JOINT INSTITUTE FOR NUCLEAR RESEARCH
Frank Laboratory of Neutron Physics

FINAL REPORT

INTEREST PROGRAMME

"Introductory course: MD-simulation research (from atomic fragments to molecular compound)"

Supervisor:
Prof. Kholmirzo Kholmurodov

Student:

Hayam Mohamed Elzeiny Abdelfattah Abdelgawad

M.Sc in Chemistry, Beni-Suef University

Participation period:
13 June– 22 July,
Wave 7
Dubna, 2022.

Contents:

1. Abstract

Molecular dynamics (MD) simulation acts as the bridge between the microscopic length and time scales and the macroscopic world of the laboratory. It is useful when actual life experiments are improbable and expensive. This technique is largely used in chemical physics, materials science, and biophysics. This report outlines the fundamental principles of molecular dynamics that have been acquired throughout our participation in this course. We studied the fundamental equations and strategies for using force fields to simulate different molecular systems in different filed such as material science and biology.

2. Introduction

Molecular dynamics (MD) is a computer simulation method for analyzing the physical movements of atoms and molecules. The atoms and molecules are allowed to interact for a fixed period of time, giving a view of the dynamic "evolution" of the system. In the most common version, the trajectories of atoms and molecules are determined by numerically solving Newton's equations of motion for a system of interacting particles, where forces between the particles and their potential energies are often calculated using interatomic potentials or molecular mechanics force fields. The method is applied mostly in chemical physics, materials science, and biophysics.

The obvious advantage of MD is that it gives a route to dynamical properties of the system: transport coefficients, time-dependent responses to perturbations, rheological properties and spectra. Computer simulations bring new insights into atomic change in the structure over a given period of time and can even be used to discover and design new molecules.

Computer simulations act as a bridge (Figure. 1) between microscopic length and time scales and the macroscopic world of the laboratory: we provide a guess at the interactions between molecules and obtain ‘exact’ predictions of bulk properties. At the same time, the hidden detail behind bulk measurements can be revealed. Simulations act as a bridge in another sense: between theory and experiment. We may test a theory by conducting a simulation using the same model. We may test the model by comparing with experimental results. We may also carry out simulations on the

computer that are difficult or impossible in the laboratory (for example, working at extremes of temperature or pressure)[1]. Ultimately, we may want to make direct comparisons with experimental measurements made on specific materials, in which case a good model of molecular interactions is essential. The aim of so-called ab initio molecular dynamics is to reduce the amount of fitting and guesswork in this process to a minimum. On the other hand, we may be interested in phenomena of a rather generic nature, or we may simply want to discriminate between good and bad theories. When it comes to aims of this kind, it is not necessary to have a perfectly realistic molecular model; one that contains the essential physics may be quite suitable[1, 2].

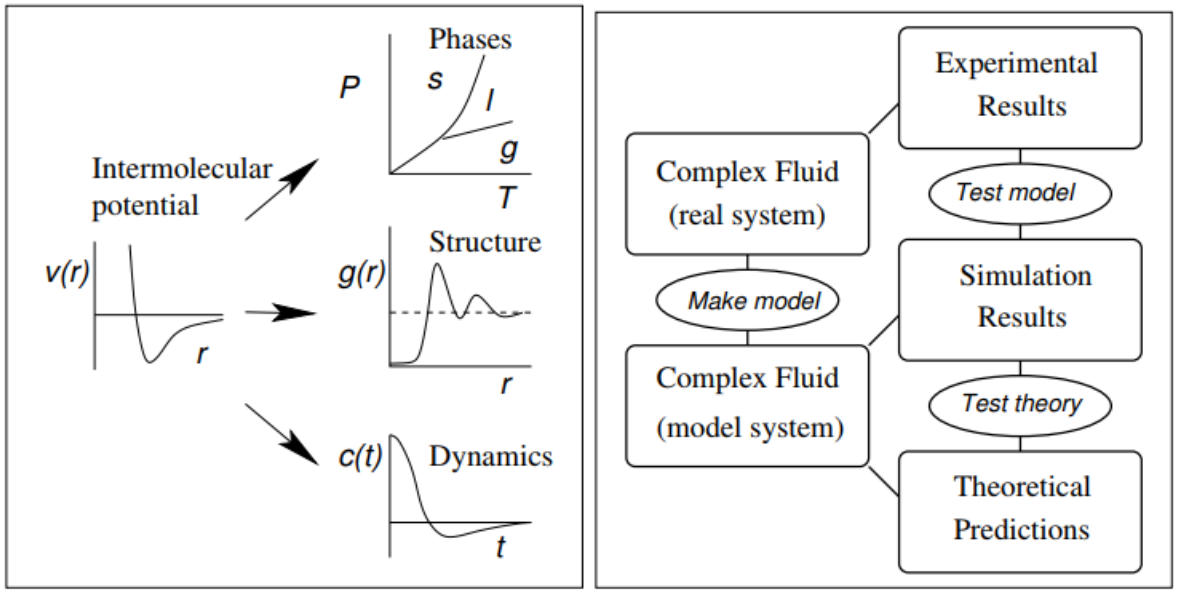

Figure 1. Simulations as a bridge between (a) microscopic and macroscopic; (b) theory and experiment.

3. Project goals

The goal of this course is to study the molecular modeling and computer design of chemical nanostructures, systems, and compounds, and this can be achieved by doing the following tasks:

1. Studying the basic equations, potentials, and simulation techniques.
2. Description the computer code for simulation of liquid model (Lenard-Jones potential).
3. Using a general-purpose code for the simulation of ionic, polymeric, and biochemical molecular systems.

4. Studying the theory of the basics of hybrid MD approach (classical quantum-chemistry potentials simulation methods).
5. MD test modeling.

4. Theoretical Background

4.1. The basic equations, potentials, and simulation techniques.

Molecular dynamics simulation is based on II Newton's law

$$F_i = \mathrm{m}_i a_i \ , \ F_i(r) = \ m_i \frac{d^2 r_i(t)}{dt^2} \qquad (1)$$

for each atom of a system of N atoms. where m_i is the atom's mass, $a_i = \frac{d^2 r_i(t)}{dt^2}$ is its acceleration, r is constant and F_i is the force acting upon it due to the interaction with other atoms.

If the initial positions and velocities of the particles are determined the solution of Newton's law gives the time evolution of a set of particles.

From Newton's I law, we know that a particle at rest will remain at rest and a particle in motion with certain velocity will continue to move with that velocity until an external force act on it.

The force can be written as the gradient of the potential energy

$$F_i = -\nabla i U \qquad (2)$$

where $\mathrm{U}(r_1, \dots, r_n)$ represents the potential energy of N interacting atoms as a function of their positions $r_i = (x_i, y_i, z_i)$.

Force Field can be understood as an empirical set of energy functions. It is typically the summation of bonded and non-bonded terms or covalent and non-covalent interactions among atoms and molecules.

$$V_{(r)} = E_{bonded} - E_{non-bonded} \qquad (3)$$

Any molecule is characterized by the presence of:

- Bonded atoms: include three types of interactions
  - Bond stretching (between two atoms)
  - Angle bending between three atoms)
  - Fixed torsion (between four atoms).

- Nonbonded atoms: include two types of interactions

- Van-der Waals interactions (Lennard-Jones interactions)
- Electrostatic interactions (Coulomb interactions) as shown in Figure 2.

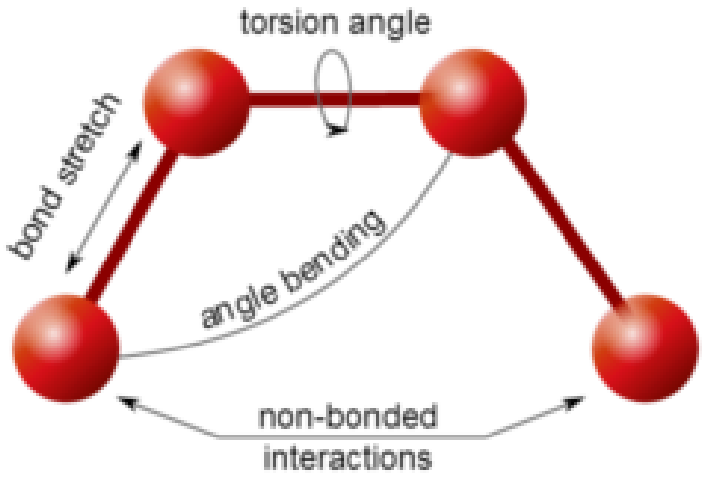

Figure 2: Different types of interactions

Then the total potential energy is described by the following equation:

$$U_{(r)} = U_b + U_\theta + U_\varphi + U_\omega + U_{LJ} + U_{el} + U_{HB} + \cdots \quad (4)$$

Where:

Valence Length potential $\boldsymbol{U_b}$: $U_b = \frac{1}{2}\sum_b k_b(r - b_0)^2$ (5)

Valence Angle Potential $\boldsymbol{U\theta}$: $U_\theta = \frac{1}{2}\sum_\theta k_\theta(\theta - \theta_0)^2$ (6)

Torsion Dihedral Potential $\boldsymbol{U\varphi}$: $U_\varphi = \frac{1}{2}\sum_\varphi k_\varphi \left[\cos(n_\varphi - \delta) + 1\right]$ (7)

Electrostatic potential (Coulombic potential) $\boldsymbol{U_{el}}$:

$$\boldsymbol{U_{el}} = \sum_{i,j} \frac{\boldsymbol{q_i\, q_j}}{\boldsymbol{r_{ij}}} \quad (8)$$

Van der Waals Interaction Potential $\boldsymbol{U_{LJ}}$: (Lennard- Jones potential)

$$\boldsymbol{U_{LJ}} = \sum_{i,j}\left(\frac{\boldsymbol{A}}{\boldsymbol{r_{ij}^{12}}} - \frac{\boldsymbol{B}}{\boldsymbol{r_{ij}^{6}}}\right) \quad (9)$$

Hydrogen Bonding Potential U_{HB}**:**

$$U_{HB} = \sum_{i,j}\left(\frac{A\prime}{r_{ij}^{12}} - \frac{B\prime}{r_{ij}^{10}}\right) \qquad (10)$$

4.2. The simulation of liquid model (Lenard-Jones potential)

The Lennard-Jones potential is an intermolecular pair potential. Among the intermolecular potentials, the Lennard-Jones potential is the potential that has been studied most extensively and most thoroughly. It is considered an archetype model for simple yet realistic intermolecular interactions. The commonly used expression for the Lennard-Jones potential is:

$$U_{LJ} = \sum_{i,j}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right) \qquad (11)$$

where ε= unit of energy and σ= unit of length, r_{ij}= distance between centres of two molecules $i\ and\ j$. both ε and σ parameters have different values for each element. The respective graph is shown below, in Figure 3.

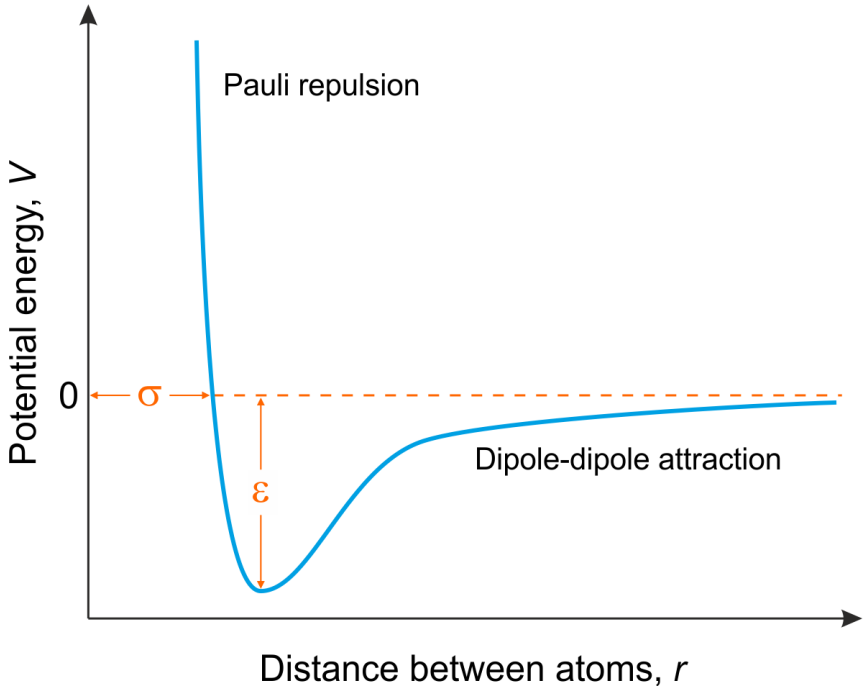

Figure 3: Lenard-Jones potential

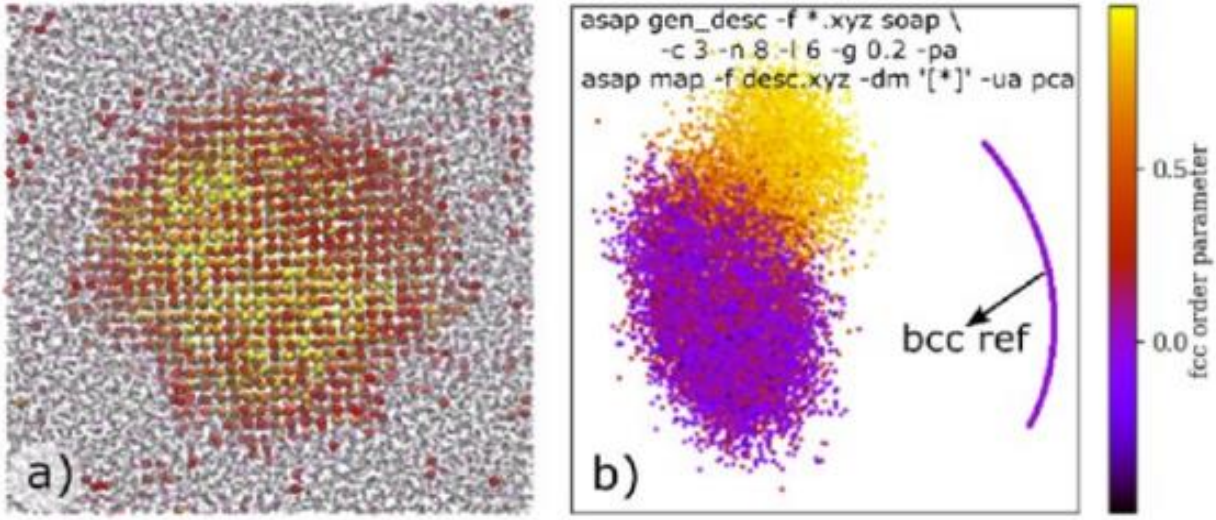

Figure 4. Snapshot of a Lennar Jones system of 23 328 atoms containing a solid nucleus surrounded by undercooled liquid with different configuration [2].

5. MD Simulation packages

There are many software programs for MD simulation programs which include both classical and quantum chemical methods and algorithms, are presented below:

(1) **AMBER** (www.ambermd.org) consists of a set of force fields for modeling macromolecular structures (proteins, nucleic acids and a number of other classes of molecules).

(2) **CHARMM** (www.charmm.org) (Chemistry at HARvard Macromolecular mechanics) software package for molecular modeling of a wide range of systems - from small molecules to biological macromolecules, using various energy functions and models.

(3) **GROMACS** (www.gromacs.org) for fast simulation of the dynamics of large molecular systems (from thousands to millions of particles). Works in Linux environment and is free.

(4) **LAMMPS** (lammps.sandia.gov) (Large scale Atomic Molecular Massively Parallel Simulator) uses classical molecular dynamics methods for modeling and calculating polymers, biomolecules, solids (metals, semiconductors, etc.).

(5) **NAMD** (www.ks.uiuc.edu/Research/namd/) An object-oriented program for calculations in the field of interactive molecular dynamics, for modeling large biomolecular systems that require significant resources.

(6) **DL_POLY** (www.cse.scitech.ac.uk/ccg/software/DL_POLY/) is used for modeling the molecular dynamics of complex systems with both sequential and parallel calculations. Adapted for graphics game processors, GPU (Graphical Processing Units), using the CUDA language. Freely available for research and educational purposes.

DL_POLY is the main simulation software used in this project, which developed at Daresbury Laboratory by I.T. Todorov, W. Smith, A.M. Elena and others[3].

In the DL_POLY code there are three input and three output files.

- **Input files**

1. CONFIG (Places and velocities of atoms in x, y, and z coordinates as well as boundary conditions.)
2. CONTROL: Contains data on (Thermodynamic parameters such as temperature, pressure,)
3. FIELD: (potential energy function of atoms)

- **Output files**

1. OUTPUT: energy, temperature, pressure in the final configuration of the system.
2. REVCON: intermediate system configuration and a restart configuration file – final.
3. HISTORY: atomic coordinates, velocities and forces containing system dynamics data, it is needed for visualization.

6. Application of molecular dynamics

Molecular dynamic simulation becomes an integral in-silico tool in many scientific fields such as structure biology, molecular biology, medicinal chemistry and drug design, material science, drug formulations.

6.1 Molecular dynamics in biology

These simulations can capture a wide variety of important biomolecular processes, including conformational change, ligand binding, and protein folding, revealing the positions of all the atoms

at femtosecond temporal resolution. Importantly, such simulations can also predict how biomolecules will respond at an atomic level to perturbations such as mutation, phosphorylation, protonation, or the addition or removal of a ligand.

MD simulations are often used in combination with a wide variety of experimental structural biology techniques, including x-ray crystallography, cryo-electron microscopy (cryo-EM), nuclear magnetic resonance (NMR), electron paramagnetic resonance (EPR), and Förster resonance energy transfer (FRET).

6.1.1. Molecular Dynamics Simulation of Proteins:

To get a dynamic picture for protein-ligand complexes and analyze their interactions with time, we need:

1. Set parameters (Temperature, pressure, time step)
2. Starting static structure: contains
   - Initial position (PDB coordinate file), from the Protein Data Bank
   - Initial velocities
     - V=0 at simulation start
3. Calculate energy
4. Calculate force
5. Calculate Acceleration
6. Determine final position and velocity of atoms
7. Obtain final trajectory of atoms

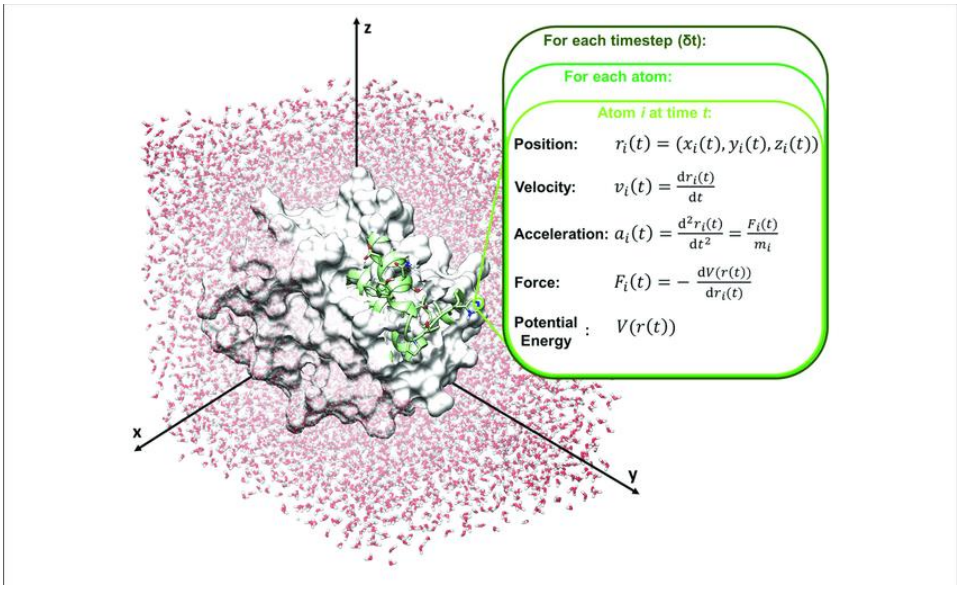

Figure 5. Schematic representation of a molecular dynamics cycle[4].

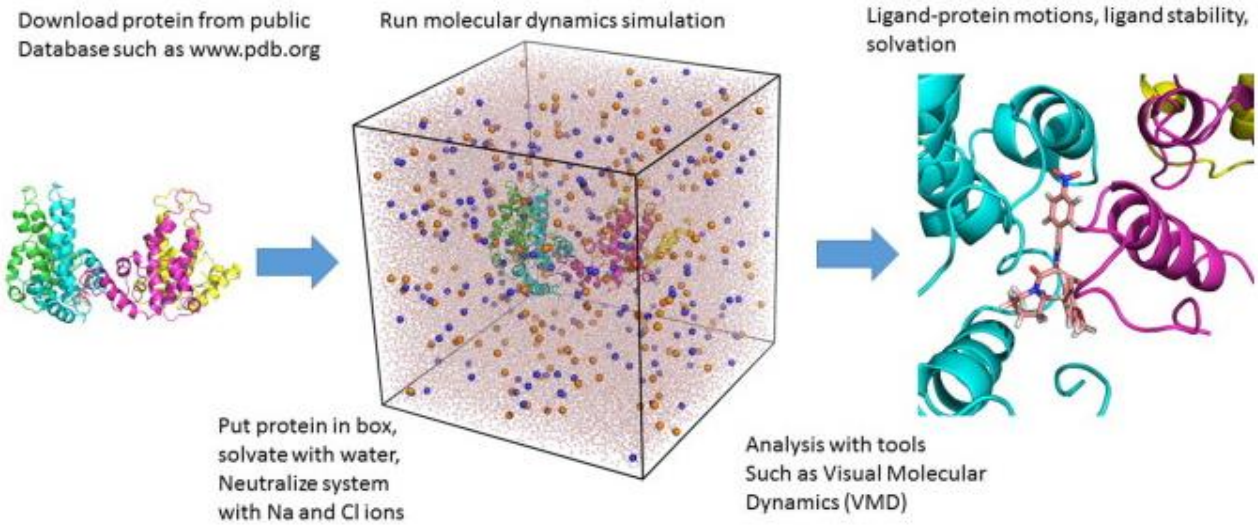

Figure 6. General steps for performing MD simulations. Starting from an experimental or modelled protein structure, a simulation system is prepared by setting up a simulation box, solvating the protein with water and neutralizing the total charge of the system with counterions. Then, the simulations are performed, visualized, and analyzed in terms of their protein motions and energetic features[5].

7. Conclusion:

The molecular modeling and computer design of chemical nanostructures, systems, and compounds have been studied. During the course, the following tasks were investigated: the basic equations, potentials, and simulation techniques, the simulation of a liquid model (Lenard-Jones potential), the use of selected general-purpose code for the simulation of ionic, polymeric, and biochemical molecular systems, the theory of the basics of hybrid MD approach (classical quantum-chemistry potentials simulation methods), methods of MD test modeling and its applications in protein, and drug design simulation.

8. Future Goals:

After I have finished my master's degree in the field of drug delivery, I became very interested in molecular dynamic simulation modelling. I am planning to get my PhD degree on relative computational areas such as, studying its role in targeting drug delivery research, controlling release of drugs, nanomedicine, drug design, protein-protein interactions, and so on… Honestly this course gave me the starting spark.

9. Acknowledgements

First of all, I wish to offer my heartfelt thanks to our supervisor Prof. Kholmirzo Kholmurodov for allowing me to be a part of such an interesting project, and I want to express my deep gratitude for his explanation, guidance, and support. All the knowledge and experience that I acquired throughout the duration of the project would greatly benefit my future studies. In addition, I would like to thank the Joint Institute for Nuclear Research (JINR) Centre and INTEREST program facilitators for providing such remote online programs under prominent supervisors for young researchers to extend their knowledge of nuclear research and its applications.

10. References:

1. Allen, M.P.J.C.s.m.f.s.p.t.p., *Introduction to molecular dynamics simulation.* 2004. **23**(1): p. 1-28.
2. Kholmurodov, K.J.M.S.S.i.M. and B. Sciences, *BY MOLECULAR DYNAMICS SIMULATIONS.* 2007: p. 129.
3. Smith, W., T. Forester, and I.J.S. Todorov, STFC Daresbury Laboratory, Daresbury, Warrington, Cheshire, WA4 4AD, United Kingdom, version, *The DL POLY Classic user manual.* 2012. **1**.
4. Salmaso, V. and S.J.F.i.p. Moro, *Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: An overview.* 2018. **9**: p. 923.
5. Morgnanesi, D., et al., *A computational chemistry perspective on the current status and future direction of hepatitis B antiviral drug discovery.* 2015. **123**: p. 204-215.