2.1 Computational chemistry

Computational chemistry is the combination of all numerical methods which are based on molecular mechanics (MM), molecular dynamics (MD), Monte Carlo (MC) and quantum chemistry (QC) simulations. These simulations are particularly employed for predicting structure, electronic and optical properties of materials.[1-9] The basic principle of these simulations is to determine accurately the total energy of an investigated system. The methods are briefly discussed in the following sections.

2.1.1 Molecular mechanics

Molecular mechanics or force field methods use classical laws of physics and offer us an extremely powerful tool for analyzing the structural, mechanistic and energetic properties of molecules. The negative of the first derivative of the potential energy of a particle with respect to displacement along some direction is the force on the particle, hence the term force field arises. A force field E (x, y, z, coordinates of atoms) can be differentiated to give the force on each atom. In molecular mechanics the electronic degrees of freedom of the molecules are ignored and perform computations based upon nuclei interactions. However, electronic effects have been implicitly included in force fields through parameterizations. These parameterizations are not used for caring out for transition states reactions, thereby, hindering the study of reaction mechanism using molecular mechanics. The approximations that are adopted in molecular mechanical calculation make the computations inexpensive. Hence, this method can be applied for systems containing thousands of atoms, such as bio-molecules. The existing drawbacks of molecular mechanics are: first, this method is appropriate only for those classes of molecules for which the force field is parameterized. Secondly, system having prominent electronic effects can't be applied molecular mechanics methods. For example, this method cannot describe the chemical reaction involving bond formation or bond breaking.

2.1.2 Quantum mechanics

2.1.2.1 The Schrödinger equation

Quantum chemical approach has become important and widespread to determine the electronic structure of an atoms or molecules. The electronic structure and total

electronic energy of atoms, molecules and crystals can be obtained by solving the time-independent, non-relativistic Schrödinger equation.

$$\hat{H}\psi = E\Psi \tag{2.1}$$

In the wave mechanics formulation, $\hat{H}$ is the Hamiltonian operator, Ψ is a wave function and E is a scalar value representing the system energy. The Hamilton operator of a system is expressed as

$$\hat{H} = -\frac{\hbar^2}{2}\sum_i \frac{1}{m_i}\left(\frac{\delta^2}{\delta x_i^2}+\frac{\delta^2}{\delta y_i^2}+\frac{\delta^2}{\delta z_i^2}\right)+\sum_{i<j}\frac{e_i e_j}{r_{ij}} \tag{2.2}$$

The equation (2.2) is the time independent Schrödinger equation, where the first quantity on right hand side account for the kinetic energies and the second term accounts for the potential energies. The exact solution of Schrödinger equation is not possible, even for the smallest systems. Numerous mathematical approximations have been applied to find out the solution of Schrödinger equation.

2.1.2.2 Born-Oppenheimer approximation

The Schrödinger equation can be solved analytically for systems having one electron. Noticeably, the solution of this equation becomes increasingly difficult as more electrons are present. Consequently, several approximations need to be made to carry out calculations on systems containing more than two particles. One of the fundamental approximations used in this context is the Born-Oppenheimer approximation in which, the motion of electrons and nuclei are separated.

$$H_{total} = H_{electronic} + H_{nuclear} \tag{2.3}$$

$$H_{electronic} = -\frac{\hbar^2}{2}\sum_i \frac{1}{m_i}\left(\frac{\delta^2}{\delta x_i^2}+\frac{\delta^2}{\delta y_i^2}+\frac{\delta^2}{\delta z_i^2}\right)+\sum_{i<j}\frac{e_i e_j}{r_{ij}} \tag{2.4}$$

The total molecular wave function is simply the product of electronic and nuclear wave functions, which results in the simplified Schrödinger equation (2.6). The approximation allows us to calculate the wave function for electrons moving in a fixed potential field of the nuclei

$$\Psi_{total} = \Psi_{electronic}\Psi_{nulear} \tag{2.5}$$

$$H_{electronic}\Psi_{electronic} = E_{electronic}\Psi_{electronic} \quad (2.6)$$

The total energy of a system can now be calculated from the electronic energy and the nuclear-nuclear repulsion remains constant for a given geometry.

2.1.2.3 Hartree-Fock method

Schrödinger equation gives an exact solution for the hydrogen atom. But when atoms other than hydrogen are considered, the number of inter-electron repulsion in the Hamiltonian will increase, thereby, interpreting the exact solution of the Schrödinger equation becomes extremely complicated. Hence, more powerful methods for calculating the ground state energy and wave functions of many electron atoms or ions are needed. In 1927, the British physicist D. R. Hartree had suggested that the wave function of an N-electron atom can be written as the product of N one-electron wavefunctions.[10] The product is typically referred to as Hartree wave function, Ψ^{HP}

$$\Psi^{HP}(r_1 r_2 r_3 ..... r_N) = \phi(r_1)\phi(r_2)\phi(r_3)..... \phi(r_N) \quad (2.7)$$

Where $\phi(r_i)$ is normalized and mutually orthogonal one-electron wave function.

But the major shortcomings associated with this wavefunction is that it fails to satisfy the antisymmetry principle, which states that a wavefunction describing fermions should be antisymmetric with respect to the interchange of any set of space-spin coordinates. By space-spin coordinates, we mean that fermions not only have three spatial degrees of freedom, but also have an intrinsic spin coordinate, α or β. Hartree wave function provided in equation (2.7) is only the product of the N one-electron spatial wave functions. Thus, in order to improve the wave function we have to include both spatial part as well the electron spin part and are defined by equation (2.8)

$$\Psi^{HP}(r_1, r_2, r_3 ..... r_N) = \chi(r_1)\chi(r_2)\phi(r_3)..... \chi(r_N) \quad (2.8)$$

Where $\chi = \phi(r_i) \times \alpha$ or $\chi = \phi(r_i) \times \beta$

The Hartree total wave function given in equation (2.8) does not satisfy the antisymmetry principle. The antisymmetry principle can be satisfied by rewriting the above functional form with slater determinant as

$$\psi^{HP}(r_1, r_2, r_3 \ldots\ldots r_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(r_1) & \chi_2(r_1) & \ldots & \chi_N(r_1) \\ \chi_1(r_2) & \chi_2(r_2) & \ldots & \chi_N(r_2) \\ . & . & \ldots & . \\ . & . & \ldots & . \\ . & . & \ldots & . \\ \chi_1(r_N) & \chi_2(r_N) & .... & \chi_N(r_N) \end{vmatrix} \quad (2.9)$$

This determinant of spin orbital is so called the slater determinant has a desired effect since interchanging positions of two electrons changes its sign. Thus, in Hartree-Fock approach, the N-electron wave function is the antisymmetric product of individual electron spin-orbital.

2.1.2.4 Density functional theory

Density functional theory also known as DFT provides an alternative way to solve the Schrödinger equation. In recent years, DFT becomes the most successful and promising approaches to compute the electronic structure of matter. Applicability of DFT ranges from atoms, molecules and solids to nuclei and quantum and classical fluids. The original formulation of DFT is that the ground-state electronic energy is determined completely by the electron density in the system of non-interacting electrons. The origins of DFT are found in the Thomas-Fermi model in 1920s, where electronic energy was attempted to be calculated in terms of the electron density.[11] Thomas-Fermi model failed to describe molecular bonding, therefore, rendered this method impractical for any real system. However, the first real improvement in the use of DFT for molecules arose from the two Hohenberg-Kohn theorems, developed in 1964.[12]

The first Hohenberg-Kohn theorem

First theorem states that the external potential $v_{ext}(\vec{r})$ is a unique functional of $\rho(\vec{r})$. Their theory is found to be different from the traditional quantum chemical methods that are based on wavefunctions. In wavefunction based methods, the wavefunction governs everything, and the electron density results from it. But, in DFT, one-to-one correspondence between the electron density of a system and the energy is exist, that

is, there is a map ($v_{ext}(\vec{r}) \leftrightarrow \chi(\vec{r}) \leftrightarrow \rho\ (\vec{r})$). The ground-state density $\rho(\vec{r})$ has uniquely determined the external potential $v_{ext}(\vec{r})$, the ground-state wave function $\psi(\rho)$ and hence, all the properties of the ground state, for example the kinetic energy $T_s[\rho]$ the potential energy $v[\rho]$ and the total energy $E[\rho]$. Now the total energy can be written as

$$E[\rho] = E_{ext}[\rho] + T_s[\rho] + E_{ee}[\rho] \tag{2.10}$$

$E_{ee}[\rho]$ denotes electron-electron interaction term while $E_{ext}[\rho]$ is for the nuclei-electron interaction. The energy expression can be subdivided into two parts: the part independent of the system along with external perturbation if any, that is $T_s[\rho] + E_{ee}[\rho]$ which is collectively referred to as Hohenberg-Kohn functional $f_{HK}[\rho]$

$$f_{HK}[\rho] = T_s[\rho] + E_{ee}[\rho] \tag{2.11}$$

And the other one which depend on the system, *i.e.*, due to nuclei-electron attraction and perturbation,

$$E_{ext}[\rho] = \int \rho(\vec{r}) v_{ext} d\vec{r} \tag{2.12}$$

Equation (2.10) can be rewritten as

$$E[\rho] = F_{HK}[\rho] + \int \rho(\vec{r}) v_{ext} d\vec{r} \tag{2.13}$$

If Hohenberg-Kohn functional $f_{HK}[\rho]$ is known explicitly, then the Schrödinger equation can be solved exactly for hydrogen atom as well as for gigantic molecules such as DNA.

The second Hohenberg-Kohn theorem

The second Hohenberg-Kohn theorem states that the Hohenberg-Kohn functional $f_{HK}[\rho]$, attains its ground state energy E_0 with respect to all allowed densities if and only if the input density is the true ground state density, that is

$$E_0 \leq E[\tilde{\rho}] = F_{HK}[\tilde{\rho}] + \int \tilde{\rho}(\vec{r}) v_{ext} d\vec{r} \tag{2.14}$$

The meaning of the equation (2.14) is that for any trial density $\tilde{\rho}(\vec{r})$, which has satisfied the necessary boundary conditions such as $\tilde{\rho}(\vec{r}) \geq 0$, $\int \tilde{\rho}(\vec{r}) v_{ext} d\vec{r} = N$ and which is associated with some external potential $\tilde{v}_{ext}$, the energy obtained from the

functional of equation (2.13) represents an upper bound to the true ground state energy E_0.

Kohn-Sham approach

In DFT, quantum chemical calculations are performed very conveniently in terms of single particle orbitals within the Kohn-Sham formalism.[13] The Kohn-Sham orbitals $\{\chi_i(\vec{r})\}$ equation is given as

$$\hat{H}\chi_i(\vec{r}) \equiv \left[-\frac{1}{2}\nabla^2 + v(\vec{r})\right]\chi_i(\vec{r}) = \varepsilon_i \chi_i(\vec{r}) \tag{2.15}$$

where, r is the space vector and $v(\vec{r})$ is the external potential, $\chi(\vec{r})$ is the wave function of orbital occupied by each electron, ε_i be the eigen value corresponding to that orbital. Accurate Kohn-Sham orbitals $\chi_i(\vec{r})$ once constructed, the electron density $\rho(\vec{r})$ can be obtained exactly using the following formula by summing over all the occupied orbitals:

$$\rho(\vec{r}) = \sum_i f_i \left|\chi_i\right|^2 \tag{2.16}$$

The kinetic, exchange, and correlation energies can be calculated from

$$E^{\mathrm{KS}} = \left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle = T_{\mathrm{s}} + V + W_{\mathrm{H}} + E_{\mathrm{x}} \tag{2.17}$$

$$T_{\mathrm{s}} = -\frac{1}{2}\sum_{\sigma}\sum_{i=1}^{N_\sigma}\int \chi_{i\sigma}^{*}(r)\nabla^2\chi_{i\sigma}(r)dr \tag{2.18}$$

$$E_{\mathrm{x}} = -\frac{1}{2}\sum_{\sigma\sigma'}\sum_{i=1}^{N_\sigma}\sum_{j=1}^{N_\sigma}\iint \frac{\chi_{i\sigma}^{*}(r)\chi_{j\sigma}(r)\chi_{j\sigma'}^{*}(r')\chi_{i\sigma'}(r')}{\left|r-r'\right|}drdr' \tag{2.19}$$

$$E_{\mathrm{c}} = E_{\mathrm{xc}} - E_{\mathrm{x}} = E - E^{\mathrm{KS}} \tag{2.20}$$

The construction of the KS orbital is a necessary step, but the determination of the KS potential $v(\vec{r})$ corresponding to accurate target density $\rho(\vec{r})$ turns out to be a difficult part in the construction of the accurate KS orbital. To obtain an expression for the Kohn-Sham potential $v(\vec{r})$, the Hohenberg-Kohn functional for the interacting system is partitioned as follows:

$$F_{\text{HK}}[\rho] = T_s[\rho] + \frac{1}{2}\iint \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|} d\vec{r}_1 d\vec{r}_2 + E_{\text{xc}}[\rho] = T_s[\rho] + W_{\text{H}}[\rho] + E_{\text{xc}}[\rho] \qquad (2.21)$$

Here, the first second and third term is kinetic energy, coulomb interactions between core charges, and exchange-correlation, respectively. Therefore, the Kohn-Sham potential $v(\vec{r})$ can be subdivided into the following parts:

$$v(\vec{r}) = v_{\text{ext}}(\vec{r}) + v_{\text{H}}(\vec{r}) + v_{\text{xc}}(\vec{r}) \qquad (2.22)$$

Where, v_{ext} is the external potential which is the Coulomb field of the nuclei, v_H is the Hartree potential, which is the classical Coulomb repulsion between the electrons and the v_{xc} is the potentials corresponding to the exchange-correlation energy which is the only unknown part. It is very important to realize that if the exact forms of $v_{xc}(\vec{r})$ is known, the Kohn-Sham strategy will lead to the exact potential $v(\vec{r})$ since v_{ext} is already known and v_H calculation is straight forward for any given density. Therefore the accurate determination of the Kohn-Sham potential from the accurate electron density $\rho(\vec{r})$ allows us to judge approximations to the energy functional $E_{xc}[\rho]$ by comparing the approximate model potential with the accurate one.[14,15] In DFT, exact force of xc potential is not available, hence to solve the Kohn-Sham equation an approximation for the xc potential $v_{xc}(\vec{r})$, which should contain all the many-body effects, is required. Many different approximations have been proposed for practical application, among which the most important approximations are the local density approximation (LDA).[16] In LDA, the quantum system under study is assumed to be based upon uniform distribution of electron gas. According to Hohenberg and Kohn, if the density $\rho(\vec{r})$ varies extremely slowly with r, then $E_{xc}[\rho]$ can be written as

$$E_{xc}^{LDA}[\rho] = \int \rho(\vec{r}) \varepsilon_{xc}^{LDA}[\rho(\vec{r})] d\vec{r} \qquad (2.23)$$

Where $\varepsilon_{xc}^{LDA}[\rho(\vec{r})]$ is the exchange-correlation energy per electron of a homogenous electron gas with electron density $\rho(\vec{r})$. Although LDA might not be appropriate for real atoms and molecules but has been remarkably successful for some systems.

In order to account for the non-homogeneity of electrons, in few years back, the generalized gradient approximations (GGA) have been developed.[17] In this approach, the electron density $\rho(\vec{r})$ at a particular point r is supplemented with the gradient of density, $\nabla_\rho(\vec{r})$. Equation of the GGA can be written as

$$E_{xc}^{GGA}[\rho] = \int \rho(r)\varepsilon_{xc}^{GGA}\left[\rho(r)\left|\nabla_\rho(r)\right|\right]dr \qquad (2.24)$$

It has been observed that the accurate estimation of correlation energy in GGA receives considerable attention although the chemical significances of gradient corrections for correlation are relatively small as compared to their exchange counterparts.[18] The most popular correlation functionals are the LYP (Lee, Yang, and Parr) including both local and non-local terms[19], the P86 (Perdew 1986) functional[20] and the PW91 (Perdew and Wang 1991) functional.[21]

2.1.2.5 DFT based reactivity descriptors

Density functional theory (DFT) provides a framework in order to describe reactions in terms of changing number of electrons (N) or changing external potential, $\upsilon(\vec{r})$ due to nuclei.

Chemical potential (μ) can be defined as the first derivative of the energy (E) with respect to the number of electrons (N) at constant external potential, $\upsilon(\vec{r})$

$$\mu = \left(\frac{\delta E}{\delta N}\right)_{\nu(\vec{r})} = -\chi \qquad (2.25)$$

Here μ equals the negative of electronegativity defined by Iczkowski and Margrave[22], since it determines the energy change upon changing the total number of electrons. Parr and Pearson defined global hardness as the corresponding second derivative of energy.[23]

$$\eta = \frac{1}{2}\left(\frac{\partial^2 E}{\partial N^2}\right)_{\upsilon(\vec{r})} = \frac{1}{2}\left(\frac{\partial \mu}{\partial \mathrm{N}}\right)_{\upsilon(\vec{r})} \qquad (2.26)$$

Chemical hardness (η) and chemical potential (μ) can also be calculated using a finite difference approximation in terms of IP and EA[24]

$$\eta = \frac{IP - EA}{2} \qquad (2.27)$$

$$\mu = -\frac{IP + EA}{2} \tag{2.28}$$

Where IP and EA are the first vertical ionization potential and electron affinity, respectively, for a chemical system.

According to Koopman's approximation[25] IP and EA can be defined in terms of the energies of highest occupied molecular orbital (E_{HOMO}) and the lowest unoccupied molecular orbital (E_{LUMO}).

$$IP = -E_{HOMO} \tag{2.29}$$

$$EA = -E_{LUMO} \tag{2.30}$$

And therefore η and μ can be expressed as.

$$\eta = \frac{1}{2}(E_{LUMO} - E_{HOMO}) \tag{2.31}$$

$$\mu = \frac{1}{2}(E_{LUMO} + E_{HOMO}) \tag{2.32}$$

Parr and his coworkers expressed[26] global electrophilicity (ω) as

$$\omega = \frac{\mu^2}{2\eta} \tag{2.33}$$

The global reactivity descriptors cannot be used for studying the site-selectivity of a chemical system, therefore, appropriate local reactivity descriptors need to be defined. Parr and Yang[27] defined Fukui function $f(\vec{r})$ as the mixed second derivative of the energy of the system with respect to the number of electrons (N) at constant external potential $\upsilon(\vec{r})$

$$f(\vec{r}) = \left(\frac{\delta^2 E}{\delta N \delta v(\vec{r})}\right) = \left[\frac{\delta \mu}{\delta v(\vec{r})}\right]_N = \left(\frac{\delta \rho(\vec{r})}{\delta N}\right)_{v(\vec{r})} \tag{2.34}$$

Where $\rho(\vec{r})$ is the electron density.

To describe site selectivity or reactivity of an atom in a molecule, the Fukui function values around each atomic site should be condensed into a single value, which can be achieved by electronic population analysis. Hence, for an atom k in a molecule, there are three different types of condensed Fukui function such as f_k^+, f_k^- and f_k^0, respectively, depending upon the type of electron transfer.

By using an atomic charge partitioning scheme, the definition of condensed Fukui function for k atom undergoing nucleophilic attack, electrophilic attack and free radical attack are

For nulcleophilic attack $f_k^+ = \rho_k(N+1) - \rho_k(N)$ (2.35)

For electrophilic attack $f_k^- = \rho_k(N) - \rho_k(N-1)$ (2.36)

For free radical attack $f_k^0 = \frac{\rho_k(N+1) - \rho_k(N-1)}{2}$ (2.37)

Where $\rho_k(N)$, $\rho_k(N+1)$ and $\rho_k(N-1)$ are the electron densities of the N, N+1 and N-1 electron systems, respectively.

The condensed form of Fukui functions introduced by Yang and Mortier[28] are as follows:

For nulcleophilic attack $f_k^+ = q_k(N+1) - q_k(N)$ (2.38)

For electrophilic attack $f_k^- = q_k(N) - q_k(N-1)$ (2.39)

For free radical attack $f_k^0 = \frac{q_k(N+1) - q_k(N-1)}{2}$ (2.40)

q_k defines the populations of atom k in the molecule.

2.1.3 Quantum mechanics/molecular mechanics method (QM/MM)

It is well known that molecular mechanics (MM) is not used for studding the chemical reactions as it does not able to calculate or simulate the breaking or formation of chemical bonds. Therefore, quantum mechanical (QM) simulation has been considered for studying the interaction of biomolecules with small drug molecules. Quantum mechanical simulations can perfectly describe the hydrogen, ionic and covalent binding interactions. QM methods are based on solving the Schrödinger equation, thereby taking directly into consideration the electronic structure of a molecule and therefore allow access to chemical interactions. Unfortunately, quantum mechanical methods of high-quality are computationally very expensive and cannot be used directly for studying large molecules. In order to overcome such limitation, Morokuma and co-workers have developed a hybrid method (ONIOM) based on combination of several theoretical approaches for large biomolecular systems.[31-34] This method has been implemented in Gaussian09 program as ONIOM (our Own N-

layer Integrated molecular Orbital molecular Mechanics) which is a powerful and systematic method which divides the system into several layers and have been suggested at various levels.[35-37] According to Morokuma *et al*., the full molecular geometry of the system which include all atoms is referred to as "real" geometry and is treated with a "low"-level of theory. Chemically most important (core) region of the system, referred to as the "model" geometry are treated using both the "low"-level and "high"-level of theory. A three layer model introduces an "intermediate" model geometry which is treated with a "medium" level of theory.

In the two-layer ONIOM method, the real system energy is obtained from three independent calculations:

$$E^{ONIOM2} = E^{\text{high}}_{\text{model system}} + E^{\text{low}}_{\text{real system}} - E^{\text{low}}_{\text{model system}} \qquad (2.41)$$

The ONIOM method uses an extrapolation to calculate the total energy. Beginning with $E^{low}_{\text{model system}}$, the extrapolation to the high level calculation ($E^{high}_{\text{model system}} - E^{low}_{\text{model system}}$) and the extrapolation to the real system ($E^{low}_{real\,system} - E^{low}_{\text{model system}}$) have been assumed to give an estimate for $E^{high}_{real system}$.

In case of the three-layered ONIOM methods, the ONIOM energy can be stated by following equation.

$$E^{ONIOM3} = E^{high}_{\text{model system}} + E^{medium}_{middle\,system} - E^{medium}_{\text{model system}} + E^{low}_{real\,system} - E^{low}_{middle\,system} \qquad (2.42)$$

The *real* system contains all the atoms and calculation is performed at MM level, while the *model* system contains the part of the system that is treated at QM level. Both QM and MM calculations need to be carried out for the model system. ONIOM method has become very successful and extensively used in studies of DNA, protein interaction.

2.1.4 Molecular docking

The application of molecular modeling methods to study the formation of intermolecular complexes has become the subject of research interest for the last three decades. In the field of molecular modeling, the process of searching for a small molecule that is able to fit both geometrically and energetically to the binding site of a target macromolecule is called molecular docking (Fig.2.1).[38]

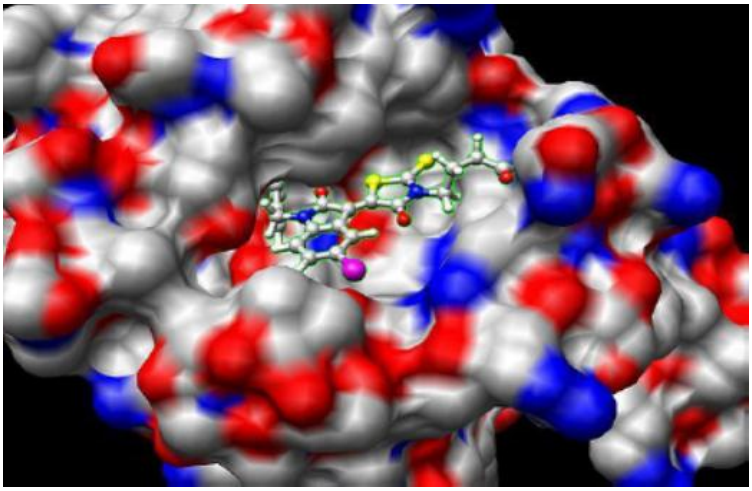

Fig.2.1 Small molecule fits to a large macromolecule[39]

2.1.4.1 Mechanics of docking

The two different components associated with the success of molecular docking are search algorithm and scoring function. The process of searching whether a given conformation and orientation of a ligand fits the active site comes under search algorithm. Search algorithm falls into two main categories: systematic and stochastic. In systematic algorithm, the outcomes of the search is deterministic and sample the search space at predefined intervals. Stochastic search methods have make random changes to the state variables until it met the user defined termination, hence, the outcome of the search varies. Systematic search algorithms are commonly used in rigid protein–rigid protein docking whereas stochastic search algorithms are more suitable for flexible ligand–protein docking.[40] Programs like DOT[41], GRAMM[42,43] and ZDOCK[44] used systematic search algorithms. Search algorithms can also be classified based on how broadly they have explored the search space, as either local or global. Local search methods search for the nearest or local minimum energy in the current conformation, whereas global methods tend to search the best or global minimum energy within the defined search space. There is another search method so called hybrid global–local search method which have been shown to perform even better than global methods alone, being more efficient and able to find lower energies.[45] AutoDock 4 uses two local search methods (Solis and Wets[46] and Pattern Search[47]), two global search methods (Monte Carlo (MC) simulated annealing (SA) [48]

and the genetic algorithm (GA) [49-51] and one hybrid global–local search method (the Lamarckian GA (LGA)).[45]

The purpose of the scoring procedure is to identify the correct binding pose by its lowest energy value and the ranking of protein-ligand complexes according to their binding affinities.[52] Scoring function usually involves simple energy calculations such as electrostatic, van der Waals, ligand strain etc. and more accurately estimated the free energy of binding (ΔG). Scoring functions are based on empirical, knowledge or molecular mechanics force fields.[53] In addition, some docking strategies have used one scoring function at the time of docking and a different one after docking to rerank the results; such retrospective scoring, however, doesn't affect the efficiency and accuracy of the primary scoring function. [54] The AutoDock scoring function is generally based on the molecular mechanics force field AMBER.[55]

2.1.4.2 Application of molecular docking

Molecular docking has a wide variety of uses and applications in drug discovery process which include structure–activity studies, drug discovery (lead optimization), virtual screening (hit identification), bioremediation, prediction of KA (biological activity), binding site identification (blind docking), protein-protein interaction or protein-nucleic acid interaction and enzyme reaction mechanism. The main goal of molecular docking is to predict the biological activity of a given ligand. A binding interaction between a drug molecule and an enzyme may result in activation or inhibition of the enzyme function, hence it is widely accepted that drug activity is obtained through the molecular docking method.

2.1.4.3 Summary

Molecular docking is a key tool which analyzes the structural molecular biology and computer-assisted drug design. Predicting the predominant binding modes of a ligand in the binding site of a protein of known three-dimensional structure is the main goal of ligand–protein docking. A successful docking method search high-dimensional spaces effectively and use a scoring function which ranks the docking candidates correctly. Docking can be used to perform virtual screening on large libraries of

compounds, rank the results and proposed structural hypotheses of how the ligands inhibit the target one.

References

1. L. L. Wang, A. Perera, H. P. Cheng, *Phys. Rev. B*, 2003, **68**, 115409.
2. A. D. Boese, A. Chandra, J. M. L. Martin, D. Marx, *J. Chem. Phys.*, 2003, **119**, 5965-5980.
3. C. Kobayashi, K. Baldridge, J. N. Onuchic, *J. Chem. Phys*., 2003, **119**, 3550-3558.
4. G. L. Penna, A. Mitsutake, M. Masuya, Y. Okamoto, *Chem. Phys. Lett*., 2003, **380**, 609-619.
5. R. L. C. Vink, G. T. Barkema, *Phys. Rev. B*, 2003, **67,** 245201.
6. W. Polak, A. Patrykiejew, *Phys. Rev. B*, 2003, **67**, 115402.
7. J. Quintana, E. C. Poire, H. Dominguez, J. Alejandre, *Mo. Phys*., 2002, **100**, 2597-2604.
8. F. Aussenac, M. Laguerre, J. M. Schmitter, E. J. Dufourc, *Langmuir*, 2003, **19**, 10468-10479.
9. A. Rockenbauer, A. Gaudel-Siri, D. Siri, Y. Berchadsky, F. L. Moigne, G. Olive, P. Tordo, *J. Phys. Chem*., 2003, **107**, 3851-3857.
10. D. R. Hartree, *Proc. Camb. Philos. Soc*., 1928, **24**, 89-110.
11. L. H. Thomas, *Proc. Camb. Phil. Soc*., 1927, **23**, 542–548.
12. P. Hohenberg, W. Kohn, *Phys. Rev*., 1964, **136**, B864-B871
13. W. Kohn, L. J. Sham, *Phys. Rev*., 1965, **140**, A1133-A1138.
14. P. Sule, O. V. Gritsenko, A. Nagy, E. J. Baerends, *J. Chem Phys*., 1995, 103, 10085-10094
15. O. V. Gritsenko, R. V. Leeuwen, E. J. Baerends, *J. Chem. Phys*., 1996, 104: 8535-8545.
16. J. C. Slater, *The self-consistent Filed of Molecular and Solids in Quantum Theory of Molecular and Solids*, Vol. 4, McGraw-Hill: New York, 1974.
17. R. V. Leeuwen, E. J. Baerends, *Phys. Rev. A*, 1994, **49**, 2421-2431.
18. A. Becke, *J. Chem. Phys*., 1992, **96**, 2155-2160.
19. C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785-789.
20. J. P. Perdew, Y. Wang, *Phys. Rev. B*., 1986, **33**, 8800-8802.
21. J. P. Perdew, Y. Wang., *Phys. Rev. B*, 1992, **45**, 13244-13249.
22. R. P. Iczkowski, J. L. Margrave, *J. Am.Chem. Soc*., 1961, **83**, 3547-3551.

23. R. G. Parr, R. G. Pearson, *J. Am. Chem. Soc*., 1983, **105**, 7512-7516.

24. R. G. Parr, W. Yang, *Density Functional Theory of Atoms and Molecules*, Oxford University Press: Oxford, U.K., 1989.

25. A. T. Koopmans, *Physica*., 1993,**1**,104-113.

26. R. G. Parr, L. V. Szentpaly, S. Liu, *J. Am. Chem. Soc*., 1999, **121**, 1922-1924.

27. R. G. Parr, W. Yang, *J. Am. Chem. Soc*., 1984, **106**, 4049-4050.

28. W. Yang, W. Mortier, *J. Am. Chem. Soc*., 1986, **108**, 5708-5711.

29. W. Yang, R. G. Parr, *Proc. Natl. Acad. Sci. U.S.A.* 1985, **82**, 6723-6726.

30. R. K. Roy, S. Krishnamurti, P. Geerlings., S. Pal, *J. Phys. Chem. A*, 1998, **102**, 3746-3755.

31. S. Dapprich, I. Komaromi, K. S. Byun., K. Morokuma, M. J. Frisch, *J. Mol. Struct. (THEOCHEM)*, 1999, **1**, 461-462.

32. M. Svensson, S. Humbel, K. Morokuma, *J. Chem Phys* 1996, **105**, 3654-3661.

33. P. B. Karadakov, K. Morokuma, *Chem. Phys. Lett*., 2000, **317**, 589-596.

34. T. Vreven, K. Morokuma, *J. Comput. Chem*., 2000, **21**, 1419-1432.

35. S. Humbel, S. Sieber, K. Morokuma, *J. Chem. Phys*, 1996, **105**, 1959-1967.

36. M. Svensson., S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber, K. Morokuma, *J. Phys. Chem*., 1996, **100**, 19357-19363.

37. R. D. J. Froese, K. Morokuma, *In Encyclopedia of Computational Chemistry*, Vol. 2; P. V. R. Schleyer, N. L. Allinger, P. A. Kollman, T. Clark, III H. F. Schaefer, J. Gasteiger, P. R. Schreiner, Eds.;Wiley: Chichester, 1998; pp 1244-1257.

38. T. Lengauer, M. Rarey, *Curr. Opin. Struct.Biol*., 1996, **6**, 402-406..

39. Molecular docking, *Wikipedia the free encyclopedia.*

40. G. M. Morris, M. L.-Wilby, *Methods in Molecular Biology*, 2008, **443**, 365-382.

41. L. F.Ten Eyck, J. Mandell, V. A. Roberts, M. E. Pique, *Surveying Molecular Interactions with DOT*, 1995.

42. I. A. Vakser, *Proteins*, 1997, **1**, 226–230.

43. A. Tovchigrechko, I. A.Vakser, *Nucleic Acids Res*, 2006, **34**, W310–W314.

44. R. Chen, Z. Weng, *Proteins*, 2002, **47**, 281–294.

45. G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Comput. Chem*, 1998, **19**, 1639–1662.

46. F. J. Solis, R. J. B. Wets, *Math. operations. res*., 1981, **6**, 19–30.

47. A. R. Conn, N. I. M Gould, P. L. Toint, *SIAM J. Nu. Anal*., 1991, **28**, 545–572.

48. S. C. D. Kirkpatrick, J. Gelatt, M. P. Vecchi, *Science*, 1983, **220**, 671–680.

49. J. H. Holland, *Adaptation in natural and artificial systems*, 1992, Cambridge, M A: The MIT Press. 211.

50. D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning,*1st ed.1989, Boston, MA: Addison-Wesley Longman Publishing Co., Inc. 372.

51. Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Program*, 3rd ed. 1996, London, UK: Springer-Verlag. 387.

52. I. Muegge, M. Rarey, *Rev. Comput. Chem*., 2001, **17**, 1-60.

53. R. D. Taylor, P. J. Jewsbury, J. W. Essex, *J. Comput. Aided. Mol. Des*., 2002, **16**, 151–166.

54. V. Mohan, A. C. Gibbs, M. D. Cummings, E. P. Jaeger, R. L. DesJarlais, *Curr. Pharm. Des*., 2005, **11**, 323–333.

55. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, J. Merz, M. Kenneth, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc*., 1995, **117**, 5179–5197.