Review on Molecular Modeling and Docking

Sinem Gürkan Aydın

İstanbul Gelisim University, Department of Faculty of Applied Sciences, Department of Aircraft Maintenance and Repair, İstanbul, Turkey

Abstract **- Molecular modeling is used effectively in the fields of physics, biochemistry, molecular biology, drug design and interaction, synthesis of polymers and high quality new substances, both in research and industrial terms. Molecular modeling methods are of great importance in research, especially in theoretical studies. In this study, information is given about molecular modeling methods, molecular docking and molecular docking programs.**

Keywords **– olecular modeling, molecular docking, molecular methods.**

1. Introduction

Spectroscopy is an experimental science that studies the interaction of electromagnetic waves with matter. The interaction is observed in many different ways, such as absorption, emission, reflection or scattering of electromagnetic waves, due to the quantized nature of the energy structures of materials. Spectroscopic observations are central to the development of the theory of quantum physics. Important explanations of the quantum theory of Max Planck's Blackbody Radiation, Albert Einstein's Photoelectric Effect and Niels Bohr's Atomic Structure were only possible with the interpretation of spectroscopic observations.

DOI: 10.18421/SAR54-06
https://doi.org/10.18421/SAR54-06

Corresponding author: Sinem Gürkan Aydın,
İstanbul Gelisim University, Department of Faculty of Applied Sciences, Department of Aircraft Maintenance and Repair, İstanbul, Turkey.
Email: sgurkan@gelisim.edu.tr

Received: 26 September 2022.
Revised: 26 November 2022.
Accepted: 02 December 2022.
Published: 26 December 2022.

The difference between the levels in the energetic structure of atoms, molecules and ions is directly related to which wavelengths in the electromagnetic spectrum they interact with photons. The interaction of quantized energy levels and photons in spectroscopy is based on the concept of resonance.

In a quantum system (atom, molecule, crystal, ion, and etc.), resonance can occur when the difference between the energies of two stable quantum levels and the energies of the oscillating electromagnetic waves are of the same magnitude, and thus the system transitions from one stable level to another. However, these transitions are not possible between all quantum levels. The quantum mechanical selection rule determines whether a spectroscopic transition between levels is possible.

When it is desired to obtain the spectra of atoms, molecules and ions by using sources of different wavelengths, the size of the difference between the quantum energy levels in that system determines which spectroscopic technique should be used.

Today, spectroscopic techniques are used intensively to determine the chemical and physical properties of materials in many branches of science, from analytical chemistry to astrophysics, from remote sensing to drug design. Some of these techniques are listed below. This ranking is not a complete list, but perhaps the most widely used or best known techniques. Each spectroscopy technique in the list contains many variations within itself. Spectroscopic techniques and their interactions with quantum energy levels; Visible-Ultraviolet Absorption and Molecular Fluorescence Emitting Spectroscopy when it comes to studying the transitions of electrons between orbitals in molecules, Infrared (IR) Absorption and Raman Scattering Spectroscopy when it comes to examining the transitions between vibrational energy levels of chemical bonds of molecules.

Microwave Absorption Spectroscopy is used when it comes to studying the transitions between rotational energy levels of chemical bonds of molecules. Nuclear Magnetic Resonance (NMR) Spectroscopy is particularly useful in the nuclei of some atoms (H, C13, F19, Cl35) in examining the transitions between spin energy levels. If there are

unpaired electrons in molecular orbitals, Electron Spin Resonance (ESR) Spectroscopy is a suitable technique when it comes to examining the transitions between electron spin energy levels. Auger Electron Emitting Spectroscopy is a suitable technique when researching the electronic structure of the material surface through the study of electrons ejected from a surface.

In addition to the analytical applications of these experimental spectroscopic techniques in many fields, it has made many positive contributions to the development of the theories of basic sciences such as physics, chemistry, astronomy and materials science. Spectroscopic techniques have become an indispensable tool in the development of basic sciences, when theoretical studies are at the forefront, and when the proposed theory needs to be supported by experimental evidence. Sometimes, the scientific explanation of an experimental observation was made possible by a later proposed theory. As an experimental science, spectroscopy is closely associated with theories to explain the structure of matter and its interaction with electromagnetic waves.

Immediately after the development of quantum theory, the laws of quantum mechanics began to be applied to atoms and molecules. In principle, all chemical properties of a molecule can be calculated with quantum theory. In fact, the structure and chemistry of a compound can be determined by experimental methods. However, making predictions by computation is very useful and has found many applications. It is widely used, for example, in the development of new drugs in pharmacology.

Molecular modeling software is very helpful to physicists and chemists. Molecules can be viewed from different angles by rotating them on the computer screen, their geometric and isomeric structures can be determined, and their energies can be determined by means of these programs. IR, UV, NMR spectra can be drawn, MO diagrams can be obtained.

Theoretical chemistry describes chemistry with mathematical methods. It tries to explain chemical structures and reactions based on the basic laws of physics. Computational chemistry, on the other hand, it applies mathematical methods developed by theoretical chemists and interprets the results. Thus, it bridges the gap between experimental chemistry and theoretical chemistry. The same is true for the department of atomic molecule physics.

With the computational field, it is possible to study not only stable molecules, but also short-lived unstable intermediates and transition states. In this way, information about molecules and reactions that cannot be obtained through observation is obtained. The qualitative or quantitative results obtained with these calculations provide very useful predictions.

Options for researchers who will use computational methods to support experimental studies or to predict the results to be obtained without experimental studies are explained in this study.

There are many possible methods and applications of molecular modeling: structure formation, structure visualization, conformation analysis, derivatization of bioactive conformations, molecular superposition and superposition, derivatization of pharmacophoric patterns, receptor mapping, prediction of inhibition activities, and molecular interactions. Molecular modeling is used in areas such as docking, calculation of molecular properties, energy calculations.

1.1. Hartree-Fock (HF) Method

The Hartree-Fock method is the most common Ab Inito calculation used to determine the state functions of multi-electron atoms and also to study the structure of the atom. This method is an iterative variation method used to approximate the total wave function. In the Hartree-Fock method, which is based on the central area approximation, the potentials of the electrons are determined only by their distance from the nucleus.Since it is thought that any electron will move within the mean spherical potential field created by its nucleus and other electrons, the Coloumb electron-electron repulsion potential is initially not taken into account and then taken into account as electron correlation [1].

1.2. Density Functional Theory (DFT)

Density Functional Theory (DFT) is a quantum mechanics-based method widely used in the study of the electronic structure of a multi-particle system. DFT calculations determine properties energetically, vibrationally and structurally. On the other hand, it predicts the electronic, magnetic and optical properties of solid phases by DFT method. It can be studied on many systems and processes (chemical reactions, restructuring of the surface, and absorption of molecules) from small molecules to complexes, periodic and amorphous solids. The results obtained with DFT depend on the variation potential used. This has an important conceptual difference with the HF-based method.

Since the simpler electron density is taken into account instead of the complex n-electron wave function (Ψ) in DFT, the calculation results are faster than the ab initio methods. According to DFT, the total energy of a system is a functional of electron density [2].

DFT has common aspects with the Hartree-Fock (HF) method. In DFT, the total electron density is decomposed into one-electron densities consisting of one-electron wave functions. These one-electron wavefunctions are similar to the wavefunctions of HF theory and require a molecular orbital (MO) descriptor compared to the DFT, HF approach for molecular systems [3].

1.3. B3LYP Mixed Density Function Theory

HF theory based on wave mechanics does not work well for exchange correlation but is suitable for kinetic energy. The correlation and exchange energies are better calculated with DFT models. Thus, instead of pure DFT and pure HF models for full energy expression, mixed models were formed as a result of using the energy expressions of both in the total electronic energy expression. These models calculate many magnitudes such as bond lengths, ionization energies and total energies better than pure models [4], [5].

1.4. Gaussian 09 and Gaussian View Programs

Gaussian 09 is a package program that is used to calculate quantum chemical values such as force fields, molecular geometry, HOMO, LUMO, IR, NMR intensities related to a molecular. In the program, calculations are made using empirical, semi-empirical and ab-initio methods. The 'GaussView' program, which allows the properties of any molecule to be defined visually and to make changes on these values, and to start calculations by creating input data, is also working with this program.

The GaussView program visualizes the calculated results for a molecule studied in the Gaussian program. It is a program that is used to visualize the wave numbers and modes obtained as a result of the calculation. When using the program, it is necessary to determine a theory level first. There are many levels of theory in the Gaussian 09 program. The most commonly used ones with their abbreviations are given below.

Abbreviation Methods

HF - Hartree Fock Self Harmonious Field Theory
B3LYP - Becketype3-parameter Density Function Theory
MP2 - 2nd order Moller Plesset Perturbation Theory
MP4 - 4th order Moller Plesset Perturbation Theory
QCISD(T) - 2nd order Configuration Interaction

2. Molecular Docking

In the field of molecular modeling, docking is a method that predicts the preferred orientation of a molecule when it binds to a second molecule, and when it binds together to form a stable complex. The preferred orientation information can be used to predict the coupling or binding affinity between two molecules. Molecular docking is widely used in multidisciplinary fields such as biology, physics, chemistry, and the pharmaceutical industry. It plays an important role in signal transduction, especially between related molecules such as lipids, nucleic acids, peptides and proteins. In addition, the type of signal produced varies depending on the two interacting partners. Therefore, docking is useful for estimating the type and strength of the generated signal.

Molecular docking is one of the most commonly used methods in structure-based drug design because of its ability to predict the binding-conformation of small molecule ligands to the appropriate target binding site. Characterization of binding behavior plays an important role in the rational design of drugs and elucidation of basic biochemical processes.

Docking is a method used to predict the preferred positions of atoms in space as a result of a molecule forming a stable complex with another molecule.

One of the main areas in which the docking method is used is to examine the nature of the binding of potential drugs to target molecules. Docking, which is a molecular mechanics method, has two types of applications as small molecule-protein and protein-protein docking. It is used to find the relationships between molecules of great and biological importance such as proteins, nucleic acids and fats.

Protein-protein docking involves the binding/interaction of two protein molecules simulated by computer programs.

In these cases, the interactions are rigid compared to ligand-protein docking. As for ligand-protein docking; the ligand is free, the protein (receptor) is stable (rigid), and both ligand.

It can also be examined by considering the cases where the protein is flexible. In addition, ligand DNA and RNA docking studies are carried out by some researchers today with the docking method, it is possible to categorize this part as ligand-DNA/RNA docking with further studies in this field [6], [7].

To understand the docking process, the key-lock analogy used to describe molecules that bind to proteins can be used [8]. Here, the protein can be thought of as the key and the ligand to be bound as the key. In the same way, the possibility of binding the substrate or inhibitor to the appropriate gaps in the enzyme is examined in the process.

In short, the docking method is the simulation of this process in the computer environment. What is desired is the position where the key has to stop to open the lock. When identifying these locations, the binding (or active) site of a protein is the site where the ligand will bind to the protein. This region is usually a gap on the protein surface and is located where the inhibitor is bound in proteins with known crystal structure.

However, when defining the binding site of protein-ligand docking, a bond is not mentioned; preferably such as the Ligand's position (position) or binding mode.

It is better to use definitions. Because it is a term that expresses the geometric pose of the ligand's conformation, orientation and position, which is mentioned as the binding site of the ligand. Protein and ligand, which can move continuously, will stay in their lowest potential energy positions as a result of these movements.

To calculate it, an optimization process is used, just as in nature. By changing the position of the ligand relative to the protein, its energy is calculated and this process is continued until the lowest energy is found. In other words, the aim is to find the conformation with the lowest free energy of the protein-ligand system.

2.1. Molecular Docking Programs

Much software is available to perform the molecular docking process. The software used is listed below.

2.1.1. Autodock

Autodock software is software that predicts the binding of small molecules to receptors with a 3D structure. Autodock is free software for academic use.

2.1.2. Gaussian

It is Gaussian software that can estimate energy, molecular structure and vibrational frequencies of molecular systems based on basic types of computation. It can be used to understand molecules, reactions, and short-lived transition molecules that are difficult or impossible to observe experimentally. Gaussian is paid software for academic use.

2.1.3. Autodock Vina

A multi-core capability option used for molecular docking, drug discovery and virtual screening is Autodock Vina.

Autodock Vina can be used in Windows XP and later versions. However, in order to run Autodock Vina program in Windows environment, a simulator program called Cygwin is required. Autodock Vina is also available for Linux and Mac platforms.

AutoDock is easy-to-use software with high performance, improved accuracy, using quasi-experimental free energy force fields, the interactions of inhibitors with pdb formats obtained from x-ray structures of biomolecules consisting of protein and nucleic acid or prepared with computer aided programs. Autodock Vina was designed and developed by Dr.Oleg Trott. It is free software for academic use.

With the auxiliary software developed for Autodock, Vina software provides ease of use in terms of determining the area that can dock, preparing files and examining the results. In addition, there is no need for manual selection of atom types for grid maps, calculation of grid map files with AutoGrid, and clustering of results after docking; because Vina does them herself automatically. Without exposing the user, he can also take the results himself [9].

It is a molecular docking program that calculates the energies of conformations, that is, used to predict the strength of coupling or bonding affinity between two molecules. In AutoDock calculations, respectively; Coordinate files are prepared using AutoDockTools, Affinity and inhibition activity of small molecule using AutoGrid, A preliminary calculation is made to determine, Docking is performed with AutoDock, Analyzes of results with AutoDockTools.

3. Conclusion

Molecular modeling is the spectroscopic, thermodynamic, biological, catalytic, and dynamic of structures that arise from the combination of atoms, molecules, ions and crystals alone or together, in the computer environment, including calculations based on quantum and classical mechanics theories. It refers to the whole of the computational simulation techniques used to determine its properties. These techniques are widely used in many fields such as computational chemistry, structural biology, drug design, and materials science.

Molecular modeling, structural, spectroscopic, thermodynamic, biological, catalytic, dynamical, etc. of atoms, molecules, ions and crystals in computer

environment, including calculations based on quantum theory. Molecular modeling enables quantum theory-based calculations of atoms, molecules, ions, and crystals. These calculations include spectroscopic, thermodynamic, biological, catalytic, dynamic, etc. example can be given. It refers to the whole of the computation / simulation techniques used to determine the properties. These techniques are widely used in many fields such as computational chemistry and biology, drug design, materials science. Its application range covers a wide range from calculating the IR spectrum of a triatomic (suppose water, H2O) molecule to modeling the dynamics of conformational change in the domain movement of a biological macromolecule (protein) with 50,000 atoms.

In molecular modeling, the molecular system is usually considered at the atomic level. In cases where the system contains too many atoms (e.g. protein), there is also a Coarse-Grained model approach, where the alpha carbon atom in each amino acid in the system is taken into account, but not every single atom.

If the sum of all the charges in the nuclei and electrons of the atoms in a system is considered as a point charge and their interactions with each other are modeled with a potential function, this is called the Molecular Mechanics approach. The potential function is the sum of several terms. The interaction between neighboring atoms (vibration, bending, and rotation) is modeled by the mass spring system (chemical bond) and the Lennard-Jones potential. Electrostatic interactions are modeled by the Coulomb potential. IR and Raman spectroscopy are frequently used to determine the chemical bond constants used in the models.

If the effect of nuclei and electrons on the interactions between atoms in the modeled system is taken into account one by one, it is called the Quantum Chemistry approach. Quantum chemistry calculations include many sub-approaches within themselves. If calculation time is saved by using some empirical data during calculation, it is called Semi-empirical approach. If the calculations are done without using experimental data, this can be called Ab Inito or the first principle approach. These approaches have many variations within themselves.

Molecular Dynamics Simulation occupies an important place among the modeling techniques. The dynamics of a molecular system is based on the numerical solution of the integration of Newton's 2nd law with respect to time by means of an assigned molecular dynamics potential function. This modeling technique is widely used in solving many problems, especially in the field of materials science and biophysics.

References

[1]. Köksal, F., & Gümüş, H. (1999). Atom ve molekül fiziği. *Ondokuz Mayıs Üniversitesi Yayınları, Samsun*.

[2]. Hohenberg, P., & Kohn, W. (1964). Inhomogeneous electron gas. *Physical review*, *136*(3B), B864.

[3]. Çiçek, E. (2016). *Histon deasetilaz (HDAC) inhibitörü olan bazı hidroksamik asit türevlerinin bağlanma özelliklerinin moleküler kenetlenme yöntemiyle teorik olarak incelenmesi* (Master's thesis, Recep Tayyip Erdoğan Üniversitesi/Fen Bilimleri Enstitüsü/Kimya Anabilim Dalı).

[4]. Becke, A. (2007). *The quantum theory of atoms in molecules: from solid state to DNA and drug design*. John Wiley & Sons.

[5]. Cramer, C. J. (2004). *Essentials of Computational Chemistry: Theories and Models*. John Wiley & Sons.

[6]. Güner, O. F. (Ed.). (2000). *Pharmacophore perception, development, and use in drug design* (Vol. 2). Internat'l University Line.

[7]. Thiriot, E., & Monard, G. (2009). Combining a genetic algorithm with a linear scaling semiempirical method for protein–ligand docking. *Journal of Molecular Structure: THEOCHEM*, *898*(1-3), 31-41.

[8]. Jorgensen, W. L. (1991). Rusting of the lock and key model for protein-ligand binding. *Science*, *254*(5034), 954-955.

[9]. Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, *31*(2), 455-461.