume of a box of 1000 Lennard-Jones particles (NPT ensemble:  $T = 325$  K,  $P = 1$  atm, LJ9-6 potential) as function of  $\epsilon$  and  $\sigma$ .

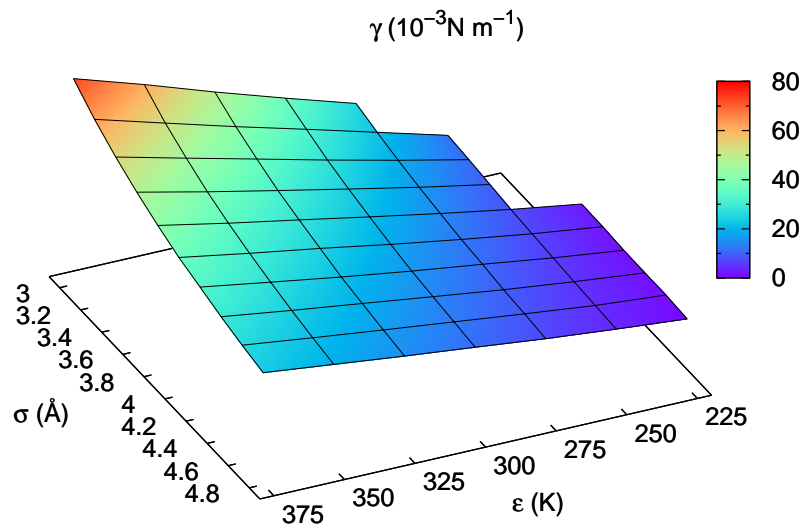
the impossibility to use density as a target property.

Therefore to compute the interactions between alkanes and water, or alkanes and methanol, a slightly different approach was used. In particular the interfacial tension and interfacial separation derived from AA simulations were used as target properties. The systems were built as described before, by simulating alternate slabs of the same thickness (about  $40 \text{ \AA}$ ) of two non-miscible species in an NVT ensemble for 5 ns. From these simulations the value of  $\sigma$  parameter of the CG LJ potential, was determined by taking the average distance between the two species at the interface.<sup>1</sup> Once determined  $\sigma$ ,  $\epsilon$  was obtained by imposing to the CG system to reproduce the value of AA  $\gamma$ . In Figure 2.11 some examples of the optimization of  $\epsilon$  are presented.

<sup>1</sup>Due to the lack of a generally accepted definition of interface, to the purpose of the average, only the atoms that are nearer than the first inflection point after the first maximum in the rdf, were considered.

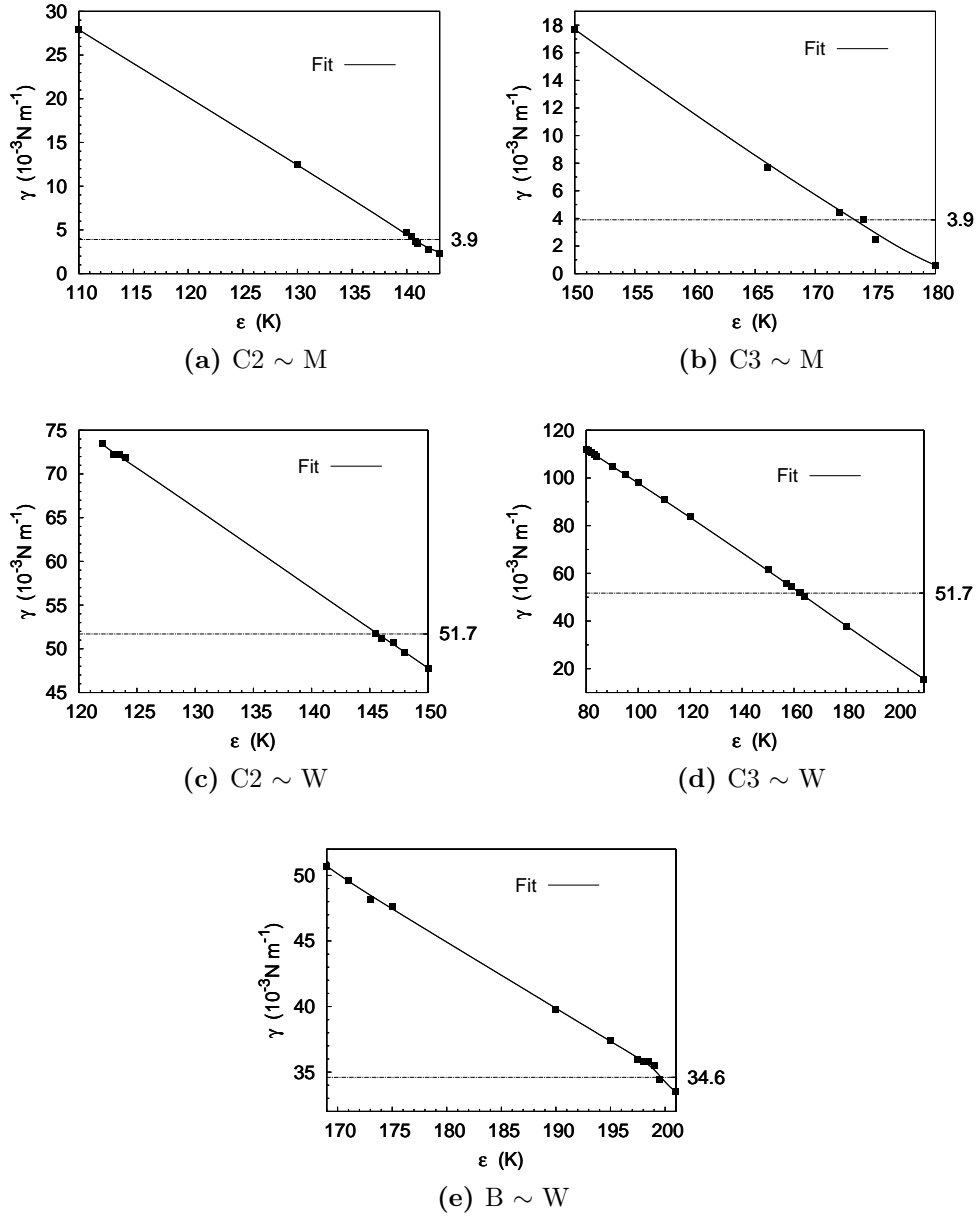


**Figure 2.9:** Isothermal compressibility of a box of 1000 Lennard-Jones particles (NPT ensemble:  $T = 325$  K,  $P = 1$  atm, LJ9-6 potential) as function of  $\epsilon$  and  $\sigma$ .



**Figure 2.10:** Surface tension of a slab of 2000 Lennard-Jones particles (canonical ensemble:  $T = 325$  K, LJ9-6 potential) as function of  $\epsilon$  and  $\sigma$ .

## FORCE FIELD PARAMETRIZATION



**Figure 2.11:** Plot of the interfacial tensions for CG species as function of  $\epsilon$  (fixed  $\sigma$ ). The dashed line in each plot corresponds to the target value set from AA simulations.

## 2.5 Charged species

The previous sections presented several methods to parametrize the CG interactions of neutral molecules and fragments. However many interesting systems (*e.g.*, protein, surfactant, membranes) also contain charged species, which need to be treated in a different way. This is because lumping different atoms together into a single site gives to the CG bead a partial charge, which is the sum of all its constituent atomic charges removing the effect of multipole interactions. For some non-bonded interactions this can be satisfactory, or the issue can be avoided by choosing a different mapping between the AA and the CG representations. For the membrane, however, this is not a viable options and other methods must be used to take into account the electrostatic effects.

### 2.5.1 Charge-charge interactions

The interaction between charged groups can be split in two contributions: electrostatic and van der Waals interactions. Electrostatic interactions need a special treatment with respect to the AA system because the dimensional reduction removes the effect of shielding due to partial charges present in neutral molecules. Thus here it is assumed that the interacting ions, of charge  $q_1$  and  $q_2$ , are dispersed in medium with an isotropic uniform dielectric constant  $\epsilon_r$ . Under this assumption, the electrostatic potential ( $U_{el}$ ) is

$$U_{el}(r) = \frac{q_1 q_2}{\epsilon_r r} = \frac{q_1^{eff} q_2^{eff}}{r} \quad (2.16)$$

where in the rightmost member of the equation the effective charge is defined as  $q_i^{eff} \equiv q_i \epsilon_r^{-1/2}$ . In other words, the charges are scaled according to the square root of an estimate of the average dielectric constant of the medium, which depends on its composition, structure and temperature.

This approximation can appears quite drastic. However, it was successfully employed in other works [36, 41, 64]. This can be understood considering that the quality of this approximation strongly depends both on the homogeneity with respect to  $\epsilon_r$ , and on the ionic strength of the system, which is related to the Debye length (the length over which mobile charge carriers screen out the electric field). Therefore in systems with high ionic charge (*e.g.*, sulfonated membranes, polyelectrolytes and surfactants) the scale of the inhomogeneity is larger than the Debye length. Hence under such conditions (which are common in the system studied in this thesis) this approximation is justified.

The VdW interactions between charged groups, *i.e.*, Na~Na, Na~S and S~S, were not parametrized but obtained either from the literature or using standard mixing rules. The rationale behind this choice is the fact that even approximate potentials are satisfactory because their contribution is small compared to the Coulombic forces acting on the beads. Therefore, to model Na~Na a LJ potential was chosen using the parameters of Smith and Dang [67], while the other interactions were obtained using the Lorentz ( $\sigma_{12} = (\sigma_1 + \sigma_2)/2$ ) Berthelot ( $\varepsilon_{12} = (\varepsilon_1\varepsilon_2)^{1/2}$ ) combining rules [68] (S~S, from S~W and W~W potentials, and Na~S from Na~Na and S~S potentials).

## 2.5.2 Free energy perturbation

The last group that has to be parametrized is the set of interactions between charged and overall neutral molecules. These potentials have to be modeled in a solvated environment and not in bulk because the electrostatic interactions between the charged groups and the neutral molecules of the solvent have to be recovered. To correctly capture these effects, and therefore include them in the force field, the first attempt was to model the CG interactions using the excess free energy of solvation  $\Delta A_{solv}$  as a target property.

Here  $\Delta A_{solv}$  is defined as the excess Helmholtz free energy needed to move a charged group or atom from vacuum into the solvent (hexane, benzene, methanol or water). To estimate  $\Delta A_{solv}$  the thermodynamic cycle depicted in Figure 2.12 was used, computing each transition with a multistage free energy perturbation method (MFEP) in a dual topology paradigm, as explained in the following sections.

### Theory

Free energy perturbation (FEP) is one of the most commonly used methods for calculating free energy differences in molecular simulations. The Helmholtz free energy associated with the generic transition from an initial ( $I$ ) to a final ( $F$ ) state is obtained from the Zwanzig equation [69]:

$$\Delta A(I \rightarrow F) = -k_B T \ln \left\langle \exp \left( -\frac{\mathcal{H}_F - \mathcal{H}_I}{k_B T} \right) \right\rangle_I \quad (2.17)$$

where  $\mathcal{H}$  is the Hamiltonian and the angle bracket represents a canonical ensemble average performed on the initial state  $I$ . Being interested only in the excess free energy, the kinetic component of the Hamiltonian can be ignored. Therefore the replacement of the  $\Delta\mathcal{H} = \mathcal{H}_F - \mathcal{H}_I$  with the potential energy difference,  $\Delta U = U_F - U_I$  is allowed. This

substitution greatly simplifies the practical implementation of this method when  $U$  has a linear dependence from one of its parameters, like LJ potential with respect to  $\varepsilon$ .

In practice, the right side of equation (2.17) is obtained by running a regular MD simulation with  $\mathcal{H}_I$ . Then at each step, the positions are used to compute the potential energy of both the unperturbed ( $I$ ) and the perturbed ( $F$ ) system, thus giving  $\Delta U$ , which is then averaged over the full trajectory to give the free energy.

Considering that the free energy is a state function, the application of the method in the two opposite directions ( $I \rightarrow F$  or  $F \rightarrow I$ ) should produce identical free energy estimates, except for the sign. However in real simulations these estimates usually differ and have different convergence speed and error. Different techniques were developed to overcome this problem and to reliably compute the free energy [70]. In this work a multistage FEP approach in conjunction with simple overlap sampling (SOS) method was chosen.

The idea behind MFEP is to split a (very) large perturbation into a series of smaller ones, which are suitable to be conducted by single-stage FEP calculation. This is often achieved with a linear parameter scaling approach, so that the Hamiltonian becomes

$$\mathcal{H}(\lambda) = (1 - \lambda)\mathcal{H}_I + \lambda\mathcal{H}_F \quad (2.18)$$

where  $\lambda$  has a value between 0 and 1. By choosing a set of intermediate states, each one corresponding to a single  $\lambda_i$ , the potential perturbation between two subsequent states can be rewritten as:

$$\Delta U = U(\lambda_{i+1}) - U(\lambda_i) \quad (2.19)$$

In other words, in MFEP the intermediate states are chosen so that at each step there is a good overlap between the initial ( $\Gamma_i$ ) and final ( $\Gamma_{i+1}$ ) important phase space<sup>2</sup>. The optimal choice of the intermediate states is usually not straightforward because the overlap between each important phase space is not known. Moreover adding *a posteriori* new intermediate states to the set requires expensive recalculation of the point of the old set and is best avoided. However, for the LJ potential for each  $i$ -th state the potential energy can be written as  $U_i(r) = \varepsilon^{(i)}\tilde{U}(r) + c$ , where  $c$  is a constant. In this case (2.17) simplifies to

$$\Delta A(i \rightarrow i + 1) = -k_B T \ln \left\langle \exp \left( -\beta(\varepsilon^{(i+1)} - \varepsilon^{(i)})\tilde{U} \right) \right\rangle_i \quad (2.20)$$

Hence for the CG system the need of recomputing  $\Delta U$  upon the insertion of new point in

---

<sup>2</sup>The *important phase space* is intuitively defined as the portion of configurations in phase space which give a significant contribution to the partition function.



the intermediate states set, is substituted by a simple scalar product, making it computationally inexpensive.

Once the proper chain of intermediate states connecting  $I$  to  $F$  is chosen, the  $\Delta A$  of each interval is computed with the FEP method. For obvious efficiency reasons the overlap between  $\Gamma_i$  and  $\Gamma_{i+1}$  is not complete. Therefore overlap sampling techniques have been developed to improve the precision and accuracy of the calculated free energy. All these methods make use of an intermediate state  $M$ , whose important phase space is a subset of both  $\Gamma_i$  and  $\Gamma_{i+1}$ .

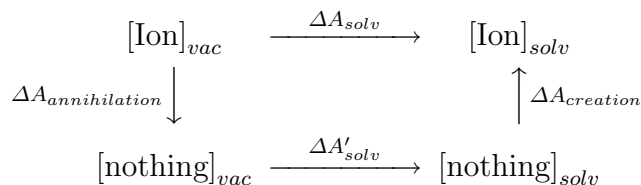
The simplest choice, called simple overlap sampling (SOS), is to select  $M$  as the middle point between  $i$  and  $i + 1$ , which leads to a free energy given as:

$$\Delta A(i \rightarrow i + 1) = (\Delta A(i \rightarrow M) - \Delta A(i + 1 \rightarrow M)) \quad (2.21)$$

Because no sampling on the intermediate sampling  $M$  is required, the calculation effort involved in SOS is identical to the direct averaging ( $\Delta A(i \rightarrow i + 1) = (\Delta A(i \rightarrow i + 1) - \Delta A(i + 1 \rightarrow i))/2$ ), but the reliability of the result can differ substantially [71, 72].

### The system

To compute  $\Delta A_{solv}$  the hypothetical (in the sense that some transformations can be carried out only *in silico*) thermodynamic cycle 2.12 was used.



**Figure 2.12:** Thermodynamic cycle used to compute  $\Delta A_{solv}$ .

Instead of using the direct route, *i.e.*, directly move the ion ( $\text{Na}^+$  or the  $\text{SO}_3^-$  fragment) into the solvent, the alternative route was chosen. It consists in annihilating (or creating) the ion both in vacuum and in the solvent, which corresponds to the vertical lines of the cycle. Using the fact that  $\Delta A'_{solv} = 0$ , the sum of these two contributions gives the desired quantity. Moreover, for all the cases where the ion is a single particle (all the CG simulation and  $\text{Na}^+$  in AA), the calculation of the excess free energy is further simplified by the fact that the creation in vacuum gives no contribution to the total energy.

The calculations were performed in a dual-topology paradigm, *i.e.*, the topology of the initial and final states coexist throughout the simulation, without interacting with each other. While this approach overcomes several complications inherent to the single-topology paradigm (where a common topology describes all the states), it unfortunately suffers from the “end-point catastrophes”. This refers to numerical instabilities near the end of the transformation, in this case when the ion interactions are so weak that, in some sampled configurations, the surrounding solvent clashes with the vanishing species.

The simpler method to avoid the consequent large fluctuations in  $\langle \Delta U \rangle$ , without doing extremely extensive sampling or small time stepping, is to use a non uniform spacing between the different states  $i$  and using a large number of narrow windows when approaching the critical points. In fact, this is equivalent to adopting a nonlinear dependence of the potential on the coupling parameter  $\lambda$ .

## Results

The target free energy was computed by dividing the transformation in 16 intervals, 7 of which are in the  $\lambda$  interval [0.9 - 1] (where  $\lambda = 1$  correspond to the annihilated ion). In each interval  $\langle \Delta U \rangle$  was computed by doing a 1 ns AA simulation in the canonical ensemble. The resulting free energy was then corrected to properly account for the fact that periodic boundary condition and Ewald summation were employed during the simulations. To compensate for the various artifacts [73] introduced by using these methods (*e.g.*, solvent reorganization, spurious interaction and finite size effects) the formula proposed by Hunenberger *et al.*[74] for the low molality limit was used

$$\Delta\Delta A = \frac{1}{2} \xi_{EW} \frac{q_B^2 - q_A^2}{\epsilon_r L} \quad (2.22)$$

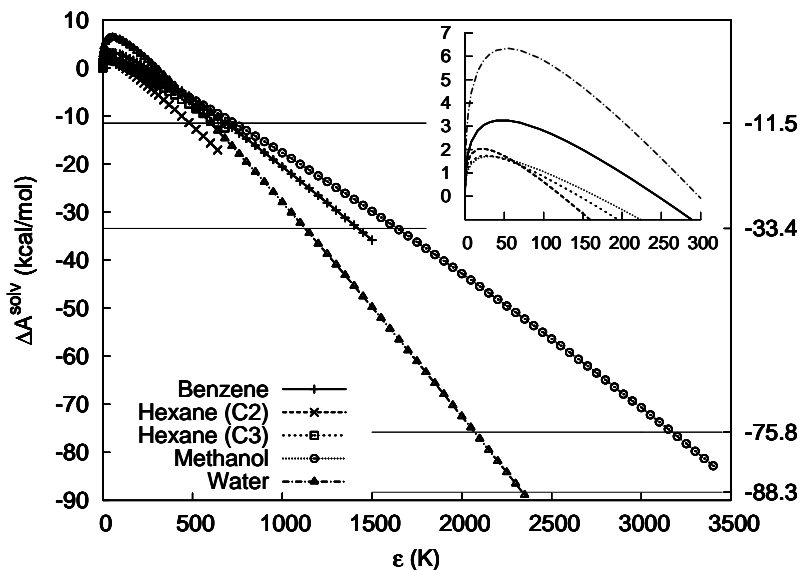
where  $\Delta\Delta A$  is the free energy correction,  $L$  is the size of the periodic box,  $\epsilon_r$  is the dielectric constant of the solvent,  $q_A$  and  $q_B$  are the total system charge of the initial and final state and  $\xi_{EW}$  is equal to -2.8327297, *i.e.*, the self-energy of a point charge in a cubic Wigner lattice. The size of the box for this simulation was about 40 Å so the low molality limit is justified because the volume of the ion is lot smaller than the total volume ( $R \ll L$ ). The final values are presented in Table 2.7.

Two steps were performed to make the LJ interaction potential reproduce the same free energy in the CG system. First  $\sigma$  was fixed by taking the shortest distance that makes the ion-solvent mean force potential of an AA simulation equal to 0. This is equivalent to

**Table 2.7:** Target  $\Delta A_{solv}$  (in Kcal/mol) of  $\text{Na}^+$  and  $\text{SO}_3^-$  in different solvent from AA simulations corrected for the periodic boundary conditions and Ewald summation artifacts. The last line lists the dielectric constant values used in equation (2.22) [75, 76, 77, 78]

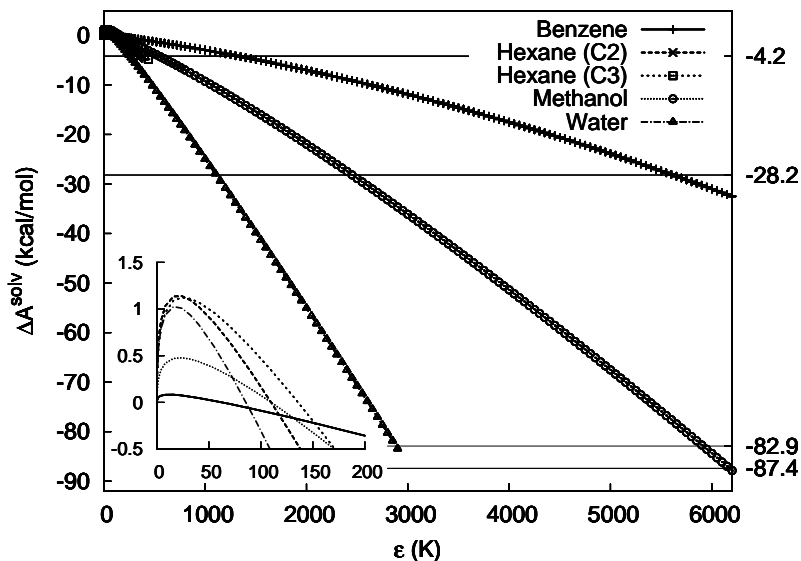
Ion	Hexane	Benzene	Methanol	Water
$\text{Na}^+$	-11.5	-33.4	-75.8	-88.3
$\text{SO}_3^-$	-4.2	-28.2	-87.4	-82.9
$\epsilon_r$	2.0	2.2	31.0	70.0

considering the shortest distance for which the ion-solvent pair radial distribution function is equal to 1 (for the purpose of computing the rdf a 4 ns AA NVT simulation was performed). Once  $\sigma$  is fixed in this way,  $\epsilon$  was obtained by computing the free energy of solvation associated to increasingly large  $\epsilon$  until the target  $\Delta A_{solv}$  was matched. The corresponding curves are plotted in Figure 2.14 and 2.13.



**Figure 2.13:**  $\Delta A_{solv}$  of  $\text{SO}_3^-$  CG bead in benzene, hexane, methanol and water as a function of the Lennard-Jones  $\epsilon$  parameter. Every point on the plot corresponds to a different simulation. On the right the target values are reported. In the inset it is visible that for small  $\epsilon$  the ion insertion in the solvent is disfavored by the free energy cost of cavity formation.

What is evident from the results is that some of the interactions, mainly the ones with



**Figure 2.14:**  $\Delta A_{solv}$  of  $\text{Na}^+$  CG bead in CG benzene, hexane, methanol and water as a function of the Lennard-Jones  $\varepsilon$  parameter. Every point on the plot corresponds to a different simulation. On the right target values are reported. From the inset it is clear that for small  $\varepsilon$  the ion solvation is disfavored, but the effect is smaller than for  $\text{SO}_3^-$ , due to the smaller size of the ion.

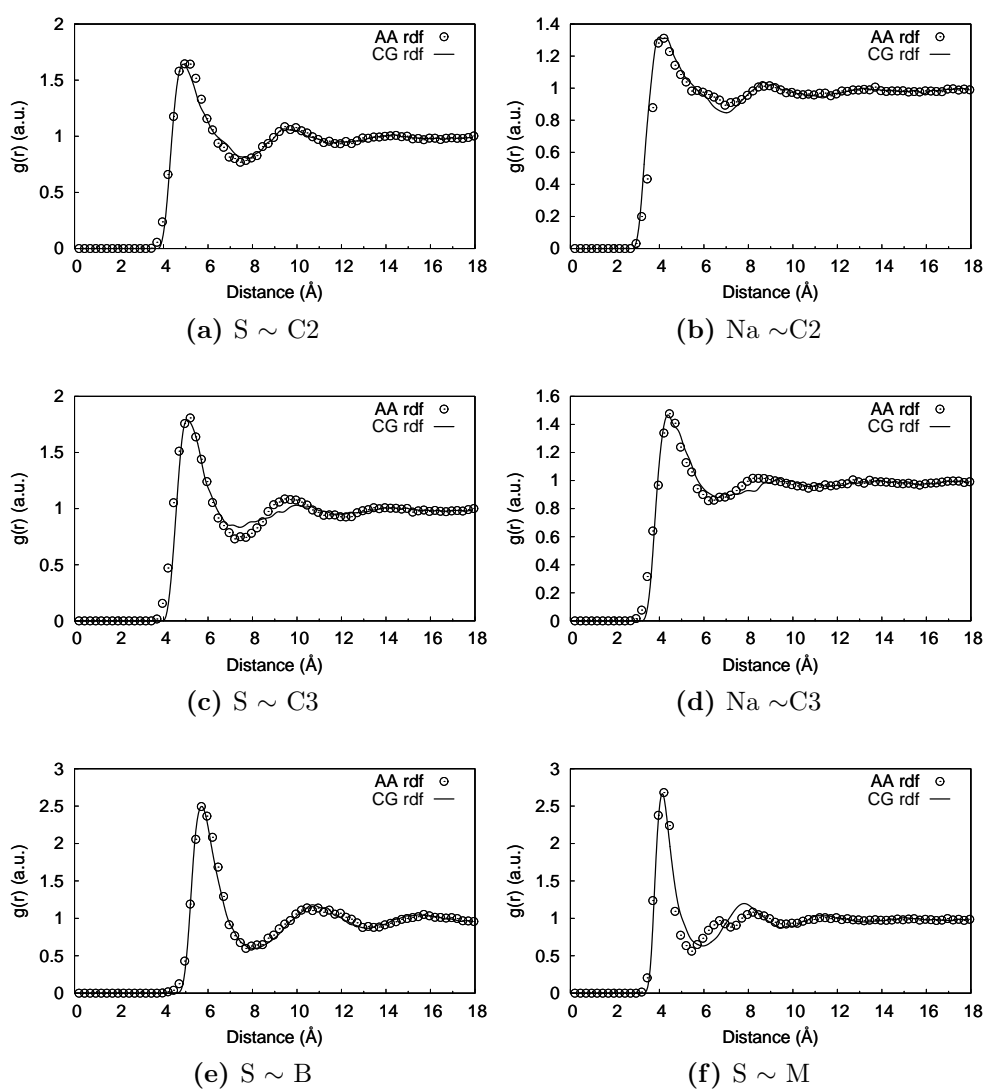
water and methanol, are quite strong, being one order of magnitude bigger than other van der Waals interactions and on the same range of values of some bonded potentials. While this is not unexpected, due to the inclusion in the potential of Coulombic forces, it is not obvious that the system will behave correctly, and therefore a few simple tests were performed.

All the systems analyzed, which included methanol-water solutions containing a low concentration of  $\text{NaHSO}_3$ , showed markedly unphysical behavior, *e.g.*, phase separation or vitrification of one solvent. The most likely reason for these results is the fact that the dipole interaction and the reorientation of the solvent around the charged group give an important contribution to the solvation energy in the AA system, but are not reproduced in our CG model.

Therefore, as an alternative approach the example of Bhargava and Klein [79] was followed, by refining the LJ9-6 parameters so that CG rdf reproduce as much as possible the AA rdfs. In order to do so, the values of  $\sigma$  were obtained in the same way used for the free energy of solvation, and then the  $\varepsilon$  was changed in order to optimize the agreement

## FORCE FIELD PARAMETRIZATION

between AA and CG rdf. A few representative examples of the obtained rdfs are plotted in Figure 2.15: although the agreement is not always excellent, the overall accordance is good and, more important, the obtained interactions do not show the problems mentioned above.



**Figure 2.15:** Comparison between AA and CG rdf for charged-neutral groups pairs.

## 2.6 Conclusions

In this chapter the details regarding the mapping between the atomic and the CG model, and the modeling of the interactions in the new low-resolution representation were presented. The forces acting on each bead were divided in four contributions, two from bonded interaction (bonds and angles) and two from non-bonded potentials (van der Waals and electrostatic). All the bonded potentials were fitted with harmonic functions, while for van der Waals interactions different approaches were used: a modified IBI method, that gives smooth and pressure-corrected potentials, was used together with the fitting of key macroscopic properties, like density and surface tension.

The standard IBI was improved by constraining the system to the target pressure at each iteration. This modification has a negligible computational cost and, together with the reduction of the numerical and statistical noise obtained by using Bézier cubic splines to fit the correction terms, gives a faster and smoother convergence. For those interaction which cannot be modeled through IBI, a LJ potential was parametrized to reproduce experimental surface tension and density, or in case of non-miscible species for which the density cannot be used information from the AA rdfs were used.

Interactions between charged groups were modeled by scaling the charge according to the square root of an estimate of the average dielectric constant. To model the interaction between charged site and overall neutral molecules, which however have a dipole or a higher order electrostatic interactions, first an approach based on the free energy of solvation was tried. In this attempt the non-bonded potential was modeled in order to reproduce the free energy of solvation of an equivalent AA system, computed with the free energy perturbation method. However, due to some inherent limitations of the CG models, like the absence of dipole-dipole interactions and solvent reorientation, the obtained potentials gave unphysical behaviors. Therefore as alternative solution LJ potentials were refined so that the CG rdfs of the ions in different solvent match as close as possible the corresponding AA rdfs.

The resulting tabulated potential are reported in Figures 2.6 and 2.7, while the parameter for the analytical potentials are reported in Appendix B.

---

# Chapter 3

## Force field testing

In the previous chapter, a systematic procedure for developing a CG FF, that reproduce desired experimental or computed properties of the system under investigation was presented. However, the dimensionality reduction and the simplification inherent to the use of two body potentials implies that the model can correctly represent only a subset of the properties of the real system [52]. While the choice of the target properties used for the force field parametrization guarantees that the key features of the system are correctly reproduced, it is important to verify its behavior with respect to other observables, so as to understand the limits of the application of the force field. Therefore, before moving to the study of more complex systems, a series of tests were performed with these potentials.

In the first series of tests, the validity of the choice of a simple harmonic potential to model bonded interactions is verified by examining the radius of gyration of linear polyethylene chains. This test is also important to check if the interactions modeled using short alkanes work properly on much longer chains (Section 3.1).

After this, IBI-obtained potentials are tested by comparing some simple properties of CG systems with experimental and AA simulated values. In particular, even if for the parametrization of the force field and all its applications the average temperature of the systems is kept at 325 K, it is interesting to examine the temperature dependence of these potentials to asses their transferability (Section 3.2).

The third series of tests concerns the behavior of methanol-water binary mixtures, which is very important due to the dependence of many membrane structural properties on the swelling solution. Albeit the interactions were modeled to reproduce some key properties of pure water and methanol, no information about their mixture, which shows a strong non-ideal behavior, was directly included in the force field. Therefore, different

experimental observables of methanol-water solution along the full composition range are compared with the ones obtained in the CG reduced representation (Section 3.3).

Finally, the last section is dedicated to some considerations about the general computational gain obtained with CG respect to AA simulations, with particular attention to the energy conservation during the dynamics (Section 3.4).

### 3.1 Polyethylene chains

As detailed in Section 2.2, the bonded interactions were modeled using the harmonic approximation. Moreover, the force field vales were parametrized over bond and bend distributions of small molecules and short alkanes. This is generally needed in order to minimize the inclusion in the CG potential of AA intramolecular non-bonded interactions, which become increasingly relevant for bigger molecules, and which reduce the transferability of the obtained potential.

Therefore, to verify the applicability of these potentials to more general cases, simulations of long linear polyethylene (PE) chains in vacuum were compared to corresponding AA runs. To this end, PE chains of different length (see Figure 3.1) were simulated for 100 ns in vacuum at three different levels of detail: first AA, second by grouping two carbon atoms (and their respective hydrogens) in one bead (C2) and third by grouping three carbon atoms in one bead (C3). All the simulations were kept at 325 K with a Nose-Hoover chain of length five; the other details of the simulation are similar to the one used in the parametrization of the force field (see Subsection 2.1.1 for further information).

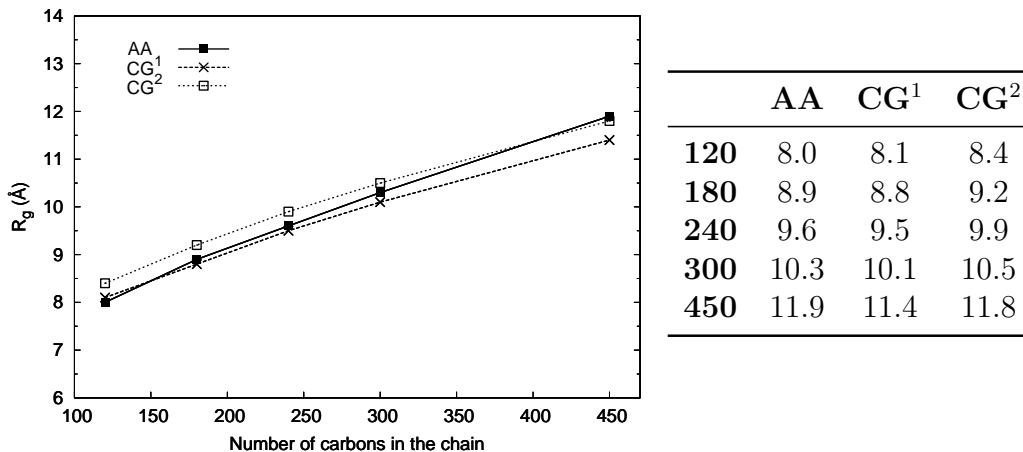
The resulting trajectories were analyzed after discarding the first nanosecond, by measuring the radius of gyration ( $R_g$ ) of each chain, *i.e.*, the mass weighted root mean square distance of each atom from the center of mass, that is:

$$R_g \equiv \sqrt{\langle R_{gyr}^2 \rangle} = \sqrt{\left\langle \frac{\sum_i m_i (\vec{r}_i - \vec{r}_{CM})^2}{\sum_i m_i} \right\rangle} \quad (3.1)$$

where  $m_i$  and  $\vec{r}_i$  are, respectively, the mass and position of each atom or bead,  $\vec{r}_{CM}$  is the position of the center of mass of the whole molecule or chain and the angular brackets are the ensemble average.

The results are presented in Figure 3.1. The plot clearly shows that the agreement between AA and CG representations is very good: within the tested range the difference





**Figure 3.1:** Plot and table of the radius of gyration of AA and CG (two different grouping) representation of linear polyethylene chains. The error on each point, obtained with the blocking analysis is smaller than  $0.05 \text{ \AA}$ . <sup>1</sup> C2 grouping <sup>2</sup> C3 grouping.

is always less than 5 % of the AA value. Therefore, despite the inability of the harmonic potential to reproduce all the features of the AA distributions, it still proves to be a good approximation.

Despite the small difference however, the curves are not identical. In particular, the trend of the AA curve suggests that for very long chains the discrepancy with the CG representations is likely to increase. Nonetheless the problem has a limited impact in real applications for at least two reasons. First, the extrapolation of the curves, suggests that even for longer chains (1000 carbons) the discrepancy between AA and C3 is likely to remain small (less than 5 %). Second, in complex systems where other species are present, the chain shape is determined by the interaction with the environment, especially if electrostatic or other long range forces are involved, and therefore a small error in the intramolecular forces can be neglected. As such, considering the target applications, the obtained level of accuracy can be considered very good, and a more detailed potential is only needed for specific studies, *e.g.*, PE chain crystallization.

Finally, it is interesting to note that the increase in the level of CG causes a solid shift of the  $R_g$  curve by about  $0.4 \text{ \AA}$ . While this effect is not general and likely related only to the functional form of the potential, this effect can be taken into account in the choice of the coarse-graining in order to better reproduce the AA behavior. Therefore, for shorter chains a more refined representation is better suited to reproduce the general shape of the

chain; on the other hand longer chains are better represented with a coarser representation.

## 3.2 Solvents

Many interactions were modeled using the IBI method, but due to the fact that the IBI procedure aims at reproducing the rdf there is no guarantee that other observables are correctly described.

Hence, to verify the properties of these potentials a series of simulations were performed to compute the density, the isothermal compressibility and the surface tensions of liquids interacting through IBI-obtained potentials. The values obtained for the CG systems are then compared with the equivalent AA system and experimental values. The comparison with the AA force field is in this case of particular interest because the target rdfs used in the IBI method were obtained from AA runs.

A box containing the molecules under study (1293 for benzene, 1169 for hexane and 1179 for ethylbenzene) was equilibrated for 5 ns in an NPT ensemble to obtain  $\rho$  and  $\beta_T$ , while  $\gamma$  was obtained by simulating two-dimensional slabs of liquid in vacuum in a canonical ensemble for 10 ns, as reported in the force field parametrization. The other details about the simulation settings are similar to the one used to fit potential against experimental data (see Subsection 2.1.1 and Section 2.4).

**Table 3.1:** Comparison of  $\rho$ ,  $\beta_T$  and  $\gamma$  of CG systems with AA and experimental values [75, 80, 81, 82, 83] at 325 K and 1 atm.

Species	$\rho$ [g/cm <sup>3</sup> ]			$\gamma$ [10 <sup>-3</sup> N/m]			$\beta_T$ [GPa <sup>-1</sup> ]		
	AA	CG	Exp.	AA	CG	Exp.	AA	CG	Exp.
Benzene (B)	0.798	0.848	0.848	19.6	40.5	25.00	1.87	3.52	1.13
Hexane (C2)	0.606	0.631	0.631	14.1	23.0	15.33	2.67	2.35	2.03
Hexane (C3)	0.606	0.637	0.631	14.1	18.0	15.33	2.67	3.21	2.03
Methanol (M) <sup>1</sup>	-	0.756	0.756	-	20.0	20.21	-	0.96	1.42
Methanol (M) <sup>2</sup>	-	0.754	0.756	-	20.2	20.21	-	2.06	1.42
Ethylbenzene (C2-B)	0.793	0.842	0.841	20.0	50.7	26.01	1.55	2.10	~1
Water (W)	-	0.990	0.988	-	67.3	67.92	-	0.37	0.442

<sup>1</sup> Coarse-grained interaction M~M modeled with LJ12-4 potential.

<sup>2</sup> Coarse-grained interaction M~M modeled with LJ9-6 potential.

The results of these tests are presented in Table 3.1. The analysis of the values shows the excellent agreement between the experimental and the CG density. Moreover is worth noticing the systematic improvement of the CG values over the AA calculations, despite the fact that the CG interactions were modeled over the AA force field. The most likely reason for these results is the additional pressure constraint imposed during the IBI procedure (Section 2.3), because the modified IBI method used in this work can be regarded as a simultaneous optimization of the potential against both local structure (AA) and global density (experimental).

The situation is different for  $\beta_T$  and  $\gamma$ . As expected, the agreement between AA and experimental values is better than that of the CG model. However, the difference between CG simulations and experiment is generally acceptable and in some cases the CG results are comparable with the values obtained from AA simulations, as for instance in the case of hexane. Benzene and ethylbenzene have the worst agreement, but this is likely related to the shape of the aromatic ring and of its orbitals: modeling such an anisotropic interaction with an isotropic potential, which averages out the angular dependencies, is bound to have some impact on the quality of the results.

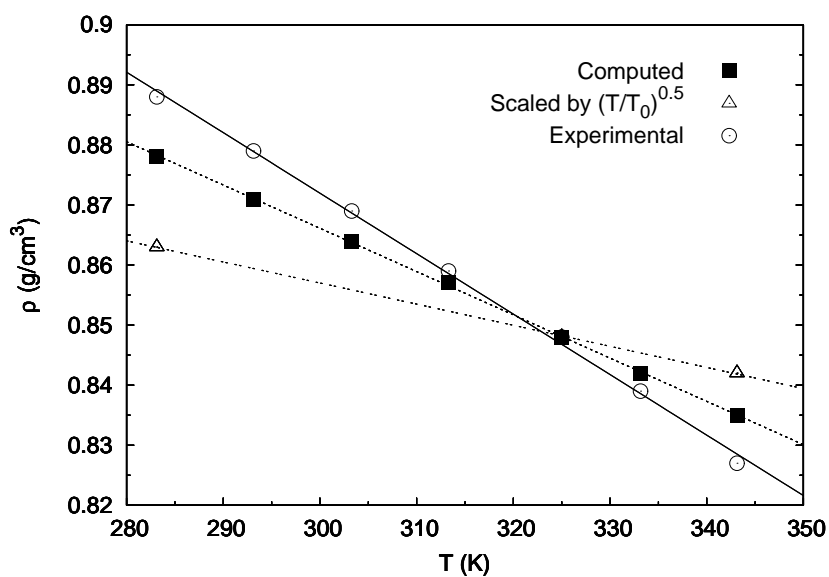
### 3.2.1 Temperature dependence

Considering the fact that all the simulations in this work were performed at 325 K, the temperature dependence of the potentials will not impact the result of this work. However, one of the factors limiting of the applicability of CG to a wider range of systems is the reduced transferability of the used force field. Therefore, it is useful to study the influence of the temperature on the results.

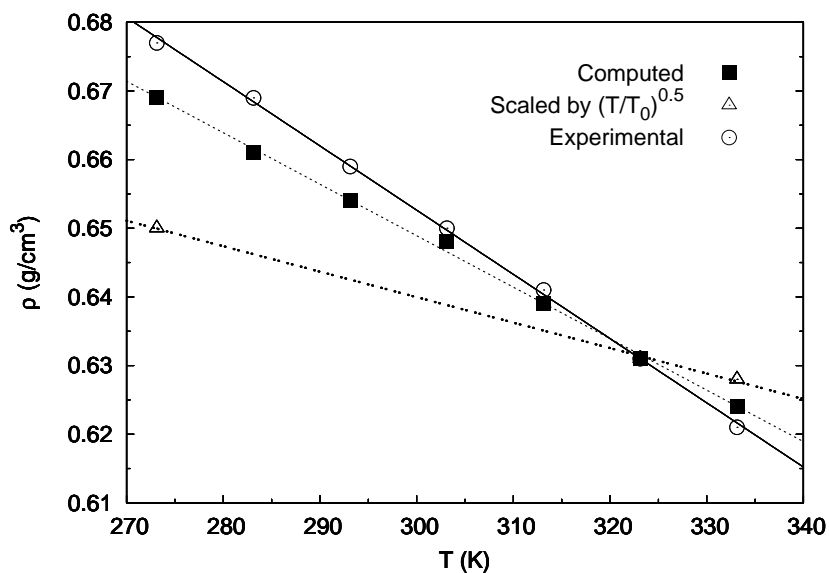
The data presented in Table 3.1 showed an excellent agreement of the CG density with the experimental values at the modeling temperature. Therefore to understand the temperature dependence of IBI-potentials, the correlation between  $\rho$  and  $T$  was analyzed for benzene and hexane (C2 grouping).

The density of benzene and hexane is presented in Figures 3.2 and 3.3 for a range of 60-70 K around 325 K. From these plots it is possible to see that the CG potentials give the right linear  $T$  dependence, but differ in slope. This limits the applicability of the CG force field to a quite narrow range of temperatures. For benzene, the fit of the experimental data gives  $-1.03 \cdot 10^{-3} \text{ g/cm}^3 \cdot \text{K}$  and  $-0.727 \cdot 10^{-3} \text{ g/cm}^3 \cdot \text{K}$  for the CG model; for hexane the experimental value is  $-0.951 \cdot 10^{-3} \text{ g/cm}^3 \cdot \text{K}$ , while the CG one is  $-0.743 \cdot 10^{-3} \text{ g/cm}^3 \cdot \text{K}$ .

### 3.2. Solvents



**Figure 3.2:** Experimental [75], CG and scaled-CG density of benzene as a function of temperature.



**Figure 3.3:** Experimental [75] and CG scaled-CG density of hexane (C2 grouping) as a function of temperature.

In case where a greater temperature range or an higher accuracy is desired, the parametrization has to be repeated at different temperatures so that the resulting CG potentials can be fitted to include the right dependence [23].

Alternatively, Qian *et al.* [59] recently proposed for similar species (but different CG reduction) a scaling factor for the potential equal to  $\sqrt{T/T_0}$ , where  $T_0$  is the temperature at which the force field was modeled. Test performed with our model however show that this factor increases the slope of the temperature dependence of the species under investigation (for benzene it becomes  $-0.357 \cdot 10^{-3} \text{ g/cm}^3 \cdot \text{K}$ , while for C2-hexane  $-0.353 \cdot 10^{-3} \text{ g/cm}^3 \cdot \text{K}$ ), markedly increasing the discrepancy with the experimental values.

### 3.3 Methanol and water mixtures

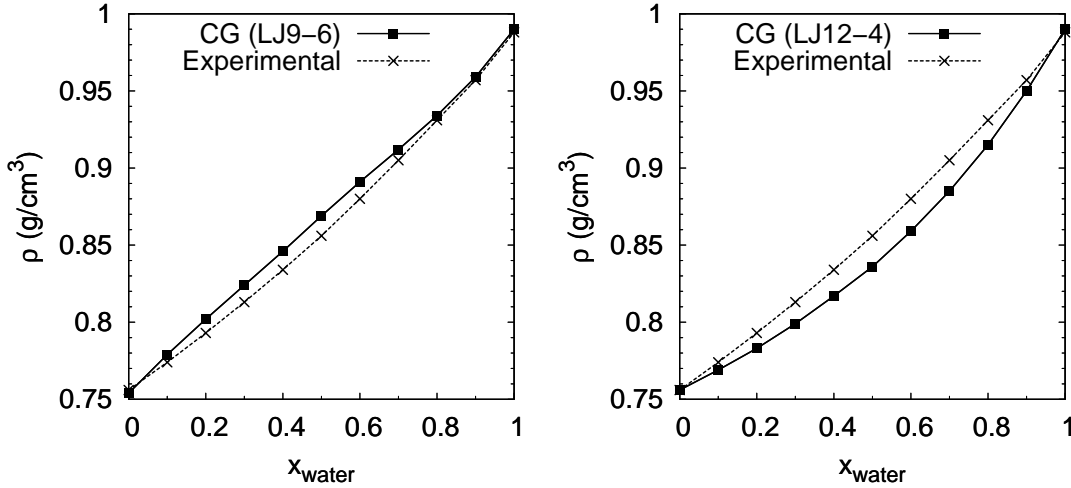
As discussed earlier (Section 2.4.1), due to the lack of general mixing rules for tabulated potentials, methanol and water interactions were modeled by optimizing LJ potentials so that the CG system matches the experimental density and surface tension under the same conditions (325 K and 1 atm). The results listed in Table 3.1 confirm that this can be done with an excellent accuracy (the difference is below 1%) but also show that this choice of target properties also gives obtaining a good approximation for the isothermal compressibility, which was not used as a target property in the FF parametrization.

During the modeling procedure nothing was assumed about the properties of methanol-water mixtures, however such solutions show a marked non-ideal behavior, since many thermodynamic properties differ significantly from the values that might be expected from the linear combination of the pure fluids [84, 85, 86, 87]. This aspect is of particular interest not only for the simulation of DMFC's membranes, also more in general because such mixtures have application in many different fields, ranging from a variety of separation processes [88, 89] to hydrogen production in solar thermal systems [90].

To asses to which extent the model and the LB combining rules are able to reproduce the non-ideal behavior of the methanol-water mixtures, a series of simulations was run to sample  $\rho$ ,  $\gamma$  and  $\beta_T$  along the whole composition range. The simulation settings are the same used in the parametrization of the force field, see Section 2.4 and Subsection 2.1.1, for the details.

In the parametrization of the FF the M~M interaction was modeled with two different potential: a LJ9-6 as recommended by Shinoda *et al.* [64] and a LJ12-4 in order to apply the LB mixing rules with the water potential. However, contrary to the findings of Shinoda,

it was found that the LJ12-4 is able to reproduce methanol properties (including the isothermal compressibility, see 3.1) as well as the LJ9-6 potential. This is likely due to the inclusion of a pressure tail correction in the present work. As the two potentials are almost equivalent, their behavior with respect to mixing properties was examined in order to chose the one better suited for our application.



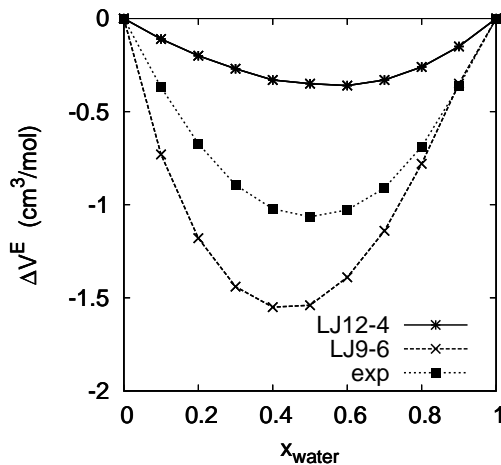
**Figure 3.4:** Comparison between CG and experimental [80] density of methanol-water solutions along the full composition range. The two plots differ in how the the  $M\sim M$  interactions are modeled: a LJ9-6 potential is used on the left, a LJ12-4 on the right. The error bar is smaller than the symbol used to represent each point. The details about the parameters used are given in Table B.3

From Figure 3.4 is clear that both potentials are able to give a very good agreement with experimental density along the full composition range. A more sensitive test to check the non-ideal behavior of the CG mixture is the measure of the deviation of the volume from ideality with the excess volume of mixing ( $\Delta V^E$ ):

$$\Delta V^E = V_M - V_{id} = V_M - (x_1 V_1 + x_2 V_2) \quad (3.2)$$

where  $V_M$  is the effective molar volume and  $V_{id}$  is the ideal molar volume of the mixture. In case of a binary liquid  $V_{id}$  can be expressed as the sum of the ideal molar volume of the pure ( $V_1$  and  $V_2$ ) weighted by the respective molar fractions.

The results are presented in Figure 3.5. The plot magnifies the difference between the two set of potentials, showing a slightly better accordance of the LJ9-6 potential with the

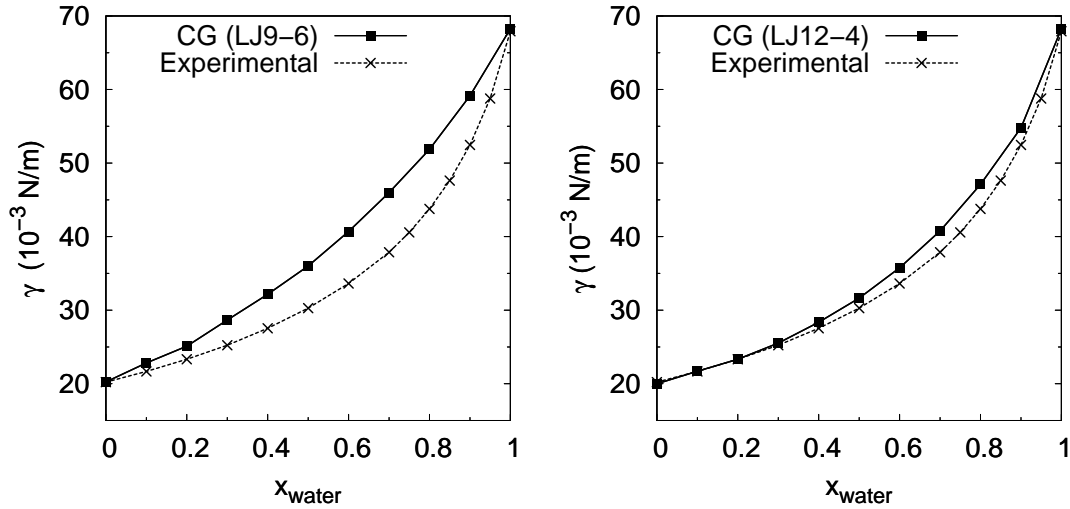


**Figure 3.5:** Excess volume for methanol-water solutions. Data obtained with different M~M interactions for CG beads (325 K, 1 atm) compared with experiment [80] (323.15 K, 0.1 MPa). The error bar is smaller than the symbol used to represent each point.

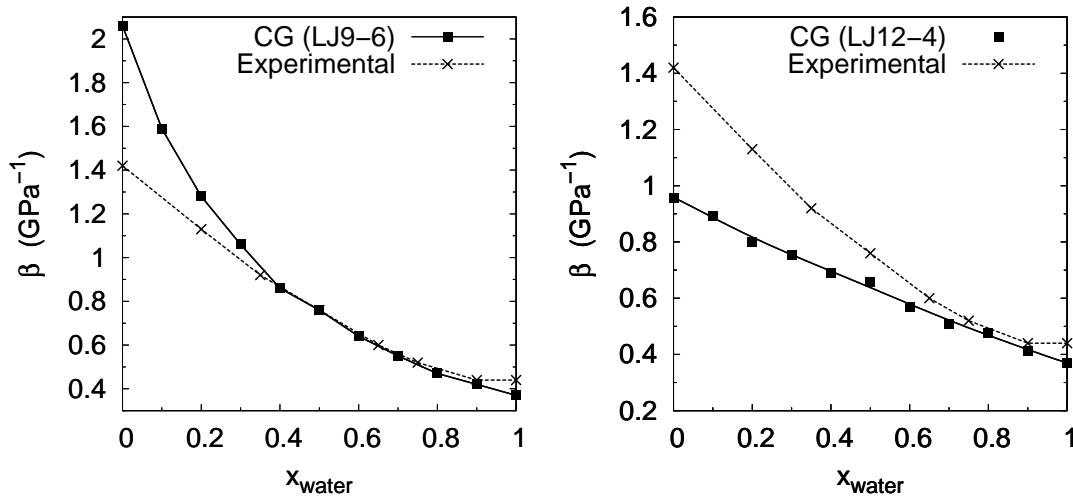
experimental values.

Figures 3.6 and 3.7 show the plots for the surface tension and the isothermal compressibility. The computed surface tension shows a general agreement with the experimental values, with the LJ12-4 potential performing better. On the other hand, for the isothermal compressibility the agreement is worse because  $\beta_T$  was not used as a fitting property. The result is however still good, in particular for the LJ9-6 potential and for water molar fractions bigger than 0.3.

In general, these results show that despite the loss of atomistic details, which is a key factor for the structural properties of such solutions where the complex hydrogen bond network plays an important role, the CG representation is able to reproduce with a good approximation even non-ideal properties that were not directly used in the FF parametrization. Moreover the differences between LJ12-4 and LJ9-6 potentials suggest that a suitable choice of exponents of the LJ potential can give even better agreements and likely improve also the accordance of  $\beta_T$  at low values of  $x_W$ . However the ultimate goal is fuel cell applications, where the main interest is in water rich solutions. Therefore, it was chosen not to increase the number of parameters, which implies additional computational costs and modeling complexity.



**Figure 3.6:** Comparison between CG and experimental [83] surface tension of methanol-water solutions along the full composition range. The two plots differ in how the the M~M interactions are modeled: a LJ9-6 potential is used on the left, a LJ12-4 on the right. The error bar is smaller than the symbol used to represent each point. The details about the parameters used are given in Table B.3



**Figure 3.7:** Comparison between CG and experimental [80] isothermal compressibility of methanol-water solutions along the full composition range. The two plots differ in how the the M~M interactions are modeled: a LJ9-6 potential is used on the left, a LJ12-4 on the right. The details about the parameters used are given in Table B.3



### 3.4 Efficiency considerations

The efficiency gain of a CG model is related to many factors. The softer potential, the lower frequency of the fastest modes and smoother free energy surface allow bigger time steps in the integration of the equation of motion, while the reduced number of interacting sites decreases both the computational cost of the forces and the one associated with various optimizations, *e.g.*, the rebuilding of neighbor lists. However many other factors both related to the simulation, *i.e.*, neighbor skin or thermostat settings, and unrelated to it, *i.e.*, the code implementation and hardware settings, influence the speed gain. In this thesis, no real systematic work was performed to measure the influence of all these factors and therefore the results reported in this paragraph should be extended with caution.

Before comparing performance obtained in the CG representation with AA runs, different time step settings were tested to verify the performance of the integrator. To this end, a series of 1 ns microcanonical simulation of 4680 beads representing sodium dodecyl benzenesulfonate in water (see Section 4.1 for further details) was performed using a simple two region RESPA configuration, with a separation distance of 12 Å. The influence of the different time step settings on the total energy drift is shown in Table 3.2.

**Table 3.2:** Absolute and relative total energy drift for the CG FF simulated with CM3D. The two values of the time step refer to the long range and short range evaluation times.

Time step (fs)	Drift (K/atom ns)	% Drift (ns <sup>-1</sup> )
1 - 1	0.08 ± 0.01	2 10 <sup>-3</sup>
4 - 2	0.08 ± 0.01	2 10 <sup>-3</sup>
8 - 4	0.69 ± 0.01	10 10 <sup>-3</sup>
10 - 2.5	0.32 ± 0.01	6 10 <sup>-3</sup>
20 - 2	0.36 ± 0.01	7 10 <sup>-3</sup>
20 - 2.5	0.89 ± 0.01	15 10 <sup>-3</sup>
30 - 2.5	3.71 ± 0.01	70 10 <sup>-3</sup>

The table shows that the inner region of the algorithm should not be updated less frequently than 2.5 fs likely due to intramolecular motions, while the outer region can be updated every 20 fs. With these settings the energy drift is still small, increasing the average temperature of the system during a microcanonical simulation of less than 4 K every 10 ns. This value is only four time bigger than the one given by updating all the system every femtosecond, but guarantees almost twenty-fold speed up, because the

### 3.4. Efficiency considerations

---

computational cost of the single step is about the same. Although bigger time steps cause an increasingly large energy drift, more complex multiple time step schemes can be devised, for example by separating intramolecular interactions from short range intermolecular ones, which can be evaluated less frequently.

Once the time step scheme was decided, the performance of different systems were compared as reported in Table 3.3. It should be stressed that the result reported for simulations performed on more than one processor were not optimized with respect to the number of CPUs, because the simulation program are very different, and are indicative.

**Table 3.3:** Performance comparison of MD simulation in the AA and CG representations, using NAMD and CM3D code respectively. The simulation is described by a letter for the system (A refers to the simulation of the S bead in methanol; B refers to S bead in water ; C refers to ten unit PSSA oligomer in a 1:3 methanol-water solution) and a number for the hardware (1 refers to an Athlon 64 x2; 2 refers to Opteron 2220 with InfiniBand QDR network connection).

Model	System	Atoms	Processors	time-cpu/step	ns/day
AA	A - 1	5480	1	0.44	0.19
AA	A - 2	5480	16	0.73	1.89
CG	A - 1	910	1	0.03	51.4
AA	B - 1	7516	1	0.56	0.15
AA	B - 2	7516	16	1.10	1.26
CG	B - 1	2505	1	0.34	5.15
AA	C - 1	13084	1	0.57	0.13
AA	C - 2	13084	16	0.76	1.82
CG	C - 1	3462	1	0.58	2.96

The reported values show that the main limiting factor in the performance of the CG simulations is related to the scalability of the CM3D code with respect to the number of atoms in the simulation box. Yet the gain is still relevant with two order of magnitude increase in the accessible time scale, which is not obtainable even with the parallelization of the AA code.

## 3.5 Conclusions

In this chapter the force field parametrized in the previous chapter was tested to verify its behavior with respect to physical quantities that were not considered during the parametrization.

To test the bonded interactions, the comparison of the radius of gyration of long linear PE chains in vacuum in different CG representations, showed that despite the absence in the force field of torsional terms and the simple form of the harmonic potential, the CG is able to provide a good approximation of the AA systems.

The second series of tests was designed to check the behavior of tabulated potentials obtained with IBI with respect to macroscopic observables like density, surface tension and isothermal compressibility. The results show a generally good agreement with experimental values, with higher discrepancy for molecules with high anisotropic shapes, like the benzene. Moreover the comparison with the density of AA simulations, suggests that the additional term added in the IBI procedure greatly improves the ability of the CG model to reproduce experimental density, with a systematic and consistent gain with respect to AA FF. The temperature dependence of the IBI potential was only briefly tested and showed that the optimal agreement is possible only in a limited range of temperatures (about 20 K), although outside this interval the accuracy is still good compared for example to the AMBER force field.

Another crucial point for the simulation of the membrane in methanol-water solution is the ability of the CG model to correctly capture the non-ideal behavior of such mixtures. Therefore, in the third part of the chapter some properties ( $\rho$ ,  $\gamma$  and  $\beta_T$ ) were analyzed along the whole composition range, showing a very good agreement with experimental values, with the only exception for isothermal compressibility at high methanol content ( $x_W < 0.3$ ), likely related to the choice of the LJ exponent used for methanol.

Finally the benchmark simulation show that generally a two order of magnitude speedup gain can be achieved although it is largely related to the code used for the simulations.

In conclusion the tests show that the force field gives a good overall reduced representation of the atomic system, being able to correctly reproduce different properties and at the same time allowing a relevant gain in computational time.



---

# Chapter 4

## FF Applications

### 4.1 Surfactants

Surfactants have long been of importance in a wide range of fields like detergents, paints, paper coatings, food and pharmaceutical. The reason for such wide usage is the tendency of surfactants, an abbreviation for surface active agent, to adsorb at interfaces. This behavior is driven by the corresponding reduction of the interfacial tension and is strictly related to their amphiphilic structure. Generally speaking, amphiphiles are composed of two parts, one that is soluble in water, *i.e.*, hydrophilic part or *head*, and one that is insoluble, *i.e.*, hydrophobic part or *tail*.

Due to this structure, when the interface is saturated they tend to aggregate in the bulk in order to reduce the contact area between hydrophobic and hydrophilic components, resulting in different aggregation structures. This critical concentration and the bulk phase behavior depend on many factors besides the composition of the phases and the nature of the surfactant itself, *e.g.*, temperature and salt concentration.

Despite their relatively simple chemical structure, surfactants are still actively researched due to the big variety of possible amphiphilic molecules and the wide range of applications. Moreover, in recent years environmental and health concerns due to the suspected toxicity, low biodegradability and high bioaccumulation, which seem to be related to allergies and skin irritation, have been the driving force towards a better understanding of the interaction between such molecules and biological systems.

Surfactants have been extensively studied [91] using different techniques, like NMR [92], fluorescence, EPR [93], light scattering, resonance Raman scattering, SANS [94] and second harmonic generation, but many atomic-scale details are still elusive. On the other hand,

this kind of information is difficult to produce with AA MD [95] or MC [96] because of the time-scales involved in the equilibration and reorganization dynamics, thus making this a perfect subject to be studied with CG models.

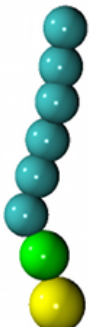
In this section a preliminary study of the aggregation behavior of one anionic surfactant, namely sodium dodecylbenzene sulfonate, is presented in order to strengthen the confidence in the proposed CG model but a thorough investigation of its phase diagram was not performed.

#### 4.1.1 Sodium dodecylbenzene sulfonate

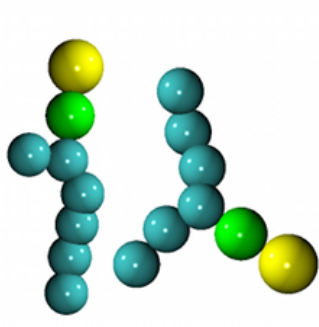
Dodecylbenzene sulfonate displays a very rich phase behavior [97] as a function of temperature, concentration, counterions isomer composition and salt concentration [98]. In particular, its sodium salt (SDBS) at 50 °C aggregates in bulk in three different structures and with increasing concentration of surfactant one finds micellar, lamellar and hydrated crystal phases [98].

In this thesis, only two isomer compositions were studied. One which is among those experimentally investigated [99], composed by 17% DBS-1, 33% DBS-2 and 50% DBS-3 (see 4.1b for details) and one for comparison purpose composed only by the linear isomer DBS-1. These mixtures were solvated in water at three different concentrations, so to have a 20% w/w, a 50% w/w and a 70% w/w concentration of surfactant, for a total of four systems, as detailed in Figure 4.1.

Label	SDBS	Isomers	Water
l20	60	100% DBS-1	4260
m20	60	17% DBS-1 33% DBS-2 50% DBS-3	4260
m50	180	17% DBS-1 33% DBS-2 50% DBS-3	3060
m70	180	17% DBS-1 33% DBS-2 50% DBS-3	1440



(a) SDBS systems composition

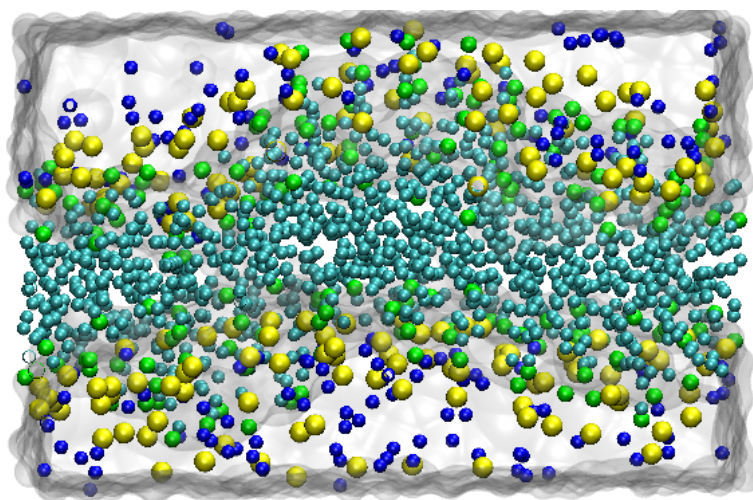


(b) DBS isomers

**Figure 4.1:** (a) Details about composition and name of the SDBS system studied in this section. (b) The three CG isomers of DBS: from left DBS-1, DBS-2 and DBS-3. C2 beads are shown in cyan, benzene in green and S group in yellow; the graphics were generated using VMD [100].

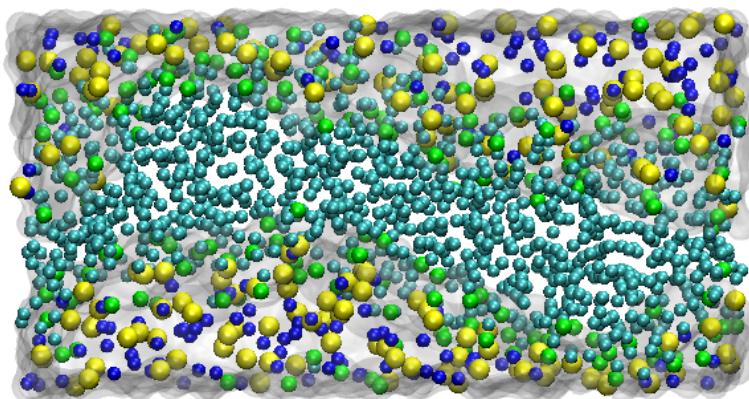
Each simulation began from a random distribution of molecules in a cubic box and the fixed angle variant of the Parrinello-Rahman constraint pressure algorithm [101] was used to equilibrate the cell. Each system was relaxed for 10 ns and the averages were accumulated during the successive 15 ns of simulation performed in the canonical ensemble. Other details are the same as the one presented in Section 2.1.1, with the exception of the time step (here a 2.5 - 20 fs RESPA scheme was used: see the results presented in Section 3.4). Finally, it should be mentioned that the times reported for CG models are just a measure of computational effort since the time scale in CG simulations is not precisely defined.

In all systems, aggregation occurred in less than 500 ps, but the full relaxation (evaluated from the total energy evolution) required longer time and especially for the two more concentrated systems, several nanoseconds were needed. In agreement with the experimental findings, the two more concentrated systems form a slightly irregular lamellar phase (Figures 4.2 and 4.3) and the 20% systems aggregates as micelles (Figures 4.4a and 4.4b).



**Figure 4.2:** Equilibrium configuration for the m50 system (about 50% w/w of surfactant in water). The color convention is the same of Figure 4.1 with the addition of blue Na ions; water is shown only as a semi-transparent box. The graphics were generated using VMD [100].

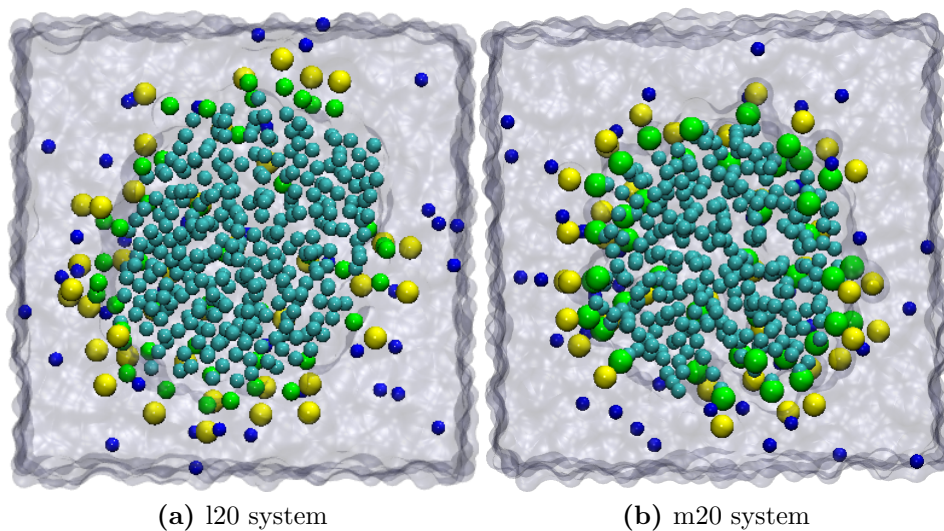
Interestingly, for the bilayers a partial swelling of the surfactant region was observed with the increase of the water content: the thickness of the bilayer increased from 19.5 Å (in the m70 system) to 22 Å (in m50). This change of the phase properties is expected for SDBS systems only above 40 °C [99] and, although the measure of the thickness both



**Figure 4.3:** Equilibrium configuration for the m70 system (about 69% w/w of surfactant in water). The color convention is the same of 4.2; the graphics were generated using VMD [100].

experimentally and computationally is complicated and therefore no direct comparison is possible, the CG simulations reproduce the correct trend.

#### 4.1.2 Micelles

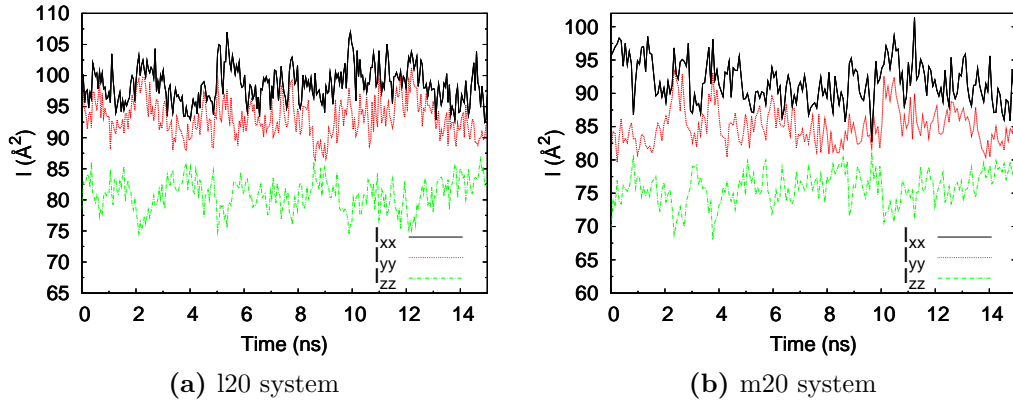


**Figure 4.4:** Section of a typical equilibrium configuration for the l20 and m20 system (see Figure 4.1a). The color convention is the same of 4.2; the graphics were generated using VMD [100].

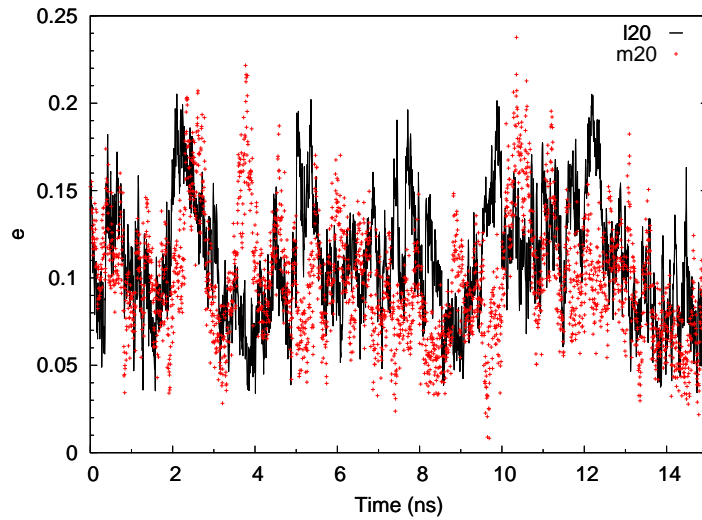
A detailed analysis of the micelle shapes reveals that for both systems they are not



spherical but assume a shape that more closely resembles an oblate spheroid, which is an ellipsoid with the polar axis shorter than the equatorial diameter. This is quantitatively confirmed by the plot of the momenta of inertia and of the eccentricity, shown in Figures 4.5 and 4.6.



**Figure 4.5:** Evolution of the principal axes of the micelle in the l20 and m20 systems.

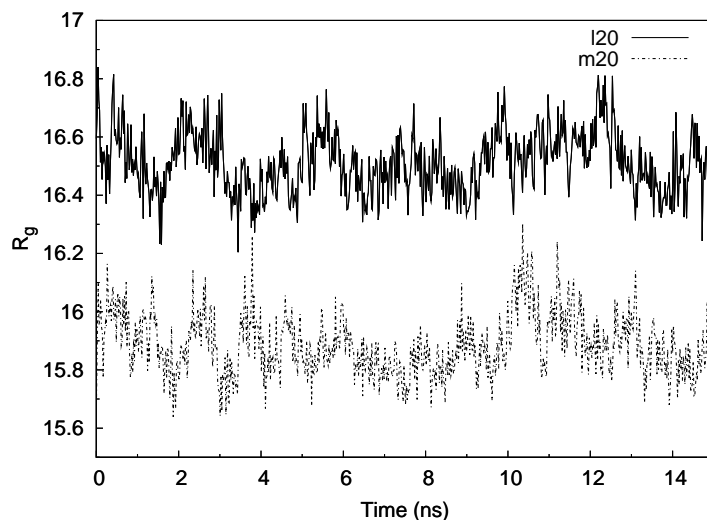


**Figure 4.6:** Eccentricity of the micelles during the simulations.

The eccentricity ( $e$ ) is defined as:

$$e = 1 - \frac{I_{min}}{I_{avg}} \quad (4.1)$$

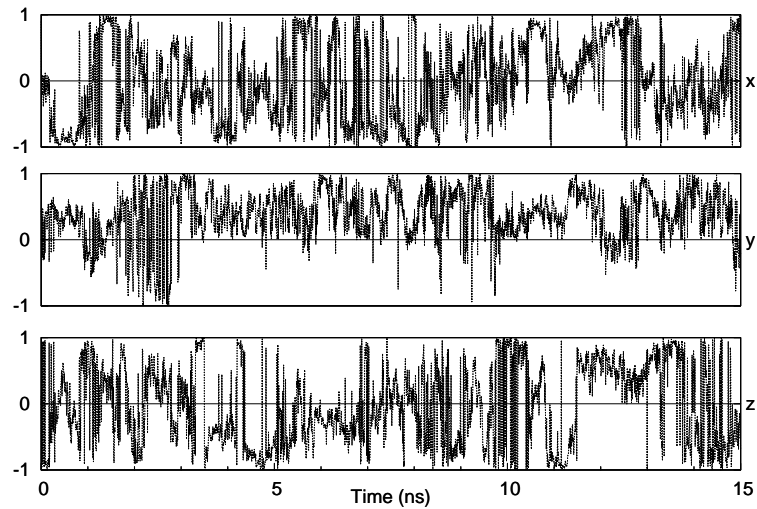
where  $I_{min}$  is the smallest momentum of inertia of the micelle and  $I_{avg}$  represent the average of the three momenta of inertia. From the plot is visible that the value of  $e$  is stable during the simulation and that it only slightly deviates ( $e = 0.1$ ) from the perfect sphere case ( $e = 0$ ) in both systems. This hints that the isomer composition does not influence the shape, even if it reflects on the size of the micelle, as one can see by comparing the scales in Figures 4.4a and 4.4b, or directly by analyzing the radius of gyration of the micelles (Figure 4.7).



**Figure 4.7:** Radius of gyration of the micelles during the simulations.

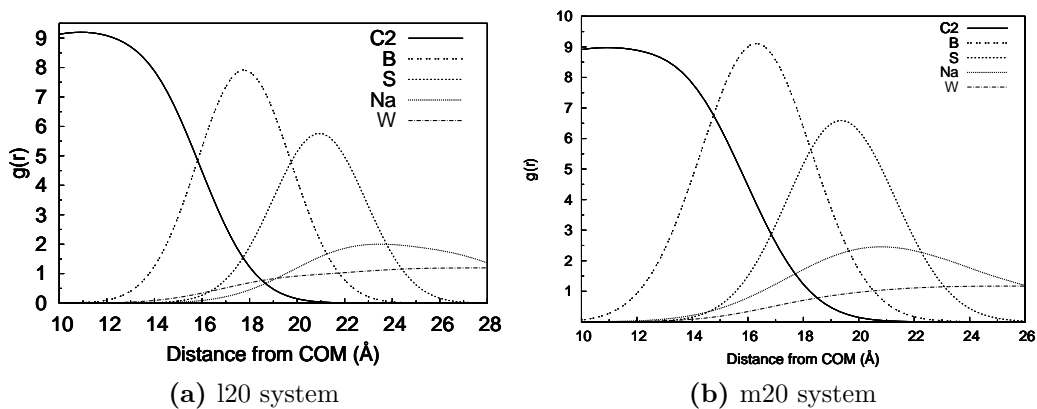
Moreover, from the plot of  $e$  it can be seen that the micelle is not a rigid structure but undergoes continuous fluctuations and rotations. This is more clearly confirmed by the analysis of the time evolution of the principal axis (see Figure 4.8 for the typical plot of the component of the versor of a principal axis of the micelle), which shows that the vectors corresponding to the axes of inertia continuously change in direction and in length.

The dynamics of the micelle shape are important because they should be taken into account when analyzing other average structural properties, like the rdfs of the DBS sites with respect to the center of mass of the micelle (COM), shown in Figure 4.9. From these plots it can be seen that the internal region is composed only by the hydrophobic part whereas the charged head of DBS is external, interacting with the water and the Stern layer of Na counterions, which is shifted about 3 Å outwards. However, the 10 Å region of overlap between the C2 and the W beads does not indicate an extended contact region between water and the alkyl chains as in SDS [102], but is mainly an effect of the spherical



**Figure 4.8:** Evolution of the components of the versor of the second eigenvector of the stress tensor for the micelle formed in the m20 system. The plots for the other axes and for the l20 system are qualitatively similar.

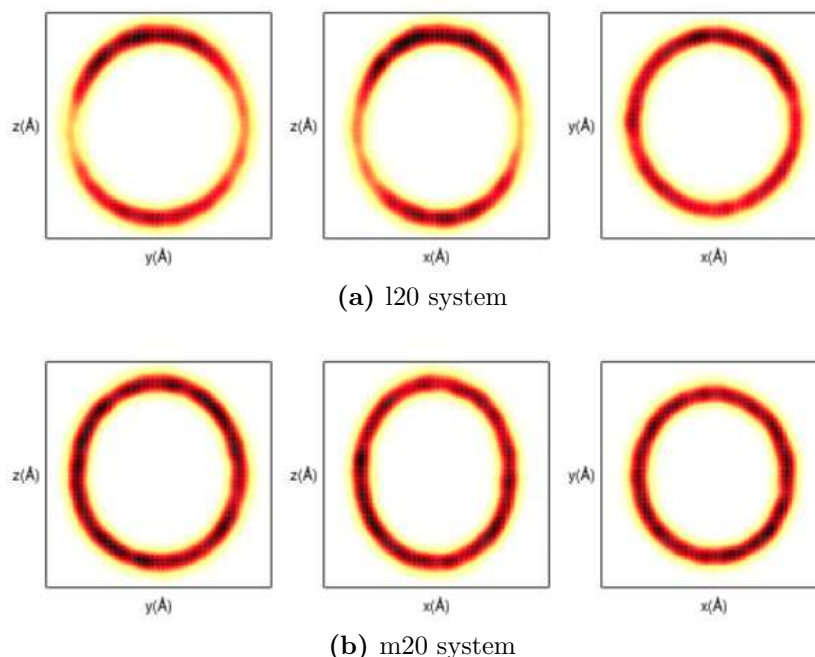
averaging of such a flexible aggregate, as can be confirmed by the visual inspection of the trajectories.



**Figure 4.9:** Radial distribution functions of different beads with respect to the micelle COM for the l20 and m20 systems.

In order to find the possible correlation between the micelle shape and the spatial isomer distribution, the average density of the S beads was analyzed. In order to do so, first the micelles were rotated so that the principal axes were aligned with the Cartesian axes and

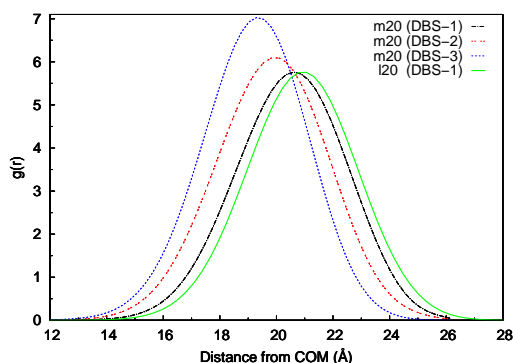
then the section of average density along the Cartesian planes containing the COM of the micelle were plotted. In Figure 4.10 it's possible to see the comparison between the l20 and m20 systems.



**Figure 4.10:** Sections of the average density of the S beads. Each section is taken on the plane which contains the COM of the micelle and lies perpendicular to the principal axes, after aligning them to the Cartesian axes. To darker color corresponds an higher density; the side of each box measures  $54 \text{ \AA}$ .

From this data it is clear that, whereas for the m20 system the distribution of the heads of SDBS is homogeneous, in the l20 system there is an evident depletion near the polar regions, suggesting that the length and shape of the alkyl chain can influence the preferred position of the isomer in the micelle. This idea is further confirmed by the radial distribution function of the S of different isomer with respect to micelle COM (see Figure 4.11), where it is clear that on average the longer the tail the farther from the COM the head prefers to stay. However, this phenomenon seems to be slightly influenced by the isomer composition, as indicated by the shift at shorter distance of the DBS-1 curve in the m20 micelle with respect to the same curve in the l20 micelle.

The mixture's effect is more clearly visible in Figure 4.12 where the average density sections of the S beads done with respect to the various isomer in the m20 system are shown. The density does not show any clear pattern in the isomer distributions except for



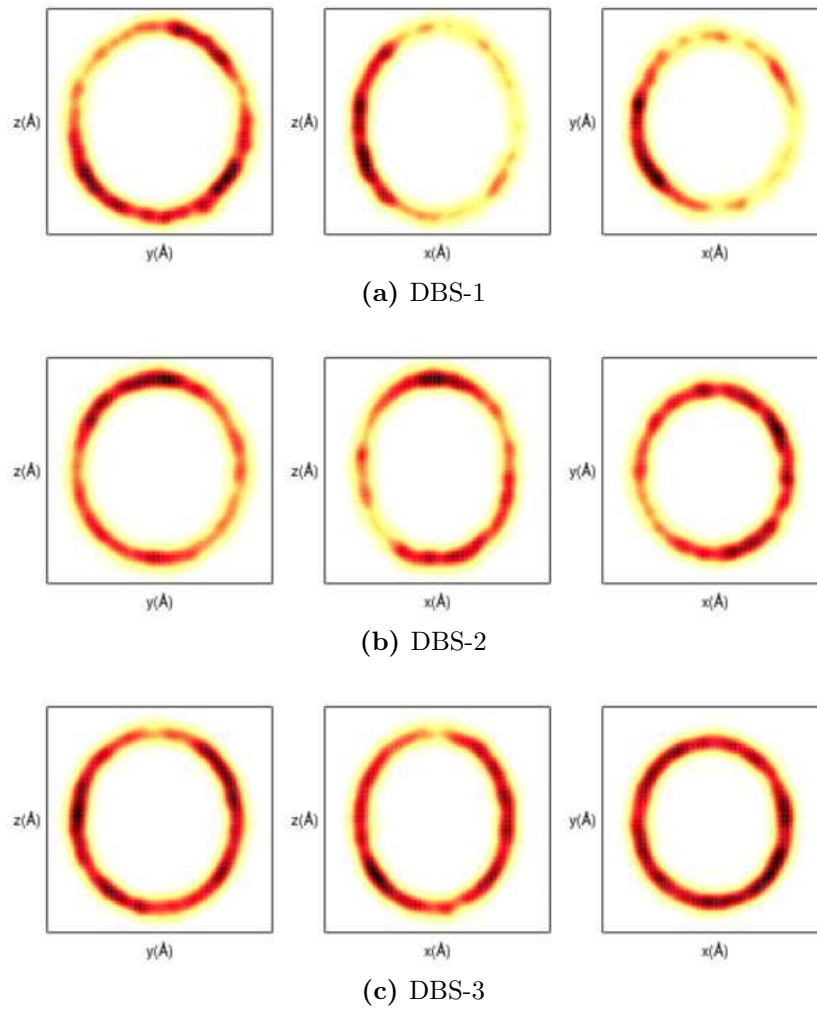
**Figure 4.11:** Radial distribution functions of S beads by isomer type with respect to the COM of the micelle.

a minimal depletion of DBS-3 in the equatorial positions.

This apparent contradiction can be explained with the different size and organization of the hydrophobic core of the micelles. In the l20 system the linear chains tend to pack more regularly (see Figure 4.4a) making difficult for the surfactant to uniformly arrange at the polar position and also giving a slightly bigger cell. On the other hand in the m20 system the different shape and length of the tails causes less steric hindrance and therefore an easier internal reorganization is possible.

### 4.1.3 Conclusions

In this section an application of the proposed CG model to a common surfactant was presented. In particular, the aggregation behavior at three different concentrations of sodium dodecylbenzene sulfonate in water was studied and the obtained results are in accordance with experimental findings. Moreover, the predicting capability of the model is not limited only to the ability to reproduce the correct phase at the different concentrations, but also to reproduce more subtle effects like bilayer swelling. Therefore, although some experimentally observed effects, like coalescence or surfactant exchange between micelles or layers were not observed in these simulations due to periodic boundary conditions and limited cell size, the proposed model proves to reproduce with good accuracy the delicate balance between hydrophobic, hydrophilic and electrostatic interactions.



**Figure 4.12:** Sections of the average density of the S beads by isomer type in the m20 system. Each section is taken on the plane which contains the COM of the micelle and lies perpendicular to the principal axes, after aligning them to the Cartesian axes. To darker color corresponds an higher density; the side of each box measures 54 Å.

## 4.2 SNOVA Membrane

As mentioned in the introduction, the goal of this work is to provide a CG FF to study the morphology of SNOVA membranes (see Section 1.3) under different working conditions in order to understand and to eventually improve their performance as an electrolyte for DMFC. After the parametrization details (Chapter 2) and the tests (Chapter 3) shown previously, the model was considered reliable and was applied to the simulation of the membrane. In this section the initial results of this study are presented, specifically those related to its behavior with respect to solvent uptake and methanol concentration.

Contrary to Nafion, for which a huge literature of experimental data and computer simulations is available, for the PE-g-PSSA membranes only a limited amount of experimental work [11, 103, 104, 105], mainly concerning the polymer preparation, was published [106, 107, 108]. This information on PE-g-PSSA is however useful to build a meaningful model for the SNOVA chain. Firstly, IR spectra confirm that the benzene ring of PS is mono-substituted and secondly it was shown that the crystalline part of the PE is almost completely destroyed in the preparation, resulting in an amorphous polymer [109]. This simplifies the planning of the simulations because no special assumption is required to take into account the non amorphous part.

Specific information for the topology of the SNOVA polymer was lacking, and therefore simulation results and proposed models for other sulfonated membranes were considered. Taking into account the wide variety of chemical structures and properties of the polymers used as electrolytes for PEMFC and the experimental limitations in obtaining clear morphologies, it is not surprising that there are several different models.

Many of them [20], however, are based on the popular early proposal of Yeager and Steck [110] and, later, of Hsu and Gierke [111]. In this model, starting from a dry membrane, first the water clusters around the acid groups then, at increasing hydration levels, inverted micelles appear in the polymer and finally the water pocket grow enough to have transitory or even stable interconnection between them. The size of these water aggregates are difficult to infer but an approximate radius of 2 nm for the clusters and a 1 nm diameter for the cylindrical channels was considered reasonable for Nafion [8].

Although this cluster-channel model (as almost all other models) was built for Nafion membranes, and despite the fact that it was also challenged by new experimental interpretations [112], it still provides a clear general picture. In particular, this model distinguishes two regimes for the membrane processes and morphology: a low hydration

state, corresponding to 2 or less water molecules per sulfonic group ( $\lambda < 2$ ), where the solvent clusters are not interconnected, and high hydration state ( $\lambda \sim 15$ ) where the channels connecting the clusters are quite stable and build a continuous network.

### 4.2.1 System setup

In the light of these data, the membrane morphology and the effect of different hydration levels and of the methanol concentration for chosen for the initial studies. To this end four systems were simulated, all consisting in eight SNOVA chains solvated in different methanol-water mixtures. Each SNOVA chain (see also Figure 1.1) consists in a main chain of linear PE, modeled by 2224 C3 beads, which is grafted by four roughly equidistant PSSA side chains, which are composed of 125 C2-B-S repeating units each. Finally due the  $-e$  charge present on the S sites, each SNOVA chain was neutralized by 500 Na atoms.

The chains were solvated in four solvent mixtures two of them consisting of pure water and the other two of a methanol-water mixture with a 2:5 ratio, corresponding to a 12 M solution. For each type of solvent two mixtures were near the maximum solvent uptake of the membrane and the other two were near the percolation limit [113].

A summary of the details for these systems is presented in Table 4.1.

**Table 4.1:** Label and composition details of the membranes studied in this section. The structure of each chain is described in the text (see also Figure 1.1). The uptake is defined as the percentile ratio between the solvent weight and weight of the dry membrane;  $\lambda$  refers to the number of solvent molecule per acid groups (S beads).

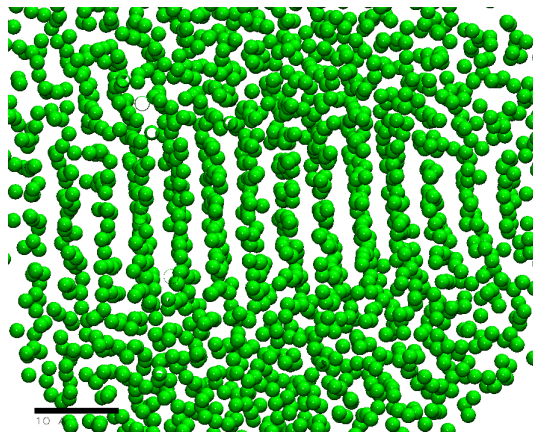
Label	# chains	W beads	M beads	Uptake	$\lambda$
8c70m3	8	33728	13504	70	11.8
8c70m0	8	47232	0	57	11.8
8c07m3	8	4288	1712	9	1.5
8c07m0	8	6000	0	7	1.5

The preparation and relaxation of such kind of systems is problematic, because long flexible polymeric chains tend to have extremely long relaxation times. Therefore, the exploration of all the conformation and structures is not possible and the sample should be prepared so that its starting configuration satisfies some constraint, *e.g.*, has a given number of entanglements or crystalline structures. It is important to note that this problem does not affect not only the simulations, although the short time scale that can be explored makes it more acute, but also the preparation of experimental samples.



Due to the lack of structural information or experimental data to replicate, the samples were prepared in the following way. Starting from the 8c70m3 system and after cutting each SNOVA chain in four equivalent pieces, the box was equilibrated at a low pressure (0.5 atm) in an NPT ensemble for 1 ns to ease chain reorganization. At this point the number of chain fragments were halved by joining the nearest two ends of different chains and then the system was relaxed for another 2 ns under the same conditions. After these steps, the final chain topology was built by repeating the last step and relaxing at 1 atm for 5 ns.

The other systems were prepared starting from this configuration by removing the excess solvent or by converting the solvent beads. The four systems were first heated to 500 K for 5 ns and then annealed by reducing the temperature to 325 K in 10 ns in an NVT ensemble. Finally the four boxes were relaxed for further 20 ns at 1 atm at 325 K and then the production simulations were run in a canonical ensemble for 25 ns.



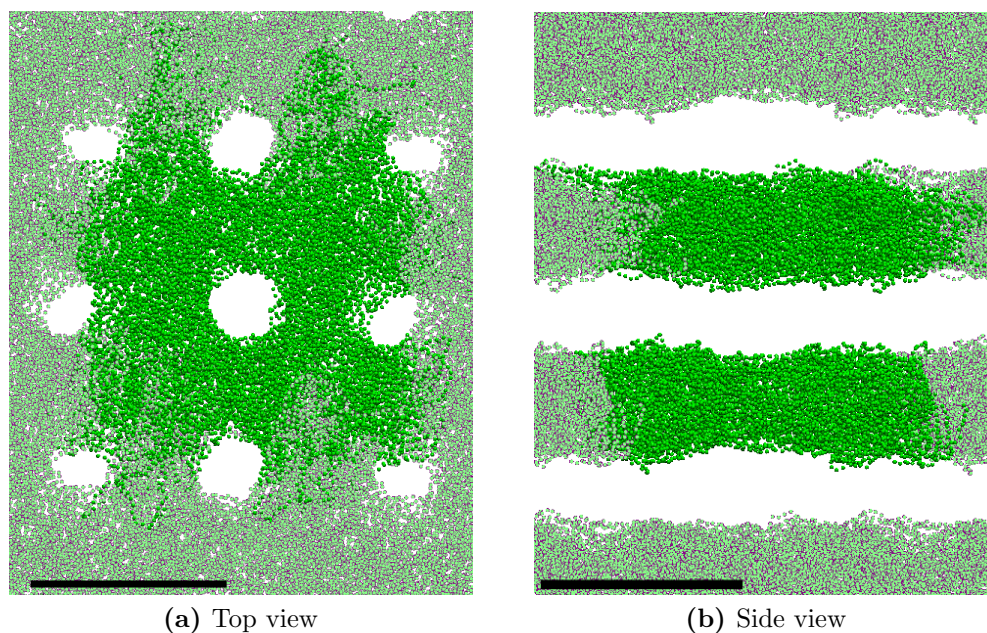
**Figure 4.13:** Details of a crystalline region of the PE main chain in the 8c70m3 system, present during the relaxation; the lammellae formed by the chain are only partially visible because are not planar. The green spheres represent the C3 groups; the black bar has a length of 10 Å. The graphics were generated using VMD [100].

It should be stressed that the main goal of this procedure is to allow some local reorganization and minimize the correlation between the configurations and not to guarantee completely different initial structures. The latter goal is not only very difficult to achieve, but also not interesting in this analysis because somewhat similar structures are needed, in order to compare structural change related to the solvent. Therefore, by starting from a low-density and low-viscosity system and building the chain by steps a low degree of entanglement was possible, while the heating and annealing guaranteed a low crystalline

content of the PE main chain, which before this procedure was quite higher (see Figure 4.13) than expected (here a 50% side chains grafting was assumed, which corresponds to a 5% of crystalline content [109]).

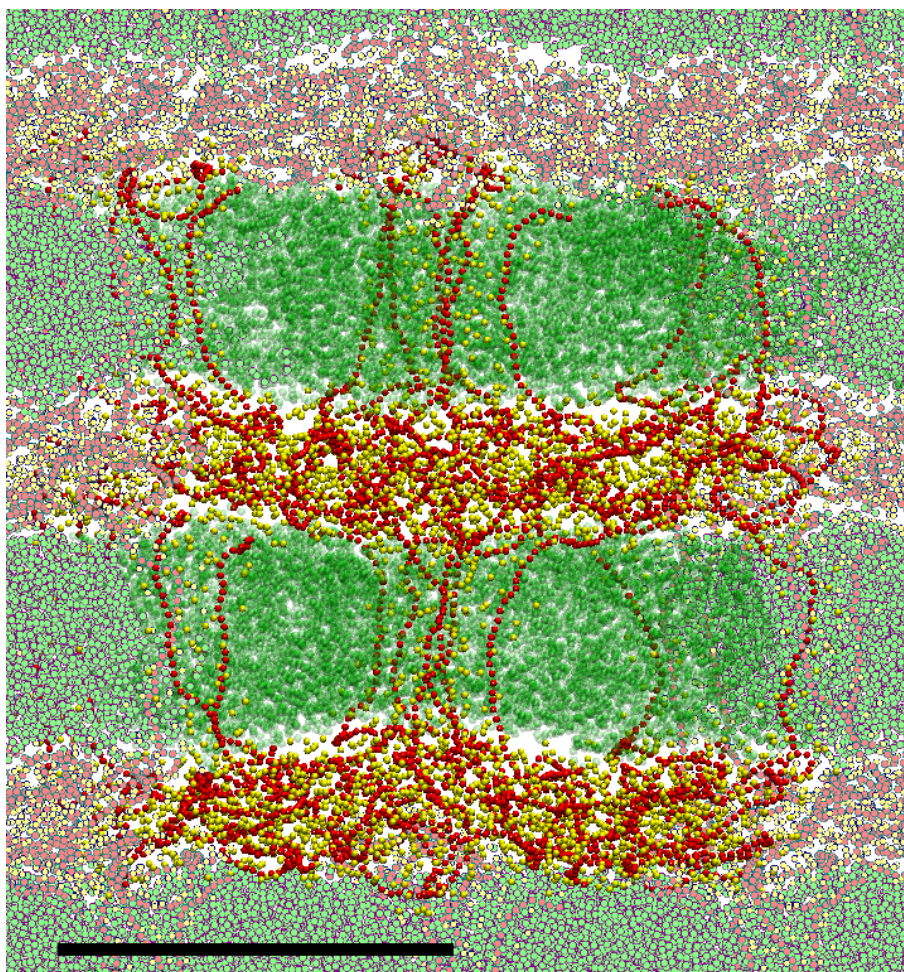
## 4.2.2 Results

As expected from the relaxation procedure and the slow relaxation times typical of these systems, the four membranes formed similar structures. In particular, the main chain, separates into an hydrophobic phase organized in layers with roughly cylindrical holes (see Figure 4.14). These holes, or channels, which are only loosely aligned between layers, are stable during the simulation and do not appreciably change or fluctuate at any solvent concentration or composition.



**Figure 4.14:** Top and side view of the C3 beads of the relaxed 8c70m3 system. The darker green represents the actual cell, whereas the pale green represents the near periodic images; the black bar on bottom has a length of 100 Å. The graphics were generated using VMD [100].

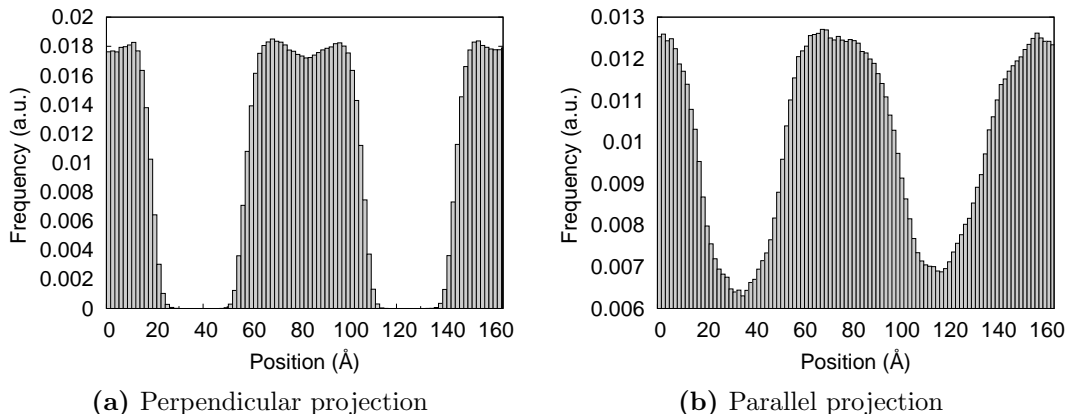
The PSSA side chain, as well as the solvent and the counterion mix together quite uniformly forming the other phase which fills the rest of the space. The PSSA side chains tend to be coiled in the interlayer phase but are almost completely stretched inside the channels as shown in Figure 4.15.



**Figure 4.15:** Section of the 8c70m3 system highlighting the position of the side PSSA chains. In red the C2 groups and in yellow the S groups; B, W, M and Na groups were hidden for clarity. As for Figure 4.14 in C3 beads are shown in green, while periodic images are reproduced in paler colors; the black bar on bottom has a length of 100 Å. The graphics were generated using VMD [100].

From the analysis of the trajectories of the four systems, it is clear that the changes in hydration level or in the solvent composition do not qualitatively change the aggregation structure, although the distance between the layers and the channel diameter are influenced by these factors.

To measure these quantities the average distribution of the C3 groups along the Cartesian axis were computed: the projection on the z-axis, which is perpendicular to the layers, is useful to measure their distance, while the other two distributions can be used to estimate the average channel diameter. Two typical distributions are shown in Figure 4.16.



**Figure 4.16:** Trajectory average of the C3 group distribution along perpendicular and parallel axes with respect to the membrane plane for the 8c70m3 system.

The distances, which are listed in Table 4.2, were computed by averaging the separation at half height between two successive peaks. Despite the fact that thermal fluctuations and the partial alignment of the holes in the two layers tends to flatten the histogram, therefore making the measure of the pore diameter only partially accurate, the results are still interesting.

**Table 4.2:** Channel diameter and layer distances for the SNOVA systems. All the values have an uncertainty of 1 Å.

System	8c70m3	8c70m0	8c07m3	8c07m0
Layer distance (Å)	38	33	24	22
Channel diameter (Å)	36	31	25	22

First, the numbers show an expected decrease in the layer distance and in the channel size at lower hydration level, but surprisingly the two quantities decrease in a similar way. This suggests that there is no substantial difference between these structures and they likely form in the same way, but can end up in layer or channel just due to the local surroundings giving a model that resembles more the one proposed by Kreuer for Nafion [114] rather than the channel-cluster model.

More surprising is the contraction associated with the substitution of methanol with water at high hydration level (8c70m3 and 8c70m0 systems), because such a marked decrease in the hydrophilic phase volume cannot be ascribed to the change of volume



of solvent. The reason for this phenomena is not clear at the moment, albeit an improved packing of the liquid phase seems to be the most likely explanation.

### 4.2.3 Conclusions

In this last section, some results from the simulation of the SNOVA membrane using the proposed CG model were presented. The systems were prepared and relaxed with a long procedure in order to have comparable but independent systems swollen by different methanol-water solutions. All the systems tend to give the same kind of aggregate: the hydrophilic phase, which includes the PSSA side chains as well as the solvent and the sodium ions, separates in layers that are connected by stable channels. Although the channel diameter and interlayer distance decreases at lower hydration level the channels are always stable and the general structure does not change even for  $\lambda = 1.5$ . Therefore, besides the details of these configurations, which do not likely represent all the relevant possible aggregates present in the membrane, the results suggest that the systems do not follow the cluster-channel model, but build a more uniform and continuous phase even at very low level of hydration.



---

# Chapter 5

## Thesis summary

Computational techniques are becoming increasingly mature as scientific tools that support and complement experimental investigations. The development and refinement of new procedures, the increase of computational power and, ultimately, their widespread successful application are the main reasons for the gradual transition of these methods from a scientific niche to an accepted general investigation tool. Already many research fields benefit from computational techniques, but still a big effort is dedicated to widen the range of their applications and to shed light on phenomena that are otherwise difficult to interpret. For example, there is a wide variety of systems that are difficult to study experimentally simply because the direct observation of phenomena in highly inhomogeneous samples at the relevant scale is very difficult or expensive, if at all possible. These systems include proteins, polymers and membranes, which are of extreme scientific and technological interest but that are in many respects still not completely understood.

On the other hand, such kinds of systems are difficult to simulate because several of the important phenomena they exhibit entail large length- and time-scales. However, many details on the atomistic scale are often irrelevant because their behavior is often determined by a small number of key properties, *e.g.*, hydrophobicity or flexibility of chains. This combination of factors makes these kinds of systems one of the preferred subject for the application of coarse-grained (CG) models in the soft matter field.

For these reasons, we decided to use coarse-grain techniques to study the behavior of a PE-g-PSSA membrane called SNOVA, which is one of the possible candidates to substitute Nafion in direct methanol fuel cells. Thus, first we built and tested the force field (FF) for a new reduced representation of the system and then we applied it to the study of these membranes.

---

Due to the impossibility of a CG model to fully reproduce all the properties of a fully atomistic representations, particular care is needed in the choice of both the degrees of freedom that are removed and in the FF building. Therefore, we kept a more detailed description of the hydrophilic parts of the polymer, because they more likely form complex structures with the solvent, and selected different parametrization methodologies in order to better capture the key properties of the real membrane (Section 2.1).

Bonded interactions were approximated with harmonic potentials, with particular attention to reduce the inclusion in these interaction of spurious terms and to obtain a fast and reproducible convergence (Section 2.2). Despite the simple analytical form of these potentials and the fact that only small molecules were used during the parametrization, the comparison with AA predictions of the radius of gyration of very long PE chains shows that the interactions are modeled with good accuracy (Section 3.1).

Non-bonded interactions between neutral species were modeled in two different ways: for all the species that give homogeneous phases, we used a modified iterative Boltzmann inversion (IBI) method (Section 2.3), whereas for the other system, as well as for water and methanol, we fitted LJ potentials to match macroscopic properties, like density and surface tension (Section 2.4).

The standard IBI procedure was improved by constraining the system to a target pressure at every iteration and by reducing the statistical and numerical noise of the tabulated potentials by using Bézier cubic splines to fit all the correction terms. These modifications have a negligible computational cost but are shown to give a faster and smoother convergence. Moreover, when testing the ability of these potentials to correctly reproduce other observables, like density, isothermal compressibility and surface tension, we obtained a very good general agreement with experimental values. The higher discrepancies are found in molecules with highly anisotropic interactions, *e.g.*, benzene or styrene, but the differences are not very large and therefore the CG potentials are useful. Our tests show that the modified IBI procedure builds interactions that have a systematic and consistently improved agreement with experimental densities with respect to the AA FF, which were used to compute the target pair radial distribution functions, likely due to the additional pressure constraint.

For water and methanol we choose not to use IBI procedure, because tabulated potentials cannot be combined to give cross-interaction potentials. This fact, added to the density dependence of the tabulated potential, would have forced us to recompute methanol-water interactions for each concentration. We instead opted to build a correla-



tion map between some experimental properties (density, surface tension and isothermal compressibility) and the parameter of LJ potentials, making it very easy to parametrize not only these liquids but also to extend this procedure to any liquid without any further calculation. This approach guarantees a reduced computational effort in computing methanol-water interactions and an excellent agreement between CG and experimental properties of the pure solvents. However, an even more important point for the simulation of the membrane in methanol-water solutions is the ability of the CG FF to reproduce the non-ideal behavior of such mixtures along the whole composition range. Despite not having been directly considered in the parametrization procedure, the tests show that different key properties of these solutions are correctly reproduced, with the only exception of isothermal compressibility at high methanol content ( $x_W < 0.3$ ). The comparison of different potentials shows also that these discrepancies can likely be reduced with a different choice of LJ exponents. However, considering the good agreement achieved at high water molar fraction, which represent the normal conditions for the fuel cell membranes, we did not further refine these interactions.

Charge-charge interactions were simply scaled by the estimate of the average dielectric constant, whereas to model the interactions between charged groups and overall neutral beads that have dipole or higher order electrostatic interactions, we tried different procedures (Section 2.5). At first, we tuned the LJ potentials parameter in order to reproduce the free energy of solvation of the charged species in various solvents obtained in the AA representation. However, the interactions obtained in this manner result in incorrect or unphysical behaviors, like solvent vitrification or water-methanol separation. This possibly happens due to long range phenomena that give crucial contribution to the free energy of solvation, *e.g.*, dipole-dipole interactions and solvent reorganization, that cannot be reproduced because of the inherent limitations of the CG model. Therefore as an alternative solution, we refined the potentials so that the CG rdfs of the ions in different solvents match the corresponding AA rdfs as closely as possible.

After parameterizing and testing single parts of the CG interactions, we tested the global behavior of our CG model by studying the aggregation behavior of sodium dodecylbenzene sulfonate, a common surfactant, in water (Section 4.1). The simulations showed the ability of the FF to reproduce not only the experimental aggregation phase at different concentrations, but also more complex phenomena like the swelling of the bilayers at lower SDBS content. Moreover, the model is accurate enough to reveal the effect of different isomer surfactant compositions, which are reflected for example in the

---

shape the internal morphologies of the micelles, potentially giving important insights to structures and dynamics that are extremely difficult to study experimentally.

Finally, we applied our model to the study of the SNOVA membrane and to the effect of different hydration levels and methanol content on its morphology (Section 4.2). In these simulations, we observe the formation of a two phase system, where the solvent phase build layers connected by stable channels, which have a diameter of 2-4 nm depending on the solvent uptake. The fact that this structure does not substantially change even at very low hydration levels, *i.e.*, 1.5 solvent molecules per acid group, and that the contraction in the hydrophilic phase is the same between layers and channels, strongly suggests that the solvent does not build a cluster-channel structure but a more uniform phase which is continuous even at low hydration levels.

In conclusion, we present a modeling method that allows fast parametrization and easy customization of CG FF, by combining different methodologies so as to better tune the reproduced properties. The model obtained not only gives a two orders of magnitude computational gain with respect to AA simulations, but is also able to reproduce key aspects of complex systems, like non-ideal mixtures, polyelectrolytes and polymeric membranes, therefore proving to be a useful tool to predict and to shed light on aspects that are outside the reach of the actual experimental techniques.

---

# Appendix A

## Bond and bend fitting tables

The following tables show all the fitting parameters of AA distance and angle distributions of single molecules in vacuum (more details in section 2.2). Note that some of the values are reported for comparison only and were not actually used to determine the force field.

**Table A.1:** Gaussian fitting parameters of angle distributions generated from AA MD simulations of single molecules in vacuum.

Interaction	Molecule	$k_a$ (K rad <sup>-2</sup> )	$\theta_0$ (deg)
B-C2-C2	<i>n</i> -butylbenzene	4(.5) 10 <sup>2</sup>	180(.5)
C2-B-S	4-ethyl-benzenesulfonate	4.22(5) 10 <sup>4</sup>	170(.5)
C2-C2-C2	<i>n</i> -hexane	3.5(5) 10 <sup>3</sup>	180(.5)
C2-C2-C2	<i>n</i> -octane	3.8(5) 10 <sup>3</sup>	180(.5)
C2-C2-C2	<i>n</i> -decane	3.7(5) 10 <sup>3</sup>	180(.5)
C2-C2-C2	<i>n</i> -dodecane	3.8(5) 10 <sup>3</sup>	180(.5)
C2-C2-C2	C <sub>60</sub> H <sub>122</sub>	2.3(5) 10 <sup>3</sup>	180(.5)
C2-C2-C2	C <sub>120</sub> H <sub>242</sub>	2.4(5) 10 <sup>3</sup>	180(.5)
C2-C2-C2	C <sub>300</sub> H <sub>602</sub>	2.4(5) 10 <sup>3</sup>	180(.5)
C2-C2-C3	2-propyl-1-butylether	4.0(5) 10 <sup>3</sup>	142(.5)
C2-C3-C3	ethyl-5- <i>n</i> -nonylether	1.13(5) 10 <sup>3</sup>	95(.5)
C3-C3-C3	<i>n</i> -nonane	1.2(5) 10 <sup>3</sup>	180(.5)
C3-C3-C3	<i>n</i> -dodecane	1.2(5) 10 <sup>3</sup>	180(.5)
C3-C3-C3	<i>n</i> -pentadecane	1.1(5) 10 <sup>3</sup>	180(.5)
C3-C3-C3	ethyl-5- <i>n</i> -nonylether	3.4(5) 10 <sup>3</sup>	144(.5)
C3-C3-C3	C <sub>60</sub> H <sub>122</sub>	1.6(5) 10 <sup>3</sup>	152(.5)
C3-C3-C3	C <sub>120</sub> H <sub>242</sub>	1.9(5) 10 <sup>3</sup>	153(.5)
C3-C3-C3	C <sub>300</sub> H <sub>602</sub>	1.1(5) 10 <sup>3</sup>	180(.5)

**Table A.2:** Gaussian fitting parameters of distance distributions generated from AA MD simulations of single molecules in vacuum.

Interaction	Molecule	$k_b$ (K $\text{\AA}^{-2}$ )	$r_0$ ( $\text{\AA}$ )
B–C2	ethylbenzene	$8.31(5) \cdot 10^4$	$3.41(5)$
B–C2	<i>n</i> -butylbenzene	$8.28(5) \cdot 10^4$	$3.38(5)$
B–C2	4-ethylbenzenesulfonate	$6.97(5) \cdot 10^4$	$3.38(5)$
B–S	benzenesulfonate	$1.47(5) \cdot 10^5$	$3.44(5)$
B–S	4-ethylbenzenesulfonate	$1.47(5) \cdot 10^5$	$3.41(5)$
C2–C2	<i>n</i> -butane	$1.16(5) \cdot 10^4$	$2.56(5)$
C2–C2	<i>n</i> -hexane	$1.31(5) \cdot 10^4$	$2.53(5)$
C2–C2	<i>n</i> -octane	$1.32(5) \cdot 10^4$	$2.52(5)$
C2–C2	<i>n</i> -decane	$1.19(5) \cdot 10^4$	$2.50(5)$
C2–C2	<i>n</i> -dodecane	$1.58(5) \cdot 10^4$	$2.54(5)$
C2–C2	<i>n</i> -butylbenzene	$1.49(5) \cdot 10^4$	$2.55(5)$
C2–C2	$C_{60}H_{122}$	$1.01(5) \cdot 10^4$	$2.46(5)$
C2–C2	$C_{120}H_{242}$	$1.20(5) \cdot 10^4$	$2.49(5)$
C2–C2	$C_{300}H_{602}$	$1.20(5) \cdot 10^4$	$2.50(5)$
C2–C3	ethyl-2-propylether	$9.27(5) \cdot 10^4$	$3.01(5)$
C2–C3	2-propyl-1- <i>n</i> -butylether	$9.17(5) \cdot 10^4$	$2.98(5)$
C2–C3	ethyl-5- <i>n</i> -nonylether	$7.00(5) \cdot 10^4$	$3.01(5)$
C3–C3	<i>n</i> -hexane	$5.5(5) \cdot 10^3$	$3.68(5)$
C3–C3	<i>n</i> -nonane	$6.0(5) \cdot 10^3$	$3.64(5)$
C3–C3	<i>n</i> -dodcane	$5.6(5) \cdot 10^3$	$3.63(5)$
C3–C3	<i>n</i> -pentadecane	$5.9(5) \cdot 10^3$	$3.62(5)$
C3–C3	ethyl-5- <i>n</i> -nonylether	$6.2(5) \cdot 10^3$	$3.59(5)$
C3–C3	$C_{60}H_{122}$	$5.1(5) \cdot 10^3$	$3.55(5)$
C3–C3	$C_{120}H_{242}$	$6.4(5) \cdot 10^3$	$3.57(5)$
C3–C3	$C_{300}H_{602}$	$5.4(5) \cdot 10^3$	$3.63(5)$

---

# Appendix B

## Force field summary tables

**Table B.1:** Bonding potential parameters.

Bond	System	$k_b$ (K $\text{\AA}^{-2}$ )	$r_0$ ( $\text{\AA}$ )
B–C2	ethylbenzene, <i>n</i> -butylbenzene, 4-ethylbenzene sulfonate	$8.33(5) \cdot 10^4$	$3.41(5)$
B–S	benzene sulfonate, 4-ethylbenzene sulfonate	$1.47(5) \cdot 10^5$	$3.42(5)$
C2–C2	<i>n</i> -butane, <i>n</i> -hexane, <i>n</i> -octane, <i>n</i> -decane, <i>n</i> -dodecane	$1.20(5) \cdot 10^4$	$2.50(5)$
C2–C3	ethyl-2-propylether, 2-propyl-1- <i>n</i> -butylether, ethyl-5- <i>n</i> -nonylether	$8.40(5) \cdot 10^4$	$3.00(5)$
C3–C3	<i>n</i> -hexane, <i>n</i> -nonane, <i>n</i> -dodecane, <i>n</i> -pentadecane	$5.7(5) \cdot 10^3$	$3.60(5)$

**Table B.2:** Bending potential parameters.

Angle	System	$k_a$ (K $\text{rad}^{-2}$ )	$\theta_0$ (deg)
B–C2–C2	<i>n</i> -butylbenzene	$4(.5) \cdot 10^2$	$180(.5)$
C2–B–S	4-ethylbenzene sulfonate	$4.22(5) \cdot 10^4$	$170(.5)$
C2–C2–C2	<i>n</i> -hexane, <i>n</i> -octane, <i>n</i> -decane, <i>n</i> -dodecane	$3.8(5) \cdot 10^3$	$180(.5)$
C2–C2–C3	2-propyl-1-butylether	$4.0(5) \cdot 10^3$	$142(.5)$
C2–C3–C3	ethyl-5- <i>n</i> -nonylether	$1.13(5) \cdot 10^4$	$95(.5)$
C3–C3–C3	<i>n</i> -nonane, <i>n</i> -dodecane, <i>n</i> -pentadecane	$1.1(5) \cdot 10^3$	$180(.5)$

**Table B.3:** Non-bonded potential parameters

Species		Method	Pot.	$\sigma$ (Å)	$\epsilon$ (K)
B	B	IBI		<i>Tabulated</i>	
B	C2	IBI		<i>Tabulated</i>	
B	C3	IBI		<i>Tabulated</i>	
B	M	IBI		<i>Tabulated</i>	
B	Na	Rdf comparison	LJ9-6	2.38(5)	325(.5)
B	S	Rdf comparison	LJ9-6	5.62(5)	180(.5)
B	W	Interfacial tension	LJ12-4	3.25(5)	199(.5)
C2	C2	IBI		<i>Tabulated</i>	
C2	C3	IBI		<i>Tabulated</i>	
C2	M	Interfacial tension	LJ9-6	4.88(5)	140(.5)
C2	Na	Rdf comparison	LJ9-6	4.65(5)	5(.5)
C2	S	Rdf comparison	LJ9-6	4.53(5)	140(.5)
C2	W	Interfacial tension	LJ9-6	4.22(5)	145(.5)
C3	C3	IBI		<i>Tabulated</i>	
C3	M	Interfacial tension	LJ9-6	5.03(5)	174(.5)
C3	Na	Rdf comparison	LJ9-6	4.40(5)	40(.5)
C3	S	Rdf comparison	LJ9-6	4.70(5)	185(.5)
C3	W	Interfacial tension	LJ9-6	4.58(5)	162(.5)
M	M	Density, surface tension	LJ12-4	4.012(5)	207(.5)
M	M	Density, surface tension	LJ9-6	3.720(5)	292(.5)
M	Na	Rdf comparison	LJ9-6	2.75(5)	675(.5)
M	S	Rdf comparison	LJ9-6	3.68(5)	585(.5)
M	W	LB rules	LJ12-4	3.52	212
Na	Na	see ref [67]	LJ12-4	2.35	65
Na	S	LB rules	LJ12-4	2.96	57
Na	W	Rdf comparison	LJ12-4	2.75(5)	675(.5)
S	S	LB rules	LJ 12-4	3.56	52
S	W	Rdf comparison	LJ 9-6	3.55(5)	660(.5)
W	W	Density, surface tension	LJ12-4	3.037(5)	217(.5)

---

# Abbreviation and symbols

AA	All atom
CG	Coarse-grain
COM	Center of mass
DMFC	Direct methanol fuel cell
DPD	Dissipative particle dynamics
FEP	Free energy perturbation
FC	Fuel cell
IBI	Iterative Boltzmann inversion
IEC	Ion exchange capacity
LB	Lorentz-Berthelot (combining rules)
LJ	Lennard-Jones
MC	Monte Carlo
MD	Molecular dynamics
MFEP	Multi-windows free energy perturbation
NPT	Isothermal-isobaric ensemble
NVE	Microcanonical ensemble
NVT	Canonical ensemble
PAFC	Phosphoric acid fuel cell
PE	Polyethylene
PEMFC	Proton exchange membrane fuel cell
PMF	Potential of mean force
PS	Polystyrene
PSSA	Polystyrene sulfonic acid
rdf	Radial distribution function
RESPA	Reversible reference system propagator algorithms
SDBS	Sodium dodecylbenzene sulfonate

---

SDS Sodium dodecyl sulfate, also know as sodium lauryl sulfate  
VdW Van der Waals

$\Delta A_{solv}$	Excess free energy
$\mathcal{H}$	Hamiltonian of the system
$L$	Side length of the simulation box ( $\text{\AA}$ )
$k_B$	Boltzmann constant ( $1.3806504 \text{ J/K}$ )
$k_a$	Harmonic constant for the angle potential
$k_b$	Harmonic constant for the bond potential
$P$	Pressure ( <i>atm</i> )
$\mathcal{P}$	Membrane permeability ( $\text{mol/s} \cdot \text{cm}^2$ )
$P(r), P(\theta)$	Distance and angle distributions
$q$	Electric charge of the atom or bead
$r_0, \theta_0$	Equilibrium distance and angle for harmonic potential
$R_g$	Radius of gyration; compact form for $\sqrt{\langle R_{gyr}^2 \rangle}$
$T$	Temperature ( $K$ or $^\circ\text{C}$ )
$U$	Potential energy
$V$	Volume
$\Delta V^E$	Excess volume of mixing
$\beta$	$1/k_B T$ , inverse of the thermal energy
$\beta_T$	Isothermal compressibility ( $1/\text{GPa}$ )
$\gamma$	Surface tension ( $10^{-3} \text{ N/m}$ )
$\Gamma$	Important phase space
$\varepsilon$	Parameter of LJ potentials related to the interaction strength (K)
$\epsilon_r$	Relative dielectric constant
$\xi_{EW}$	Self-energy of a point charge in a cubic Wigner lattice
$\rho$	Numeral or volumetric density
$\sigma$	Parameter of LJ potentials related to the average distance ( $\text{\AA}$ )
$\sigma_H$	Proton conductivity ( $S/\text{cm}$ )
$\chi^2$	Chi square distribution
$\chi_s^2$	Average of the residuals



---

# Acknowledgments

Completing a PhD is truly a long and painful odyssey, and I would not have been able to complete this journey without the aid and support of countless people over the past years. I must first express my gratitude to my supervisor Professor Michele Parrinello for giving me the opportunity to do this work and for being patient during the delays and difficulties I encountered during these years. A very special thanks goes to Professor Jim Pfaendtner for providing a wealth of advice, for being always optimistic and encouraging and for the time “well beyond the call of duty” he spent helping me. I would also like to express my gratitude to Dr. Meher Prakash Ayalasomayajula who even in the most hectic moments always found the time and the patience to help me.

I would also like to express my thanks to the people who have been very helpful to me during the time it took me to write this thesis: Dr. Alessandro Barducci for his critical questions and the refreshing insight on the most different topics, Massimiliano Bonomi for his irony, common sense and practical point of view, Gloria Mazzone for patiently standing my swinging mood and having always an encouraging word for me, and Michele Ceriotti for helping me with the most improbable and strange problems.

Over the years, I have enjoyed the aid and friendship of several people that I would like to thank: Dr. Matteo Masetti, Daniela Wirz, Dr. Fabien Bruneval, Giacomo Miceli, Sebastiano Caravati, Dr. Davide Branduardi, Dr. Davide Donadio and Dr. Paolo Raiteri. Finally, I would also like to extend my thanks to all those who worked or visited the group in the last five years providing an always stimulating and friendly environment.



---

# List of Figures

1.1	Chemical structure and properties of SNOVA . . . . .	8
2.1	Dodecane bond distribution . . . . .	16
2.2	Dodecane angle distribution . . . . .	17
2.3	Pressure convergence in a hexane box for the IBI . . . . .	24
2.4	Convergence comparison with different optimization of the IBI . . . . .	25
2.5	Convergence of the potential for CG hexane with the IBI method . . . . .	27
2.6	Tabulated potential for CG alkanes non-bonded interactions . . . . .	28
2.7	Tabulated potential for CG benzene interaction . . . . .	29
2.8	Volume of a LJ96 liquid as function of $\varepsilon$ and $\sigma$ . . . . .	31
2.9	Isothermal compressibility of a LJ96 liquid as function of $\varepsilon$ and $\sigma$ . . . . .	32
2.10	Surface tension of a LJ96 liquid as function of $\varepsilon$ and $\sigma$ . . . . .	32
2.11	Interfacial tension for CG species as function of $\varepsilon$ . . . . .	33
2.12	Thermodynamic cycle used to compute $\Delta A_{solv}$ . . . . .	37
2.13	Solvation free energy . . . . .	39
2.14	$\Delta A_{solv}$ of $\text{Na}^+$ CG beads as function of $\varepsilon$ . . . . .	40
2.15	AA and CG rdfs for charged-neutral site interactions . . . . .	41
3.1	Radius of gyration of AA and CG PE chains . . . . .	45
3.2	Experimental and CG density of benzene as a function of temperature. . . . .	48
3.3	Experimental and CG density of hexane as a function of temperature. . . . .	48
3.4	Density of methanol-water mixture . . . . .	50
3.5	Excess volume for CG methanol-water solutions . . . . .	51
3.6	Surface tension of methanol-water mixture . . . . .	52
3.7	Density of methanol-water mixture . . . . .	52
4.1	Composition of the SDBS systems . . . . .	58

## LIST OF FIGURES

---

4.2	Equilibrium configuration for the m50 system . . . . .	59
4.3	Equilibrium configuration for the m70 system . . . . .	60
4.4	Section of a equilibrium configuration for the l20 and m20 systems . . . . .	60
4.5	Evolution of the principal axes of the micelle . . . . .	61
4.6	Eccentricity of the micelles during the simulation . . . . .	61
4.7	Radius of gyration of the micelles during the simulations . . . . .	62
4.8	Scaled components of a principal axes of the micelle (m20 system) . . . . .	63
4.9	Rdfs with respect to micelle COM for the l20 and m20 systems . . . . .	63
4.10	Sections of the average density of the S beads . . . . .	64
4.11	Rdf of S beads with respect to micelle COM . . . . .	65
4.12	Sections of the average density of the S beads by isomer type . . . . .	66
4.13	Crystalline part of the main chain . . . . .	69
4.14	Structure of C3 beads in the 8c70m3 system . . . . .	70
4.15	Structure of the 8c70m3 system . . . . .	71
4.16	Projection of the C3 frequency along parallel and perpendicular axes . . . . .	72

---

# List of Tables

1.1	Data for different types of fuel cell . . . . .	2
2.1	Mapping between AA and CG units . . . . .	12
2.2	Results of different fitting methods for the C2–C2–C2 potential . . . . .	15
2.3	Consistency test of the CG parameters . . . . .	15
2.4	System setting used to test different IBI method corrections . . . . .	23
2.5	Pressure, convergence and drift for different IBI methods . . . . .	24
2.6	IBI system settings . . . . .	26
2.7	$\Delta A_{solv}$ of $\text{Na}^+$ and $\text{SO}_3^-$ in different solvents . . . . .	39
3.1	Comparison of $\rho$ , $\beta_T$ and $\gamma$ of CG, AA and experimental systems . . . . .	46
3.2	Total energy drift as function of the time step . . . . .	53
3.3	Performance of MD simulation in the AA and CG representations . . . . .	54
4.1	Composition of the simulated membranes . . . . .	68
4.2	Channel diameter and layer distances for the SNOVA systems . . . . .	72
A.1	Full list of angle distributions fitting parameters . . . . .	79
A.2	Full list of distance distribution fitting parameters . . . . .	80
B.1	Bonding potential parameters. . . . .	81
B.2	Bending potential parameters. . . . .	81
B.3	Non-bonded potential parameters . . . . .	82



---

# Bibliography

- [1] Larminie, J. and Dicks, A. (2003) Fuel cell systems explained, Wiley New York, 1st edition.
- [2] Sammes, N. M., Smirnova, A., and Vasylyev, O. (2005) Fuel Cell Technologies: State and Perspectives, Springer, 1st edition.
- [3] Basu, S. (2007) Recent trends in fuel cell science and technology, Springer, New York.
- [4] Hoogers, G. (2002) Fuel Cell Technology Handbook, CRC, 1st edition.
- [5] Grove, W. (1839) *Philosophical Magazine* **14**, 127–130.
- [6] Grove, W. (1842) *Philosophical Magazine* **21**, 417–420.
- [7] Bossel, U. (2000) The birth of the fuel cell : 1835 - 1845., European Fuel Cell Forum, Oberrohrdorf.
- [8] Mauritz, K. A. and Moore, R. B. (2004) *Chemical Reviews* **104(10)**, 4535–4586.
- [9] Wasmus, S. and Küver, A. (1999) *Journal of Electroanalytical Chemistry* **461(1-2)**, 14–31.
- [10] Carrette, L., Friedrich, K. A., and Stimming, U. (2001) *Fuel Cells* **1(1)**, 5–39.
- [11] Hatanaka, T., Hasegawa, N., Kamiya, A., Kawasumi, M., Morimoto, Y., and Kawahara, K. (2002) *Fuel* **81(17)**, 2173–2176.
- [12] Casciola, M., Alberti, G., Sganappa, M., and Narducci, R. (2006) *Journal of Power Sources* **162(1)**, 141–145.
- [13] Weber, A. Z. and Newman, J. (2004) *Chemical Reviews* **104(10)**, 4679–4726.
- [14] Berning, T. and Djilali, N. (2003) *Journal of Power Sources* **124(2)**, 440–452.
- [15] Natarajan, D. and Nguyen, T. V. (2003) *Journal of Power Sources* **115(1)**, 66–80.
- [16] Bernardi, D. M. and Verbrugge, M. W. (1992) *Journal of The Electrochemical Society* **139(9)**, 2477–2491.

## BIBLIOGRAPHY

---

- [17] Verbrugge, M. W. and Hill, R. F. (1990) *Journal of The Electrochemical Society* **137(3)**, 893–899.
- [18] Janssen, G. J. M. (2001) *Journal of The Electrochemical Society* **148(12)**, A1313–A1323.
- [19] Rowe, A. and Li, X. (2001) *Journal of Power Sources* **102(1-2)**, 82–96.
- [20] Weber, A. Z. and Newman, J. (2003) *Journal of The Electrochemical Society* **151(2)**, A311–A325.
- [21] Eikerling, M., Kornyshev, A. A., and Stimming, U. (1997) *The Journal of Physical Chemistry B* **101(50)**, 10807–10820.
- [22] Ladd, A. J. C. (1993) *Physical Review Letters* **70(9)**, 1339.
- [23] Fukunaga, H., Takimoto, J., and Doi, M. (2002) *The Journal of Chemical Physics* **116(18)**, 8183–8190.
- [24] Sun, Q. and Faller, R. (2007) *The Journal of Chemical Physics* **126(14)**, 144908–4.
- [25] Hoogerbrugge, P. J. and Koelman, J. (1992) *Europhysics Letters* **19(3)**, 155–160.
- [26] Koelman, J. M. V. A. and Hoogerbrugge, P. J. (1993) *Europhysics Letters* **21(3)**, 363–368.
- [27] Flekkøy, E. G. and Coveney, P. V. (1999) *Physical Review Letters* **83(9)**, 1775.
- [28] Groot, R. and Rabone, K. (2001) *Biophysical Journal* **81(2)**, 725–736.
- [29] Avalos, J. B. and Mackie, A. D. (1997) *Europhysics Letters* **40(2)**, 141–146.
- [30] Espanol, P. (1997) *Europhysics Letters* **40(6)**, 631–636.
- [31] Malevanets, A. and Kapral, R. (2000) *The Journal of Chemical Physics* **112(16)**, 7260–7269.
- [32] Tozzini, V. (2005) *Current Opinion in Structural Biology* **15(2)**, 144–150.
- [33] Cagin, T., Wang, G., Martin, R., Zamanakos, G., Vaidehi, N., Mainz, D. T., and Goddard, W. A. (2001) *Computational and Theoretical Polymer Science* **11(5)**, 345–356.
- [34] Izvekov, S., Parrinello, M., Burnham, C. J., and Voth, G. A. (2004) *The Journal of Chemical Physics* **120(23)**, 10896–10913.
- [35] Reith, D., Meyer, H., and Müller-Plathe, F. (2001) *Macromolecules* **34(7)**, 2335–2345.



## BIBLIOGRAPHY

---

- [36] Marrink, S. J., deVries, A. H., and Mark, A. E. (2004) *Journal of Physical Chemistry B* **108**(2), 750–760.
- [37] Zaopo, A. SUPERNOVA concept (2006) Pirelli Labs SpA - private communication.
- [38] Sun, Q. and Faller, R. (2005) *Computers & Chemical Engineering* **29**(11-12), 2380–2385.
- [39] Sun, Q. and Faller, R. (2006) *Journal of Chemical Theory and Computation* **2**(3), 607–615.
- [40] Lyubartsev, A. P. and Laaksonen, A. (1995) *Physical Review E* **52**(4), 3730–3737.
- [41] Lyubartsev, A. P. and Laaksonen, A. (1999) *The Journal of Chemical Physics* **111**(24), 11207–11215.
- [42] Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., and Schulten, K. (2005) *Journal of Computational Chemistry* **26**(16), 1781–1802.
- [43] Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) *Journal of Computational Chemistry* **25**(9), 1157–1174.
- [44] Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., and Kollman, P. (2003) *Journal of Computational Chemistry* **24**(16), 1999–2012.
- [45] Tuckerman, M., Berne, B. J., and Martyna, G. J. (1992) *The Journal of Chemical Physics* **97**(3), 1990–2001.
- [46] Darden, T., York, D., and Pedersen, L. (1993) *The Journal of Chemical Physics* **98**(12), 10089–10092.
- [47] Martyna, G. J., Tobias, D. J., and Klein, M. L. (1994) *The Journal of Chemical Physics* **101**(5), 4177–4189.
- [48] Feller, S. E., Zhang, Y., Pastor, R. W., and Brooks, B. R. (1995) *The Journal of Chemical Physics* **103**(11), 4613–4621.
- [49] CM3D <http://hydrogen.usip.edu/twiki/bin/view/CM3D/WebHome>.
- [50] Nose, S. (1984) *The Journal of Chemical Physics* **81**(1), 511–519.
- [51] Martyna, G. J., Klein, M. L., and Tuckerman, M. (1992) *The Journal of Chemical Physics* **97**(4), 2635–2643.
- [52] Louis, A. A. (2002) *Journal of Physics: Condensed Matter* **14**(40), 9187–9206.

## BIBLIOGRAPHY

---

- [53] Soper, A. K. (1996) *Chemical Physics* **202(2-3)**, 295–306.
- [54] Tozzini, V., Trylska, J., Chang, C., and McCammon, J. A. (2007) *Journal of Structural Biology* **157(3)**, 606–615.
- [55] Chan, E. R., Striolo, A., McCabe, C., Cummings, P. T., and Glotzer, S. C. (2007) *The Journal of Chemical Physics* **127(11)**, 114102–15.
- [56] Henderson, R. L. (1974) *Physics Letters A* **49(3)**, 197–198.
- [57] Jain, S., Garde, S., and Kumar, S. (2006) *Industrial & Engineering Chemistry Research* **45(16)**, 5614–5618.
- [58] Reith, D., Pütz, M., and Müller-Plathe, F. (2003) *Journal of Computational Chemistry* **24(13)**, 1624–1636.
- [59] Qian, H., Carbone, P., Chen, X., Karimi-Varzaneh, H. A., Liew, C. C., and Müller-Plathe, F. (2008) *Macromolecules* **41(24)**, 9919–9929.
- [60] Akima, H. (1970) *Journal of the ACM* **17(4)**, 589–602.
- [61] Lorentz, G. G. (1986) *Bernstein Polynomials*, Chelsea Publishing Co, New York 2nd edition.
- [62] Sun, Q., Pon, F. R., and Faller, R. (2007) *Fluid Phase Equilibria* **261(1-2)**, 35–40.
- [63] Nielsen, S. O., Lopez, C. F., Srinivas, G., and Klein, M. L. (2003) *The Journal of Chemical Physics* **119(14)**, 7043–7049.
- [64] Shinoda, W., DeVane, R., and Klein, M. L. (2007) *Molecular Simulation* **33(1)**, 27–36.
- [65] Allen, M. P. and Tildesley, D. J. (1987) *Computer Simulation of Liquids*, Oxford University Press, Clarendon.
- [66] Frenkel, D. and Smit, B. (2002) *Understanding Molecular Simulation: From Algorithms to Applications*, Academic Press, San Diego 2nd edition.
- [67] Smith, D. E. and Dang, L. X. (1994) *The Journal of Chemical Physics* **100(5)**, 3757–3766.
- [68] Hansen, J. P. and McDonald, I. R. (1986) *Theory of simple liquids*, Academic Press, London 2nd edition.
- [69] Zwanzig, R. (1954) *The Journal of Chemical Physics* **22(8)**, 1420–1426.
- [70] Chipot, C. and Pohorille, A. (2007) *Free Energy Calculations.*, Springer, Berlin, Heidelberg, 1st edition.

## BIBLIOGRAPHY

---

- [71] Lu, N., Singh, J. K., and Kofke, D. A. (2003) *The Journal of Chemical Physics* **118**(7), 2977–2984.
- [72] Lu, N. and Woolf, T. B. (2004) *Molecular Physics* **102**(2), 173.
- [73] Hummer, G., Pratt, L. R., and Garcia, A. E. (1997) *The Journal of Chemical Physics* **107**(21), 9275–9277.
- [74] Hunenberger, P. H. and McCammon, J. A. (1999) *The Journal of Chemical Physics* **110**(4), 1856–1872.
- [75] Lide, D. R. (2004) *Handbook of Chemistry and Physics a Ready-Reference Pocket Book of Chemical and Physical Data*, CRC Press, Boca Raton 85th edition.
- [76] Albright, P. S. and Gosting, L. J. (1946) *Journal of the American Chemical Society* **68**(6), 1061–1063.
- [77] Mardolcar, U. V., deCastro, C. A. N., and Santos, F. J. V. (1992) *Fluid Phase Equilibria* **79**, 255–264.
- [78] Harrington, E. A. (1916) *Physical Review* **8**(6), 581.
- [79] Bhargava, B. L. and Klein, M. L. (2009) *The Journal of Physical Chemistry A* **113**(10), 1898–1904.
- [80] Easteal, A. J. and Woolf, L. A. (1985) *The Journal of Chemical Thermodynamics* **17**(1), 49–62.
- [81] Sun, T. F., Schouten, J. A., Trappeniers, N. J., and Biswas, S. N. (1988) *The Journal of Chemical Thermodynamics* **20**(9), 1089–1096.
- [82] Matilla, A. D., Aicart, E., Peña, M. D., and Tardajos, G. (1989) *Journal of Solution Chemistry* **18**(2), 143–150.
- [83] Vazquez, G., Alvarez, E., and Navaza, J. M. (1995) *Journal of Chemical & Engineering Data* **40**(3), 611–614.
- [84] Derlacki, Z. J., Easteal, A. J., Edge, A. V. J., Woolf, L. A., and Roksandic, Z. (1985) *The Journal of Physical Chemistry* **89**(24), 5318–5322.
- [85] Mukherjee, A. and Bagchi, B. (2001) *The Journal of Physical Chemistry B* **105**(39), 9581–9585.
- [86] Soper, A. K., Dougan, L., Crain, J., and Finney, J. L. (2006) *The Journal of Physical Chemistry. B* **110**(8), 3472–6.
- [87] Gonzalez-Salgado, D. and Nezbeda, I. (2006) *Fluid Phase Equilibria* **240**(2), 161–166.

## BIBLIOGRAPHY

---

- [88] Smith, M. A. and Davies, P. J. (1985) *Plant Physiology* **78(1)**, 89–91 PMID: PMC1064682.
- [89] Wong, S., Knight, J., Hopfer, S., Zaharia, O., Leach, C., and Sunderman, F. (1987) *Clin Chem* **33(2)**, 214–220.
- [90] Duffie, J. A. and Beckman, W. A. (2006) *Solar Engineering of Thermal Processes*, Wiley, 3th edition.
- [91] Holmberg, B. J. K. (2003) Phase behaviour of concentrated surfactant systems In *Surfactants and Polymers in Aqueous Solution* pp. 67–96 John Wiley & Sons 2 nd edition.
- [92] Hedin, N., Furo, I., and Eriksson, P. O. (2000) *The Journal of Physical Chemistry B* **104(35)**, 8544–8547.
- [93] Closs, G. L., Forbes, M. D. E., and Norris, J. R. (1987) *The Journal of Physical Chemistry* **91(13)**, 3592–3599.
- [94] Vass, S. (2001) *The Journal of Physical Chemistry B* **105(2)**, 455–461.
- [95] Bruce, C., Berkowitz, M., Perera, L., and Forbes, M. (2002) *Journal of Physical Chemistry B* **106(15)**, 3788–3793.
- [96] Floriano, M. A., Caponetti, E., and Panagiotopoulos, A. Z. (1999) *Langmuir* **15(9)**, 3143–3151.
- [97] Sein, A., Engberts, J. B. F. N., van derLinden, E., and van dePas, J. C. (1996) *Langmuir* **12(12)**, 2913–2923.
- [98] Jang, S., Lin, S., Maiti, P., Blanco, M., Goddard, W., Shuler, P., and Tang, Y. (2004) *Journal of Physical Chemistry B* **108(32)**, 12130–12140.
- [99] Stewart, J., Saiani, A., Bayly, A., and Tiddy, G. (2009) *Colloids and Surfaces A: Physicochemical and Engineering Aspects* **338(1-3)**, 155–161.
- [100] Humphrey, W., Dalke, A., and Schulten, K. (1996) *Journal of Molecular Graphics* **14(1)**, 33–38.
- [101] Parrinello, M. and Rahman, A. (1981) *Journal of Applied Physics* **52(12)**, 7182–7190.
- [102] Bruce, C., Senapati, S., Berkowitz, M., Perera, L., and Forbes, M. (2002) *Journal of Physical Chemistry B* **106(42)**, 10902–10907.
- [103] Saarinen, V., Kallio, T., Paronen, M., Tikkanen, P., Rauhala, E., and Kontturi, K. (2005) *Electrochimica Acta* **50(16-17)**, 3453–3460.

## BIBLIOGRAPHY

---

- [104] Tricoli, V., Carretta, N., and Bartolozzi, M. (2000) *Journal of The Electrochemical Society* **147(4)**, 1286–1290.
- [105] Scott, K., Taama, W. M., and Argyropoulos, P. (2000) *Journal of Membrane Science* **171(1)**, 119–130.
- [106] Nasef, M. M., Saidi, H., and Yahaya, A. H. (2005) *J. Applied Membrane Science and Technology* **2**, 47–63.
- [107] Choi, Y. J., Kang, M. S., Kim, S. H., Cho, J., and Moon, S. H. (2003) *Journal of Membrane Science* **223(1-2)**, 201–215.
- [108] Shen, M., Roy, S., Kuhlmann, J. W., Scott, K., Lovell, K., and Horsfall, J. A. (2005) *Journal of membrane science* **251(1-2)**, 121–130.
- [109] Nasef, M. and Saidi, H. (2004) *International Journal of Polymeric Materials* **53**, 1027–1043.
- [110] Yeager, H. L. and Steck, A. (1981) *Journal of The Electrochemical Society* **128(9)**, 1880–1884.
- [111] Hsu, W. Y. and Gierke, T. D. (1983) *Journal of Membrane Science* **13**, 307–326.
- [112] Schmidt-Rohr, K. and Chen, Q. (2008) *Nat Mater* **7(1)**, 75–83.
- [113] Jan, N. (1999) *Physica A: Statistical and Theoretical Physics* **266(1-4)**, 72–75.
- [114] Kreuer, K. D. (2003) Hydrocarbon membranes In *Handbook of Fuel Cells - Fundamentals, Technology and Applications* volume **3**, pp. 420–435 John Wiley & Sons.

