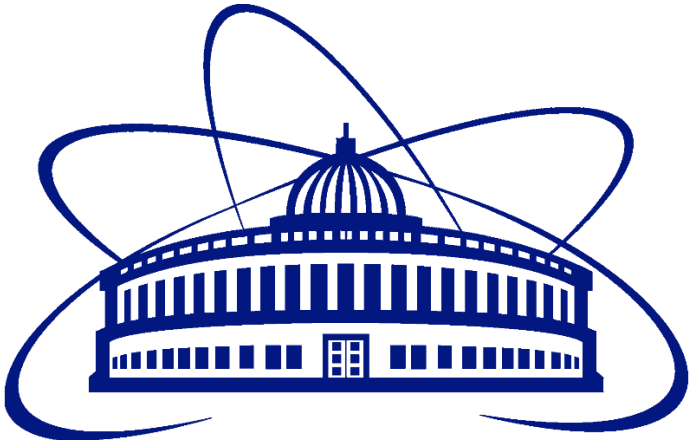

Joint Institute of Nuclear Research
Frank Laboratory of Nuclear Physics

Final Report on:

Molecular Dynamics Simulation Research (From Atomic Fragments to Molecular Compounds)

Alexis Achacoso
MSc Chemistry Student
University of Debrecen, Hungary

Supervisor: Professor Kholmirzo Kholmurodov
Wave 7 – JINR Interest Program

June 2022

Table of Contents

Introduction

A method for examining atom placement in space is molecular dynamics (MD). In this method, a dynamic model that forces the nuclear system into motion takes the role of a single-point model. The traditional Newtonian dynamic equations are numerically solved to simulate the motion. For a particular molecule, the collection of potential atom sites, for instance, provides the conformational ensemble profile. MD can also reveal the thermodynamic and dynamic characteristics of the molecules. The MD may be used to modify X-ray structures and simulate protein forms [1]. This method can be advantageous in synthesizing drugs by analyzing the possible interactions of drugs and other compounds.

The Joint Institute of Nuclear Research – Frank Laboratory of Nuclear Physics has offered a training course on Molecular Dynamics Simulation Research (From Atomic Fragments to Molecular Compounds) on its 7th Project Wave of the JINR-Interest Program. The course Molecular Dynamics Simulation Research is aimed at computer molecular dynamics of nanoscale phenomena, exploring new drugs and materials, MD developments, and recent applications. During this course, Prof. Kholmirzo Kholmurodov has explained the basics of molecular dynamics, its basic equations, and the simulation packages/coding used in such studies. This report summarizes the covered topics and their impact on my future research.

Molecular Dynamics Simulation

Molecular Dynamics

For large molecules like proteins, molecular dynamics is frequently the choice approach. It may be used to explore conformational space. In molecular dynamics, the energy surface is investigated by resolving the system's Newton's equations of motion. When scanning conformational space, it's usual practice to simulate a very high, physically improbable temperature since doing so increases the system's capacity to overcome energy barriers. Then, structures from the trajectory are picked at regular intervals for the following energy minimization [2]. Molecules move naturally, and molecular dynamics replicate this motion. A molecular dynamics technique supplies energy that enables atoms to move and even clash with one another. The molecule can traverse the energy barriers separating local minima on that molecule's conformational potential energy surface if enough thermal energy is supplied, making this a sort of conformational searching. The cycle must be repeated several times to obtain and analyze various conformations. With an explicit protein and solvent environment model, molecular dynamics simulations on a 10-100 nanosecond time scale function well. [3]

As shown in Figure 1, the various simulation methodologies include quantum physics, DFT methods, MD simulations (e.g., All-atom (AA), Coarse-grain (CG), Implicit solvent (IS), and Brownian dynamics (BD). Tens of thousands of atoms make up the usual structure of a biomolecule, and MD modeling enables us to establish the parameters that govern how those atoms interact.

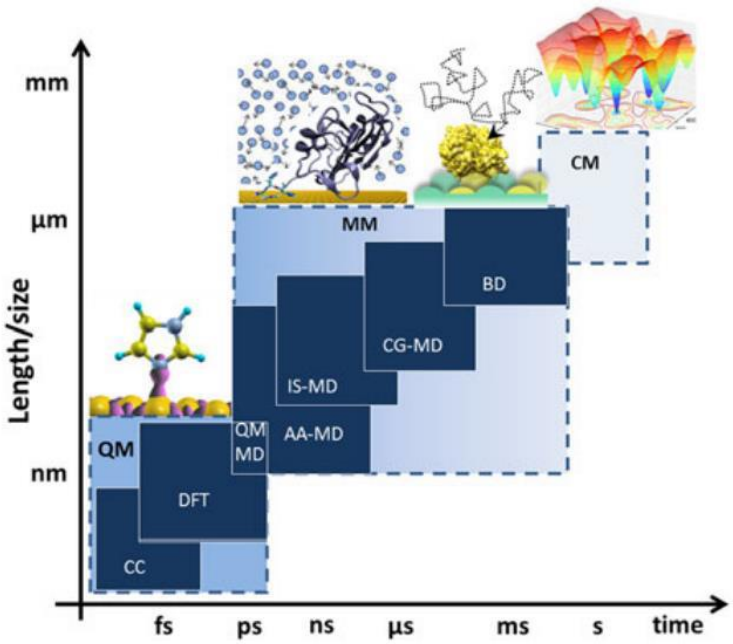

Fig 1. Typical time and length scales of several simulation techniques [4]

Basic Equations and Principles

Molecular dynamic is mainly based on the equation of Newton's second law:

$$m_i \frac{d^2 r_i(t)}{dt^2} = F_i(r)$$

The lowest free energy structure must be found to determine the molecule's geometrically ideal conformation. The equilibrium free energy of a molecular structure is calculated using molecular force fields. A mathematical function called a molecular force field explains how a molecule's potential energy depends on the positions of its atoms. Electrostatic forces and potentials (Coulomb interactions) also exist if the atoms have a charge, chemical bonds, and unbound Van der Waals interactions. The remaining energy is potential energy:

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

Using such a force field model, macromolecules are reduced to a collection of atoms held together by fundamental harmonic forces, Coulombic interactions, and Van der

Waals interactions. The force field has to be easy to evaluate quickly while still being accurate enough to represent actual structural characteristics in calculations. As shown in Figure 2, a molecule has a bond stretching between two atoms, three-atom angle bending, and four-atom fixed torsion.

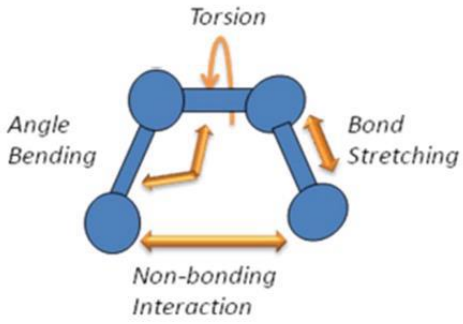

Fig 2. Different types of interactions in a molecule [4]

The molecular force field's first term describes the covalent bond extension. Bond stretching is sometimes described as a simple harmonic function that adjusts the length of covalent bonds. The second force field term describes the bond angle distortion. When there is a chemical bond between A and B and between B and C, the energy required to bend the angle created by at least three atoms—A-B-C—is known as bond angle distortion.

$$U_b = \frac{1}{2}\sum_b K_b(r - b_0)^2 \quad U_\theta = \frac{1}{2}\sum_\theta K_\theta(\theta - \theta_0)^2$$

The third force field term describes the distortion of dihedral angles from their desired values. The dihedral term must be considered in the force field if a molecule contains more than four atoms in a row, which is a given in macromolecules.

$$U_\varphi = \frac{1}{2}\sum_\varphi K_\varphi[\cos(n\varphi - \delta) + 1]$$

The fourth term describes the electrostatic forces that develop between atoms having ionic charges. Positive and negative ion interactions known as salt bridges are crucial for protein structure stability. The last term of the force field describes van der Waals forces. As electrons circle the atomic nucleus, they produce an electric dipole moment; a short-range attractive attraction forms between the non-bonded atoms due to this dipole's ability to polarize neighboring atoms. All the terms are summarized in the figure shown below.

$$U_{el} = \sum_{i,j} \frac{q_i q_j}{r_{ij}} \qquad U_{LJ} = \sum_{i,j} \left(\frac{A}{r_{ij}^{12}} - \frac{B}{r_{ij}^{6}} \right)$$

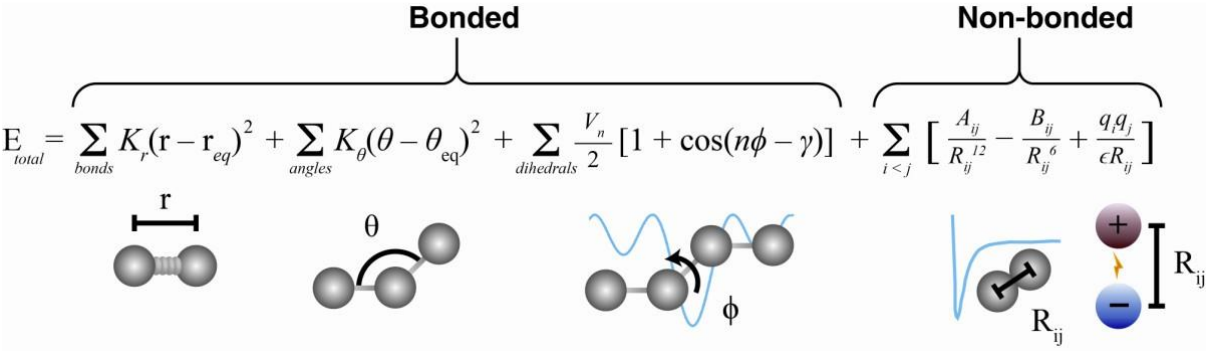

Fig. 3. Approximation of molecular movement

Lennard-Jones Potential

The Lennard-Jones potential is an intermolecular pair potential. The intermolecular potential that has undergone the most thorough and meticulous investigation is the Lennard-Jones potential. The Lennard-Jones potential describes the Van der Waals nonbonding contact between two atoms. Lennard-Jones potential is represented by the equation below:

$$V(r) = 4\varepsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right]$$

Where r is the distance between the centers of two atoms, is the energy unit, V(r) is the Lennard-Jones potential and is the average atomic density.

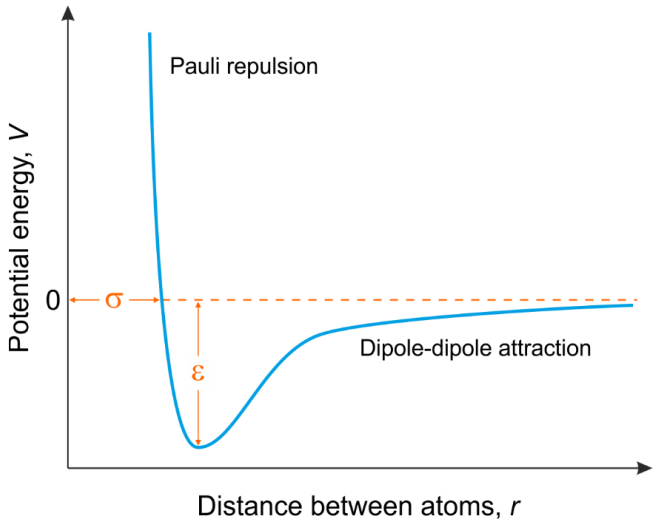

Fig 4. Graph of the Lennard Jones potential function

The Radial Distribution Function (RDF) in MD modeling is frequently used to track system equilibrium states. In this equation, N stands for the total number of atoms, for atomic density, r_{ij} for the distance between atom centers, and g for distances greater than or equal to one atomic diameter:

$$\rho g(r) = \frac{1}{N} < \sum_{i}^{N} \sum_{j \neq i}^{N} \delta[r - r_{ij}] >$$

The order parameter is a set of the following function-parameter γ is used to distinguish between equilibrium states:

$$\gamma_x = \frac{1}{N} \sum \cos\left(\frac{4\pi x_i}{a}\right)$$

$$\gamma_y = \frac{1}{N} \sum \cos\left(\frac{4\pi y_i}{a}\right)$$

$$\gamma_z = \frac{1}{N} \sum \cos\left(\frac{4\pi z_i}{a}\right)$$

$$\gamma = \frac{1}{3} [\gamma_x + \gamma_y + \gamma_z]$$

Lastly, the Boltzmann distribution is used for the monitoring of the equilibrium Hx

$$H_x(t) = \int f(v_x) ln f(v_x) dv_x$$

All of these simulations should also follow PBC or periodic boundary conditions to avoid problems with the movement of the atoms inside a given space.

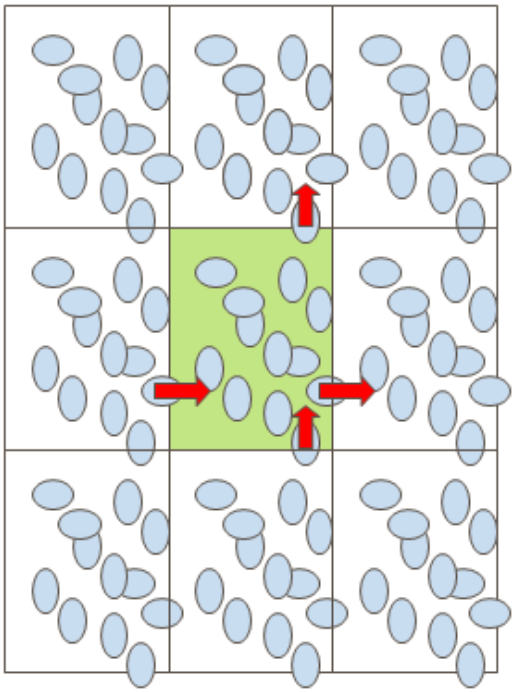

Fig. 5. Periodic Boundary Conditions

MD-Simulation Packages

MD can model protein flexibility using a variety of software tools. It was initially intended to refine NMR structures when the first MD package, Amber (Assisted Model Building with Energy Refinement), was created. The MD software suite and the force fields used to simulate biomolecules go by "Amber." Amber94, Amber99SB, and Amber03 are the most often employed versions of force fields. The following one, CHARMM (Chemistry at HARvard Macromolecular Mechanics) (Brooks et al., 2009), likewise refers to force fields and the software. Harvard University's Martin Karplus initially created it in the United States, the earliest

biomolecular MD package. CHARMM-GUI provides a web-based graphical interface for configuring input files for the simulation. [5] Lastly, IT Todorov, W. Smith, A.M. Elena, and others created DL_POLY, a general-purpose classical molecular dynamics (MD) simulation program, at Daresbury Laboratory. One version of the DL_POLY program, DL_POLY_4 (LGPL_v3 published), is currently being developed, and support is only offered to UK academic institutions. The DL_POLY_2 version, now known as DL_POLY_Classic, is made available as open-source software under the BSD license.

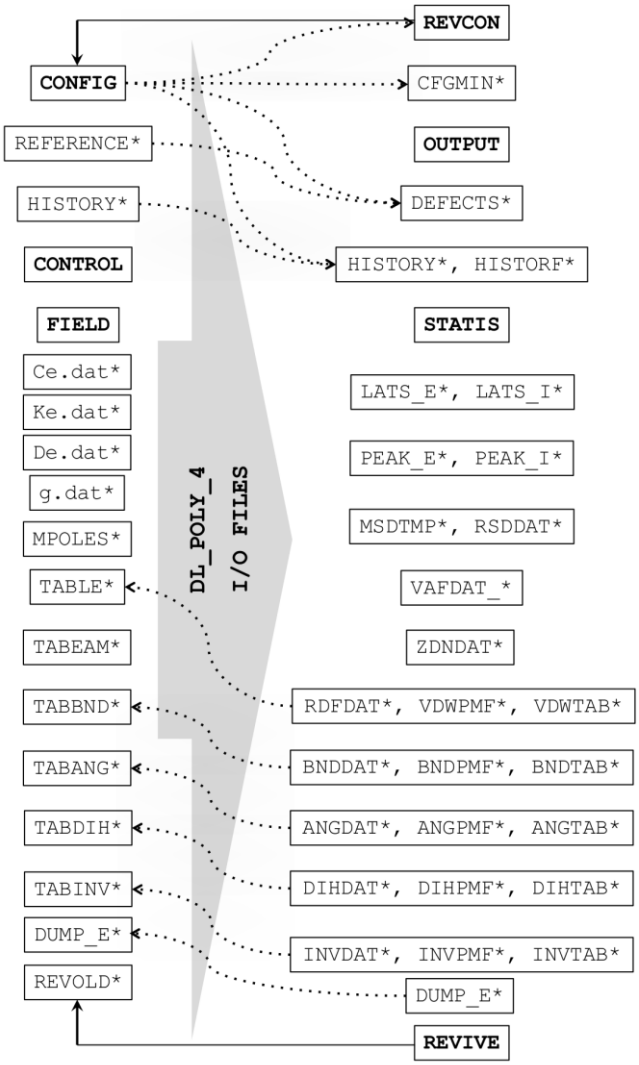

Fig. 6. DL POLY 4 input (left) and output (right) files.

Prospect Project:

Designing PET-CA using Molecular Dynamics for Imaging β-amyloid peptide in-vivo

Introduction

The enormous number of resources used for late-life neurologic illnesses like Alzheimer's disease (AD) will only get worse over the next few decades. Around 24 million people worldwide already have AD, but that figure might rise to more than 80 million by 2040. A significant component of the diagnosis of Alzheimer's disease (AD) is the presence of the β-amyloid-peptide in the brain, which is almost certainly essential to the illness's progression [6]. To effectively diagnose this disease, there is a need to develop an in-vivo imaging technique, which can be achieved by 52Mn Positron Emission Tomography (PET) or Magnetic Resonance Imaging (MRI).

Methodology

The most widely used amyloid imaging agent is Pittsburgh Compound B (PiB, 2-[4'-(Methylamino)phenyl]-6-hydroxybenzothiazole; a derivative of the amyloid dye, thioflavin T. On PET scans, regions of the brain prone to A deposition in AD can be shown to have increased 11C-PiB retention [6]. PiB (Figure 7) can be synthesized to be bound to a pyclen-type macrocycle containing a Mn metal ion center (Figure 8) to produce a contrast agent (CA) for the AD susceptible brain regions. The MD simulation will be used to design the most efficient pyclen-type ligand that will give the fastest relaxation time for better imaging. The MD simulation will also be used to study the equilibrium kinetics of the contrast and the amyloid to remove the said CA successfully after imaging.

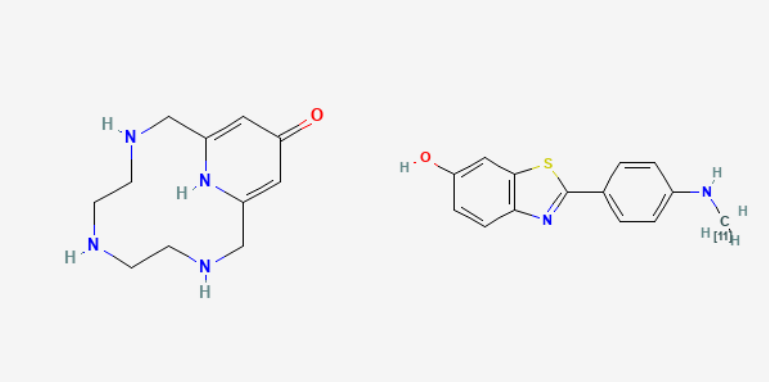

Fig. 7 and 8. Pyclen-type ligand (left) and PiB (right) (PubChem)

Acknowledgment

I want to thank Prof. Kholmirzo Kholmurodov for this invaluable effort in his lectures for the JINR-Interest program. I would also like to thank the JINR for this opportunity. This training will be beneficial for me in pursuing my career in drug design.

References

[1] A. Leach, "Ligand-Based Approaches: Core Molecular Modeling," in *Comprehensive Medicinal Chemistry II*, 2007, pp. 87-117.

[2] J. Polanski, "Chemoinformatics," in *Comprehensive Chemometrics*, 2020, pp. 459-506.

[3] R. I. M. Melissa R. Pitman, "Methods for Protein Homology Modelling," in *Applied Mycology and Biotechnology*, 2006, pp. 37-59.

[4] K. Kholmurodov, Computer Design for New Drugs and Materials: Molecular Dynamics of Nanoscale Phenomena, New York: Nova Science Publishers, 2017.

[5] X. Han, "Study of the Variability of the Native Protein Structure," in *Encyclopedia of Bioinformatics and Computational Biology*, 2019, pp. 606-619.

[6] B. TL, "Postmortem Pittsburgh Compound B (PiB) binding increases with Alzheimer's disease progression," *J Alzheimers Dis,* vol. 2, no. 32, pp. 127-138, 2012.