

POLITECNICO DI TORINO



Course of study in Mechanical Engineering

Department of Mechanical and Aerospace Engineering

Master Degree Thesis

**Mercaptosuccinic Acid Ligands
in Polar and Non-Polar
Solvents: a Molecular
Dynamics Study**

Supervisors

Matteo Fasano

Annalisa Cardellini

Candidate

Davide LUNGI

ANNO ACCADEMICO 2019-2020

“Computers are incredibly fast, accurate, and stupid. Human beings are incredibly slow, inaccurate, and brilliant. Together they are powerful beyond imagination.”

Albert Einstein

Contents

List of Tables	6
List of Figures	7
1 Introduction	12
1.1 State of the art	12
1.2 Study purpose	15
2 Materials and methods	16
2.1 Molecular Dynamics	16
2.1.1 Classical and quantum mechanics	16
2.1.2 The Molecular Dynamics approach	18
2.2 Gromacs	23
2.3 Force Fields	24
2.4 Control Volume Configuration	25
2.4.1 The pdb and topology files	25
2.4.2 Procedure and boxes creation	32
2.5 Energy Minimization and Dynamics Runs	37
2.5.1 Energy Minimization	38
2.5.2 Equilibrations	39
2.6 Molecular Dynamics run	42
3 Results	44
3.1 Solvents	44
3.1.1 Temperature and Density	44
3.1.2 Radial Distribution Function	47
3.2 Mercaptosuccinic Acid in solvents	51
3.3 Mercaptosuccinic Acid interactions	54
3.3.1 Coulomb and Lennard Jones Energies	54

3.3.2	Forces	58
3.3.3	Potential of Mean Forces	65
4	Conclusions	72
	Bibliography	74

List of Tables

2.1	Solvents' dielectric constants	28
3.1	Solvents' density comparison between experimental and literature data	46
3.2	Colomb and Lennard-Jones energies acting between one Mercaptosuccinic Acid molecule and each solvent during MD simulation	52
3.3	Minimum energies acting between two Mercaptosuccinic Acid molecules immersed in solvents	57
3.4	Minimum force acting on one Mercaptosuccinic Acid molecule along z-axis direction and corresponding ligands COM distance	62
3.5	Minimum force acting on one Mercaptosuccinic Acid molecule along z-axis direction and corresponding ligands COM distance in presence of electric field	64
3.6	Minimum potential of mean force acting on one Mercaptosuccinic Acid molecule along z-axis direction and corresponding ligands COM distance	67
3.7	Minimum potential of mean force acting on one Mercaptosuccinic Acid molecule along z-axis direction and corresponding ligands COM distance in presence of electric field	70

List of Figures

2.1	Newtonian mechanics applied to Molecular Dynamics	19
2.2	Multi-level parallelism in GROMACS	23
2.3	Protein Data Bank	25
2.4	List of atoms with their spatial position, image taken from Mercaptosuccinic Acid pdb file	26
2.5	Atomtypes of Mercaptosuccinic Acid, image taken from topol- ogy file	27
2.6	Molecule of Chlorobenzene	28
2.7	Molecule of Cyclohexane	29
2.8	Molecule of Hexane	30
2.9	Molecule of Toluene	30
2.10	Molecule of Mercatposuccinic Acid	31
2.11	Molecule of Mercatposuccinid Acid in the box center	33
2.12	Creation of the configurations used for PMF computation	35
2.13	Force acting on Mercaptosuccinic Acid molecules along z-axis	36
2.14	Decrease of Potential Energy during Energy Minimization	38
3.1	Solvents' temperature during MD simulation	45
3.2	Solvents' density during MD simulation	46
3.3	Solvents' Radial Distribution Function from MD simulation	48
3.4	Solvents' Radial Distribution Function from MD simulation in presence of electric field	49
3.5	Oxigen-Oxigen Radial Distribution Function of Water SPC/E model from MD simulation and literature	50
3.6	Oxigen-Oxigen Radial Distribution Function of Water SPC/E model obtained from MD simulation with electric field	50
3.7	Coulomb energy trend value between a Mercaptosuccinic Acid molecule and toluene during MD simulation	51

3.8	Percentage values of Coulomb and Lennard-Jones energies acting between one Mercaptosuccinic Acid molecule and the solvents	52
3.9	Colomb and Lennard-Jones energies acting in a box containing one Mercaptosuccinic Acid molecule and a solvent	53
3.10	Colomb energy averages acting at different COM distances between two Mercaptosuccinic Acid molecules immersed in different solvents	55
3.11	Lennard Jones energy averages acting at different COM distances between two Mercaptosuccinic Acid molecules immersed in different solvents	56
3.12	Coulomb plus Lennard Jones energy average acting at different COM distances between two Mercaptosuccinic Acid molecules immersed in different solvents	57
3.13	Force acting on one Mercaptosuccinic Acid molecule along z axis direction in a box containing two ligand molecules and chlorobenzene	59
3.14	Force acting on one Mercaptosuccinic Acid molecule along z axis direction in a box containing two ligand molecules and chlorobenzene	59
3.15	Comparison between mean forces acting on the two Mercaptosuccinic Acid molecules along z axis direction at different ligands COM distance, in a box containing two ligand molecules and toluene	60
3.16	Mean forces and standard deviations acting on one Mercaptosuccinic Acid molecule along z axis direction at different ligands COM distance, in a box containing two ligand molecules and non-polar solvents	61
3.17	Mean forces and standard deviations acting on one Mercaptosuccinic Acid molecule along z axis direction at different ligands COM distance, in a box containing two ligand molecules and water	62
3.18	Mean forces and standard deviations acting on one Mercaptosuccinic Acid molecule along z axis direction at different ligands COM distance, in a box containing two ligand molecules and non-polar solvents with electric field	63

3.19	Mean forces and standard deviations acting on one Mercaptosuccinic Acid molecule along z axis direction at different ligands COM distance, in a box containing two ligand molecules and water with electric field	64
3.20	Potential of mean force acting on one Mercaptosuccinic Acid molecule along z axis direction in a box containing two ligand molecules and non-polar solvents	66
3.21	Potential of mean forces acting on one Mercaptosuccinic Acid molecule along z axis direction at different ligands COM distance, in a box containing two ligand molecules and water	67
3.22	Potential of mean force acting on one Mercaptosuccinic Acid molecule along z axis direction in a box containing two ligand molecules and non-polar solvents with electric field	69
3.23	Potential of mean forces acting on one Mercaptosuccinic Acid molecule along z axis direction at different ligands COM distance, in a box containing two ligand molecules and water with electric field	70

Abstract

Gold nanorods represent an interesting object of study due to their applications in the nanoscale world because of their physical properties and size. The nanorods are used in organized structures called assemblies. Many are the fields in which these materials are used: in medicine, due to the small size, nanorods take part during non-invasive treatments for targeted therapies, for example to kill cancerous cells. In micro and nano robots gold nanorods constitute the fundamental part of the powering for the movimentation, while in chemistry and biology the nanorods are used to recognize molecules and organic structures. Experimental work was conducted in Limerick University about gold nanorods coated with Mercaptosuccinic Acid ligands. The functionalized nanorods were immersed in four solvents: anhydrous toluene, chlorobenzene, n-hexane and cyclohexane and different assemblies were obtained according with the solvent used. In this work the behaviour of Mercaptosuccinic Acid solvated in polar and non-polar solvents is investigated using Molecular Dynamics simulations. First at all, the anhydrous toluene, chlorobenzene, n-hexane and cyclohexane molecules were modelled according with the opls-aa force field and water SPC/E model was chosen as polar solvent. MD simulations of boxes containing only one type of solvent were

carried out and the results were compared with literature data. The interactions between one Mercaptosuccinic Acid molecule and the solvents were studied computing the Coulomb and Lennard-Jones energies. The ligand-solvent interactions resulted to be much higher in case of solvation in polar solvent. Thirty-two configurations were created using Packmol in order to understand how two ligand molecules could interact depending on their distance and the solvent used. Within these configurations, one ligand molecule was kept fixed in space, while the other was shifting along the z-axis direction. The boxes were then filled with the polar and non-polar solvents and simulated. The Coulomb and Lennard-Jones energies were extracted and the potential of mean force (PMF) was computed from the mean forces calculation using the trapezoidal rule. The comparison between these energies showed that entropic effects were given by solvent molecules located between the two ligands, screening their interactions. In particular, more intense screening was given by water molecules. Simulations conducted in presence of electric field proved that the screening effect was reduced by the electric influence. With the purpose of better understanding the Mercaptosuccinic Acid behaviour in polar and non polar solvents, further analysis can be carried on. MD simulations of two line of bonded ligands can explain how the interactions results in different assemblies. Entire coated nanorods can be simulated with the coarse-graining approach in order to define a model according which the nanorods form assemblies.

Chapter 1

Introduction

1.1 State of the art

In the last decades several studies have been conducted about the assembly of gold nanorods. The nanorods size ranges typically between 1 and 100 nanometers, and it represents a non expensive material that can be used for applications in nanorobots, sensors and medical fields.

In medicine nanorods are particularly appreciated for their size that allows non-invasive treatments through targeted therapies. In particular, the nanorods can be used for localized thermal therapy: during this process, the nanorods, due to their high surface aspect ratio and gold composition, are able to convert light absorption to heat through a process called plasmonic heating. With this method it is possible to raise the temperature over 43°C and kill cancerous cells [1].

The miniaturization of robots to the micro and nano scale brought issues regarding powering the control, sensing and actuation. In particular, the use of micro robots for minimal invasive medical purpose require a locomotion

in fluid environment [2], that can be carried on through the usage of chiral nanoparticles. More in detail, thanks to the plasmonic heating previously mentioned, the nanoparticles are able to activate thermo-responsive polymers that, through repetitive shape deformation, determine the motion of the micro robots [1].

Gold nanorods are used in chemical sensing to recognise small molecules, such as H^+ , Cl^- , K^+ , Fe^{3+} , Cu^{2+} , through optical changes induced by aggregation. In biological sensing it is possible to individuate amino acids, peptides, DNA, proteins and toxins with the usage of the Frequency Resonance Energy Transfer (FRET). In this method nanorods aggregate to the target structure and, since the light absorbed by the nanorods quench the fluorescence, it is possible to individuate the desired particles [3].

In 2012, Wurtzite nanorods, thiol capped, have been assembled in 3D ordered clusters in the Material and Surface Science Institute and Departemnt of Chemical and Enviromental Sciences in SFI-Strategic Research Cluster in Solar Energy Research, in University of Limerick, Irland. Researchers stated that these structures were not toxic, had high radiation stability and high absorption coefficients. Furthermore, this study showed that the formation of these assemblies is driven by the ligand environment [4].

In depth studies were developed about the factors that had a relevant influence on the nanorod assembly [5]. These factors were: rod concentrations, solvent nature and surface charge. The paper showed that computing the total energy existing between two nanorods, it was possible to determine the best rod distance for the assembly, but it was difficult to link this with a correct nanorod concentration in the solution. Then, experimental tests were

conducted to determine a suitable concentration range: below or above these values the assemblies brought to the formation of non ordered superstructures. The solvent had to be able to disperse efficaciously the nanorods, thus during this work toluene, cyclohexane, benzene and chloroform were used due to their suitable properties. The surface charge study was conducted using Molecular Dynamic simulations: four functionalized nanorods were simulated and the orientation angles were analyzed. The nanorods, with electric charge equivalent to +0.2 times the electron charge, were firstly positioned aligned along the same axis. After the simulation, two of the four nanorods were aligned in the opposite direction, thus suggesting an anti-ferromagnetic property of the material. A brief but complete information about how molecular dynamics simulations developed from the 70s until nowadays, is given by A. Hospital et. al., outlining the potentiality of this means for a study for dynamic properties of molecules [7]. Another factor that strongly influence the formation of the assemblies, is the presence of an electric field. In depth studies about the usage of electric field during assembly were conducted by P.Liu et al., in 2017 [6]. In this study, cadmium chalcogenide nanorods in a toluene solution were used. A uniform DC electric field with intensity of 612 V/cm permitted to achieve the formation of samples with thickness of 800 nm. The layers formed were constituted by nanorods aligned vertically and orthogonal to the substrate, with a high packing ratio. Such superstructures are suitable in the laser technology due to the the low cost and high performance.

1.2 Study purpose

An experimental work about nanorods assembly was conducted in Limerick University. The nanorods were made of gold and coated with a particular ligand: the Mercaptosuccinic Acid. This is a dicarboxylic acid containing a thiol functional group. The nanorods were assembled using a DC electric field, in presence of different solvents: anhydrous toluene, chlorobenzene, n-hexane and cyclohexane. The experiments led to different geometries of assembly, according with the solvent used. The purpose of this work is to study the reason why different assemblies are formed according with the solvent used. In fact, the different rods orientation can be explained looking at the polarity of the solvents, how this can be modified in presence of the electric field, and how the ligand reacts once immersed in the solvent.

Anhydrous toluene, chlorobenzene, n-hexane and cyclohexane and water, were simulated using Molecular Dynamics (MD) in order to determine their physical properties and, in particular, their interaction energies. Then, boxes containing a Mercaptosuccinic Acid molecule for each solvent were simulated in order to look at the interaction occurring between the solvents and the molecule under study. Finally, the interactions between two Mercaptosuccinic Acid molecules were studied in function of the ligands COM distance and of the solvent in which the ligands were immersed. Since the experimental work conducted in Limerick was done with an electric field driven assembly, the configurations were simulated with and without electric field.

Chapter 2

Materials and methods

2.1 Molecular Dynamics

2.1.1 Classical and quantum mechanics

The term classical mechanics include all the mechanical theories developed until XX century. The main theory to which classical mechanics is referred to was developed in the seventies by Isaac Newton: his three formulations are now a milestone for any field of scientific study.

Newton's laws of motion [8]:

1. Every object in a state of uniform motion will remain in that state of motion unless an external force acts on it
2. The change of motion of a body is proportional to the motive force impressed and is made in the direction of the right straight line in which that force is impressed
3. When a force acts on a body due to another body, then an equal and

opposite force acts simultaneously on that body

Later, other theories have been developed and had relevant importance on the world stage. Among the others we can mention Lagrangian mechanics, useful for dynamic studies with holonomic constraints and Hamiltonian mechanics, which main point is to include the total energy of a system in a function called with his name.

The main characteristic that group all these theories under the name of classical mechanics is that it works with continuum bodies in a deterministic approach. It can be applied to describe the motion of bodies with a certain size range and for a certain speed range. Under these assumptions, the laws of classical mechanics, whichever among these theories we apply, are able to determine how the body is going to move in the future and how it has moved in the past.

We state that the limitations of the classical approach are defined within two ranges: one massive and one relative to the speed. In fact, describing with classical mechanics the motion of a body which moves with velocities close to the speed of light, leads to wrong results. For this kind of study, special theory of relativity, described by Albert Einstein in 1905 is used. Again, the swiss physicist, in 1916, developed the general theory of relativity, nowadays used to study bodies with enormous mass, as the planets. Finally, in order to describe the motion of infinitesimal objects, like atoms or molecules, quantum mechanics play the role.

The reason why classical mechanics fails describing the motion of infinitesimal scale objects is that it considers particles and waves as separated entities. The uncertainty principle, stated by Heisenberg in 1927, put the focus on the

impossibility of measuring precisely, at the same time, two complementary variables. In particular, he stated the impossibility of measuring the position and the momentum of a particle at a certain time instant. With the Heisenberg's principle, it became relevant the wave-particle duality. So, it means that both, object and radiations, are not only waves or only particles, but they have, intrinsically, characteristics of both. This is the basis of quantum mechanics, and it explains why it works representing the reality not only in the microscopical level, but also for the objects represented by classical mechanics.

2.1.2 The Molecular Dynamics approach

According to what we outlined in the previous paragraph, a study at molecular size would be accurate if developed with quantum mechanics, while it would bring to inaccurate results if classical mechanics is used. Nevertheless, computing power for simulations involving quantum mechanics are too heavy, and it is impossible to simulate control volumes containing hundreds of atoms using this approach. Because of this, classical mechanics, and, in particular Newtonian mechanics is used. This bring to less accurate results, but the software is able to limit the error using coefficients during the simulations.

Finally, the second Newton's law is used. The formula [8]

$$\vec{F} = m \cdot \vec{a}$$

is used to determine the position and the velocities of every atom contained in the control volume.

In figure 2.1 it is shown how Newtonian mechanics is applied in Molecular Dynamics simulations. For simplicity, a 2D case is illustrated.

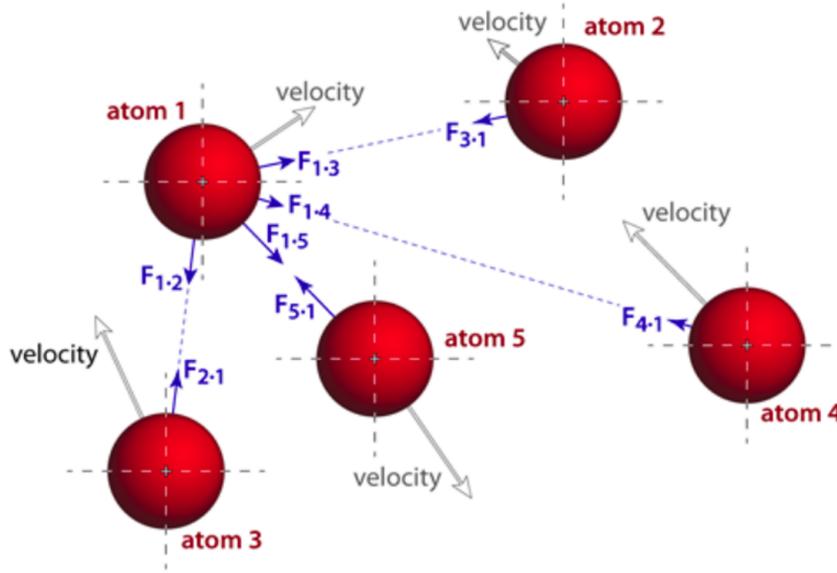


Figure 2.1. Newtonian mechanics applied to Molecular Dynamics [9]

Integration schemes

The integration of Newton's equation can be computed through two algorithms: leap-frog [10] and Verlet [11]. The equations used in these two schemes are the following [12]:

- Leap-frog algorithm:

$$\vec{v}\left(t + \frac{1}{2}\Delta t\right) = \vec{v}\left(t - \frac{1}{2}\Delta t\right) + \frac{\Delta t}{m} \cdot \vec{F}(t) \quad (2.1)$$

$$\vec{r}(t + \Delta t) = \vec{r}(t) + \Delta t \cdot \vec{v}\left(t + \frac{1}{2}\Delta t\right) \quad (2.2)$$

- Verlet algorithm:

$$\vec{v}\left(t + \frac{1}{2}\Delta t\right) = \vec{v}(t) + \frac{\Delta t}{2m} \cdot \vec{F}(t) \quad (2.3)$$

$$\vec{r}(t + \Delta t) = \vec{r}(t) + \Delta t \cdot \vec{v}\left(t + \frac{1}{2}\Delta t\right) \quad (2.4)$$

$$\vec{r}(t + \Delta t) = \vec{v}\left(t + \frac{1}{2}\Delta t\right) + \frac{\Delta t}{2m} \cdot \vec{F}(t + \Delta t) \quad (2.5)$$

The leap-frog integrator is the default one used in Gromacs. This algorithm is of third order and it is time-reversible. When temperature and pressure are coupled, the equations of motion are modified and extended in order to conserve the constraints. The use of Verlet scheme is required when an extremely accurate integration is necessary due to temperature and pressure coupling.

Temperature and pressure coupling

In Gromacs it is possible to use different types of temperature and pressure coupling, listed below.

Thermostats used for temperature coupling [12]:

- Berendsen: the Berendsen thermostat couple the system with a bath temperature set by the user. In the control volume it is possible to form different groups and to link each group to a different time constant and

temperature value [13]. The algorithm used is:

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau} \quad (2.6)$$

Where T_0 is the reference temperature, T is the temperature of the system and τ is the time constant.

- Nose-Hoover: in this thermostat the time constant controls the period of the fluctuations of temperature at equilibrium. Again, the temperature reference value is set by the user and different groups can be coupled separately [14] [15]. The strength of the coupling is set by a constant called *mass parameter*, defined through the equation:

$$Q = \frac{\tau_T^2 \cdot T_0}{4\pi^2} \quad (2.7)$$

This constant modifies the equation of motion through the formulation:

$$\frac{d^2\vec{r}_i}{dt^2} = \frac{\vec{F}_i}{m_i} - \frac{p_\zeta}{Q} \cdot \frac{d\vec{r}_i}{dt} \quad (2.8)$$

where

$$\frac{dp_\zeta}{dt} = (T - T_0) \quad (2.9)$$

is the equation used to describe the motion of ζ , the heat bath parameter.

- Andresen: the coupling takes place with some of the particles, chosen randomly between the particles of a part of the system for each timestep.
- Andersen-Massive: similar to the previous one, but the random selection takes place between all the particles in infrequent timesteps.

- V-rescale: it is similar to the Berendsen thermostat, but it uses a stochastic term for velocity rescaling.

Barostats used for pressure coupling [12]:

- Berendsen: as for the Berendsen thermostat, it couple the system to a pressure bath. This barostat rescales the box vector in every step, so that the reference pressure P_0 is reached. The pressure variation for each time step is calculated during the entire dynamics run through the equation:

$$\frac{dP}{dt} = \frac{P_0 - P}{\tau_p} \quad (2.10)$$

- Parrinello-Rahman: if pressure fluctuations are too big, it is necessary to use the Parrinello-Rahman barostat. In order to reach the reference pressure value, it computes different algorithms that are described in detail in the references [16] [17].
- MTTK: the Martyna-Tuckerman-Tobias-Klein method combine both the temperature and pressure coupling. It has some restrictions, since only the velocity Verlet allows the NPT calculation with this thermostat. Because of this, the MTTK can be combined with the Nose-Hoover thermostat only. Detailed equations can be found in the references [18] [19].

Meanwhile the MTTK barostat must to be used together with the Nose-Hoover thermostat, the Berendsen and Parrinello-Rahman barostats can be used with any of the thermostats described above.

2.2 Gromacs

Gromacs is the most widely used software for Molecular Dynamics. The name Gromacs comes from GRONingen MACHine for Chemical Simulations, it is an open source and free software created in Groningen University, Netherlands, in 1991. Since 2001 it is developed from the Royal Institute of Technology and the Uppsala University, in Sweden. The software is based on inner loops written in C language using intrinsic functions, that the compiler converts to Single Input-Multiple Data instructions. It works on parallel dedicated systems, based on bus computer architecture. In figure 2.2 it is possible to see the multi-level of parallelism: the starting point is the ensemble, so the control volume of our system, plus the dynamic information about the system and the simulation specifications. These are elaborated by the domain decomposition algorithm, that splits the whole calculation into smaller problems and coordinates in order to reach the solution correctly. The numerical problems formulated by the domain decomposition are solved in the Graphic Processing Units and in the Central Processing Units using Single Input-Multiple Data instructions [20].

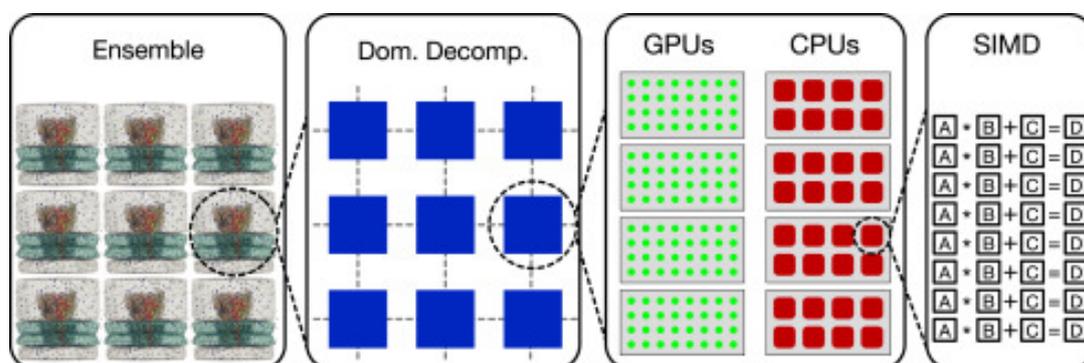


Figure 2.2. Multi-level parallelism in GROMACS [20]

Just to give an idea about the complexity of a typical Molecular Dynamic system, a control volume with 150,000 atom has about thirty million particle–particle interactions for each MD step [20].

2.3 Force Fields

One of the main feature used for Molecular Dynamics simulation is the utilization of force fields. The force fields used in molecular modelling have the purpose of describing intermolecular and intramolecular potential energies through equations called *potential functions*. Besides the potential functions, every force field contains all the parameters necessaries in order to compute the potential energies acting on each of the atoms contained in the control volume. The force fields supported by Gromacs are:

- AMBER: *Amber94* [21], *Amber96* [22], *Amber99* [23], *Amber99SB* [24], *Amber99SB-ILDN* [25], *Amber03* [26], *Ambergs* [27]
- CHARMM: *Charmm19*, *Charmm22*, *Charmm27*, *Charmm36* [28] [29] [30] [31] [32] [33]
- GROMOS *Gromos96 43a1*, *Gromos96 43a2*, *Gromos96 45a3*, *Gromos96 53a5*, *Gromos96 53a6*, *Gromos96 54a7* [35] [36]
- OPLS *Opls-ua*, *Opls-aa*
- Coarse Grained Force Fields: *Martini* [37], *Plum* [38] [39]

Each of these force filed has its own potential functions and parameters, obtained by the research works given in the citations. The reason why many

force fields are implemented is that different molecule structures require different parameters. So, for example, Gromos96 is suitable to simulate small molecules, but not for long alkanes and lipids. Opls-aa force field would be rather used for lipid simulations [40].

For this thesis work, research was done in order to determine the force field that better could model the Mercaptosuccinid Acid immersed in the polar and non-polar solvents studied.

Finally, works about the study of aliphatic and aromatic chloro derivatives [41], cyclohexane-water distribution coefficients [42] and others [43], [44] suggested to use the Opls-aa force field.

2.4 Control Volume Configuration

2.4.1 The pdb and topology files

In Molecular Dynamics, the control volume consists in a box in which a discrete number of atoms and molecules are defined. The atoms' and bonds' information are contained in the force field folder, while the molecule composition is listed in a file not present in Gromacs and introduced by the user: the pdb file. The word pdb is an acronym for Protein Data Bank (figure 2.3), a worldwide database containing more than hundred thousand molecule structures.



Figure 2.3. Protein Data Bank [45]

The pdb file can be downloaded from the Protein Data Bank website. It can also be found in other research works or it can be created by the user itself after appropriate studies about the molecule geometry. Each pdb file contains a list of atoms and its position in space (figure 2.4). The atom name listed

1	SM	LIG	1	19.730	19.461	17.355	1.00	0.00
2	OH1	LIG	1	18.021	19.747	20.824	1.00	0.00
3	OH2	LIG	1	21.073	21.307	21.769	1.00	0.00
4	OC1	LIG	1	18.031	21.218	19.103	1.00	0.00
5	OC2	LIG	1	21.055	19.038	21.872	1.00	0.00
6	CS	LIG	1	19.730	19.461	19.184	1.00	0.00
7	CC	LIG	1	21.014	20.104	19.704	1.00	0.00
8	C01	LIG	1	18.514	20.239	19.658	1.00	0.00
9	C02	LIG	1	21.059	20.070	21.213	1.00	0.00
10	HC	LIG	1	19.639	18.420	19.512	1.00	0.00
11	HC1	LIG	1	21.909	19.585	19.340	1.00	0.00
12	HC2	LIG	1	21.090	21.150	19.377	1.00	0.00
13	HS	LIG	1	20.805	18.673	17.202	1.00	0.00
14	H01	LIG	1	17.247	20.261	21.140	1.00	0.00
15	H02	LIG	1	21.089	21.268	22.749	1.00	0.00

Figure 2.4. List of atoms with their spatial position, image taken from Mercaptosuccinic Acid pdb file

in the pdb file is then referred to an atom type in another very important file: the topology. Since the force field defines many atoms according with their bonded and non-bonded parameters, the topology file links the atom with the respective atomtypes (figure 2.5).

```
[ atoms ]
; nr      type  resnr residue  atom  cgnr    charge      mass  typeB    chargeB
; residue 1  LIG  rtp  LIG  q  0.0
  1  opls_200  1  LIG  SM      1  -0.335    32.06  ; qtot -0.335
  2  opls_268  1  LIG  OH1     2  -0.53    15.9994 ; qtot -0.865
  3  opls_268  1  LIG  OH2     3  -0.53    15.9994 ; qtot -1.395
  4  opls_269  1  LIG  OC1     4  -0.44    15.9994 ; qtot -1.835
  5  opls_269  1  LIG  OC2     5  -0.44    15.9994 ; qtot -2.275
  6  opls_207  1  LIG  CS      6   0.12    12.011  ; qtot -2.155
  7  opls_136  1  LIG  CC      7  -0.12    12.011  ; qtot -2.275
  8  opls_267  1  LIG  C01     8   0.52    12.011  ; qtot -1.755
  9  opls_267  1  LIG  C02     9   0.52    12.011  ; qtot -1.235
 10  opls_140  1  LIG  HC     10  0.06     1.008   ; qtot -1.175
 11  opls_140  1  LIG  HC1    11  0.06     1.008   ; qtot -1.115
 12  opls_140  1  LIG  HC2    11  0.06     1.008   ; qtot -1.055
 13  opls_204  1  LIG  HS     12  0.155    1.008   ; qtot -0.9
 14  opls_270  1  LIG  H01    13  0.45     1.008   ; qtot -0.45
 15  opls_270  1  LIG  H02    14  0.45     1.008   ; qtot 0
```

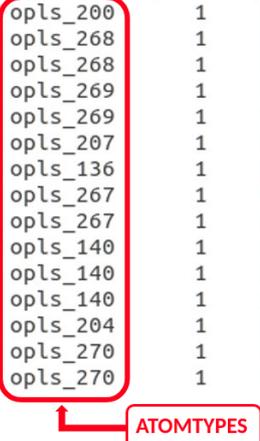


Figure 2.5. Atomtypes of Mercaptosuccinic Acid, image taken from topology file

Moreover, the topology file specify bonds, pairs, angles and dihedrals existing between the atoms of the molecule. In addition, the restraint files, a file containing the structure of the solvent and the number of each molecule contained in the control volume are listed.

The Molecules Models

Before proceeding with the molecules modelling description, it is useful to point out the difference between polar and non-polar solvents. The polarity of a molecule is given by a number called dielectric constant and a threshold is set in order to define if a molecule is polar or non-polar. For dielectric constants higher than $15 \frac{C^2}{N \cdot m^2}$ the solvent is said to be polar, otherwise it is non-polar. The table 2.1 shows the dielectric constants of the solvents used in this thesis work.

SOLVENT	Dielectric Constant, $\frac{C^2}{N \cdot m^2}$
CHLOROBENZENE	5.62
CYCLOHEXANE	2.02
HEXANE	1.89
TOLUENE	2.38
WATER	78.54

Table 2.1. Solvents' dielectric constants [46]

It is now possible to state that water is the only polar solvent, while chlorobenzene, cyclohexane, hexane and toluene are all non-polar solvents. It is now possible to proceed with the description of the models used to describe the solvents.

Chlorobenzene is an aromatic compound composed by six carbon atoms forming a simple aromatic ring, to which five hydrogen atoms and one chlorine atom are attached.

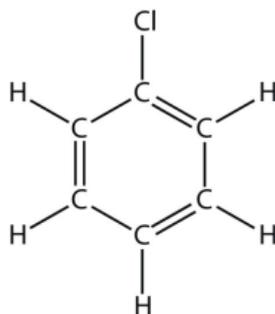


Figure 2.6. Chlorobenzene

For this molecule, the atomtypes "opls_145" and "opls_146", specifically designed for the benzene molecule, have been used to define, respectively, the

five carbon and the five hydrogen atoms composing part of the aromatic ring. Then, chlorine and the last carbon atoms have been written as "opls_164" and "opls_163", precisely listed in the opls-aa force field for the composition of chlorobenzene.

Cyclohexane is a cycloalkane and it is composed by six carbon atoms forming a ring and six hydrogen atoms attached to them.

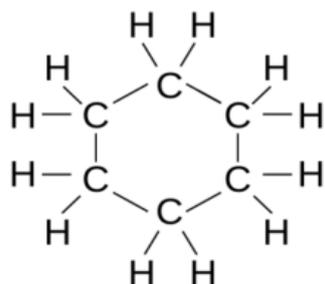


Figure 2.7. Cyclohexane

The atomtypes chosen for the cyclohexane are "opls_136" for carbon and "opls_140" for hydrogen, both specifically designed for alkane groups.

N-hexane (or hexane) is a linear molecule, belonging to the alkane group. It is composed by six carbon atoms in a row, to which are attached two hydrogen atoms each, plus an hydrogen atom for the first and the last carbon atoms forming the molecule.

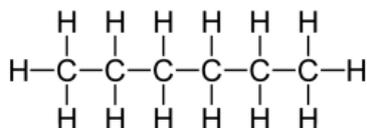


Figure 2.8. Hexane

For the characterization of the hexane molecule, the atomtypes "*opls_135*" and "*opls_136*" have been chosen, respectively, for the external and the central carbon atoms. In fact, these are the atomtypes designed for the CH_3 and CH_2 groups in alkane molecules. Then "*opls_140*", generical atomtypes for hydrogen atoms in alkane groups, have been chosen for all the hydrogen atoms.

Toluene is formed by a simple aromatic ring, to which a methyl group CH_3 is attached.

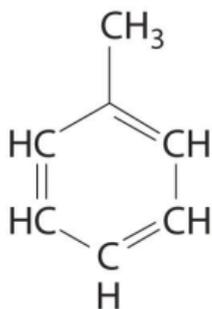


Figure 2.9. Toluene

As for the chlorobenzene molecule aromatic ring, also in this case the "*opls_145*" and "*opls_146*" have been used for this purpose. Then "*opls_148*", carbon

atom for toluene methyl group, and "opls_148", hydrogen atom for alkanes have been chosen.

The Mercaptosuccinic Acid molecule is a dicarboxylic acid, thus an organic compound containing two functional $-COOH$ groups, linked to a thiol functional group $-SH$; Its formula is $SH(CH)CH_2_2(COOH)$.

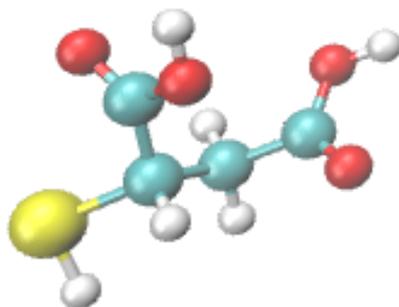


Figure 2.10. Mercatposuccinid Acid: sulfur atom is in yellow, carbons in grey, oxygens in red and hydrogens in white

"Opls_200, "opls_204 and "opls_207, all specific for thiols groups, have been chosen to define the sulphur, hydrogen and carbon atoms of this functional group. "Opls_267, "opls_268, "opls_269 and "opls_270, proper for $-COOH$ groups, describe the carbons, oxygens and hydrogens atoms of these groups. Finally, "opls_136 and "opls_140 have been used for the remaining carbon atom, the hydrogens attached to it and the last hydrogen attached to the thiol group. In figure 2.10 a representation of the Mercaptosuccinic Acid is shown, with various colors assigned to different atoms. From the image it is possible to observe the thiol functional group, composed by sulfur and the two atoms attached to it, and the two $-COOH$ groups, on the top and on the right of the molecule. The others atoms have been modelled with atoms

proper for alkanes.

After the definition of all the atomtypes, it was possible to proceed with the creation of the topology file and the control volume boxes.

2.4.2 Procedure and boxes creation

Polar and Non-polar Solvents

First at all, boxes of small dimensions, containing only one type of solvent each were created. Each box was cubic and measured 5.5 *nm* for side. The 166.375 *nm*³ space of the boxes was containing between 572 (in the case of hexane) and 742 (for chlorobenzene) molecules.

The purpose of these boxes was to verify that the atomtypes used were correct, thus that the simulations of these solvents would bring to appropriate values of density and energies. Furthermore, the electric field influence on the solvents could be observed.

In order to have a comparison for the solvents required for this study, a box with the same dimension containing water was created. This have been done for two reasons:

1. Water is already present in the force fields as a default solvent and it is often object of study, thus resulting in many works existing in literature. The presence of many research works gives a good possibility of comparison between the results obtained. Because of this, the study of water represents a reference point in the study of the solvents used in this work.

2. The final purpose of this study is to observe the behaviour of the Mercaptosuccinic Acid according with the changes of solvents and electric field. Since chlorobenzene, cyclohexane, hexane and toluene are all non-polar solvents, to use a polar solvent could give important information about the behaviour of the Mercaptosuccinic Acid related with electric field and polarisation.

In Gromacs, opls-aa force field contains many water types. The *SPC/E*, thus extended simple point charge, water model was used.

Solvation of Mercaptosuccinic Acid

The second main step of the study was to look at the influence between the solvents and a single Mercaptosuccinic Acid (or "*ligand*") molecule. In order to reach this purpose, a ligand molecule was placed in the center of a cubic box with side of 4 nm, as shown in figure 2.11.

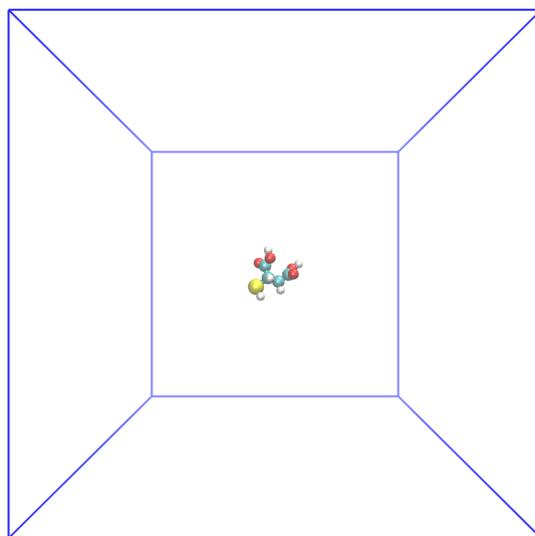


Figure 2.11. Mercatposuccinid Acid in the box center

Five boxes with this dimension were created, with a ligand molecule in the center, and filled with the solvents previously described and water. Then, the boxes were simulated with and without electric field.

Potential of Mean Force setup

The final and more elaborated main step of this thesis work was to study how two ligand molecules interact in function of their distance and of the solvent used. In order to make a study depending from the distance, the goal was to calculate the potential that gives the average force over the configurations, containing a different ligands distance each. This potential is called Potential of Mean Force (PMF) and it is usually calculated using the umbrella sampling approach. The umbrella sampling method plans to generate many configurations along a reaction coordinate. Through the windowing technique the configurations are selected in order to form an appropriate distance spacing. Than the configurations extracted are equilibrated, simulated and, finally, the Weighted Histogram Analysis Method (WHAM) [47] is used to compute the Potential of Mean Force [48].

Unfortunately, this method brought to wrong results because during the configurations generation the two molecules were moving too far away from the starting configuration, thus the impossibility of forming appropriate sampling windows. Modifying the parameters used for the umbrella sampling procedure still brought to unsatisfying results, so a different approach was used. Thirty-one configurations were then created manually using Packmol [49], an open source package specifically designed to create initial configurations for Molecular Dynamics simulations. The two ligand molecules were

placed specular with respect to a symmetry plane orthogonal to the line connecting the center of mass (COM) of the two molecules. Then, keeping one molecule (Ligand 1) fixed in space and shifting the second molecule (Ligand 2) along the z-axis, 31 configurations were created. This set of configurations was created so that the line connecting the COM of the two molecules was coincident with the z-axis, in order to have the same molecule orientation for each configuration. The spacing between the two ligand molecules was initially set to 1 nm, so that the COM distance between the two ligands was ranging from 3.4 nm to 0.4 nm through the 31 configurations.

Then, each of this configuration was placed in a rectangular box of dimension 4x4x8 nm, with the bigger length oriented along the z-axis. The COM of the fixed ligand molecule (Ligand 1) was placed at coordinates 2x2x2 nm, while the COM of the second ligand molecule was shifting along the z-axis, with fixed x and y coordinates of 2 nm both. The shifting of Ligand 2 inside the box is shown in figure 2.12.

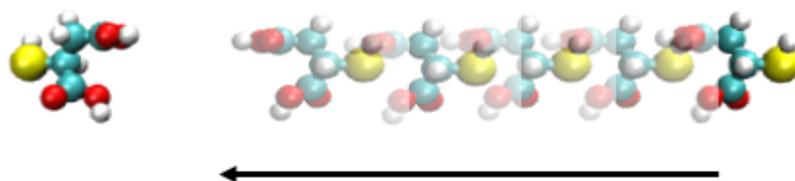


Figure 2.12. Procedure used to create the configurations used for PMF computation: the image shows the first configuration, with maximum ligands COM distance and, in transparency, the shifting of Ligand 2 along the z-axis

Each of this box was then filled with a solvents and simulations were carried on for all the solvents used. After the post processing, it was clear that the configurations with COM distance bigger than about 1.5 nm were not

relevant, since they were having a null force value. Because of this, the spacing was shortened to 0.5 nm and 32 configurations were created with a COM distance ranging from 0.2 to 1.8 nm. The procedure described allowed to bypass the umbrella sampling windowing technique and permitted to have the configurations necessary to compute carry on the analysis. In order to compute the PMF of the systems, reminding that the force is equal to the negative gradient of the potential:

$$\vec{F} = -\vec{\nabla}U \quad (2.11)$$

Considering moreover that the object of study was the interaction between the two ligands placed along the z-axis, the component of the force acting along the z-axis was extracted (figure 2.13).



Figure 2.13. Force acting on Mercaptosuccinic Acid molecules along z-axis

The mean force was then computed for each configuration and the PMF was calculated using the trapezoidal rule [50]:

$$\int_a^b F(x) dx = \frac{h}{2} \left[f(a) + 2 \sum_{i=1}^{m-1} f(a + ih) + f(b) \right] + R_m(f) \quad (2.12)$$

Where a and b define the extremes of integration. In the domain $[a, b]$ $F(x)$ is discretized into m equally spaced panels of length $h = \frac{b-a}{m}$. The trapezoidal

rule approximates the area situated under the graph of a function as a sum of trapezoids with an error $R_m(f)$ of order h^2 . In this thesis work, the function is represented by the mean force plot and the trapezoids are constructed according with the ligands COM distance.

2.5 Energy Minimization and Dynamics Runs

In this section are explained the following steps used for the MD study of the configurations previously described. All the control volumes, regardless if containing only solvents or ligands and solvents, were simulated using the following procedure:

1. Energy Minimization
2. NVT
3. NPT
4. MD run

Since all these steps require a high computational power, the procedure implemented in Gromacs consists in the following passages: first at all the user sets the run parameters in a *.mdp* file. Then this file, together with the information coming from the previous ensemble, are translated by Gromacs through the command *gmx grompp* in a *.tpr* file, written in a binary code. This last file contains all the necessary information for the molecular dynamics run and it is then given to the software through the *gmx mdrun* command for the proper run. In the following subsections the steps necessary for the box equilibration are explained.

2.5.1 Energy Minimization

Once the control volumes were assembled, before proceeding with dynamics simulations, it is necessary to reduce the system to a lower potential energy value. This process is necessary since, during the box assembly, it is possible that molecules were placed in positions with high potential, thus non according with a natural disposition of a real system. The process used to reduce the potential energy is called Energy Minimization. In order to reach the purpose of reducing the system potential energy, Gromacs individuates the molecules with higher force values and shifts them to a position with lower force values. This process is repeated in an iterative way until the maximum force is lower than a threshold number set by the user. A maximum number of minimization steps is also set, so that, if the target force is not reached, the algorithm stops without a calculator crash. In figure 2.14 it is possible to see the decreasing of potential energy in function of the Energy Minimization Steps for a box of water containing lysozyme.

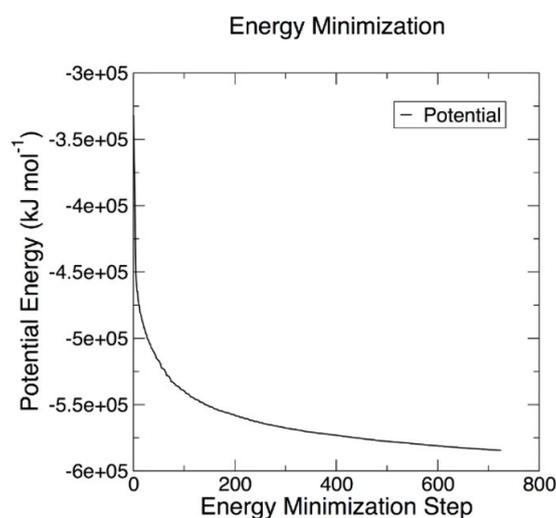


Figure 2.14. Decrease of Potential Energy during Energy Minimization steps [51]

The algorithms implemented in Gromacs for Energy Minimization are:

1. Steepest Descent
2. Conjugate Gradient
3. L-BFGS

Since the steepest descent algorithm is the most robust and efficient, it has been chosen for all the Energy Minimization used in this thesis work. More in detail, for boxes containing one solvent each a maximum force of $400 \frac{kJ}{mol*nm}$ and a maximum of 50000 steps were set. The maximum force threshold was then reached after few hundreds of iterations. In boxes containing one and two ligands immersed in solvents, the maximum force was raised to $1000 \frac{kJ}{mol*nm}$, while the maximum steps number was maintained at 50000. Because of the dimension and the complexity of these systems, the number of steps required to reach the threshold force number was bigger, but still lower than the maximum.

2.5.2 Equilibrations

Once the systems were brought to a lower potential energy value, the dynamics simulations could start. The Molecular Dynamics approach requires the two following steps, also called ensembles:

1. NVT: during this run, the particles number N, the volume V and the temperature T are kept constant. In particular, the internal energy and pressure can change, while temperature is maintained constant through a temperature coupling.
2. NPT: in this ensemble, particle number N, pressure P and temperature

T are kept constant, while the volume is free the change. In particular, the pressure is maintained constant through a pressure coupling.

Temperature and pressure are kept constant through the use of thermostats and barostats. Both works to bring temperature and pressure to reference values, even if fluctuations are present due to the dynamics of the system.

During the NVT and NPT steps, since the system is still not equilibrated, restraints are used to keep atoms in a fixed positions. It is not necessary to restraint all the atoms located inside the box, but a selection can be done according with the simulation purpose. The restraints are set as force values and written for each direction of each atom of interest in an *.itp* file. This file has to be included in the topology and written in the *.mdp* file if used. In this thesis work, no restraints were used for the equilibration of the boxes containing only solvents. Restraints of $1000 \frac{\text{kJ}}{\text{mol} \cdot \text{nm}}$ were applied to all the atoms of the ligands, both for boxes containing one and two ligands. Moreover, all the thermostat were set to the temperature of 300 K and all the barostat were set to the pressure of 1 bar. Let's now describe the other parameters used for NVT and NPT equilibrations, in particular the duration of the simulations, the time step used and the time constant set for temperature and pressure couplings.

Polar and Non-polar Solvents

For boxes with only solvents, the NVT ensemble was lasting 150 ps, with a time step of 2 fs. As thermostat, V-rescale scheme was used, with a time constant of 0.1 ps. For the NPT run, the same duration and time step of

150 ps and 2 fs were chosen. The Nose-Hoover thermostat and Parrinello-Rahman barostat were chosen, with temperature time constant of 0.4 ps and pressure time constant of 2 ps. Since these boxes were relatively simple, since were containing only one molecule type each, the equilibration phases didn't presented particular issues.

Mercaptosuccinic Acid in solvents

Boxes containing one ligand molecule immersed in solvents, instead, were treated with a different approach: after the NVT run, two NPT simulations were carried on in order to have more equilibrated results. For NVT a time step of 1 fs and 100 ps of simulation duration were chosen. The thermostat used was again the V-rescale, with two groups for coupling: one for the solvent and one for the ligand. For both the coupling a time constant of 1 fs was set. As previously stated, two NPT simulations were run, both with the same duration and time step of 100 ps and 1 fs. For the first NPT run V-rescale thermostat and Berendsen barostat were used, with temperature time constant of 20 fs and pressure time constant of 2 ps while Nose-Hoover and Parrinello-Rahman were chosen for the second NPT run, with temperature and time constants equal to the first NPT simulation.

Two Mercaptosuccinic Acid molecules in solvents

The boxes with two ligands immersed in solvents were simulated for longer time: both NVT and NPT were lasting 1 ns, with a time step of 1 fs. For the NVT runs, V-rescale thermostat was used, while for NPT simulations Nose-Hoover thermostat and Parrinello-Rahman barostat were chosen. For both the runs, two groups coupling for temperature bath were chosen and

the time constants used were 0.1 ps for NVT and 0.2 ps for NPT. The time constant in pressure coupling was, instead, of 2 ps. Moreover, as outlined in the previous chapter, despite the restraints the two ligands were shifting too much during the dynamics runs. In order to avoid this, a freezing was introduced along the three spatial directions acting on the sulfur atom and on the carbon atom attached to it. These two atoms were listed in the index file in the groups *SC_1* for the ligand 1, thus the each one maintained fixed through all the configurations, and *SC_2* for ligand 2, the shifting one.

Both NVT and NPT, even if are dynamics runs, are still a preparation phase for the proper Molecular Dynamics simulation, that is the last phase of the dynamics simulation.

2.6 Molecular Dynamics run

After the NVT and NPT steps, the system reached the equilibration at the required temperature and pressure. In this last dynamics run, also called MD run, it is then possible to remove the restraints used during the equilibration and to collect data for the post processing.

In the MD run carried on during this thesis work, as in the equilibration phases, the reference temperatures and pressures were set to 300 K and 1 bar.

The simulations of boxes containing solvents only were lasting 200 ps with a time step of 2 fs. V-rescale thermostat and Parrinello-Rahman barostat were used, with time constants of 0.1 ps and 2 ps respectively. A second MD run was then performed, with the same parameters as the previous simulation,

plus the addition of an electric field acting along the z-axis direction with intensity of $1 \frac{V}{nm}$.

The boxes containing one ligand molecule immersed in solvents were simulated for 1 ns with 1 fs as time step. Nose-Hoover thermostat was used and, as for the equilibration runs, two coupling groups for the ligand (now not restrained as it was during the equilibration process) and for the solvent were chosen and the time constants were set for both at 20 fs. Parrinello-Rahman was chosen as barostat with a time constant of 2 ps. As for the boxes with solvents only, a second MD simulation was run with the addition of an electric field with the same direction and intensity as in the previous case.

The simulations of boxes containing two ligand molecules immersed in solvents were lasting 2 ns with a time step of 0.5 fs. Nose-Hoover thermostat with two groups coupling and 0.2 ps of time constant were chosen. Parrinello-Rahman barostat with 2 ps as time constant was set. Again, simulations were carried on also in presence of an electric field.

Chapter 3

Results

3.1 Solvents

The main purpose of simulating control volumes containing one type of solvent was to validate the models used to describe the molecules and to verify if equilibrations and MD runs brought to successful results.

3.1.1 Temperature and Density

The temperature of the solvents were ranging around 300 K during the entire MD runs, with fluctuations of ± 5 K as visible in figure 3.1. Oscillations of this magnitude are acceptable for molecular dynamics simulations; it is then possible to state that the boxes were well equilibrated under a thermal point of view.

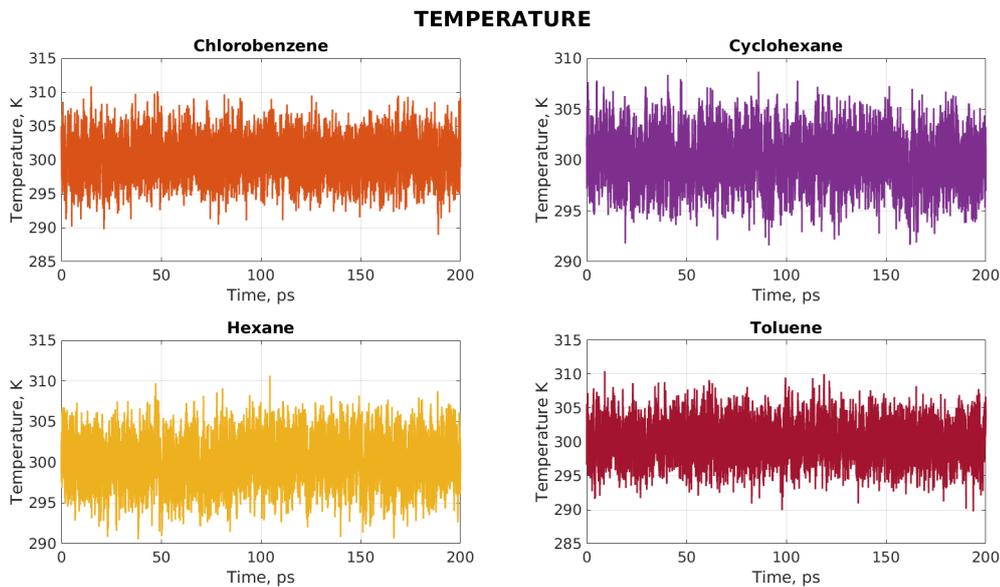


Figure 3.1. Solvents' temperature during MD simulation

A relevant measure that could give important information about the accuracy of the atomtypes and force field used, was the density value. The densities of the solvents during the entire simulations were then extracted from the *.edr* file and plotted, figure 3.2.

These trends are more variable if compared to the temperature plots, but still no relevant changes, that would mean a not stable simulation, occur during the run. Similar plots of temperatures and densities were obtained from simulations in presence of electric field with intensity of $1 \frac{V}{nm}$. Also in case of electric field, the temperature mean values were ranging at 300 K (\pm)5 K and the densities shown comparable fluctuations.

The averages of the densities were then calculated for all the solvents, both for simulations with and without electric field. Thereafter these numbers,

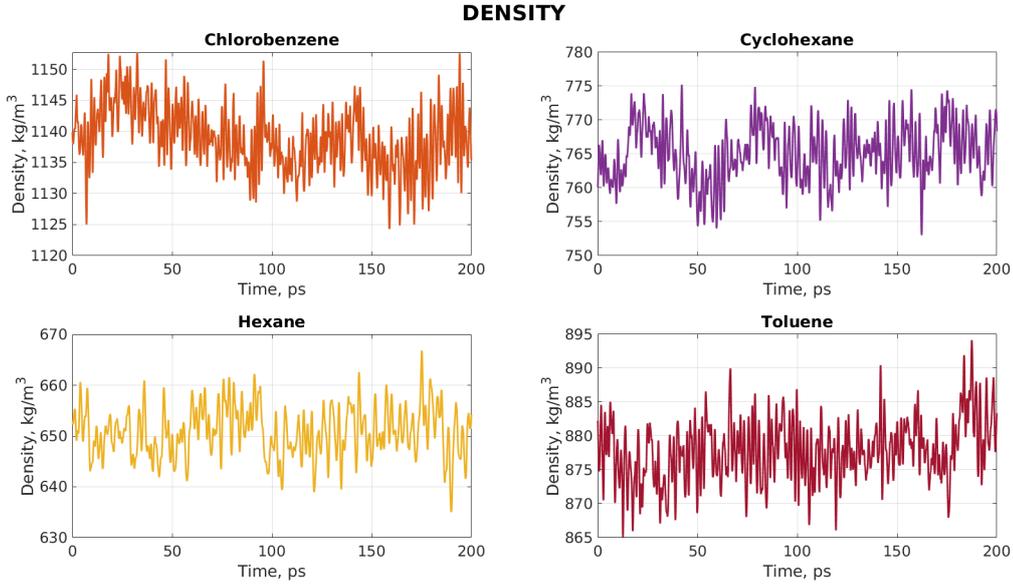


Figure 3.2. Solvents' density during MD simulation

with corresponding standard deviation, were compared with density values existing in literature, as shown in table 3.1.

DENSITIES

SOLVENT	ρ , $\frac{kg}{m^3}$ (Ex.)	ρ^* , $\frac{kg}{m^3}$ (Ex.*)	ρ , $\frac{kg}{m^3}$ (Lit.)
CHLOROBENZENE	1139.0 ± 5.0	1144.3 ± 4.6	1096
CYCLOHEXANE	765.1 ± 4.1	763.9 ± 4.2	769
HEXANE	650.8 ± 4.8	651.5 ± 4.2	652
TOLUENE	877.8 ± 4.4	880.8 ± 5.3	820
WATER	996.9 ± 4.2	997.3 ± 4.5	1000

Table 3.1. Solvents' density comparison between Experimental (Ex.) and Literature (Lit.) data [52], [53]. The * means that the simulation was run in presence of electric field

As it is possible to observe from the table, the density values obtained from

the MD simulations were similar to the each one found in literature. In particular, the biggest gap between experimental and literature values was reached by chlorobenzene, which average from the simulation was $1139.0 \frac{kg}{m^3}$ while literature was asserting $1096 \frac{kg}{m^3}$. The reason why the gap is bigger in this solvent rather than in the others is probably due to the higher number of chlorobenzene molecules contained in the control volume box. Still, this difference is quite small, with a relative error of 1.93%, absolutely acceptable for a dynamic simulation run. The solvent which density average is closer to the literature one is, as expected, water. This happens since water is one of the most used solvents for MD simulations and its model, already contained in the force field folder, is more accurate. The solvents density mean values obtained in presence of electric field do not vary significantly from the other simulations, with cyclohexane, hexane and water densities that become closer to the literature values while chlorobenzene and toluene density gaps slightly increase. In conclusion, this analysis gives a remarkable indication about the accuracy of the models used to describe the solvents molecules.

3.1.2 Radial Distribution Function

Another important data extracted from these simulations is the radial distribution function, indicated with $g(r)$. It describes the variation of density in function of a given point. In order to find the radial distribution function, for example, of A with respect to B, Gromacs uses this formula [12]:

$$4\pi r^2 g_{AB}(r) = V \sum_{i \in A}^{N_A} \sum_{j \in B}^{N_B} P(r)$$

where V is the volume of the box and $P(r)$ is the probability to find a B

atom at the distance r from the position of atom A.

The radial distribution functions were calculated giving as reference the center of mass of each molecule contained in the control volume. These functions were extracted from the trajectory file through the Gromacs command `gmx rdf` and plotted. In figure 3.3 it is possible to see the function of each solvent.

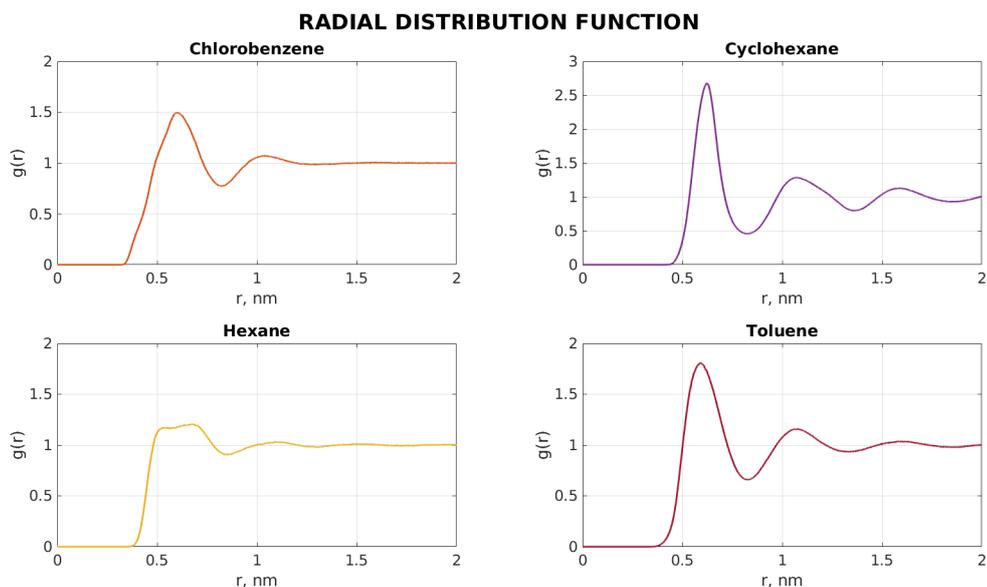


Figure 3.3. Solvents' Radial Distribution Function from MD simulation

As it is possible to observe, all the functions have a peak at the distance of roughly 0.6 nm and smaller peaks with at higher distances. This means that the molecules distribute themselves in order to have preferably a distance of 0.6 nm between their centers of mass. From the graphs it is clear that the hexane radial distribution function is the one that most differs from the others. This happens since it is the only molecule with a linear structure, while all the other molecules have a circular shape due to the presence of

an aromatic ring. Because of this reason, if compared with the other solvents, hexane have a different density variation inside the box. The radial distribution functions with center of mass reference were computed also for simulations in presence of electric field, but these plots did not show a significant difference from the each one already shown, as visible in figure 3.4

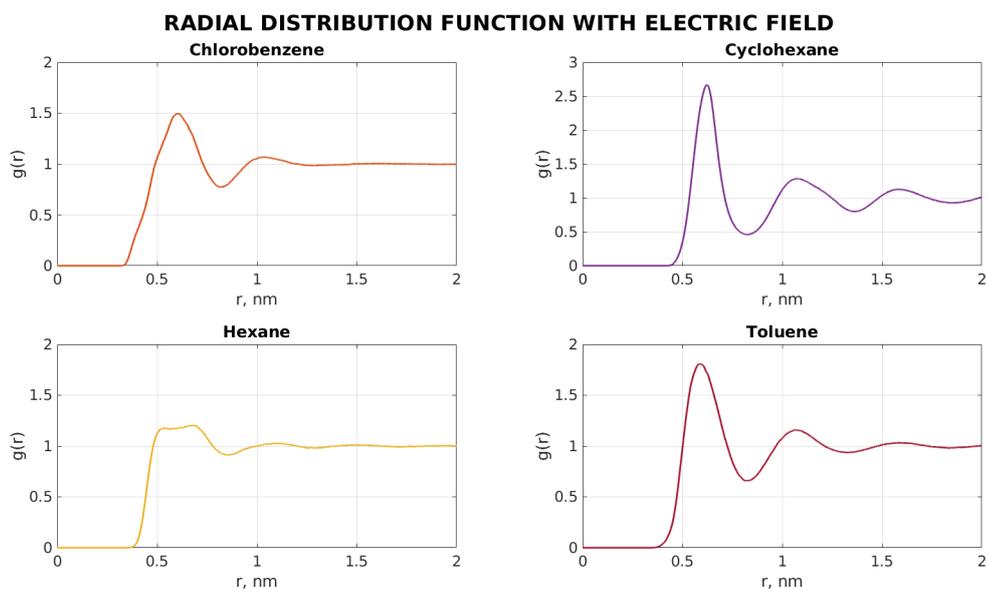


Figure 3.4. Solvents' Radial Distribution Function from MD simulation in presence of electric field

The radial distribution function of water was also calculated and compared with images found in literature. The similarities of the two images confirmed the correct procedure and parameters used during the equilibration and the final MD run. In figure 3.5 it is possible to observe the comparison of oxygen-oxygen radial distribution function water SPC/E model obtained from MD simulation and literature.

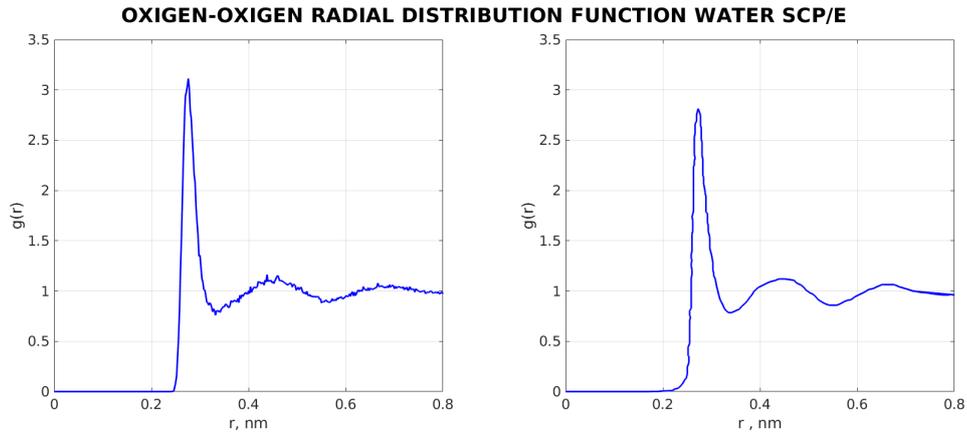


Figure 3.5. Oxygen-Oxygen Radial Distribution Function of Water SPC/E model from MD simulation (left) and literature (right) [54]

A similar plot was obtained from water MD simulation run in presence of electric field, as visible in figure 3.6.

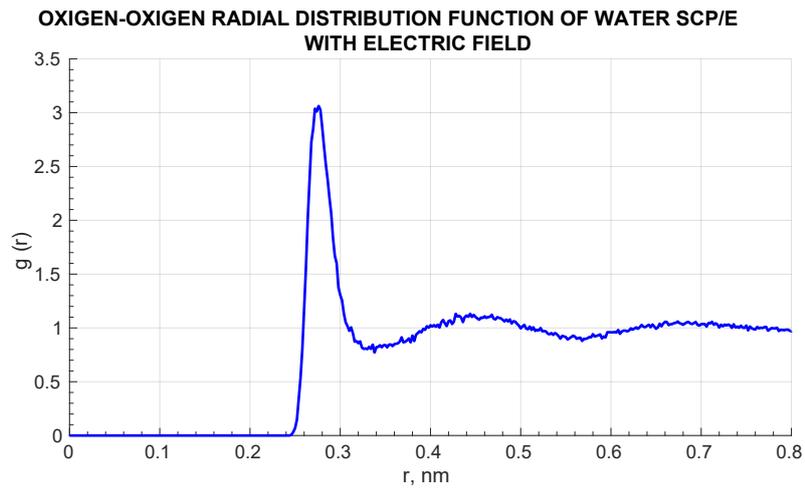


Figure 3.6. Oxygen-Oxygen Radial Distribution Function of Water SPC/E model obtained from MD simulation with electric field

3.2 Mercaptosuccinic Acid in solvents

Firstly, a check on temperature and density values was done. Once verified that the results obtained were reliable, it was possible to assert that the simulations were stable and correctly conducted and a further analysis was carried on.

In order to understand how the Mercaptosuccinic Acid interacts with the various solvents, the energies acting between a ligand molecule and the solvents were analyzed. Gromacs compute the calculation of Coulomb and Lennard-Jones energies in each time step of the simulation and these energies can be extracted from the *.edr* file once the simulation ends. The Coulomb energy trend value in case of ligand molecule immersed in toluene is shown in figure 3.7.

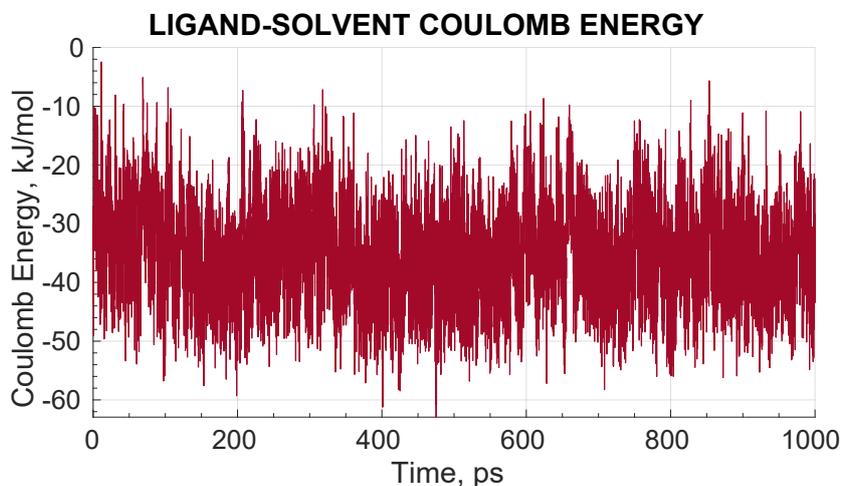


Figure 3.7. Coulomb energy trend value between a Mercaptosuccinic Acid molecule and toluene during MD simulation

The averages over time of the Coulomb and Lennard-Jones energies acting between one ligand molecule and the solvents were calculated and their values

are shown in the table 3.2

SOLVENT	Coul. Energy, $\frac{kJ}{mol}$	L-J Energy, $\frac{kJ}{mol}$
CHLOROBENZENE	-21.36	-93.88
CYCLOHEXANE	-0.97	-79.35
HEXANE	-1.02	-71.25
TOLUENE	-35.18	-85.64
WATER	-130.02	-33.88

Table 3.2. Colomb and Lennard-Jones energies acting between one Mercaptosuccinic Acid molecule and each solvent during MD simulation

In order to better understand the Coulomb and Lennard-Jones energies composition in the various solvents, the percentages were computed and represented as pie charts in figure 3.8.

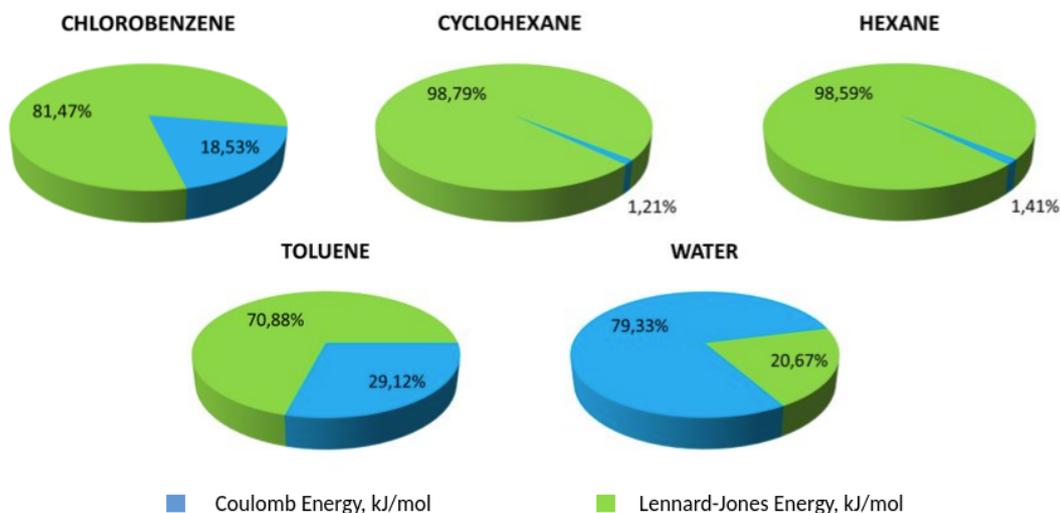


Figure 3.8. Percentage values of Coulomb and Lennard-Jones energies acting between one Mercaptosuccinic Acid molecule and the solvents

As it is possible to observe from the pie charts, for non-polar solvents the Lennard-Jones influence is much higher than Coulombian one. In particular,

the Coulomb energy in case of ligand immersed in cyclohexane and hexane is almost null. On the contrary water, the only polar solvent, show a clear prevalence of Coulombian energy. This result proves that the polarity of the solvents have a great influence on the ligand-solvent interactions.

The Coulomb and Lennard Jones energies were calculated also for the entire system. Meanwhile in the previous case only the forces acting between the ligand and the solvent molecules were extracted, the image 3.9 shows the energies calculated for the whole system, thus also the each one acting between the solvent molecules.

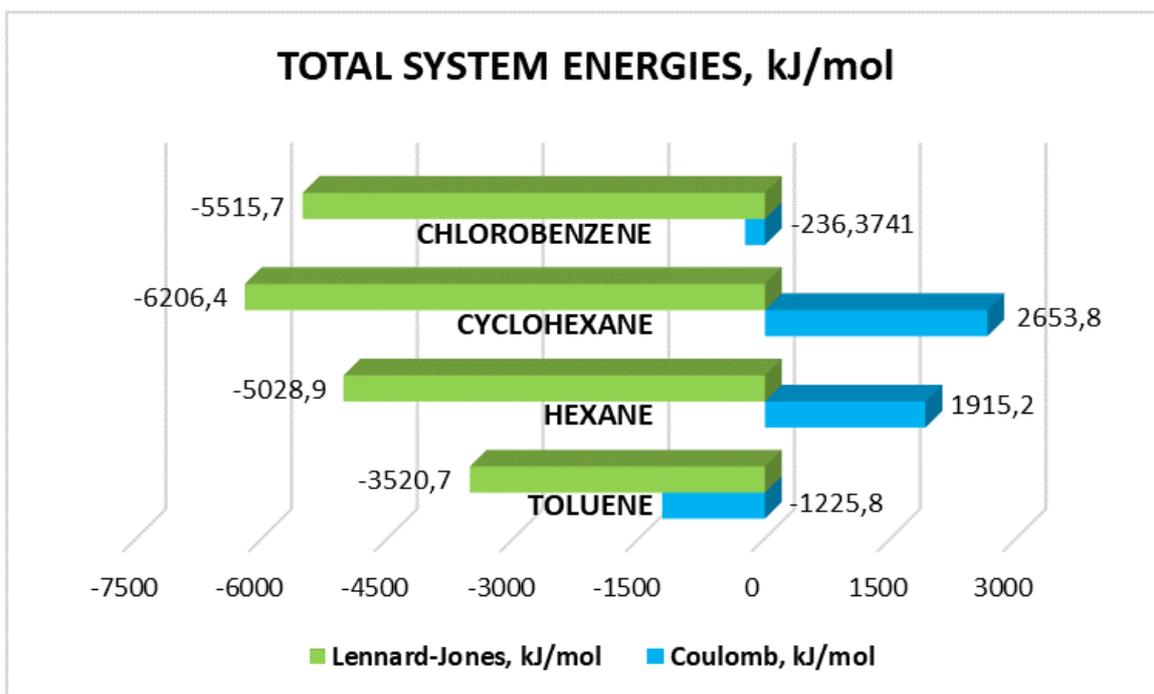


Figure 3.9. Colomb and Lennard-Jones energies acting in a box containing one Mercaptosuccinic Acid molecule and a solvent

As it is possible to notice, highly negative values of Lennard-Jones were

obtained for all the solvents. Great difference is instead notable for Coulomb energies: these are slightly negative for chlorobenzene, highly positive in cyclohexane and hexane and highly negative for toluene. Much different values were obtained in the calculation of total energies acting in case of ligand immersed in water: the coulomb energy was $-129750 \frac{kJ}{mol}$ and the Lennard-Jones energy was $21451 \frac{kJ}{mol}$. So, again a huge difference exists between the simulations of polar and non-polar solvents.

3.3 Mercaptosuccinic Acid interactions

The purpose of this analysis is to understand how two Mercaptosuccinic Acid molecules interact between them. In particular, the main goal is to understand how the center of mass distance between two ligand molecules influence their interactions. Obviously high interest is put in the comparison between the solvents in which the ligands are immersed.

3.3.1 Coulomb and Lennard Jones Energies

First at all, the Coulomb and Lennard-Jones energies acting between the two ligand molecules were extracted from the *.edr* files of each of the 32 configurations created. The mean values were then calculated and plotted in function of the center of mass distance between the two ligands. A negative energy value is attractive, while positive numbers correspond to repulsion. In figures 3.10 and 3.11 the ligand-ligand Coulomb and Lennard-Jones energies computed for ligands immersed in the solvents studied are shown with the relative standard deviations.

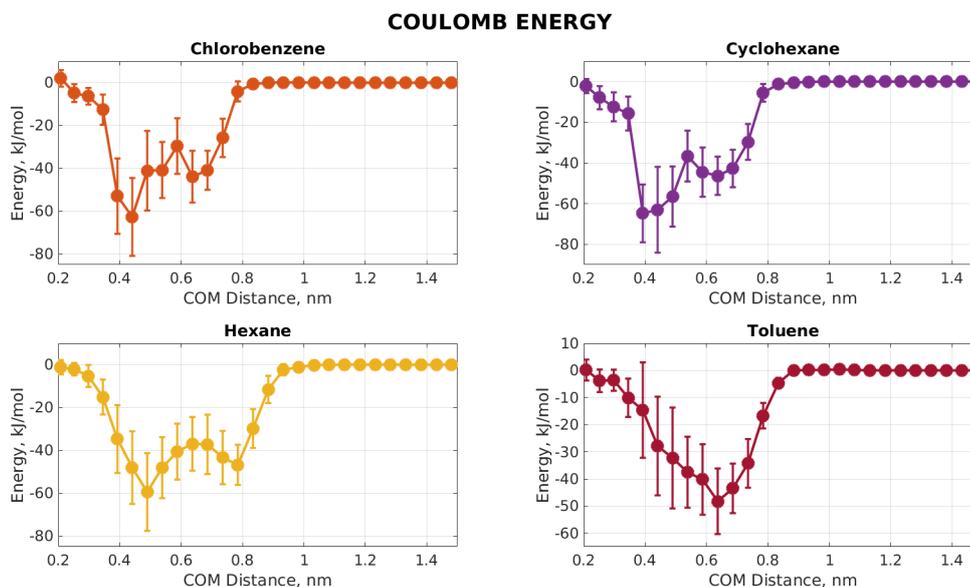


Figure 3.10. Colomb energy averages acting at different COM distances between two Mercaptosuccinic Acid molecules immersed in different solvents

In the images the colored dots represent the mean energy values calculated for each configuration, while the continuous line links the dots in order to define a continuous function. The vertical bars represent the standard deviation for each mean force calculated. As it is possible to observe from the graphs, both the Coulomb and Lennard Jones energies are null for center of mass distance higher than about 0.9 nm. Then, for Coulomb energies, the energies become negative and define clearly minimum points. For ligands immersed in toluene, the trough is clear, with a minimum force value of $-48.25 \frac{kJ}{mol}$ at a COM distance of 0.64 nm. In the other plots, instead, the trend is more fluctuating and relative minimums are identifiable. The Coulomb energies become then close to a null value for ligands COM distance close to 0.2 nm. The plots representing Lennard-Jones energies, instead, show an uncertain

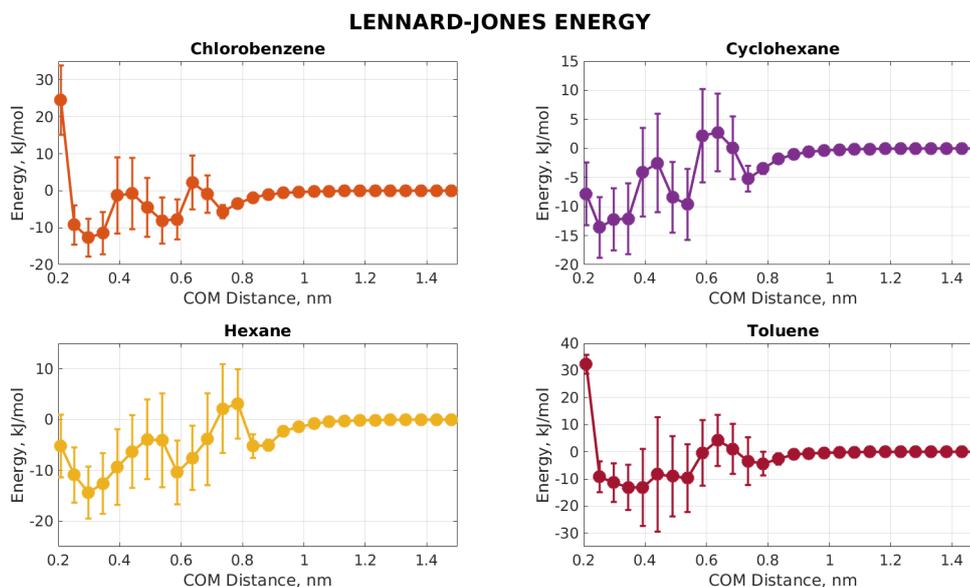


Figure 3.11. Lennard-Jones energy averages acting at different COM distances between two Mercaptosuccinic Acid molecules immersed in different solvents

evolution with decreasing and increasing values, some also positive. Still, it is possible to individuate minimum values at different COM distances. The point with minimum value corresponds to the distance where the energy is higher, thus where the two ligand molecules more interact one with the other. In order to understand which is the distance where the maximum energy play the role, the Coulomb and Lennard-Jones energies were summed. The plots that take in consideration both the energies are shown, with the relative standard deviation, in figure 3.12.

From the figure it is possible notice that the plots where Coulomb and Lennard-Jones energies are summed show a well-defined trough. The points with higher interaction between the two ligand molecules show different minimum energy values and COM distances according to the solvents in which

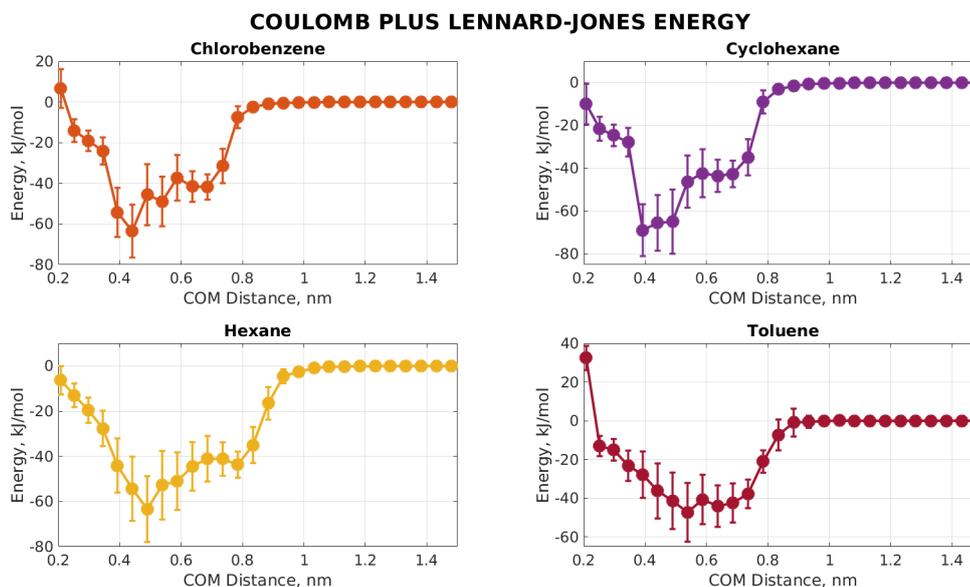


Figure 3.12. Coulomb plus Lennard-Jones energy average acting at different COM distances between two Mercaptosuccinic Acid molecules immersed in different solvents

the ligands are immersed. The table 3.3 lists the minimum energy values, calculated over the sum of the Coulomb and Lennard-Jones energies, and the relative standard deviations, together with the corresponding ligands COM distances.

SOLVENTS	Minimum Energy, $\frac{kJ}{mol}$	COM Distance, nm
CHLOROBENZENE	-63.49	0.44
CYCLOHEXANE	-68.86	0.39
HEXANE	-63.34	0.49
TOLUENE	-47.28	0.54

Table 3.3. Minimum energies acting between two Mercaptosuccinic Acid molecules immersed in solvents and corresponding COM ligand molecules distances

It has to be taken into account that the minimum energy comes from the simulations of the configurations created. Thus, the energies are exact values computed at the molecules COM distance and, since the configurations were created with a COM distance sampling of 0.05 nm, the real minimum can be situated in an approximation of 0.1 nm. For example, the actual minimum energy for chlorobenzene is situated at 0.44 nm \pm 0.05 nm. From the results listed in table 3.3 it is possible to observe that the higher energy is acting on the two ligands immersed in cyclohexane, at the distance of 0.39 nm. All the others minimum energy points are located at higher COM distances, in particular toluene, whose the minimum energy point, if compared with the ones of the other solvents, is located at the higher distance of 0.54 nm. Furthermore, it has also the higher energy value: $-47.28 \frac{kJ}{mol}$. Thus, the ligand molecules interactions, when immersed in toluene, are lower than when immersed in the other solvents.

3.3.2 Forces

A study similar to the each one described in the previous paragraph was done for the forces analysis. The forces acting on each atom of every molecule contained in the control volume were calculated in every dynamics simulation step. The forces could be extracted with the command *gmx traj* from the *.trr* file: Gromacs produced an *.xvg* file containing, for each time step, three values corresponding to the forces acting on the three spatial directions. The forces acting on the first ligand, the fixed one through all the configurations, were extracted and the force values acting along the z-axis direction were selected. According with the molecules disposition, as visible in figure 3.13, a positive force value correspond to attractive force between the two ligand

molecules.

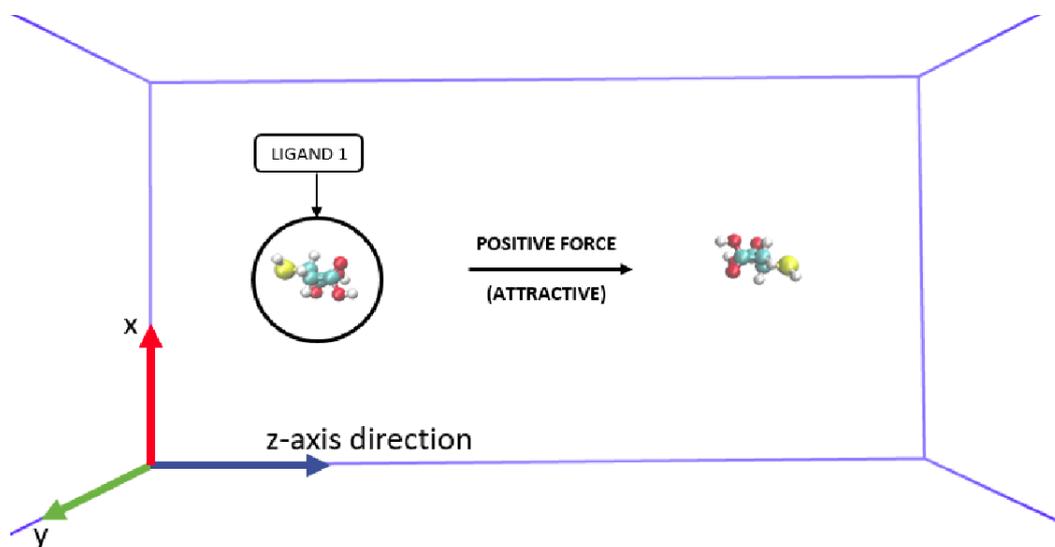


Figure 3.13. Force acting on one Mercaptosuccinic Acid molecule along z axis direction in a box containing two ligand molecules and chlorobenzene

A plot of the force acting during the entire MD simulation in case of ligands immersed in chlorobenzene is shown in figure 3.14.

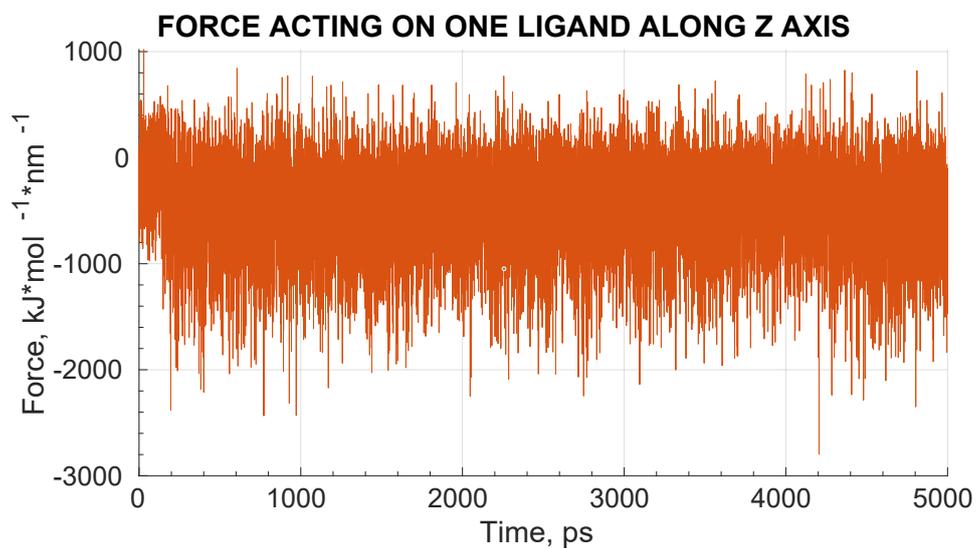


Figure 3.14. Force acting on one Mercaptosuccinic Acid molecule along z axis direction in a box containing two ligand molecules and chlorobenzene

As it is possible to observe in figure 3.14, the variation is very high, ranging roughly from $-2000 \frac{kJ}{mol \cdot nm}$ to $500 \frac{kJ}{mol \cdot nm}$. Since the variation was so important, the mean values computed had a very high standard deviation, making the mean force value not reliable. So a higher sampling during MD simulation was chosen, in order to reduce the standard deviation. In order to have a check, for some of the configurations the average of the forces acting on the second ligand molecule along the z axis direction were calculated. The averages computed on the two ligands of the same configuration had, as expected, the same absolute value with opposite sign, as visible in figure 3.15, thus meaning that not asymmetric inaccuracy occurred.

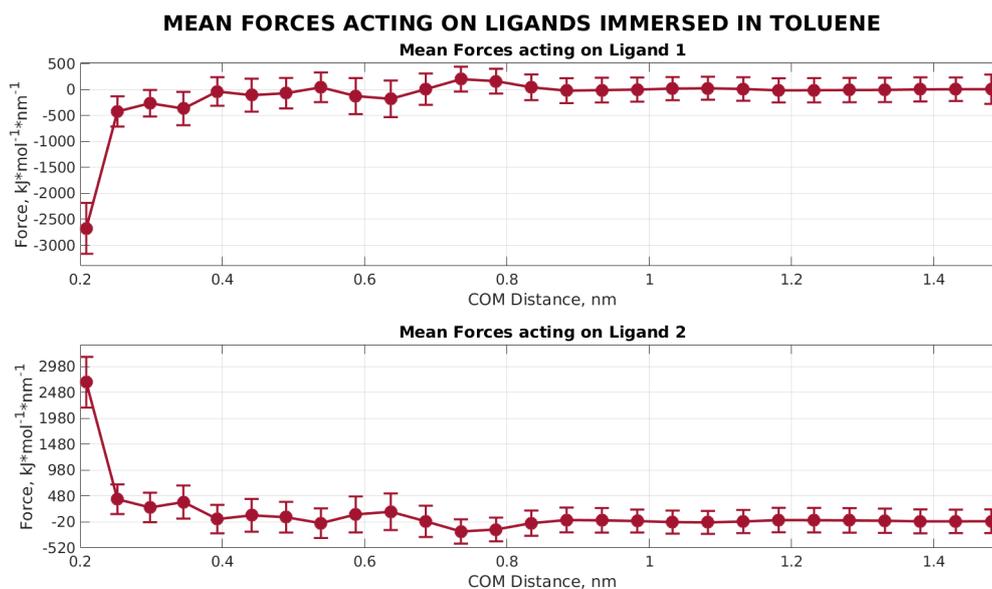


Figure 3.15. Comparison between mean forces acting on the fixed (above) and shifting (below) Mercaptosuccinic Acid molecule along z axis direction at different ligands COM distance, in a box containing two ligand molecules and toluene

Finally, the force averages and the standard deviations were performed for

each of the configurations. The mean forces and the related standard deviation for each configuration were plotted in function of the ligands COM distance, as shown in figure 3.16.

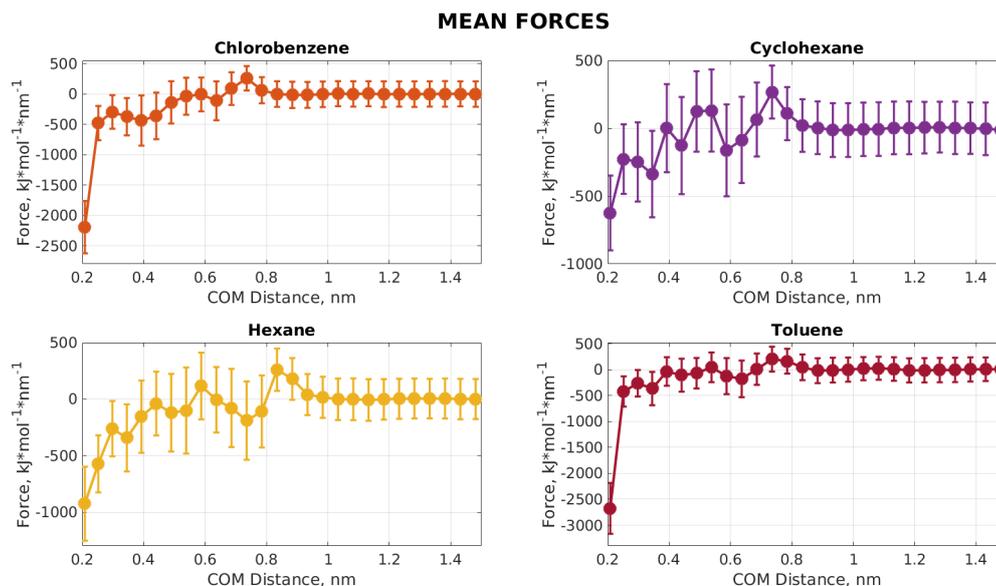


Figure 3.16. Mean forces and standard deviations acting on one Mercaptosuccinic Acid molecule along z axis direction at different ligands COM distance, in a box containing two ligand molecules and non-polar solvents

From the plots, it is possible to observe that the forces were almost null for distances higher than roughly 0.9 nm. Then, the forces had an uncertain trend, with positive and negative values. Highly negative numbers were obtained in configurations where the COM distance was very short, thus meaning highly repulsive forces between the ligands were occurring. The same observation could be done in case of MD simulation of ligands immersed in water whose mean forces plot is shown in figure 3.17.

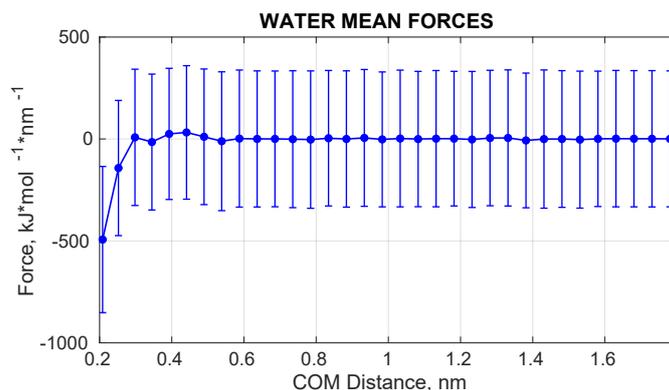


Figure 3.17. Mean forces and standard deviations acting on one Mercaptosuccinic Acid molecule along z axis direction at different ligands COM distance, in a box containing two ligand molecules and water

Nevertheless, neglecting the highly negative values located at very close COM distances, it was possible to find the minimum mean force values, listed in table 3.4 together with the corresponding ligands COM distances.

SOLVENTS	Minimum Force, $\frac{kJ}{mol*nm}$	COM Distance, nm
CHLOROBENZENE	-437.78	0.39
CYCLOHEXANE	-336.85	0.34
HEXANE	-343.39	0.34
TOLUENE	-366.42	0.34
WATER	-15.08	0.34

Table 3.4. Minimum force acting on one Mercaptosuccinic Acid molecule along z -axis direction and corresponding ligands COM distance

From the table it is clear that the minimum force values obtained in case of ligands immersed in non-polar solvents were much lower compared with the scenario of ligands solvated in water. Furthermore, it is possible to notice that the minimum force values were situated at the ligands COM distance of 0.34 nm for ligands immersed in all the solvents, except for ligands solvated

in chlorobenzene, where the minimum force was occurring at 0.39 nm.

The computation of mean forces was done also in MD simulations in presence of electric field. The plots obtained are shown in figure 3.18.

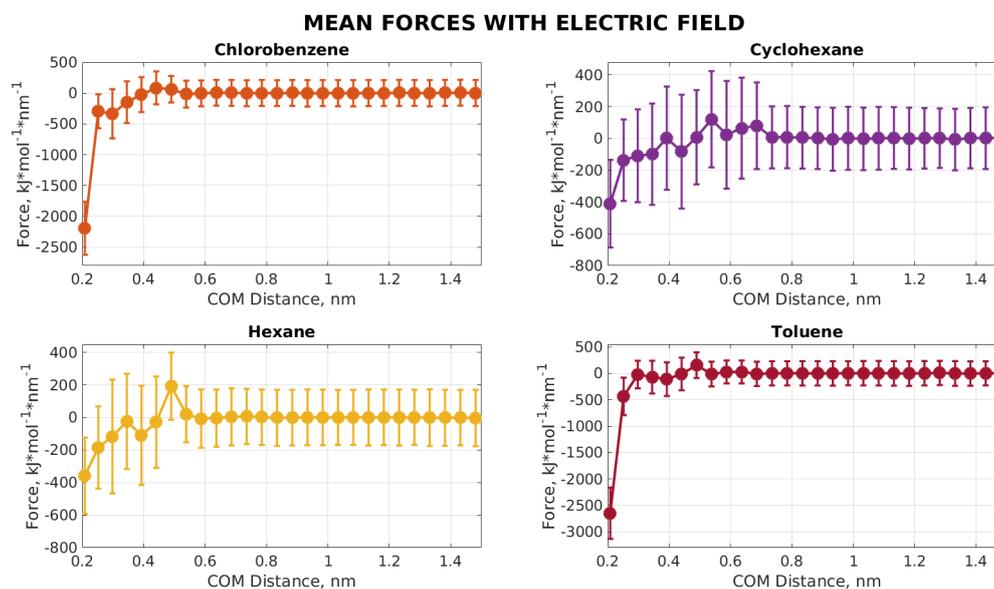


Figure 3.18. Mean forces and standard deviations acting on one Mercaptosuccinic Acid molecule along z axis direction at different ligands COM distance, in a box containing two ligand molecules and non-polar solvents with electric field

Again, as for MD simulations without electric field, the plots show null forces for ligands COM distances longer that 0.9 nm and highly repulsive forces for short ligands COM distances. The plot of ligands immersed in water in presence of electric field is show in figure 3.19

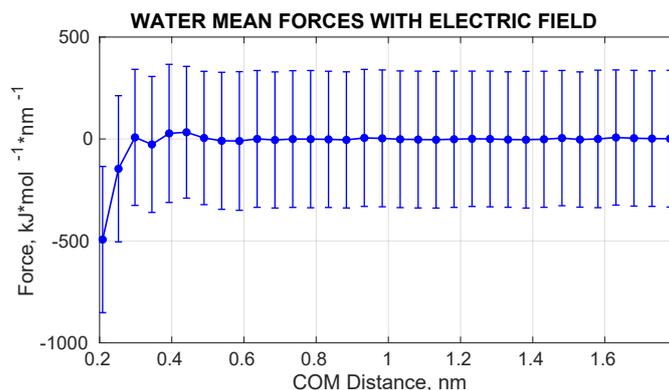


Figure 3.19. Mean forces and standard deviations acting on one Mercaptosuccinic Acid molecule along z axis direction at different ligands COM distance, in a box containing two ligand molecules and water with electric field

The table 3.5 lists the minimum forces acting on the first ligand molecule and corresponding ligands COM distance.

SOLVENTS	Minimum Force, $\frac{kJ}{mol*nm}$	COM Distance, nm
CHLOROBENZENE	-336.43	0.35
CYCLOHEXANE	-107.24	0.44
HEXANE	-109.67	0.44
TOLUENE	-110.46	0.44
WATER	-26.83	0.34

Table 3.5. Minimum force acting on one Mercaptosuccinic Acid molecule along z-axis direction and corresponding ligands COM distance in presence of electric field

The results changed significantly from the simulations without electric field. The minimum force value was lower for simulations with electric field in case of ligands immersed in polar solvents, thus water. Instead in ligands solvated in non-polar solvents scenarios, the minimum force values were much higher if electric field was present. In particular for cyclohexane, hexane

and toluene, the minimum force values raised from less than $-300 \frac{kJ}{mol*nm}$ to slightly less than $-100 \frac{kJ}{mol*nm}$. Furthermore, for these three solvents, the ligands COM distance shifted from 0.34 nm for simulations without electric field to 0.44 nm when the electric field was present. In case of simulations with chlorobenzene, the minimum force still raised, from $-437.78 \frac{kJ}{mol*nm}$ to $-336.43 \frac{kJ}{mol*nm}$, but the COM distance, on the contrary with respect of the other solvents, diminished from 0.39 nm to 0.35 nm. In case of ligands immersed in water, instead, the minimum force value was 0.34 nm both for simulations with and without electric field. Finally, it is possible to state that the influence of electric field significantly modified the forces acting between the two ligands. In particular, ligands immersed in the polar solvent increased the repulsive forces, while ligands solvated in non-polar solvents diminished it, with a higher electric field influence in case of the solvents cyclohexane, hexane and toluene, with respect to chlorobenzene.

3.3.3 Potential of Mean Forces

In order to evaluate the potential value that generate the forces acting during the MD run, the potential of mean force, also called PMF, was calculated. Reminding that the force is given by:

$$\vec{F} = -\vec{\nabla}U \quad (3.1)$$

A decreasing PMF trend for positive mean force values and increasing evolution for negative mean forces are expected.

The plot of PMF values for two ligand molecules immersed in non-polar solvents is shown in figure 3.20.

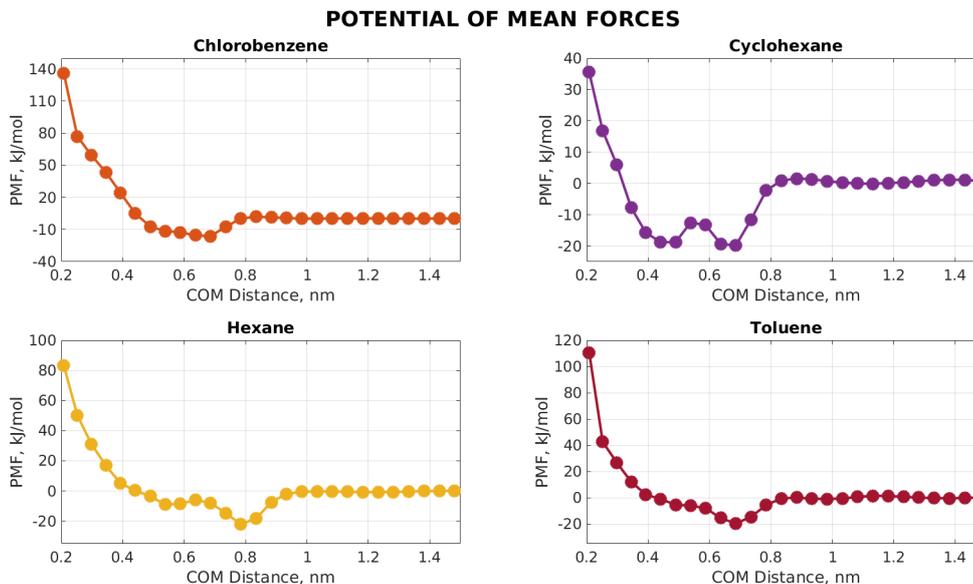


Figure 3.20. Potential of mean force acting on one Mercaptosuccinic Acid molecule along z axis direction in a box containing two ligand molecules and non-polar solvents

As expected from the force analysis, also the potential of mean force is almost null for configurations with ligands COM distance higher than 0.9 nm. After that, a trough is defined by the potentials calculated in each COM distance step. Finally, the increasing PMF values situated at shorter ligands COM distance highly increases. These correspond to the minimum force values: when the ligands COM are very close, the two molecules initiate overlapping one over the other, thus resulting into highly repulsive forces. In figure 3.21 the PMF of two ligand molecules solvated in water is shown.

The values of minimum PMF with related COM distance, both for polar and non-polar solvents, are listed in table 3.6.

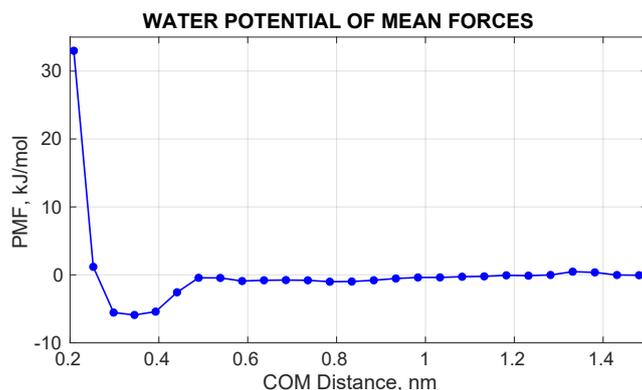


Figure 3.21. Potential of mean forces acting on one Mercaptosuccinic Acid molecule along z axis direction at different ligands COM distance, in a box containing two ligand molecules and water

SOLVENTS	Minimum PMF, $\frac{kJ}{mol}$	COM Distance, nm
CHLOROBENZENE	-16.10	0.69
CYCLOHEXANE	-19.83	0.69
HEXANE	-22.04	0.78
TOLUENE	-19.69	0.69
WATER	-5.90	0.34

Table 3.6. Minimum potential of mean force acting on one Mercaptosuccinic Acid molecule along z -axis direction and corresponding ligands COM distance

The minimum PMF values and the corresponding COM distances for ligands immersed in non-polar solvents were similar. In particular, for ligands molecules immersed in chlorobenzene, cyclohexane and toluene, the minimum PMF was computed in the configuration corresponding to a ligands COM distance of 0.69 nm. The minimum value of PMF was acting in case of ligands immersed in hexane, with $-22.04 \frac{kJ}{mol}$ at the ligands COM distance of 0.78 nm.

Given that the Coulomb and Lennard-Jones energies take in consideration

enthalpic effects, while the PMF computations consider the free energy, and reminding the Gibbs free energy equation:

$$\Delta G = \Delta H - T\Delta S \quad (3.2)$$

a comparison between the Coulomb and Lennard-Jones energy values listed in table 3.3 (page 57) and the PMF results shown in table 3.6 is significant. From the results, it is possible to state that a reduction of attractive interaction between the ligand molecules is evident in the PMF computations, resulting in entropic effects of the solvent molecules situated between the two ligands, which screened the attractive interactions. Furthermore it is notable that for ligands immersed in non-polar solvents, the minimum Coulomb and Lennard-Jones were obtained for ligands COM distances situated between 0.4 nm and 0.5 nm (table 3.3), while the minimum PMF values were computed at ligands COM distances of roughly 0.7 nm (table 3.6). This difference is due to the solvents molecules obstruction that do not allow the ligands molecules to get closer. In order to have a ligands self-assembly it is then necessary to overcome the solvents barrier that, as visible from the solvents radial distribution functions shown in figure 3.3 (page 48), is prevalent at a distance of roughly 0.6 nm. The PMF results also highlighted the difference between polar and non-polar solvents: the minimum PMF values were about $-20 \frac{kJ}{mol}$ for ligands immersed in non-polar solvents and only $-5.90 \frac{kJ}{mol}$ in ligands solvated in water scenario. The lower value obtained for ligands immersed in water signifies that the polar solvent screened more intensely the ligands interactions. Moreover, the minimum PMF in case of ligands immersed in water is localized at a ligands COM distance of 0.34 nm, according

with the radial distribution function obtained for water molecules.

The same procedure was used to elaborate the forces obtained from simulations in presence of electric field. In figure 3.22 the PMF of ligands immersed in non-polar solvents are shown.

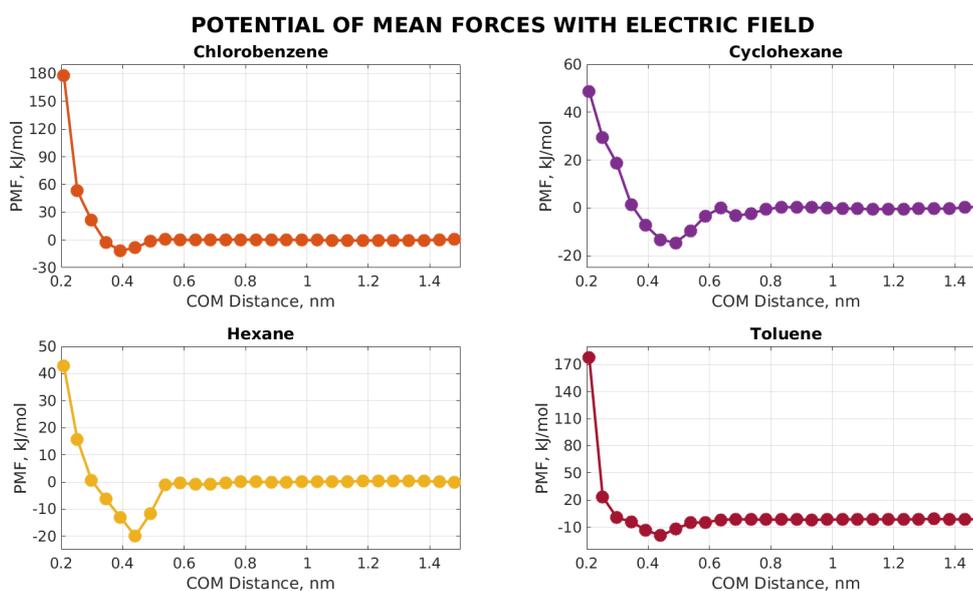


Figure 3.22. Potential of mean force acting on one Mercaptosuccinic Acid molecule along z axis direction in a box containing two ligand molecules and non-polar solvents with electric field

As for the previous plot, almost null values were computed for ligands COM distance higher than 0.9 nm. The PMF values then show a trough, situated between 0.4 nm and 0.5 nm, and increased significantly for shorter ligands COM distances. The PMF trend computed in case of ligands immersed in water in presence of electric field is represented in figure 3.23.

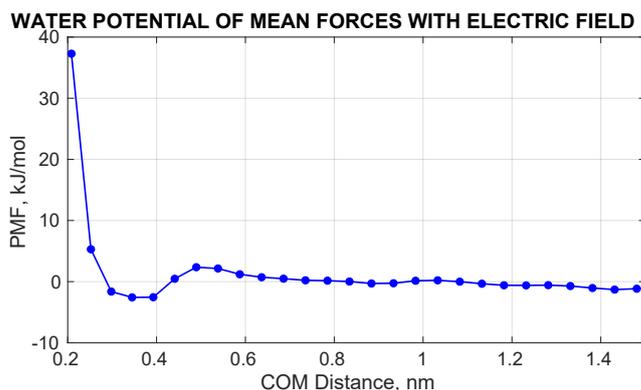


Figure 3.23. Potential of mean forces acting on one Mercaptosuccinic Acid molecule along z axis direction at different ligands COM distance, in a box containing two ligand molecules and water with electric field

In table 3.7 are listed the minimum PMF values and the related ligands COM distances for ligands immersed in polar and non-polar solvents.

SOLVENTS	Minimum PMF, $\frac{kJ}{mol}$	COM Distance, nm
CHLOROBENZENE	-11.42	0.44
CYCLOHEXANE	-14.59	0.49
HEXANE	-19.95	0.49
TOLUENE	-19.30	0.49
WATER	-2.57	0.34

Table 3.7. Minimum potential of mean force acting on one Mercaptosuccinic Acid molecule along z -axis direction and corresponding ligands COM distance in presence of electric field

As for the minimum force computation, also in the potential of mean force study the minimum values obtained from of simulations with electric field are higher with respect to simulations without it. In particular, a relevant increase is notable in case of chlorobenzene: its minimum PMF raised from $-16.10 \frac{kJ}{mol}$ to $-11.42 \frac{kJ}{mol}$.

For ligands immersed in cyclohexane, hexane and toluene, the minimum PMF values were situated at a ligands COM distance of 0.49 nm, while for ligands solvated in chlorobenzene it was located at 0.44 nm. These values are significantly different if compared with the minimum PMF values obtained without electric field, where the ligands COM distances for minimum PMF values were located at roughly 0.7 nm. Thus meaning that the electric field permitted the ligands to overcome the obstruction given by the presence of non-polar solvents molecules. In case of ligands immersed in polar solvent, instead, the ligands COM distance related with the minimum PMF value was 0.34 nm both for simulations with and without electric field. Remarkable difference is instead notable in the minimum PMF values: the minimum PMF obtained from simulation with electric field was halve the minimum PMF computed during simulation without electric field, thus resulting in a minor entropy level when electric field was acting.

Chapter 4

Conclusions

The main purpose of this thesis was to perform a primary analysis about the behaviour of Mercaptosuccinic Acid molecules immersed in polar and non-polar solvents.

First, MD simulations of boxes containing chlorobenzene, cyclohexane, hexane, toluene and water were carried out. The successful comparison between the results and the literature data proved that the solvent models and the force field chosen in this work were reliable. The solvents behaviour under the electric field was also studied, but no relevant differences were noticed in the molecules distribution. Interactions between one Mercaptosuccinic Acid molecule and the solvents were then analyzed. The results have shown a remarkable difference between polar and non-polar solvents: the attractive interactions for non-polar solvents were mainly due to Lennard Jones energy, with a minor component given by Coulomb energy; instead for polar solvent the ligand-solvent interactions were much higher and Coulomb energy played a major role over Lennard-Jones energy.

The simulations of two Mercaptosuccinic Acid molecules immersed in polar

and non-polar solvents showed that the interactions were null for ligands COM distance higher than 0.9 nm, independently from the solvent in which the ligands were immersed. The simulations of ligands immersed in non-polar solvents highlighted that the attractive energy due to Coulomb and Lennard-Jones potentials was much higher than the calculated PMF interactions. Through the Gibbs free energy equations $\Delta G = \Delta H - T\Delta S$, it was possible to state that the differences between enthalpic and free energies were due to entropy. The entropic effects were given by the solvents molecules located between the two ligands and screened their attractive interactions. Also, from simulations of ligands immersed in non-polar solvents, the maximum attractive Coulomb and Lennard-Jones energies corresponded to the ligands COM distances located between 0.4 nm and 0.5 nm, while the minimum PMF values were obtained at ligands COM distances of roughly 0.7 nm. This difference was again related to the solvents molecules obstruction, which did not allowed the ligands to spontaneously get closer. The ligands, in order to self-assembly, must overcome the solvents barrier that, as suggested from the rdf calculation, was located at 0.6 nm. Comparing the PMF values computed from simulations of ligands immersed in polar and non-polar solvents, it resulted that water had a more intense screening effect on the interaction energies acting between the ligands. Finally, the presence of electric field reduced the ligands free energy interactions both in simulations with polar and non-polar solvents, with higher influence for ligands immersed in water. In fact, in this case the free energy value halved the each one obtained from simulations without electric field.

Bibliography

- [1] H. Jeong, E. Choi, E. Ellis, T. Lee, Recent advances in gold nanoparticles for biomedical applications: from hybrid structures to multi-functionality, *J. Mater. Chem.*, v.7, pp.3480-3496, 2019.
- [2] S. Palagi, P. Fischer, Bioinspired microrobots, *Nature Reviews Materials*, v.3, c.6, pp.113–124, 2018.
- [3] J. Wang, H. Z. Zhang, R. S. Li, C. Z. Huang, Localized surface plasmon resonance of gold nanorods and assemblies in the view of biomedical analysis, *Trends in Analytical Chemistry*, v.80, pp. 429–443, 2016.,
- [4] A. Singh, C. Coughlan, F. Laffir and K. M. Ryan, Assembly of $CuIn_{1-x}Ga_xS_2$ Nanorods into Highly Ordered 2D and 3D Superstructures, *ACS Nano*, v.6, n.8, pp.6977-6983, 2012.
- [5] A. Singh, R. D. Gunning, S. Ahmed, C. A. Barrett, N. J. English, J. A. Garate, K. M. Ryan, Controlled semiconductors nanorod assembly from solution: influence of concentration, charge and solvent nature, *Journal of Materials Chemistry*, v. 22, pp. 1562-1569, 2012.
- [6] P. Liu, S. Singh, Y. Guo, J. J. Wang, H. Xu, C. Silien, N. Liu, K. M. Ryan, Assembling Ordered Nanorod Superstructures and Their Application as Microcavity Lasers, *Scientific Reports*, srep 43884, 2017.
- [7] A. Hospital, J. R. Goñi, M. Orozco, J. L. Gelpi, Molecular dynamics simulations: advances and applications *Advanced and applications in bioinformatics and chemistry*, v.8 , pp.37-47, 2015.
- [8] L. Eisenbud, On the Classical Laws of Motion, *Am. J. Phys.*, v.26, p.145, 1958.
- [9] "Atoms In Motion - Chapter 5 - MD", atomsinmotion.com, Web, 21 March 2020.
- [10] R. W. Hockney, P. Goel, J. Eastwood, Quiet High Resolution Computer Models of a Plasma, *J. Comp. Phys.*, v.14, pp.148–158, 1974.

- [11] W. C. Swope, H. C. Andersen, P. H. Berens, K. R. A. Wilson, Computer-simulation method for the calculation of equilibrium-constants for the formation of physical clusters of molecules: Application to small water clusters, *J. Chem. Phys.*, v.76, pp.637–649, 1982.
- [12] M.J. Abraham, D. van der Spoel, E. Lindahl, B. Hess, and the GRO-MACS development team, GROMACS User Manual version 2018, 2018.
- [13] H. J. C. Berendsen, J. P. M. Postma, A. DiNola, J. R. Haak, Molecular dynamics with coupling to an external bath, *J. Chem. Phys.*, v.81, pp.3684–3690, 1984.
- [14] S. Nosé, A molecular dynamics method for simulations in the canonical ensemble, *Mol. Phys.*, v.52, pp.255–268, 1984.
- [15] W. G. Hoover, Canonical dynamics: equilibrium phase-space distributions, *Phys. Rev. A*, v.31, pp.1695–1697, 1985.
- [16] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, *J. Appl. Phys.*, v.52, pp.7182–7190, 1981.
- [17] S. Nosé, M. L. Klein, Constant pressure molecular dynamics for molecular systems, *Mol. Phys.*, v.50, pp.1055–1076, 1983.
- [18] G.J. Martyna, M.E. Tuckerman, D.J. Tobias, M.L. Klein, Explicit reversible integrators for extended systems dynamics, *Mol. Phys.*, v.87, pp.1117–1157, 1996.
- [19] M.E. Tuckerman, J. Alejandre, R. López-Rendón, A.L. Jochim, and G.J. Martyna, A Liouville-operator derived measure-preserving integrator for molecular dynamics simulations in the isothermal-isobaric ensemble, *J. Phys. A.*, v.59, pp.5629–5651, 2006.
- [20] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C.Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers *Elsevier, SoftwareX* 1–2, v.19, n.25, 2015.
- [21] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.*, v.117, c.19, pp.5179–5197, 1995.
- [22] P.A. Kollman, Advances and Continuing Challenges in Achieving Realistic and Predictive Simulations of the Properties of Organic and Biological Molecules. *Acc. Chem. Res.*, v.29 c.10 pp.461–469, 1996.
- [23] J. M. Wang, P. Cieplak, P. A. Kollman, How Well Does a Restrained

- Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules, *J. Comp. Chem.*, v.21 c.12 pp.1049–1074, 2000.
- [24] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters, *PROTEINS: Struct. Funct. Gen.*, v.15 c.65, pp.712–725, 2006.
- [25] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dorr, D. E. Shaw, Improved side-chain torsion potentials for the AMBER ff99SB protein force field, *PROTEINS: Struct. Funct. Gen.*, v.78, pp.1950–1958, 2010.
- [26] Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P. A. Kollman, Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations, *J. Comp. Chem.*, v.24 c.16 pp.1999–2012, 2003.
- [27] A. E. García, K. Y. Sanbonmatsu, α -Helical stabilization by side chain shielding of backbone hydrogen bonds, *Proc. Natl. Acad. Sci. USA*, v.99 c.5, pp.2782–2787, 2002.
- [28] J. A. D. MacKerell, M. Feig, C. L. Brooks III, Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations, *J. Comp. Chem.*, v.25 c.11, pp.1400–15, 2004.
- [29] A. D. MacKerell, D. Bashford, Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B.*, v.102 c.18 pp.3586–3616, 1998.
- [30] S. E. Feller, A. D. MacKerell, An improved empirical potential energy function for molecular simulations of phospholipids, *J. Phys. Chem. B.*, v.104 c.31 pp.7510–7515, 2000.
- [31] N. Foloppe, A. D. MacKerell, All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data, *J. Comp. Chem.*, v.21 c.2 pp.86–104, 2000.

- [32] A. D. MacKerell, N. K. Banavali, All-atom empirical force field for nucleic acids: II. application to molecular dynamics simulations of DNA and RNA in solution, *J. Comp. Chem.*, v.21 c.2 pp.105–120, 2000.
- [33] P. Bjelkmar, P. Larsson, M. A. Cuendet, B. Hess, E. Lindahl, Implementation of the CHARMM force field in GROMACS: Analysis of protein stability effects from correction maps, virtual interaction sites, and water models, *J. Chem. Theory Comput.*, v.6 pp.459–466, 2010.
- [34] P. Larsson, E. A. Lindahl, High-Performance Parallel-Generalized Born Implementation Enabled by Tabulated Interaction Rescaling, *J. Comp. Chem.* v.31 c.14 pp.2593–2600, 2010.
- [35] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, I. G. Tironi, Biomolecular Simulation: The GROMOS96 manual and user guide, *Zürich, Switzerland: Hochschulverlag AG an der ETH Zürich*, 1996.
- [36] C. Oostenbrink, A. Villa, A. E. Mark, W. F. Van Gunsteren, A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6, *Journal of Computational Chemistry*, v.25 c.13 pp.1656–1676, 2004.
- [37] V. Rühle, C. Junghans, A. Lukyanov, K. Kremer, D. Andrienko, Versatile Object-Oriented toolkit for Coarse-Graining applications, *J. Chem. Theory Comput.*, v.5 c.12 pp.3211–3223, 2009.
- [38] T. Bereau, Z. J. Wang, M. Deserno, Solvent-free coarse-grained model for unbiased high-resolution protein-lipid interactions, submitted.
- [39] Z. J. Wang, M. Deserno, A systematically coarse-grained solvent-free model for quantitative phospholipid bilayer simulations, *J. Phys. Chem. B.* v.114 c.34 pp.11207–11220, 2010.
- [40] W. Kuliga, M. Pasenkiewicz-Gierulab, T. Róg, Topologies, structures and parameter files for lipid simulations in GROMACS with the OPLS-aa force field: DPPC, POPC, DOPC, PEPC, and cholesterol, *Chemistry and Physics of Lipids*, v.195 pp.12–20, 2016.
- [41] O. Beckstein, B. I. Iorga, Prediction of hydration free energies for aliphatic and aromatic chloro derivatives using molecular dynamics simulations with the OPLS-AA force field, *J Comput Aided Mol Des*, v.26, pp.635–645, 2012.
- [42] I. M. Kenney, O. Beckstein, B. I. Iorga, Prediction of cyclohexane-water distribution coefficients for the SAMPL5 data set using molecular dynamics simulations with the OPLS-AA force field, *J Comput Aided Mol Des*, v.30, pp.1045–1058 2016.

- [43] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids, *J. Am. Chem. Soc.*, v.118 pp.11225-11236, 1996.
- [44] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules, *J. Am. Chem. Soc.*, v.117, pp.5179-51977, 1995.
- [45] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, *The Protein Data Bank Nucleic Acids Research*, v.28, pp.235-242, 2000.
- [46] B. S. Furniss, A. J. Hannaford, P. W. G. Smith, A. R. Tatchell, *Textbook of practical organic chemistry*, V, New York, Vogel, 1989.
- [47] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, P. A. Kollman, The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, *Journal of Computational Chemistry*, v.13, c.8, pp.1011-1021, 1992.
- [48] Lemkul, Justin, "GROMACS Tutorial-Tutorial 3: Umbrella Sampling", mdtutorials.com, Virginia Tech Department of Biochemistry, Royal Institute of Technology and the Uppsala University, Web, 20 March 2020.
- [49] L. Martínez, R. Andrade, E. G. Birgin, J. M. Martínez, Packmol: A package for building initial configurations for molecular dynamics simulations, *Journal of Computational Chemistry*, v.30 c.13, pp.2157-2164, 2009.
- [50] G. Monegato, *Metodi e algoritmi per il CALCOLO NUMERICO*, 9788879922654, Torino, Edizioni C.L.U.T., 2008.
- [51] Lemkul, Justin, "GROMACS Tutorial-Tutorial 1: Lysozyme in Water", mdtutorials.com, Virginia Tech Department of Biochemistry, Royal Institute of Technology and the Uppsala University, Web, 20 March 2020.
- [52] T. Jarusuwannapoom, W. Hongrojjanawiwat, S. Jitjaicham, L. Wannatong, M. Nithitanakul, C. Pattamaprom, P. Koombhongse, R. Rangkupan, P. Supaphol, Effect of solvents on electro-spinnability of polystyrene solutions and morphological appearance of resulting electro-spun polystyrene fibers, *European Polymer Journal*, v.41, pp.409-421, 2005.
- [53] C. Vega, C. McBride, E. Sanz, J. L. F. Abascal, Radial distribution functions and densities for the SPC/E, TIP4P and TIP5P models for liquid water and ices Ih, Ic, II, III, IV, V, VI, VII, VIII, IX, XI and XII,

- Phys. Chem. Chem. Phys.*, v.7, c.7, pp.1450–1456, 2015.
- [54] G. N. I. Clark, C. D. Cappa, J. D. Smith, R. J. Saykally, T. Head-Gordon, The structure of ambient water, *Molecular Physics*, v.108, c.11, pp.1415–1433, 2010.

Appendix A

```
title          = ENERGY MINIMIZATION OF TWO LIGAND MOLECULES IMMERSSED IN HEXANE
; Parameters describing what to do, when to stop and what to save
integrator     = steep      ; Algorithm (steep = steepest descent minimization)
emtol         = 1000.0     ; Stop minimization when the maximum force < 1000.0 kJ/mol/nm
emstep        = 0.01      ; Energy step size
nsteps        = 50000     ; Maximum number of (minimization) steps to perform

; Parameters describing how to find the neighbors of each atom and how to calculate the interactions
nstlist       = 1         ; Frequency to update the neighbor list and long range forces
cutoff-scheme = Verlet    ; Buffered neighbor searching
ns_type       = grid      ; Method to determine neighbor list (simple, grid)
rlist         = 1.4       ; Cut-off for making neighbor list (short range forces)
coulombtype   = PME       ; Treatment of long range electrostatic interactions
rcoulomb      = 1.4       ; Short-range electrostatic cut-off
rvdw          = 1.4       ; Short-range Van der Waals cut-off
pbc           = xyz       ; Periodic Boundary Conditions
```

Appendix B

```
;title = NVT EQUILIBRATION OF TWO LIGAND MOLECULES IMMERSSED IN HEXANE
define = -DPOSRES_LIG_1 -DPOSRES_LIG_2 ; position restrain the ligands
; Run parameters
integrator = md ; leap-frog integrator
nsteps = 1000000 ; 0.001 * 1000000 = 1 ns
dt = 0.001 ; 1 fs
;Output control
nstxout = 1000 ; save coordinates every 1 ps
nstvout = 1000 ; save velocities every 1 ps
nstenergy = 1000 ; save energies every 1 ps
nstlog = 1000 ; update log file every 1 ps
nstcalcenergy = 1000 ; calculate energies every 1 ps
; Bond parameters
continuation = no ; first dynamics run
constraint_algorithm = lincs ; holonomic constraints
constraints = h-bonds ; bonds involving H are constrained
lincs_iter = 1 ; accuracy of LINCS
lincs_order = 4 ; also related to accuracy
; Nonbonded settings
cutoff-scheme = Verlet ; Buffered neighbor searching
ns_type = grid ; search neighboring grid cells
nstlist = 10 ; 10 fs, largely irrelevant with Verlet
rcoulomb = 1.0 ; short-range electrostatic cutoff (in nm)
rvdw = 1.0 ; short-range van der Waals cutoff (in nm)
DispCorr = EnerPres ; account for cut-off vdW scheme
; Electrostatics
coulombtype = PME ; Particle Mesh Ewald for long-range electrostatics
pme_order = 4 ; cubic interpolation
fourierspacing = 0.16 ; grid spacing for FFT
; Temperature coupling is on
tcoupl = V-rescale ; V-rescale thermostat
tc-grps = LIG HEX ; two coupling groups - more accurate
tau_t = 0.1 0.1 ; time constant, in ps
ref_t = 300 300 ; reference temperature, one for each group, in K
; Pressure coupling is off
pcoupl = no ; no pressure coupling in NVT
; Periodic boundary conditions
pbc = xyz ; 3-D PBC
; Velocity generation
gen_vel = yes ; assign velocities from Maxwell distribution
gen_temp = 300 ; temperature for Maxwell distribution
gen_seed = -1 ; generate a random seed
; Freezing
freezegrps = SC_1 SC_2 ; freezing groups
freezedim = Y Y Y Y Y Y ; directions of freezing
```

Appendix C

```
title = NPT EQUILIBRATION OF TWO LIGAND MOLECULES IMMERSED IN HEXANE
define = -DPOSRES_LIG_1 -DPOSRES_LIG_2 ; position restrain the ligands
; Run parameters
integrator = md ; leap-frog integrator
nsteps = 1000000 ; 0.001 * 1000000 = 1 ns
dt = 0.001 ; 1 fs
; Output control
nstxout = 1000 ; save coordinates every 1 ps
nstvout = 1000 ; save velocities every 1 ps
nstenergy = 1000 ; save energies every 1 ps
nstlog = 1000 ; update log file every 1 ps
nstcalcenergy = 1000 ; calculate energies every 1 ps
; Bond parameters
continuation = yes ; Restarting after NVT
constraint_algorithm = lincs ; holonomic constraints
constraints = h-bonds ; bonds involving H are constrained
lincs_iter = 1 ; accuracy of LINCS
lincs_order = 4 ; also related to accuracy
; Nonbonded settings
cutoff-scheme = Verlet ; Buffered neighbor searching
ns_type = grid ; search neighboring grid cells
nstlist = 10 ; 10 fs, largely irrelevant with Verlet scheme
rcoulomb = 1.0 ; short-range electrostatic cutoff (in nm)
rvdw = 1.0 ; short-range van der Waals cutoff (in nm)
DispCorr = EnerPres ; account for cut-off vdW scheme
; Electrostatics
coulombtype = PME ; Particle Mesh Ewald for long-range electrostatics
pme_order = 4 ; cubic interpolation
fourierspacing = 0.16 ; grid spacing for FFT
; Temperature coupling is on
tcoupl = Nose-Hoover ; Nose-Hoover thermostat
tc-grps = LIG HEX ; two coupling groups - more accurate
tau_t = 0.2 0.2 ; time constants, in ps
ref_t = 300 300 ; reference temperature, in K
; Pressure coupling is on
pcoupl = Parrinello-Rahman ; Pressure coupling on in NPT
pcoupltype = isotropic ; uniform scaling of box vectors
tau_p = 2.0 ; time constant, in ps
ref_p = 1.0 ; reference pressure, in bar
compressibility = 4.5e-5 ; isothermal compressibility, bar^-1
refcoord_scaling = com ; scaling of coordinate references
; Periodic boundary conditions are in all directions
pbc = xyz ; 3-D PBC
; Velocity generation
gen_vel = no ; Velocity generation is off
; Freezing
freezegrps = SC_1 SC_2 ; freezing groups
freezedim = Y Y Y Y Y ; directions of freezing
```

Appendix D

```
title = MD RUN OF TWO LIGAND MOLECULES IMMERSERD IN HEXANE
; Run parameters
integrator = md ; leap-frog integrator
nsteps = 4000000 ; 0.0005 * 4000000 = 2 ns
dt = 0.0005 ; 0.5 fs
; Output control
nstxout = 10 ; save coordinates every 0.005 ps
nstvout = 10 ; save velocities every 0.005 ps
nstenergy = 10 ; save energies every 0.005 ps
nstlog = 10 ; update log file every 0.005 ps
nstfout = 10 ; save every 0.005 ps
nstxout-compressed = 10 ; save every 0.005 ps
nstcalenergy = 10
; Bond parameters
continuation = yes ; Restarting after NVT
constraint_algorithm = lincs ; holonomic constraints
constraints = h-bonds ; bonds involving H are constrained
lincs_iter = 1 ; accuracy of LINCS
lincs_order = 4 ; also related to accuracy
; Nonbonded settings
cutoff-scheme = Verlet ; Buffered neighbor searching
ns_type = grid ; search neighboring grid cells
nstlist = 10 ; 20 fs, largely irrelevant with Verlet scheme
rcoulomb = 1.0 ; short-range electrostatic cutoff (in nm)
rvdw = 1.0 ; short-range van der Waals cutoff (in nm)
; Long-range dispersion correction
DispCorr = EnerPres ; account for cut-off vdW scheme
; Electrostatics
coulombtype = PME ; Particle Mesh Ewald for long-range electrostatics
pme_order = 4 ; cubic interpolation
fourierspacing = 0.12 ; grid spacing for FFT
; Temperature coupling is on
tcoupl = Nose-Hoover ; Nose-Hoover thermostat
tc-grps = LIG HEX ; two coupling groups - more accurate
tau_t = 0.2 0.2 ; time constants, in ps
ref_t = 300 300 ; reference temperature, in K
; Pressure coupling is on
pcoupl = Parrinello-Rahman ; Pressure coupling on in NPT
pcoupltype = isotropic ; uniform scaling of box vectors
tau_p = 2.0 ; time constant, in ps
ref_p = 1.0 ; reference pressure, in bar
compressibility = 4.5e-5 ; isothermal compressibility of water, bar^-1
refcoord_scaling = com
; Periodic boundary conditions are in all directions
pbc = xyz ; 3-D PBC
; Velocity generation
gen_vel = no ; Velocity generation is off
; Energy groups
energygrps = HEX Ligand_1 Ligand_2 ; energy groups
; Freezing
freeze-grps = SC_1 SC_2 ; freezing groups
freezedim = Y Y Y Y Y Y ; directions of freezing
; Electric field
E-z 1 1 0 ; electric field strength, in V/nm
```

Acknowledgements

First at all, I would like to thank the person that most followed me during my thesis work, Annalisa. Thank you for your patient and for the time that you spent in order to allow me to work properly on this project.

A big thank goes, of course, to mum and dad. Their efforts and support have been fundamentals during the studies, so as during my entire life. Thank you to my brother Matteo. Despite the physical distance, I know how much you love me and how much you care about me.

I would like to thank all the people with who I shared hard times studying for the exams. University studies are finally over, I can't count the times we were hoping to reach this moment.

A big thank goes to my friends of Pavia, since always part of my life despite the many travels and transfers. I appreciate you so much for the perseverance shown in these years to keep loving me and to be always present.

Last but not least, I would like to thank my girlfriend Elena. Always present since the day we met, you found the way to help me even for this thesis.

Thank you all.