CHAPTER 2

Computational methodology

Chapter 2. Computational methodology

2. Computational methodology

Development of the drug has always been time consuming and money spending. However, there exist rising many questions in achieving targeted drugs as well as therapy within required time and budget. With progress in computational research, computational simulations are becoming more convincing than experimental techniques [1]. In case of simulations, the behaviour of the system is studied using computer programming with a model of the system and then the desired results are reproduced with experiments. New scientific insight is obtained by observing such computer experiments; often for controlled conditions those are not accessible in the laboratory. The *in silico* optimization of novel drug delivery systems can significantly increase accuracy and easiness of the application [2].

In this chapter, we describe the computational methods that have been applied in the current thesis. We carried out computer simulations with the hope of understanding properties of assemblies of molecules in terms of their structure and the microscopic interactions between them. It was pointed out in the introduction that the present work focused on the anticancer drug design and to pursue the study, we performed molecular dynamics simulation (MDs) [3] and density functional theory (DFT) based first principle calculations [4] to gives a route to dynamical properties of the system such as: transport coefficients, time-dependent responses to perturbations and rheological properties such as electronic, structural and vibrational properties.

2.1 Theoretical approaches

The purpose of this section is to describe the theory and explain the methodology of the calculations. Ab-initio techniques are applied to solve the many body problem so that the detailed study of the ground state properties of the system could be under taken. The DFT simulation package GAUSSIAN enables to computes an approximate solution to the many-body Schrödinger equation, either within density functional theory (DFT), solving the Kohn-Sham equations, or within the Hartree-Fock (HF) approximation. In GAUSSIAN, central quantities like the one-electron orbitals, the electronic charge density, and the local potential are expressed in localized (orbital like) wave basis sets.

2.1.1 The Many body Schrödinger Equation

This section presents an introduction to density –functional theory (DFT) calculating ground-state properties of electronic systems [5-8] and briefly presents the basics of density functional theory. The many-body problem, we consider is an N-electron system (atom, molecule, or solid) in the Born-Oppenheimer and non-relativistic approximations. The electronic Hamiltonian of the system in atomic unit system is represented as;

$$H(r_1,r_2,...,r_N) = -\frac{1}{2}\sum_{i=1}^{N}\nabla_{ri}^{2} + \frac{1}{2}\sum_{i}^{N}\sum_{i\neq j}^{N}\frac{1}{r_{i-r_j}} + \sum_{i}^{N}\upsilon_{ne^{r_i}} \qquad (2.1)$$

Where, $\upsilon_{(ne)} = -\sum_{\alpha} Z_{\alpha}/|r_i - R_{\alpha}|$ is the nuclei-electron interaction (R_{α} and Z_{α} are the position and charges of the nuclei). The first and second term represents kinetic energy of electron and coulomb interaction respectively. The stationary electronic states are determined by the time-independent Schrödinger equation,

$$H(r_1, r_2, ..., r_N)\Psi(x_1, x_2, ..., x_N) = E\Psi(x_1, x_2, ..., x_N) \qquad (2.2)$$

Where $\Psi(x_1, x_2, ..., x_N)$ is a wave function written with space-spin coordinates $xi = (r_i, \sigma_i)$ (with $r_i \in r_3$ and $\sigma_i =\uparrow$ or $\downarrow$) which is antisymmetric with respect to the exchange of two coordinates, and E is the associated energy. Using Dirac notations, the above equation can be rewritten in a representation independent formalism

$$\hat{H}\,|\Psi_i\rangle = E|\Psi_i\rangle \qquad (2.3)$$

where the Hamiltonian is formally written as;

$$\hat{H} = \hat{T} + \hat{W}_{ee} + \hat{V}_{ne} \qquad (2.4)$$

with the kinetic-energy operator $\hat{T}$, the electron-electron interaction operator $\hat{W}_{ee}$, and the nuclei-electron interaction operator $\hat{V}_{ee}$. The quantity of primary interest is the ground-state energy E_0. The variational theorem establishes that the E_0 can be obtained by the following minimization

$$E_0 = min_{\Psi}\langle\Psi|\hat{H}|\Psi\rangle \qquad (2.5)$$

where the search is over all N-electron antisymmetric wave functions Ψ, normalized to unity $\langle\Psi|\hat{H}|\Psi\rangle = 1$. DFT is based on a reformulation of the variational theorem in terms of the one electron density defined as:

$$n(r) = N\int \ldots\ldots \int |\Psi_1, \Psi_2, \ldots.\Psi_N|^2 \, d\sigma dx_2 \ldots\ldots.. dx_N \qquad (2.6)$$

which is normalized to the electron number, $\int n\,(r)dr = N$ [5-8]. Therefore, the only way to make the things manageable is to resort to some approximations.

2.1.2 The Born-Oppenheimer Approximation

The very first approximation is Born Oppenheimer approximation. Even the simplest molecule, H_2^+, consist of three particles, hence its Schrödinger equation is quite difficult to solve analytically. To overcome this barrier, we apply the Born-Oppenheimer approximation, by taking the advantage of relatively low velocity of nuclei as compare to electron due to nucleus' relatively heavy mass. Due to this difference, electrons can respond almost instantaneously to displacement of the nucleus. Thus instead of solving rigorous Schrodinger equation for all the particles at once, we consider nuclei to be fixed at its position and solve the Schrodinger equation for electron and the static electric potential.

The Born-Oppenheimer approximation is very reliable for ground electronic state, but it is less reliable for excited state. Applying this approximation to the many body Schrodinger equation, allows to solve the electronic structure problem in an efficient way,

$$\hat{H}|\Psi >= E|\Psi_i > \qquad (2.8)$$

with $$\hat{H} = \sum_{i=1}^{N}\left[-\frac{1}{2}\nabla_i^2 + V_{ext}(r_i)\right] + \sum_{i\neq j}^{N}\frac{1}{|\vec{r}_i-\vec{r}_j|} \qquad (2.9)$$

where, ψ_n is the many body wave functions for the *N*-electron eigen states. This Fermionic wave function has antisymmetric nature and changes sign with interchange of coordinates of any two particles. *E* is the total energy. Due to approximation of fixed position of nuclei, the second term of Eqn:

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \frac{1}{2}\sum_{A=1}^{M}\frac{1}{M_A}\nabla^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{|\vec{r}_i-\vec{r}_j|}$$

$$+ \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{|\vec{r}_i-\vec{r}_j|} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{|\vec{R}_A-\vec{R}_B|} \qquad (2.10)$$

becomes zero and the last term appears as a constant in total energy expression.

2.2 Fundamentals of DFT

2.2.1 Introduction

DFT can be stated as an exact reformulation of the electronic structure problem in terms of electronic density n(r) rather than the many-body wavefunction Ψ ($r_1\sigma_1$, $r_2\sigma_2$, ..., $r_{(n)}$ $\sigma_{(n)}$), The major strength of the formulation is to provide an excellent compromise between accuracy and computational cost. Although in principle, the theory requires a universal functional which is in general unknown and must be approximated. In the following text, the basic of DFT which comprises of Hohenberg-Kohn theorems and the Kohn-Sham equations are discussed. These concepts would form the building blocks of the computational implementation of density functional theory.

Solving a quantum mechanically interacting many body system to deduce the electronic structure of a solid is a very hefty task and is unsolvable even with the modern computing technology, however using Born-Oppenheimer approximation and considering only valence electrons and ions will simplify the problem a bit further, but it still remains unsolvable. Earlier the Thomas-Fermi method which is the precursor of DFT was used to describe electron systems.

$$\hat{\mathrm{H}}\Psi(\mathrm{r}) = (\hat{\mathrm{T}} + \hat{\mathrm{V}}_{\mathrm{ne}} + \hat{\mathrm{V}}_{\mathrm{ee}})\, \Psi(\mathrm{r}) = \mathrm{E}\Psi(\mathrm{r}) \tag{2.11}$$

where, $\hat{\mathrm{T}} = \sum_{\mathrm{i}=1}^{\mathrm{Ne}} \hat{\mathrm{T}}_{\mathrm{i}} = \sum_{\mathrm{i}=1}^{\mathrm{Ne}} \frac{1}{2\mathrm{m}} \mathrm{p}_{\mathrm{i}}^{2}$ is the kinetic energy operator, $\hat{\mathrm{V}}_{\mathrm{ne}} = \sum_{\mathrm{i}=1}^{\mathrm{Ne}} \mathrm{v}(\mathrm{r}_{\mathrm{i}})$ is the external potential, which describes the interaction with a crystal lattice and $\hat{\mathrm{V}}_{\mathrm{ee}} = \sum_{\mathrm{i}<j} \mathrm{w}(\mathrm{r}_{\mathrm{i}}, \mathrm{r}_{\mathrm{j}})$ is the electron-electron interaction potential. Within the Thomas-Fermi model, the kinetic energy density T[n(r)] is approximated by the kinetic energy of a uniform, non-interacting electron gas taken at a given density n(r). With the Hartree expression for the electron-electron energy plus the contribution of the external potential one arrives at the Thomas-Fermi energy functional E_{TF}[n(r)]. Minimization of E_{TF}[n(r)] produces the ground state density distribution and the ground state energy. Unfortunately, this method fails to describe the shell structure of atoms and it cannot explain the existence of molecules. Modern DFT generalizes the THOMAS-FERMI approach in such a way that it behaves as an exact theory for an inhomogeneous many-body problem. The DFT based calculation uses a set of equation which is solved iteratively, and these set of equations are known as Kohn-Sham equations.

2.2.2 Hohenberg-Kohn Theorems

Consider an electronic system with an arbitrary external local potential v(r) in place of v_{ne}(r).The corresponding ground-state wave function Ψ can be obtained by solving the Schr$\ddot{o}$dinger equation, from which the associated ground-state density n(r) can be deduced. Therefore, one has a mapping from the potential v(r) to the ground-state density n(r),

$$\mathrm{v(r)} \text{ ------> } \mathrm{n(r)} \tag{2.12}$$

In 1964, Hohenberg and Kohn [4] showed this mapping can be inverted, i.e. the ground-state density n(r) determines the potential v(r) up to an arbitrary additive constant,

$$\mathrm{n(r)} \xrightarrow[Hohenberg-Kohn\ v(r)]{------\rightarrow} \mathrm{v(r) + const.} \tag{2.13}$$

The following two theorems put forward by Hohenberg and Kohn (HK) in 1964 forms the basis of DFT:[9]

Theorem 1

The ground state density n(r) of a many-body quantum system in some external potential v_{ext}(r) determines this potential uniquely. This formally enables using the three-dimensional (real) function n(r) instead of the 3N-dimensional (complex) many-body wave function Ψ {r1, r2,…. r(n)} as the basic variable. More precisely, all physical observables can in principle be expressed as functionals of the electronic charge-density. The second major ingredient is the existence of a variational principle for the energy functional.

Theorem 2

For any trial density n(r) it holds E0 ≤ E [n(r)], where E0 is the ground-state energy for the system. In other words, the minimum value of the total-energy functional E [n(r)] is the ground state energy of the system, and the density which yields its minimum value is exactly the single particle ground-state density. The explicit form of this energy functional reads,

$$\mathrm{E[n]} = \int \mathrm{v(r)n(r)dr} + \mathrm{F[n]} \qquad (2.14)$$

$$\mathrm{F[n]} = \frac{1}{2}\iint \frac{\mathrm{n(r)n(\acute{r})}}{|\mathrm{r}-\acute{\mathrm{r}}|}\mathrm{drd\acute{r}} + \mathrm{G[n]} \qquad (2.15)$$

where the universal functional is often separated into the classical Coulomb term due to the charged electrons and the new universal functional G[n].

2.2.3 Kohn-Sham Equation

The HK theorems (1964) were helpful in providing the theoretical foundation for DFT (fig. 2.1), the crucial role in simplifying the problem was played by Kohn and Sham (1965) [10]. *For any interacting ground state density n_0(r), a corresponding noninteracting system in some effective potential V_s is constructed such that the non-interacting Hamiltonian H_s = T + V_s produce the same ground state density n_0(r).*

In other words within the framework of KS approach, the intractable many-body problem of the interacting electrons in a static potential is reduced to a tractable problem of non-interacting electrons moving in an effective potential. The K.E expression from TF model is non-complying for atoms and molecule in practice. The wave function could be used to calculate K.E easily, provided that the former is known. To overcome these complexions Kohn and Sham (1965) [10] approached the problem in rather a ingenious way they coupled the wave function and density approach. They reformulated the total energy functional as follows:

$$E[n(r)] = T_0[n(r)] + \int [V_{ext}(r) + U_{Coul}(r)]n(r)d^3r + E_{xc}[n(r)] \qquad (2.16)$$

where $T_0[n(\vec{r})]$ is the kinetic energy of electrons in a system which has the same electron density as that of the real system but with no electron-electron interactions. This is the system of non-interacting electrons, but electrons are still interacting with nuclei. The Coulomb interaction is:

$$U_{Coul}(r) = \int \frac{n(r\prime)d^3r}{(r'-r)} \tag{2.17}$$

It includes the electron self-interaction explicitly, since the corresponding energy is given by

$$\int U_{Coul}[n(r)]d^3r = \frac{1}{2}\int\int \frac{n(r)n(r\prime)d^3rd^3r\prime}{(r'-r)} \tag{2.18}$$

Which represents interaction of $[n(\vec{r})]$, with itself $V_{ext}(\vec{r})$ is the external potential, i.e. the potential coming from nuclei. The last functional, $E_{xc}[n(\vec{r})]$ is called exchange-correlation (XC) energy. This includes all the remaining energy contributions i.e. electron exchange, electron correlation, a portion of kinetic energy to obtain the true kinetic energy and correction for the self-interaction introduced by the classical potential, which were not accounted by the previous terms. Applying the variational principle

$$\mu = \frac{\delta E[n(r)]}{\delta n(r)} = \frac{\delta T_0[n(r)]}{\delta n(r)} + V_{ext}(r) + U_{Coul}(r) + \frac{\delta E_{xc}[n(r)]}{\delta n(r)} \tag{2.19}$$

$$\text{or} \quad \mu = \frac{\delta E[n(r)]}{\delta n(r)} = \frac{\delta T_0[n(r)]}{\delta n(r)} + V_{eff}(r) \tag{2.20}$$

where **μ** is electronic chemical potential and

$$V_{eff}(r) = V_{eff}(r) + U_{Coul}(r) + V_{xc}(r) \tag{2.21}$$

$$\text{with} \qquad V_{xc}(r) = \frac{\delta E_{xc}[n(r)]}{\delta n(r)} \tag{2.22}$$

Schrödinger equation for non-interacting particles:

$$[-\frac{1}{2}\nabla^2 + V_{eff}(r)]\,\emptyset_i(r)^{KS} \tag{2.23}$$

where the operator on left hand side, depends only on $\vec{\mathbf{r}}$, and not upon the index of the electron which is presented by subscript "i". It is same for all the electrons. KS orbitals $\phi_i(\vec{r})^{KS}$, can be used to compute the total electron density:

$$n(\vec{r}) = \sum_{i=1}^{N} \left|\Phi_i^{KS}(\vec{r})\right|^2 \quad (2.24)$$

This can be used to calculate the improved potential $V_{eff}(\vec{r})$ self consistently. Electron density can also be used to calculate the total energy, in which kinetic energy, $T_0[n(\vec{r})]$, is calculated from corresponding orbitals, rather than density itself:

$$T_0[n(r)] = \frac{1}{2}\sum_{i=1}^{N}\langle \emptyset_i^{KS}|\nabla^2|\emptyset_i^{KS}\rangle \quad (2.25)$$

and rest part of total energy as

$$V_{eff}[n(r)] = \int V_{eff}(r)n(r)d^3r \quad (2.26)$$

Thus, in KS approach, the electrons move in effective potential which includes the external potential and the effects of the Coulomb interactions between electrons e.g. exchange and correlation effect. The crucial quantity in KS approach is the XC energy $E_{xc}[n(\vec{r})]$ containing terms coming from the Pauli's exclusion principle, from correlation due to the repulsive Columbic electron-electron interaction and from the contributions due to the kinetic energy of the interacting electrons. Even though $E_{xc}[n(\vec{r})]$ must be very complex, great progress has been made with remarkably simple approximations like Local Density Approximation (LDA), Local Spin Density Approximation (LSDA), Generalized Gradient Approximation (GGA) etc.

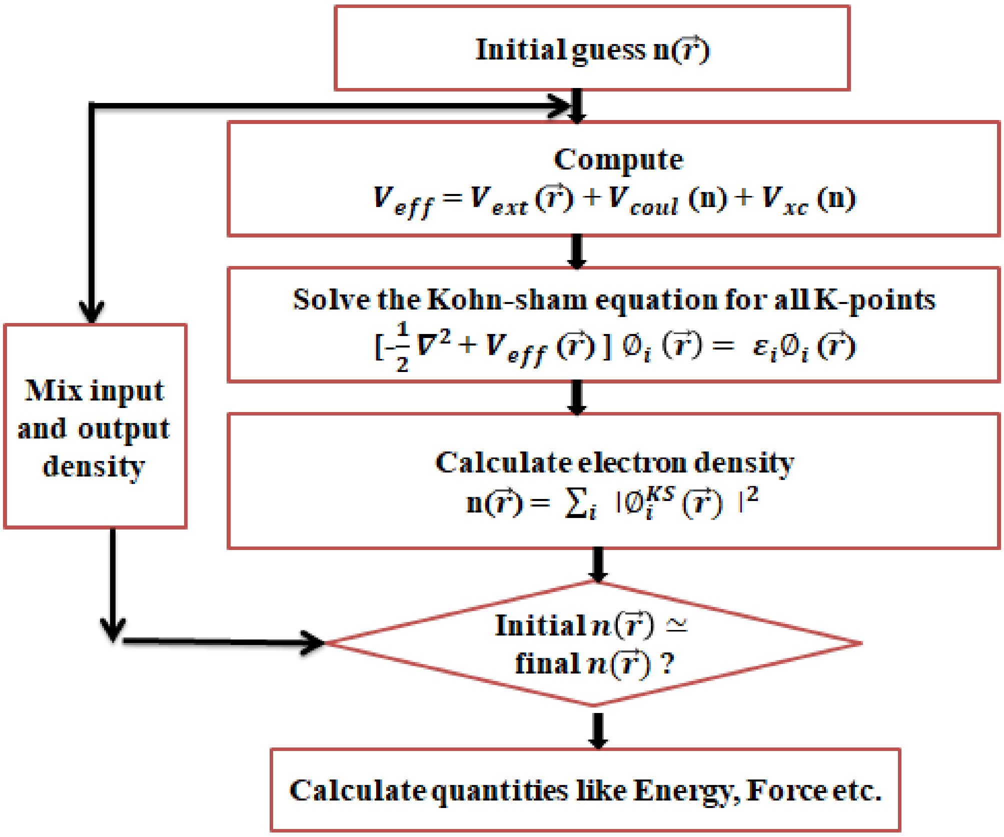

Figure 2.1. Flow chart for self-consistent density functional calculations.

2.3 Properties from the DFT calculations for present study

In the present section, we briefly describe the methodology adopted for properties investigated to understand the interaction between different molecules such as SAs and ligands. The calculated properties from the DFT are Mulliken charge [11], Molecular electrostatic potential (MEP) [12], Natural Bond Orbital (NBO), Frontier molecular orbitals (FMO), Fukui indices and vibrational spectra.

2.3.1 Mulliken charge

Mulliken charge, also referred as partial charge [13] is used to characterize the electronic charge distribution in a molecule and the bonding, anti-bonding, or nonbonding nature of the molecular orbitals for particular pairs of atoms. To develop the idea of these populations matrix, consider a real, normalized molecular orbital composed from two normalized atomic orbitals [14]. Mulliken charges can be obtained by a simple multiplication

$$P_{ij} = D_{ij}S_{ij} \tag{2.27}$$

where, P is the Mulliken population matrix, D is the density matrix, and S is the overlap matrix.

2.3.2 Molecular electrostatic potential (MEP)

MEPs are used to investigate the biochemical reactivity between target SA and ligand monovalent, bi- valent and mimic saccharide molecules [12].

$$V(r) = \sum_{\alpha} Z_{\alpha}/|R_{\alpha} - r| - \int d^3r_1 \, \rho(r_1)/|r_1 - r| \tag{2.28}$$

where, Z_α and R_α denote the charge and position of nucleus α, respectively, ρ shows the electron density. Equation renders the electro- static interaction between the unperturbed charge distribution of the molecule and a positive unit charge located at point r.

2.3.3 Natural Bond Orbital (NBO)

The NBOs [15] are one of a sequence of natural localized orbital sets that include natural atomic (NAO), hybrid (NHO), and (semi-)localized molecular orbital (NLMO) sets, intermediate between basis AOs and canonical molecular orbitals (MOs).

AOs → NAOs → NHOs → NBOs → NLMOs → MOs

In accordance with the simple bond orbital each bonding NBO σ_{AB} can be written in terms of two directed valence hybrids (NHOs) h_A, h_B on atoms A and B, with corresponding polarization coefficients C_A and C_B ,

$$\sigma_{AB} = C_A h_A + C_B h_B \tag{2.29}$$

that vary smoothly from covalent (cA = cB) to ionic (cA >> cB) limit. Each valence bonding NBO must in turn be paired with a corresponding valence anti-bonding NBO,

$$\sigma_{AB} * = C_A h_A + C_B h_B \tag{2.30}$$

As a result of each σi → σj* perturbation, the starting NBO acquires a weak anti-bonding "tail" in the final (doubly occupied) NLMO Ωi. More generally, each semi localized NLMO Ωi can be expressed as a linear combination of the parent Lewis-type NBO σi (with coefficient cii ≅ 1) and residual weak contributions (cji ≅ 0) from non-Lewis (NL) NBOs σj*

$$\Omega_i = C_{ii}\sigma_i + \sum_j^{NL} C_{ji}\sigma_j^* \tag{2.31}$$

2.3.4 Frontier molecular orbitals (FMO)

The interaction of two orbitals gives rise to a new set of orbitals. If overlap is neglected, the difference in energy before and after the interaction is given by the second-order perturbation expression.

$$E = 2\sum_{I=1}^{n/2} \lambda_i \qquad \text{if n is even} \tag{2.32}$$

$$E = 2\sum_{I=1}^{(n-1)/2} \lambda_i + \lambda_{(n+1)/2} \qquad \text{if n is odd} \tag{2.33}$$

If n is even, then the (n/2) eigenvector represents the highest (doubly) occupied molecular orbital (the HOMO), whose energy is λn/2 . The next eigenvector pertains to the lowest unoccupied molecular orbital (the LUMO), whose energy is $\frac{\lambda n}{2} + 1$. Then the HOMO– LUMO separation is given as:

$$\Delta HL = \frac{\lambda n}{2} - \frac{\lambda n}{2} + 1 \tag{2.34}$$

2.3.5 Fukui indices

The Fukui function is the derivative of the electron density with respect to the number of electrons of the system. It has been formally defined as,

$$f(r) = \left[\frac{\delta_\mu}{\delta_{v(r)}}\right]_{\mathrm{N}} \tag{2.35}$$

where, $V(r)$ is the external potential and the functional derivative must be taken at a constant number of electrons N [16].

$f^{+}(r) = \rho N + 1(r) - \rho N(r)$ (Nucleophilic attack) (2.36)

$f^{-}(r) = \rho N(r) - \rho N - 1(r)$ (Electrophilic attack) (2.37)

$f^{0}(r) = \rho N(r)$ (Radical attack) (2.38)

Where, ρN+1, ρN and ρN-1 represent the electron density of anionic, neutral and cationic residues, respectively.

2.3.6 Vibrational spectra

In the wave mechanical solution of the vibrational motion, the potential and kinetic energy of the molecule should be substituted in the Schrödinger equation. Assuming the harmonic nature of the vibrations the following expression is obtained for the vibrational energy.

$$E_0 = \sum_{i=1}^{3N-6} h\mathrm{v}_i \left(v_i + \frac{d_i}{2}\right) \tag{2.39}$$

where, $\mathrm{v}_i$is the frequency of the normal vibrations, $v_i$is the appropriate vibrational quantum number, $d_i$is the degree of degeneration.

2.4 Molecular dynamics Simulation (MDS)

As discussed previously, drug-discovery is a time-consuming process, and to overcome this; various computational methods have been employed like structure based drug discovery (SBDD) and ligand-based drug discovery (LBDD) which rely on the structure of the target and ligands respectively. Although there are several methods available, The two main families of simulation technique are molecular dynamics (MD) and Monte Carlo (MC); additionally, there is a whole range of hybrid techniques which combine features from both. Molecular dynamics (MD) simulations have turned out to be the essential technique in the field of the designing of new bioactive compounds [17-22]. MD simulations explain the crucial role of protein conformational flexibility in binding to the ligand or vice-versa. Moreover, MD simulations have been successfully employed to observe the interactions [23-32] and equilibration structures of the bio-molecules such as membranes, proteins, and nucleic acids [33-38]. The stability of the docked complexes was extrapolated by employing the Newtonian equation of motion (equation 2.40).

2.4.1 Equation of motion

$$F = m_i a_i \approx m_i \frac{\delta^2 r_i}{\delta t^2} \qquad (2.40)$$

From the above equation, F_i is the force on the i^{th}atom, m_i is the mass of the i^{th}atom, a_i is the acceleration of the i^{th} atom, and $\frac{\delta^2 r_i}{\delta t^2}$ is the change in the position of the i^{th} atom with respect to the time. The dynamic behaviour of the finalized candidates with respect to the biological environment studies were studied using Maestro Desmond Interoperability Tool (academic) [39, 40] and GROMACS 2016.4 [41]. In general, MD simulations are used to calculate the force on the i^{th} atom with respect to the time. The calculated force is defined as the first derivative of the energy with respect to the change in the position of the atom (equation 2.41).

$$F_i = \frac{\delta E(r_1 r_2, \dots\dots, r_n)}{\delta r_i} \qquad (2.41)$$

Here in equation 2.9, F_i is the force on the i^{th} atom, $\frac{\delta E(r_1 r_2, \dots\dots, r_n)}{\delta r_i}$ is the change in the potential energy of the specific conformation with respect to the position. This potential energy is calculated by employing the force field [42]. In terms of mathematics, the force field is defined as the expression encompassing the functional form the potential energy. In other words, the force field incorporates bonded and non-bonded interactions, which are formed between the different atoms as represented in Figure 2.2

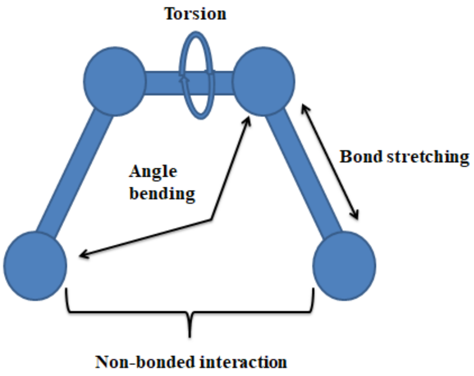

Figure 2.2. Different types of interactions observed in a molecular system

The energy of the system/model is calculated as the sum of the bonded and non-bonded interactions (equation 2.42), where bonded potential is made from stretching, bending and rotation along the bonds (torsion) (equation 2.43) and non-bonded potential is composed of pair-wise addition of electrostatic and Van der Waals potential (equation 2.44).

$$E(r) = E_{bonded} + E_{non-bonded} \tag{2.42}$$

$$E_{bonded} = E_{bond-streach} + E_{angle-bend} + E_{rotate-along-bond} \tag{2.43}$$

$$E_{non-bonded} = E_{van-der-Waals} + E_{electrostatic} \tag{2.44}$$

So far, various types of force fields have been developed for the simulations of the biological systems [43-45]. In our study, we have used the optimized potentials for liquid simulations (OPLS) force field [46-49] and GROMOS 54a7 force field [50] to calculate the potential energy of the system (equation 2.45).

$$E = \Sigma_{i<j}\left[\frac{q_i q_j e^2}{r_{ij}} + 4\,\epsilon_{ij}\left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^{6}}{r_{ij}^{6}}\right)\right] f_{ij} + \Sigma_{bonds} K_r (r - r_{eq})^2 + \Sigma_{angles} K_\theta (\theta - \theta_{eq})^2 +$$

$$\Sigma_{dihedral}[\frac{V_1}{2}(1 + cos\varphi) + \frac{V_2}{2}(1 - cos2\varphi) + \frac{V_3}{2}(1+\cos 3\varphi) + \frac{V_4}{2}(1-\cos 4\varphi)] \tag{2.45}$$

In the equation 2.45, the harmonic potential function was employed to model the angle bending and bond stretching, whereas cosine function was used to formulate the dihedrals of the system. The bond length (r) and angle (θ), having subscript eq denotes the equilibrium state of the same. The torsion term is indicated by reduced Fourier series which incorporates the summation of all the dihedrals of the system. For modelling, the non-bonded interactions, pair-wise interactions (i and j) were calculated by employing the Coulomb's law and Lennard-Jones (LJ) potentials [51]. As per geometric combining rule, LJ potential parameters that are, σ and ϵ are site-specific [48]. For the calculation of the long-range interactions (electrostatic), particle-mesh Ewald (PME) method is utilized under the periodic boundary conditions [52]. For a macromolecule, the solution of the Newtonian equation is unrealistic as the biological systems comprise of thousands of degree of freedom. Thus, to overcome this, different algorithms were developed to solve the Newton equation of the motion (equation 2.8) like Leap-Frog integrator [53], Velocity Verlet integrator [54] and Verlet integrator [55]. The solution of the algorithms provides us the trajectory of the system while performing the dynamics calculations. In our work, to calculate the position and momentum of the complex, Velocity Verlet integrator was used. Equation 2.46 represents the calculation of the change in the position, and equation 2.47 represents the change in the velocity of the system with respect to a specific time interval.

$$x(t + \Delta t) = x(t) + v(t)\Delta t + \frac{1}{2}\mathrm{a(t)}\Delta t^2 \tag{2.46}$$

$$v(t + \Delta t) = v(t) + \frac{a(t)+a(t+\Delta t)}{2}\Delta t \tag{2.47}$$

In equation 2.46 and 2.47, $x(t)$, $v(t)$ and $a(t)$ are the initial positions, velocity, and acceleration of the system, respectively. $x(t + \Delta t)$ and $v(t + \Delta t)$ are the updated position, velocity, and acceleration of the system with respect to the time interval of Δt, respectively.

To restrict the dynamic movement of the least important covalently hydrogen bonding, the integration algorithms employed are, LINCS [56], RATTLE [57] and SHAKE [58] algorithms. And to perform this step, a time-step of 1.5 to 2fs is required for the simulation of the bio-molecules [59]. In the current thesis, we applied M-SHAKE algorithm [60] and the LINCS algorithm [56] to calculate the distance constraints. It uses Lagrange multiplier to compute the magnitude of the force and predicts the changes in the position accordingly until the position constraints are satisfied.

2.4.2 MD in various ensembles

MD simulations are playing an important role in answering various biological questions as this technique can mimic the experimental situation [61]. We can control the experimental factors like the number of atoms, temperature, pressure, ionic concentration, and solvent type by employing the statistical mechanic's ensembles [62]. Depending on the type of factor/variable we wish to keep fixed, different ensembles can be generated [63, 64]. These variables are the macroscopic properties of the system like temperature (T), volume (V), pressure (P), the chemical potential of particles (μ), number of particles (N) and energy (E) [3]. Different types of ensembles are canonical (NVT), isothermal-isobaric (NPT), microcanonical (NVE) and grand canonical (μVT). Out of these four ensembles, most frequently used ensembles are NVT [65] and NPT [66]. NVT is the constant volume and temperature ensemble, and NPT is the constant pressure and temperature ensemble. NVE [67] ensemble is not realistic as it keeps the energy and volume constant. For the equilibration process, the constant energy ensemble is not recommended because it is not possible to obtain the required temperature. Therefore, in the present thesis, we have used the *NVT* ensemble followed by *NPT* to provide the algorithmic stability to the selected system. We have employed Nose-Hoover thermostat [68] to treat the pressure, and Martyna-Tuckerman-Klein barostat [69] was exploited to treat the temperature of the system. The general steps involved in conducting the MD simulations are shown in figure 2.3.2.

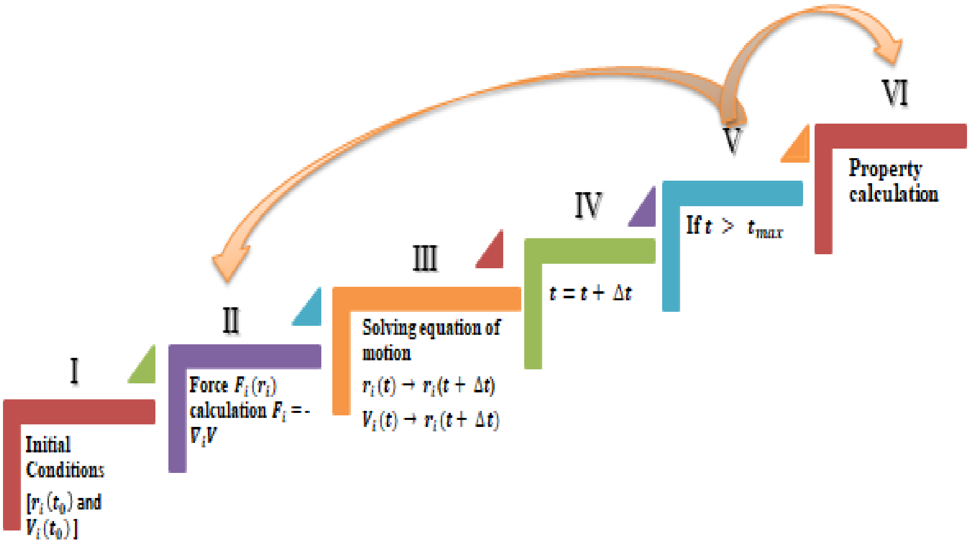

Figure 2.3. General steps involved in MD simulation.

References

[1] R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles* (CRC Press, 1988).

[2] J. Siepmann and F. Siepmann, Int. J. Pharm. **364**, 328 (2008).

[3] J. A. McCammon, B. R. Gelin, and M. Karplus, Nature **267**, 585 (1977).

[4] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[5] P. Korhonen, M. Kulmala, A. Laaksonen, Y. Viisanen, R. McGraw, and J. H. Seinfeld, J. Geophys. Res. Atmos. **104**, 26349 (1999).

[6] R. Zhang, Science (80-.). **328**, 1366 (2010).

[7] B. M. Boghosian and W. Taylor, Int. J. Mod. Phys. C **8**, 705 (1997).

[8] S. Guo, M. Hu, M. L. Zamora, J. Peng, D. Shang, J. Zheng, Z. Du, Z. Wu, M. Shao, L. Zeng, and others, Proc. Natl. Acad. Sci. **111**, 17373 (2014).

[9] R. J. Charlson, S. E. Schwartz, J. M. Hales, R. D. Cess, J. J. A. Coakley, J. E. Hansen, and D. J. Hofmann, Science (80-.). **255**, 423 (1992).

[10] F. R. Manby and P. J. Knowles, Phys. Rev. Lett. **87**, 163001 (2001).

[11] R. S. Mulliken, J. Chem. Phys. **7**, 20 (1939).

[12] D. S. Kosov and P. L. A. Popelier, J. Phys. Chem. A **104**, 7339 (2000).

[13] R. S. Mulliken, J. Chem. Phys. **3**, 573 (1935).

[14] R. Santamaria, G. Cocho, L. Corona, and E. Gonz'alez, Chem. Phys. **227**, 317 (1998).

[15] P.-O. L"owdin, Phys. Rev. **97**, 1474 (1955).

[16] R. G. Parr and W. Yang, J. Am. Chem. Soc. **106**, 4049 (1984).

[17] J. A. Alonso, Chem. Rev. **100**, 637 (2000).

[18] J. Mortier, C. Rakers, M. Bermudez, M. S. Murgueitio, S. Riniker, and G. Wolber, Drug Discov. Today **20**, 686 (2015).

[19] I. Buch, T. Giorgino, and G. De Fabritiis, Proc. Natl. Acad. Sci. **108**, 10184 (2011).

[20] P. Doruker, A. R. Atilgan, and I. Bahar, Proteins Struct. Funct. Bioinforma. **40**, 512 (2000).

[21] Y. Shan, E. T. Kim, M. P. Eastwood, R. O. Dror, M. A. Seeliger, and D. E. Shaw, J. Am. Chem. Soc. **133**, 9181 (2011).

[22] I. Bartova, M. Otyepka, Z. K\vr'\iž, and J. Koča, Protein Sci. **13**, 1449 (2004).

[23] K. Arora and C. L. Brooks Iii, J. Am. Chem. Soc. **131**, 5642 (2009).

[24] J. I. Sułkowska, J. K. Noel, and J. N. Onuchic, Proc. Natl. Acad. Sci. **109**, 17783 (2012).

[25] K. Nam, J. Pu, and M. Karplus, Proc. Natl. Acad. Sci. **111**, 17851 (2014).

[26] R. L. Hayes, J. K. Noel, U. Mohanty, P. C. Whitford, S. P. Hennelly, J. N. Onuchic, and K. Y. Sanbonmatsu, J. Am. Chem. Soc. **134**, 12043 (2012).

[27] A. Yildirim, M. Sharma, B. M. Varner, L. Fang, and M. Feig, J. Phys. Chem. B **118**, 10874 (2014).

[28] M. Sekijima, C. Motono, S. Yamasaki, K. Kaneko, and Y. Akiyama, Biophys. J. **85**, 1176 (2003).

[29] V. Gapsys, B. L. de Groot, and R. Briones, J. Comput. Aided. Mol. Des. **27**, 845 (2013).

[30] H. I. Ingolfsson, M. N. Melo, F. J. Van Eerden, C. Arnarez, C. A. Lopez, T. A. Wassenaar, X. Periole, A. H. De Vries, D. P. Tieleman, and S. J. Marrink, J. Am. Chem. Soc. 136, 14554 (2014).

[31] Z. A. Levine, R. M. Venable, M. C. Watson, M. G. Lerner, J.-E. Shea, R. W. Pastor, and F. L. H. Brown, J. Am. Chem. Soc. **136**, 13582 (2014).

[32] A. J. Sodt, M. L. Sandar, K. Gawrisch, R. W. Pastor, and E. Lyman, J. Am. Chem. Soc. **136**, 725 (2014).

[33] J. E. Kerrigan, In silico Models for Drug Discovery, Springer. **993,** 95 (2013).

[34] C. S. Tautermann, D. Seeliger, and J. M. Kriegl, Comput. Struct. Biotechnol. J. **13**, 111 (2015).

[35] A. Anwar-Mohamed, K. H. Barakat, R. Bhat, S. Y. Noskov, D. L. Tyrrell, J. A. Tuszynski, and M. Houghton, Toxicol. Lett. **230**, 382 (2014).

[36] K. Barakat, J. Mane, D. Friesen, and J. Tuszynski, J. Mol. Graph. Model. **28**, 555 (2010).

[37] K. Barakat and J. Tuszynski, J. Mol. Graph. Model. **29**, 702 (2011).

[38] K. H. Barakat, A. Anwar-Mohamed, J. A. Tuszynski, M. J. Robins, D. L. Tyrrell, and M. Houghton, J. Chem. Inf. Model. **55**, 362 (2015).

[39] K. J. Bowers, D. E. Chow, H. Xu, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I. Kolossvary, M. A. Moraes, F. D. Sacerdoti, and others, in SC '06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing, 43 (2006).

[40] M. Tuckerman, B. J. Berne, and G. J. Martyna, J. Chem. Phys. **97**, 1990 (1992).

[41] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, J. Comput. Chem. **26**, 1701 (2005).

[42] A. D. MacKerell Jr, J. Comput. Chem. **25**, 1584 (2004).

[43] A. D. MacKerell Jr, D. Bashford, M. Bellott, R. L. Dunbrack Jr, J. D. Evanseck, M. J.

Field, S. Fischer, J. Gao, H. Guo, S. Ha, and others, J. Phys. Chem. B **102**, 3586 (1998).

[44] 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, and P. A. Kollman, J. Am. Chem. Soc. **117**, 5179 (1995).

[45] X. Daura, A. E. Mark, and W. F. Van Gunsteren, J. Comput. Chem. **19**, 535 (1998).

[46] G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, and W. L. Jorgensen, J. Phys. Chem. B **105**, 6474 (2001).

[47] M. J. Robertson, J. Tirado-Rives, and W. L. Jorgensen, J. Chem. Theory Comput. **11**, 3499 (2015).

[48] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, J. Am. Chem. Soc. **118**, 11225 (1996).

[49] W. L. Jorgensen and J. Tirado-Rives, J. Am. Chem. Soc. **110**, 1657 (1988).

[50] N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark, and W. F. van Gunsteren, Eur. Biophys. J. **40**, 843 (2011).

[51] D. N. Kouetcha, H. Ramézani, and N. Cohaut, J. Supercomput. **73**, 1629 (2017).

[52] T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).

[53] M. A. Cuendet and W. F. van Gunsteren, J. Chem. Phys. **127**, 184102 (2007).

[54] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, J. Chem. Phys. **76**, 637 (1982).

[55] L. Verlet, Phys. Rev. **159**, 98 (1967).

[56] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, J. Comput. Chem. **18**, 1463 (1997).

[57] H. C. Andersen, J. Comput. Phys. **52**, 24 (1983).

[68] J.-P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, J. Comput. Phys. **23**, 327 (1977).

[59] S. Kalyaanamoorthy and Y.-P. P. Chen, Prog. Biophys. Mol. Biol. **114**, 123 (2014).

[60] S. G. Lambrakos, J. P. Boris, E. S. Oran, I. Chandrasekhar, and M. Nagumo, J. Comput. Phys. **85**, 473 (1989).

[61] A. Ganesan, M. L. Coote, and K. Barakat, Drug Discov. Today **22**, 249 (2017).

[62] D. McQuarrie, Statistical mechanics university science books. Sausalito, CA. 222 (2000).

[63] L. Wang, Y. Wu, Y. Deng, B. Kim, L. Pierce, G. Krilov, D. Lupyan, S. Robinson, M. K. Dahlgren, J. Greenwood, and others, J. Am. Chem. Soc. **137**, 2695 (2015).

[64] K. T. Gon, B. P. Zeigler, and H. Praehofer, *Theory of modeling and simulation: integrating discrete event and continuous complex dynamic systems*, San Diego, CA :

Academic Press (2000).

[65] J. M. Richardson and S. R. Brinkley Jr, J. Chem. Phys. **33**, 1467 (1960).

[66] D. S. Corti and G. Soto-Campos, J. Chem. Phys. **108**, 7959 (1998).

[67] D. A. Beard and H. Qian, *Chemical Biophysics: Quantitative Analysis of Cellular Systems*, **126,** Cambridge University Press Cambridge, 336 (2008).

[68] S. Nos'e, J. Chem. Phys. **81**, 511 (1984).

[69] G. J. Martyna, D. J. Tobias, and M. L. Klein, J. Chem. Phys. **101**, 4177 (1994).