Original Article

Molecular Docking: In Silico Approach For Structure-Based Drug Designing Discovery

Jayshri Swarnkar*[1] , Dr.Sandip Prasad tiwari[2] , Khushboo Gupta[3] , Pranjul Shrivastava[4]

[1]Assistant Professor, Faculty of Pharmacy,Kalinga University ,Naya Raipur492101 jayshree.swarnkar@kalingauniversity.ac.in
[2]Associate Professor, Faculty of Pharmacy,Kalinga University ,Naya Raipur492101 sandip.tiwari@kalingauniversity.ac.in
[3]Assistant Professor, Faculty of Pharmacy,Kalinga University ,Naya Raipur492101 khushboo.gupta@kalingauniversity.ac.in
[4]Assistant Professor, Faculty of Pharmacy,Kalinga University ,Naya Raipur492101 pranjul.shrivastava@kalingauniversity.ac.in
DOI: 10.47750/pnr.2022.13.S09.1218

Abstract

With the advancement of novel techniques in structure-based drug discovery, one of the important tool for drug discovery is molecular docking. The aim of molecular docking is to achieve an optimized conformation for both the target protein and ligand. In this review, we focus on various aspects of molecular docking including basic steps of docking, theories of docking, types of interactions, software tools used for in silico docking studies. Molecular docking study is highly relevant in order to predict potential targets of diseases as well as in designing effective drugs for pharmaceutical industry.

Keywords: docking, in silico, algorithm, protein, score

Introduction

Over the past few years, there is an increasing number of new therapeutic targets for drug discovery. Traditionally, approaches were very expensive, time consuming and inefficient to discover novel therapeutic drugs. To overcome this drawback of traditional methods, more efficient and logical methods that rely on virtual screening have been devised to address this flaw in conventional approaches (1). At the same time, due to the availability of structural data on proteins and protein-ligand complexes by techniques of chemical synthesis, purification, X-ray crystallography and Nuclear Magnetic Resonance Spectroscopy (NMR), therapeutically significant molecular targets are recognised (2).

Virtual screening involves a number of distinct computational techniques that together can be scaled down to a more manageable size. In this paper we describe the application of in silico molecular docking as a virtual screening tool for structure-based drug discovery (3). The interaction between the ligand and the target molecule is determined by the molecular docking technique. It predicts binding affinity of ligand to form a stable complex with protein by finding preferred orientation of minimal free binding energy (4). The binding involves various types of non-covalent interactions such as hydrogen bonding, ionic bonding, hydrophobic bond and van der Waal forces. Molecular docking study can be possible in between protein-protein, protein-ligand and protein-nucleotide (5).

Theory of docking

The main aim of molecular modelling in docking studies is to give a prediction of the ligand-receptor complex structure using computational methods (6). In order to perform docking, two related processes must be taken: first by sampling conformations of the ligand in the active site of the protein and next, these conformations must be ranked using a scoring function (7).

Search algorithm

This theory comprises of all possible geometric orientations and conformations between protein or receptor and ligand using search algorithm. However, each binding mode is referred to as a pose. Mostly search algorithms such as Monte Carlo (8) and Genetic algorithm are well-known stochastic methods whereas Multiple Copy Simultaneous Search (MCSS) and LUDI are fragment-based methods for the de novo design of ligands and modifications of known ligands that may enhance their binding to the target protein. Matching algorithms are based on geometry or molecular shape of a ligand which fits into an active site of a protein (9). Exhaustive methods are also used to predict different conformations of ligand.

Scoring function

The purpose of the scoring function is to determine the correct poses from incorrect poses, or binders from inactive compounds in a computational time. Scoring functions evaluate the binding affinity between the protein and ligand and through these functions, adopting various assumptions and simplifications. Scoring functions consists of force field-based, empirical and knowledge-based scoring functions, consensus scoring and descriptor based scoring functions (10).

Classical force field-based scoring functions estimate the binding energy by calculating the sum of the non-bonded interactions (electrostatic and van der Waals). electrostatic terms calculated using the Coulombic formulation. Software such as DOCK (11) , GOLD (12) and AutoDock (13) perform the function.

However, in empirical scoring functions, the binding energy is divided into several energy components, such as hydrogen bonding, ionic interaction, hydrophobic effect, and binding entropy. Each component is multiplied by a factor of and then added to get the final score. The coefficients were obtained from a regression analysis fitted to a test set of ligand-protein complexes with known binding affinities.

Knowledge-based scoring functions are evaluated using statistical analysis of ligand-protein complexes as crystal structures to obtain the interatomic contact frequencies or distances between the ligand and protein (10).

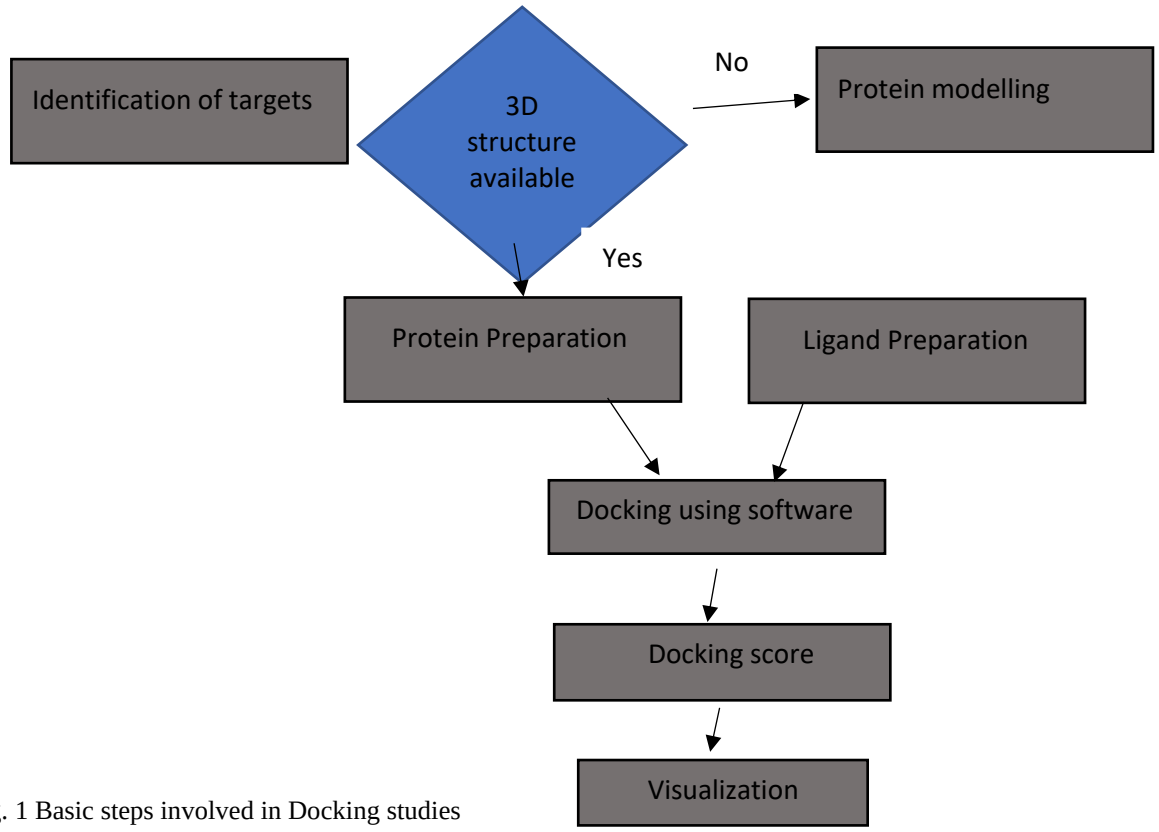

Fig. 1 Basic steps involved in Docking studies

Types of molecular docking

The basic methodology of molecular docking can be categorized into three ways:

I. Induced fit docking:
In this type of docking, ligand and receptor both are flexible in nature. Soft potential, rotamer library, receptor side chain flexibility and ensemble of protein conformations are various methods currently available to implement the receptor flexibility. The ligand binds flexibly at the active site of receptor to maximize bonding forces among them. It implements the idea of complementarity between protein and ligand.

II. Lock and key docking:
On the basis of Lock and key theory, both ligand and receptor are rigid bodies and show tight binding (14). In this case, the search space is very limited, considering only three translational and three-dimensional complementarity of freedom. Geometrical and chemical matching algorithms are used, and also the ligand-receptor complexes are often scored by accounting for steric fit, chemical complementation or pharmacophore similarity.

III. Ensemble docking:
This approach is treating as the ligand is flexible while the receptor is kept rigid during docking. So in that case both the ligand and receptor change their conformations to form a minimum energy perfect-fit complex (15, 16).

Software tools for Docking study

I. AutoDock
In molecular docking, the structure of the target macromolecule is obtained from protein data bank (PDB) and the structure of the small molecule or ligand can be obtained from a variety of computational methods.; molecular docking predicts how the small molecule will be most likely to bind to the macromolecule or protein (17).
The scoring function in this software is based on the United Atom version of the AMBER force field, in which nonpolar hydrogen atoms are eliminated to reduce the number of atoms to be simulated, and the van der Waals' radius of the heavy atom to which they are connected is increased in accordance with the necessary alteration of its partial, as well as to maintain the initial total charge. Therefore, both the ligand and the target protein are prepared with explicit polar hydrogen atoms and the nonpolar hydrogen atoms are treated implicitly.
AutoDock is using a method based on atomic solvation parameters to estimate the free energy change of solvation upon binding as well as computes the electrostatic interaction energy (18).
The preparation of macromolecule by removing water molecules and hydrogen atoms have been added. AutoDock (ADT) displays a message for the formatted ligand describing details that the charges were added, the number of non-polar hydrogens, aromatic carbons, and rotatable bonds that were present, the number of torsional degrees of freedom that were found (TORSDOF), and the amount that the total charge deviated from an integral value. ADT also detect and choose the best root for atom and marks it with a green sphere. Saving the macromolecule and ligand as a PDBQT formatted file. Preparing the grid parameter file is followed by running AutoDock (19) program for interaction using Lamarkian genetic algorithm.

II. Glide
Glide uses a series of hierarchical filters such as site point search, diameter test, subset test and greedy score to search for possible locations of the ligand in the active-site region of the receptor. Different sets of fields on a grid that enable progressively more precise scoring of the ligand posture are used to describe the structure and characteristics of the receptor (20).
The starting point for Glide scoring is the empirically based ChemScore function which can be written as

$$\Delta G_{bind} = C_0 + C_{lipo}\sum f(r_{lr}) + C_{hbond}\sum g(\Delta r)\, h(\Delta R) + C_{metal}\sum f(r_{lm}) + C_{rotb}H_{rotb}$$

Glide 2.5 employs two forms of GlideScore: (i) GlideScore 2.5 SP, used by Standard-Precision Glide; (ii) GlideScore 2.5 XP, used by Extra-Precision Glide (21). The FirstDiscovery suite14, was

used to obtain all results on an AMD Athelon MP 1800+ processor running Linux. The protein-preparation method was used to prepare all structures. The vdW radii of nonpolar protein atoms were not scaled for these computations, however the radii of nonpolar ligand atoms, which are defined as atoms with a partial charge smaller than 0.15 e-, were scaled down by a factor of 0.8.

Conformation Generation is a first step in its docking protocol followed by initial screening of ligand position and orientation over the active site of the protein. Energy is minimized using a Molecular Mechanics Scoring Function on precomputed OPLS-AA van der Waals and electrostatic grids for the receptor. The energy and gradient calculations are performed using standard three-dimensional interpolation methods. A model energy score ("Emodel") that incorporates the energy-grid score, the binding affinity predicted by GlideScore, and (for flexible docking) the internal strain energy for the model potential used to guide the conformational-search method is used to determine the best-docked structure (22, 23).

III. GOLD (Genetic Optimization for Ligand Docking)

GOLD software is based on genetic algorithm which is a computer program that mimics the process of evolution by manipulating a collection of data structures called chromosomes (12). Two binary strings were used to encode the conformational information: one for the protein and the other for the ligand. Each byte in the strings represented an angle of rotation about a rotatable bond (24). The fitness function was evaluated in six stages are as follows: (1) A conformation of the ligand and protein active site was generated. (2) The ligand was placed within the active site using a leastsquares ®tting procedure. (3) A hydrogen bonding energy H-Bond Energy, was obtained for the complex. (4) A pairwise energy, Complex Energy, was obtained for the steric energy of interaction between the protein and the ligand. (5) Molecular mechanics expressions were used to generate the term Internal Energy which was a measure of the internal energy of the ligand. (6) The energy terms were summed together to give a final fitness score (25). The advantage of this software is that it allows atomic overlapping between protein and ligand.

IV. iGEMDOCK

From preparations to post-screening analyses with pharmaceutical interactions, iGEMDOCK provides a comprehensive VS environment. First, iGEMDOCK offers interactive interfaces for setting up both the screening compound library and the target protein's binding site (14, 26). The company's internal docking tool GEMDOCK is then used to dock each compound in the library into the binding site. Then, I GEMDOCK creates profiles of electrostatic (E), hydrogen-bonding (H), and van der Waals (V) interactions between proteins and compounds. iGEMDOCK predicts the pharmacological interactions and groups the screening compounds for the post-screening analysis based on these profiles and compound structures. Finally, by merging the pharmacological interactions and energy-based scoring function of GEMDOCK, iGEMDOCK ranks and visualises the screening compounds (27).

Applications of molecular docking in drug designing

Today's study demands the use of molecular docking. If done before the experimental portion of any inquiry, it can show that any task is feasible. In several fields, molecular docking has completely changed the way that research is done. Some of the major applications of molecular docking are described below-

Lead optimization

A tiny molecule or ligand's optimal orientation on its target can be predicted by molecular docking. It can foretell several ligand binding mechanisms in the groove of the target molecule. The information gathered from these kinds of investigations could be used to create analogues that are more powerful, selective, and effective. The advantage of Glide tool is mainly lead discovery and lead optimization (28, 29).

Hit identification

Molecular docking in combination with scoring function can be used to screen huge databases for finding out potent drug candidates in silico, which can target the molecule of interest (4).

Bioremediation

Protein ligand docking can also be used to predict pollutants that can be degraded by enzymes Molecular docking leads to discovery of therapeutic drugs through multiple ways that include: identification of potential target, screening of potent drugs as activators or inhibitors against certain diseases, designing of novel drugs by lead optimization, prediction of binding mode and nature of active site as well as synthesis of chemical compounds with less time consumption (30) .

Discussion and Conclusions

In this review, the importance of molecular docking studies can be seen in a variety of computer-aided drug design applications (29). An enzyme protein may be activated or inhibited as a result of a binding contact between a small molecule ligand and the enzyme. If the protein is a receptor, agonism or antagonism may result from ligand binding. This paper has also describe the software tools for computational algorithms for docking and scoring which explore binding affinity of ligand against multiple receptors or proteins. As we consider molecular docking to be highly efficient method for the designing, synthesis and discovery of therapeutically important drugs. It can be implemented in medicinal chemistry, protein engineering, chemo informatics, bioremediation and many other biological and medicinal fields (30).

References

1. Saikia S, Bordoloi M. Molecular docking: challenges, advances and its use in drug discovery perspective. Current drug targets. 2019;20(5):501-21.
2. Meng X-Y, Zhang H-X, Mezei M, Cui M. Molecular docking: a powerful approach for structure-based drug discovery. Current computer-aided drug design. 2011;7(2):146-57.
3. Walters WP, Stahl MT, Murcko MA. Virtual screening—an overview. Drug discovery today. 1998;3(4):160-78.
4. Ferreira LG, Dos Santos RN, Oliva G, Andricopulo AD. Molecular docking and structure-based drug design strategies. Molecules. 2015;20(7):13384-421.
5. Rangaraju A, Rao A. A review on molecular docking: Novel tool in drug design and analysis. J Harmon Res Pharm. 2013;2:215-21.
6. Fischer D, Norel R, Wolfson H, Nussinov R. Surface motifs by a computer vision technique: searches, detection, and implications for protein–ligand recognition. Proteins: Structure, Function, and Bioinformatics. 1993;16(3):278-92.
7. Baxter CA, Murray CW, Waszkowycz B, Li J, Sykes RA, Bone RG, et al. New approach to molecular docking and its application to virtual screening of chemical databases. Journal of chemical information and computer sciences. 2000;40(2):254-62.
8. Hart TN, Read RJ. A multiple-start Monte Carlo docking method. Proteins: Structure, Function, and Bioinformatics. 1992;13(3):206-22.
9. Brint AT, Willett P. Algorithms for the identification of three-dimensional maximal common substructures. Journal of Chemical Information and Computer Sciences. 1987;27(4):152-8.
10. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. Nature reviews Drug discovery. 2004;3(11):935-49.
11. Ewing TJ, Makino S, Skillman AG, Kuntz ID. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. Journal of computer-aided molecular design. 2001;15(5):411-28.
12. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD. Improved protein–ligand docking using GOLD. Proteins: Structure, Function, and Bioinformatics. 2003;52(4):609-23.
13. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. Journal of computational chemistry. 1998;19(14):1639-62.
14. Agarwal S, Mehrotra R. An overview of molecular docking. JSM chem. 2016;4(2):1024-8.
15. Lorber DM, Shoichet BK. Flexible ligand docking using conformational ensembles. Protein Science. 1998;7(4):938-50.
16. Huang SY, Zou X. Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking. Proteins: Structure, Function, and Bioinformatics. 2007;66(2):399-421.
17. Forli S, Huey R, Pique ME, Sanner MF, Goodsell DS, Olson AJ. Computational protein–ligand docking and virtual drug screening with the AutoDock suite. Nature protocols. 2016;11(5):905-19.
18. Morris GM, Huey R, Olson AJ. Using autodock for ligand-receptor docking. Current protocols in bioinformatics. 2008;24(1):8.14. 1-8.. 40.
19. Huey R, Morris GM, Forli S. Using AutoDock 4 and AutoDock vina with AutoDockTools: a tutorial. The Scripps Research Institute Molecular Graphics Laboratory. 2012;10550:92037.
20. Repasky MP, Murphy RB, Banks JL, Greenwood JR, Tubert-Brohman I, Bhat S, et al. Docking performance of the glide program as evaluated on the Astex and DUD datasets: a complete set of glide SP results and selected results for a new scoring function integrating WaterMap and glide. Journal of computer-aided molecular design. 2012;26(6):787-99.
21. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, et al. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. Journal of medicinal chemistry. 2004;47(7):1750-9.

22. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. Journal of medicinal chemistry. 2004;47(7):1739-49.
23. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, et al. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein− ligand complexes. Journal of medicinal chemistry. 2006;49(21):6177-96.
24. Jones G, Willett P, Glen RC. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. Journal of molecular biology. 1995;245(1):43-53.
25. Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. Journal of molecular biology. 1997;267(3):727-48.
26. Subbaiyan A, Ravichandran K, Singh SV, Sankar M, Thomas P, Dhama K, et al. In silico molecular docking analysis targeting SARS-CoV-2 spike protein and selected herbal constituents. J Pure Appl Microbiol. 2020;14:989-98.
27. Hsu K-C, Chen Y-F, Lin S-R, Yang J-M. iGEMDOCK: a graphical environment of enhancing GEMDOCK using pharmacological interactions and post-screening analysis. BMC bioinformatics. 2011;12(1):1-11.
28. Shoichet BK, McGovern SL, Wei B, Irwin JJ. Lead discovery using molecular docking. Current opinion in chemical biology. 2002;6(4):439-46.
29. Gschwend DA, Good AC, Kuntz ID. Molecular docking towards drug discovery. Journal of Molecular Recognition: An Interdisciplinary Journal. 1996;9(2):175-86.
30. Naqvi AA, Mohammad T, Hasan GM, Hassan M. Advancements in docking and molecular dynamics simulations towards ligand-receptor interactions and structure-function relationships. Current topics in medicinal chemistry. 2018;18(20):1755-68.