Eco. Env. & Cons. 28 (January Suppl. Issue) : 2022; pp. (S127-S132)
Copyright@ EM International
ISSN 0971–765X

http://doi.org/10.53550/EEC.2022.v28i01s.018

Molecular Docking; future of Medicinal Research

Anuradha Sharma, Sahil Ahuja, Pragya Deep, Shravani, Saranya Nair, Sneha Sambhyal, Deependra Mishra, Chetan Pandey, Preet Manchanda, Krishma, Akash Deep, Lamha Kumar, Parveen Gwalia, Ria Arora, Bhupender Singh, Shubham Attri, Deepika Kumari Singh, Areeba, Muskaan Gupta and *Vivek Chopra

Department of Botany, Hindu College, University of Delhi, Delhi 110 007, India

(Received 8 April, 2021; Accepted 26 August, 2021)

ABSTRACT

Molecular docking is a Molecular modelling technique that is used to predict the interaction between two molecules such as drugs, enzymes or proteins. It predicts the structure of the interacting molecules using computational modelling. The objective is to obtain plausible three-dimensional structures of the molecules under study. The candidates produced by docking are ranked by various methods to identify the most likely naturally occurring structure. This review encompasses the various types of docking models and the mechanism of docking, drug design types and available docking software.

Key words: *Molecular docking, Drug design, Docking software, Docking mechanism*

Introduction

Molecular targets and their binding mechanism are well studied to see how perfectly they fit together. Molecular Docking is basically a technique to study how different molecular structures (ex- Drug, enzyme, or protein) best fit with each other forming a stable complex. It basically predicts the preferred orientation of one molecule with respect to the other molecule (Roy *et al.,* 2015). The biologically important molecules play a significant role in signal transduction (Walters *et al.,* 1998). This technique is frequently used in drug designing which helps Scientists to control different biochemical processes occurring and study their effect and functioning (Kitchen *et al.*, 2004). One can relate this technique with lock and key, where the lock can be opened only when there is a perfect key to its orientation. Here we can assume "protein" like lock and "ligand" as a key that would only bind to a particular protein of interest (Gohlke *et al.,* 2002).

The recent development of technology-oriented studies has helped us deeply analyze the mechanism of action of drugs and Molecular Docking is an innovative approach that is also used in recent studies of forensic sciences. A detailed understanding of the basic principle and types of molecular docking is discussed in this research paper. We have analyzed various aspects and taken into account some noteworthy features that are going to play a role in the drug industry.

This review highlights the major key concepts like the present technologies, the software, approaches, and future prospects.

Molecular Docking Approaches

Currently there are many molecular docking approaches in practice. Among these approaches two of them are particularly popular.

1. Shape complementarily

This approach uses a set of features such as molecular surface/complementary surface descriptors. The shape matching description determined by the

complementarily between receptor and Ligand surfaces helps in finding the complementary pose of docking receptor and Ligand molecules.

Simulation Approach

In this approach the ligand and the protein receptor are physically separated by some distance. The ligand binds to the protein's active site after a number of "moves" in the conformational space. These "moves" by ligand incorporates body transformations and rotations, along with the internal changes in the ligand's structure such as torsion angle rotations. Thus each "move" by ligand in the conformational space brings about a total energetic cost of the system. With every such "move" the energy of the system is calculated, (as every move in the conformational limit induces a total energy cost). After calculation of energetic costs with respect to each move, the optimal pose of binding is evaluated (Gaba *et al.,* 2010).

More specific classification of molecular docking approaches is as follows:

(1) Blind docking: This approach involves scanning of the entire surface of protein receptors. It is used for detection of possible modes and binding sites for the ligand.
(2) Distance geometry: The intermolecular or intermolecular distances can be helpful in deriving various structural information. This approach assembles these distances and accordingly calculates the three-dimensional structures.
(3) Fragment based method: This method involves dividing the ligand into fragments or separate protons. These fragments are finally linked after the docking process.
(4) Inverse docking: This approach is used in assessment of potential side effects or toxicity of the drug candidate. Knowledge of the protein targets when combined with proteomics and pharmacokinetics facilitates the assessment of toxicity/side effects of the drug.
(5) Ligand fit approach: It provides accurate and rapid approach for docking small molecule ligands into active sites of the protein, for considering shape complementarity between them.
(6) Matching approach: In this approach, ligand molecule is placed at the most favourable position in the active site of the protein. This may further require optimization.
(7) Monte Carlo approach: This approach focuses on generating different configurations of ligand in the active site. Each configuration is scored. At each step, if the new configuration scores better, then it is accepted against the previous one (Metropolis criterion).
(8) Point complementarity approach: This approach is based on evaluation of chemical and/or shape complementarity between the ligand and protein (Mani and Shankar, 2016).

Types of Molecular Docking

(1) Rigid docking

In this docking, the receptor as well as the ligand to be attached is assumed to be rigid. It is often referred as, "lock and key", as there are no conformational changes in both, i.e. the receptor molecule and the ligand molecule.

(2) Semi-flexible docking

One of the binding molecules is rigid and the other one is flexible and it often undergoes conformational changes to fit in the receptor.

(3) Flexible docking

In this type of docking, both receptor and ligand are assumed as flexible entities. Both ligand and the receptor undergo conformational changes in their shape to fit in for docking. This type of docking often referred to as induced-fit (Mani and Shankar, 2016; Salmaso *et al*., 2018; Chaudhary *et al*., 2016).

Requirements for Molecular Docking

The basic requirements for molecular docking are: - structure of a target protein, all the molecules of interest to perform docking and suitable database which has ligand and receptor or virtual compounds for docking. There is also requirement of a computational framework to allow implementation of derived docking and scoring procedure. Usually, the protein is assumed as rigid and ligand as flexible one. One more thing to be considered is that with conformational degree of freedom, the binding pose in protein's binding pocket must be considered. This can be done by placing rigid molecules or fragments into the active site of the proteins which is done by many approaches like clique search, geometric hashing or pose clustering.

Ligand Representation

First for this the structure that is better working at neutral pH is generated. Further it is adjusted by

adding or removing hydrogen of approximate pKa value. The thing which is needed to be considered is that accurate atom typing is there otherwise this can lead to serious docking error.

There may be the case where the stereochemistry of the compound which is synthesized is not known, in such cases it is beneficial to have all possible diasetreomers and dock them individually to the appropriate receptor. The software available for such case be –pipeline pilot (Hopkins *et al*., 2000), stergen (Tondi *et al.,* 1999) and stereoplex (Aronov *et al.,* 2000).

Receptor Representation

The quality of the receptor structure used plays an important role in determining the success of shipping. In general, high-resolution crystal structure used gives better results. Using ICM enemy docking has shown that they can produce known ligand binding methods to within 1x 10-10 binding bonds in cases where agreement crystal structure structures were better than 2.0x 10-10. The latest revision of the accuracy, limitations and pitfalls of the structural algorithms for protein ligand complexes generally provided critical analysis of available structures. The importance of relying on the pH of the ligand binding was highlighted earlier. The reliability of the ligand structures found in co-complexes has also been questioned. Even at high resolution the complexity of the definition of ligand positions inexplicably can be caused by differences between high-frequency dictionaries of the band, the angle and the available torsions of proteins and nucleic acid structures and those found in small molecules. Ignoring potential failures, there has been significant success reported in shipping studies using X-ray receptor structures. Recent examples of this type of study include: kinesin (Hopkins *et al*., 2000), thymidylate synthase (Tondi *et al.,* 1999), phosphoribosyltransferase (Aronov *et al*., 2000), HIV protease (Olson and Goodsell, 1998) and beta-lactamase (Tao *et al*., 2020).

Steps to Perform Docking

Step 1- building the receptor
The 3-dimension structure of the protein should be retrieved which can be download from protein data bank (PDB). Purify the protein by removing the water molecules and all the other molecules. The receptor should be biologically active and stable.

STEP 2- identification of the active site

The active site within the receptor should be identified. The receptor may have many active sites but the one of the interests should be selected.

STEP 3- ligand selection and preparation

Ligand can be obtained from various databases like Zinc, Pubchem or can be sketched using tools like Chemsketch.

Mechanics of Docking

The aim of molecular docking is to give a prediction of the conformation of ligand receptor complex structure using computational method. To accurately carry out docking one requires the high resolution achieved in two steps-one is "sampling algorithm" of ligand in the active site of the protein and second is "scoring function".

Sampling Algorithm

The first docking algorithm was developed in 1980s (Kuntz *et al.,* 1982); the receptor was approximated by a series of spheres filling its surface cleft, and the ligand by another set of spheres defining its volume. With six degrees of translational and rotational freedom as well as the conformational degrees of freedom of both ligand and protein, there are a huge number of possible binding modes between two molecules.

Various sampling algorithm have been developed and widely used in molecular docking software are molecular algorithm, incremental construction, MCSS (multiple copy simultaneous search), LUDI, genetic algorithm and molecular docking.

Scoring Function

The purpose of the scoring function is to delineate the correct poses from incorrect poses, or binders from inactive compounds in a reasonable computation time. However, scoring functions involve estimating, rather than calculating the binding affinity between the protein and ligand and through these functions, adopting various assumptions and simplifications. Scoring function is physical phenomenon i.e., entropy and electrostatic interactions are disregarded in scoring schemes (Dar *et al.,* 2017). Popular scoring functions have an adequate balance between accurate estimation of binding energy and computational cost in terms of time. Scoring functions have also been developed to predict the strength of intermolecular interactions between two proteins or between protein and DNA (Robertson and Varani, 2007). Scoring functions rely on statisti-

cal means to extract rules on preferred, and no preferred, atom pair interactions from experimentally determined protein-ligand complex.

Scoring functions can be divided in 3 major classes

1) Force field scoring function- affinities are estimated by summing the strength of intermolecular Vander Waals and electrostatic interactions between all atoms of the two molecules in the complex using a force field. The intramolecular energies (also referred to as strain energy) of the two binding partners are also frequently included.
2) Empirical scoring functions: the basis of this scoring function is that the binding energies of a complex can be approximated by a sum of individual uncorrelated terms. The coefficients of the various terms involved in calculation of binding energy are obtained from regression analysis using experimentally determined binding energies or potentially from X-ray structural information.
3) Knowledge based scoring functions: these are derived from the structural information embedded in experimentally determined atomic structures.

Applications of Molecular Docking

This technique is very useful in assisting different tasks of drug discovery programs like Hit identification and optimization. It is also used in drug repositioning and also in the case of multi-target ligand design and repositioning. Docking also helps to understand the relationships between different molecular targets involved in a given disease that can be high relevance for polypharmacology and modern drug discovery.

Moreover, it is useful in identifying protein binding sites in which ligands

Table 1. Software available for Molecular Docking

S. No.	Software	Company / Institution	License terms	Docking approach	Supported Operating System	Scoring function	Evaluation method	Accuracy (%)
1	Lig and Fit	Accelrys Inc.[1]	Commercial[1]	Monte Carlo Sampling [1]	Linux, SGI, IBM AIX [1]	Lig Score, PLP, PMF [1]	The molecular force field [2]	46
2	Dock	I. Kuntz University of California, San Francisco [1]	Free for Academic use [1]	Shape fitting (sphere sets)[1]	Unix, Linux, Sun, IBM AIX,	Mac OSX, Windows[1]	Chem Score, GB/SA solvation scoring, other [1]	The molecular force field [2]
3	Gold	Cambridge Crystallographic Data Centre [1]	Commercial Free evaluation (2 months)[1]	Genetic Algorithm[1]	Linux, SGI, Sun, IBM, Windows[1]	Gold Score, Chem Score user-defined [1]	Semi-empirical free energy [2]	- 78
4	Glide	Schrödinger Inc.[1]	Commercial[1]	Monte Carlo Sampling[1]	Unix, Linux, SGI, IBM AIX [1]	Glide Score, Glide Comp [1]	Semi-empirical free energy [2]	82
5	FRED	Open Eye Scientific Software [1]	Free for Academic use [1]	Shape fitting (Gaussian)[1]	Unix, Linux, SGI, Mac OSX, IBM AIX, Windows[1]	Screen Score, PLP, Gaussian shape score, user-defined [1]	Cognate Docking and Virtual screening performance[3]	-
6	Flex X	T. Lengauer and M. Rarey Bio SolveIT [1]	Commercial Free evaluation (6 weeks)[1]	Incremental Construction[1]	Unix, Linux, SGI, Sun Windows[1]	Flex X Score, PLP, Screen Score, Drug Score [1]	Semi-empirical free energy[2]	58
7	Auto Dock	D. S. Good sell and A. J. Olson The Scripps Research Institute [1]	Free for Academic use [1]	Genetic algorithm Lamarckian genetic algorithm Simulated Annealing[1]	Unix, Mac OSX, Linux, SGI [1]	Auto Dock (force-field methods)[1]	Semi-empirical free energy [2]	78

Table 1 shows different software available for Molecular Docking (Chaudhary *et al.*, 2016; Tao *et al.*, 2020; McGann, 2012)

could bind and to identify novel molecular targets of known ligands it is also helpful to check for the potential adverse drug reactions (ADRs) or in simple words drug reactions. Further it can also check for the ligands with novel chemotypes active against a given target or a set of desired targets.

Docking can predict the adverse drug reactions effect. It is also widely used in the field of drug design. It is also useful for predicting protein-protein interaction. The protein-ligand docking used to predict pollutants that can be degraded by enzymes. It is possible to check the accuracy of potential drug against homologous proteins through docking. Docking can also be applied for bioremediation and also to study protein engineering (Chaudhary *et al*., 2016; Lin *et al*., 2020).

Conclusion

In this review we included molecular docking models, mechanism of docking and also focussed on drug design types and available docking softwares. Molecular docking is a user friendly and an inexpensive tool; it is used in revealing 3-dimensional structures and examines the molecular properties of molecules. Molecular docking plays an important role in drug designing and analysis. Further it has revolutionized computational methods in drug designing and screening helps in improving the efficiency and precision which is a demand of time for big data era. Molecular docking has a number of strengths, one of the ability to cover a huge database system at low cost, as compared to other laboratory techniques such as HTS (High Throughput sequencing). High Flow Capacity docking has been broadly used in library designing. In the view to facilitate scientific researches, we generate a web system that confers a complete dataset to enhance the results of docking pose. Over a few decades, molecular docking is improved and contributing to the enhancement and improvements of pharmacology. Increasing demand of molecular docking has revolutionary advancement in technology. We can improve the current molecular docking technique by unifying the large biological data by scoring function. Scoring functions are used to determine binding mode and site of ligand. Currently, there are many limitations in present scoring functions and in molecular docking such as there is no any effective method to account for the energy difference between free ligand and receptor bound. New algorithms will arise to find the solutions of coming challenges of the future, such as,

To find the most reliable and robust scoring function and to find the new solutions to docking problems. Molecular docking also is extending its role in techniques like genomics, computational enzymology and proteomics search engine.

References

Roy, K., Kar, S. and Das, R.N. 2015. *Understanding the basics of QSAR for applications in pharmaceutical sciences and risk assessment*. Academic press.

Walters, W.P., Stahl, M.T. and Murcko, M.A., 1998. Virtual screening—an overview. *Drug Discovery Today*. 3(4) : 160-178.

Kitchen, D.B., Decornez, H., Furr, J.R. and Bajorath, J. 2004. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Reviews Drug Discovery*. 3(11) : 935-949.

Gohlke, H. and Klebe, G. 2002. Approaches to the description and prediction of the binding affinity of smallmolecule ligands to macromolecular receptors. *Angewandte Chemie International Edition*. 41(15) : 2644-2676.

Gaba, M., Gaba, P. and Singh, S. 2010. An overview on Molecular Docking. *International Journal of Drug Development and Research. 2* : 219-231.

Shankar, M. 2016. A Over View on Molecular Docking. *American Journal of Biological and Pharmaceutical Research. 3*. 83.

Salmaso, V. and Moro, S. 2018. Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: An overview. *Frontiers in Pharmacology*. 9 : 923.

Chaudhary, K.K. and Mishra, N. 2016. A review on molecular docking: novel tool for drug discovery. *Databases.* 3(4).

Hopkins, S.C., Vale, R.D. and Kuntz, I.D., 2000. Inhibitors of kinesin activity from structure-based computer screening. *Biochemistry*. 39(10) : 2805-2814.

Tondi, D., Slomczynska, U., Costi, M.P., Watterson, D.M., Ghelli, S. and Shoichet, B.K. 1999. Structure-based discovery and in-parallel optimization of novelcompetitive inhibitors of thymidylate synthase. *Chemistry & Biology*. 6(5) : 319-331.

Aronov, A.M., Munagala, N.R., Ortiz de Montellano, P.R., Kuntz, I.D. and Wang, C.C. 2000. Rational design of selective submicromolar inhibitors of Tritrichomonas foetus hypoxanthine-guanine-xanthine phosphoribosyltransferase. *Biochemistry*. 39(16) : 4684-4691.

Olson, A.J. and Goodsell, D.S. 1998. Automated docking and the search for HIV protease inhibitors. *SAR and QSAR in Environmental Research*. 8(3-4) : 273-285.

Powers, R.A., Morandi, F. and Shoichet, B.K. 2002. Structure-based discovery of a novel, noncovalent inhibitor of AmpC β-lactamase. *Structure*. 10(7) : 1013-1023.

Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R. and Ferrin, T.E. 1982. A geometric approach to macromolecule-ligand interactions. *Journal of Molecular biology*. 161(2) : 269-288.

Tao, X., Huang, Y., Wang, C., Chen, F., Yang, L., Ling, L., Che, Z. and Chen, X. 2020. Recent developments in molecular docking technology applied in food science: a review. *International Journal of Food Science & Technology*. 55(1) : 33-45.

Dar *et al.*, 2017 J Anal Bioanal Tech , 8:2 DOI: 10.4172/2155-9872.1000356

Robertson, T.A. and Varani, G. 2007. An allatom, distance dependent scoring function for the prediction of protein–DNA interactions from structure. *PROTEINS: Structure, Function, and Bioinformatics.* 66(2) : 359-374.

McGann, M. 2012. FRED and HYBRID docking performance on standardized datasets. *Journal of Computer-aided Molecular Design*. 26(8) : 897-906.

Lin, X., Li, X. and Lin, X. 2020. A review on applications of computational methods in drug screening and design. *Molecules.* 25(6) : 1375.