INTERNATIONAL JOURNAL OF PHARMACEUTICAL, CHEMICAL AND BIOLOGICAL SCIENCES

Available online at www.ijpcbs.com

Review Article

COMPUTER-AIDED DOCKING: AN INVALUABLE TOOL IN DRUG DISCOVERY AND MOLECULAR BIOLOGY

Nagaraju Pappula*, Ravichandra Sharabu and Pradeep Kumar Thadiboina

Hindu College of Pharmacy, Amaravathi Road,
Guntur – 522 002, Andhra Pradesh, India.

ABSTRACT

Docking is a pharmacologically important tool in the field of drugs designing and computational biology. It works with the basic understanding of structure prediction of intermolecular complex formed between drug and its target molecule. The aim of ligand-receptor docking is to identify the pivotal active binding sites of a ligand with a protein of already known three-dimensional structures. Molecular docking is a computational procedure that aims to predict the favored orientation of a ligand to its macromolecular target when these are bound to each other to form a stable complex. The present review focused on importance of docking and its applications in drug innovation. The relevant basic theories including sampling algorithms, scoring functions are summarized. The differences in and performances of available docking software are also discussed.

Keywords: Dock, Drug discovery, Scoring and Lipinski rule.

INTRODUCTION

Computational docking or computer-aided docking is a tremendously valuable tool to gain an understanding of protein-ligand interactions which is important for the drug discovery [1-3]. Docking server integrates a number of computational chemistry software specifically aimed at correctly calculating parameters[4] needed at different steps of the docking procedure, i.e. accurate ligand geometry optimization, energy minimization, charge calculation, docking calculation and protein-ligand complex representation. Thus, the use of docking server allows the user to carry out highly efficient and robust docking calculations[5] by integrating a number of admired software used in in-silico chemistry into one comprehensive web service.

PROCEDURE FOR DOCKING

Preparing the protein and ligand for docking[6]

Docking algorithms require each atom to have a charge and an atom type that describes its properties [7]. However, the PDB (Protein Data Base) structure lacks these. So, we have to prepare the protein and ligand files and to include these values along with the atomic coordinates. Furthermore, for flexible ligand docking, we should also define ligand bonds that are rotatable. All this will be done in a tool called Auto Dock Tools (ADT).

Procedure[8,9]

Step-1

Receptor building – The receptor complex is downloaded from RCSB PDB.

Step-2

After downloading the pdb format of the protein, remove the water molecules the solvent molecules and all non-interacting ions by editing the TEXT of the protein.

Step-3

Add the missing hydrogens/side chain atoms and minimized the protein complex (AMBER Program).

Step-4

Clean the minimized complex (delete all the water and the solvent molecules and all non-interacting ions).

Step-5

Separate the active site of minimized complex in macromolecule (LOCK) and ligand (KEY) and prepare the docking suitable files for LOCK and KEY (pdbqt files).

Step-6
Prepare all the needing files for docking (grid parameter file, map files, docking parameter files including GA (Genetic Algorithm) parameter likes Population size, Generations etc which are necessary for docking accuracy). Set the output path to store the prepared structure.
Step-7
The predicted and later the ligand are uploaded. Arrange all the parameters such as number of pose to be obtained and score and run the Docking.
Step-8
The docking score is calculated and even the fitness is displayed, analyze your data and select the most optimum ligand and its pose.
Step-9
The complex can be viewed and checked for the orientation of the ligand with the receptor.
Step-10
To representation and analyzing of orientation of the ligand with the receptor viewed in various formats.
Eg: Ball and stick view format

PROTEIN DATA BANK (PDB)
Research Collaborators for structural Bioinformatics Protein Data Bank (RCSB PDB) began in 1970's by group of the young crystallographers, including Edgar Meyer, Gerson Coheon and Helen M Berman[7]. The PDB archive is maintained by the members of the worldwide PDB (wwPDB) – the RCSB (Research Co-laboratory for Structural Bioinformatics) PDB, EBI-MSD (Electrically Macromolecular Structure Relational Database), PDBj (Protein Data Bank Japan) and the BMRB (British Market Research Bureau Limited). Data deposited to the archive is processed using agreed-upon standards for full validation of the data[10]. These data are forwarded to the RCSB PDB for release into the archive. WwPDB (Worldwide protein data bank) members also maintain websites that provide different views to the data.

IMPORTANCE OF DOCKING IN NEW DRUG DEVELOPMENT
Ligand-protein docking is an optimization of problem based on predicting the position of a ligand with the lowest binding energy in the active site of the receptor. The net predicted binding free energy (ΔG_{bind}) is revealed in terms of various parameters: Hydrogen bond (ΔG_{hbond}), electrostatic[11] (ΔG_{elec}), Torsional free energy (ΔG_{tor}), Dispersion and repulsion[12] (ΔG_{vdw}), Desolvation (ΔG_{desolv}), Total internal energy (ΔG_{total}) and Unbound system's energy (ΔG_{unb}).
Molecular docking is a kind of bioinformatic modeling which involves the interaction of two or more molecules to give the stable adduct. Depending upon binding properties of ligand and target, it predicts the three-dimensional structure of any complex. At present, docking technique is utilized to predict the tentative binding parameters of ligand-receptor complex beforehand. Molecular docking generates different possible adduct structures that are ranked and grouped together using scoring function in the software[13]. Molecular docking of small molecules to a target includes a pre-defined sampling of possible conformation of ligand in the particular groove of target in an order to establish the optimized conformation of the complex. This can be made possible using scoring function of software. Since the infrared spectroscopy, X-ray crystallography and Nuclear Magnetic Resonance (NMR) spectroscopy are the techniques for the investigation and establishment of three dimensional structures of any organic molecule/ bio-molecular targets[14].
A new multi-objective strategy for molecular docking, named as MoDock, is presented to further improve the docking accuracy with available scoring functions. Tests of MoDock against the GOLD test data set reveal the multi-objective strategy improves the docking accuracy over the individual scoring functions. Meanwhile, a 70% ratio of the good docking solutions with the RMSD (simply root-mean-square deviation) value below 1.0 A^{o} outperforms other six commonly used docking programs, even with a flexible receptor docking program included [15].

GOLD
PARAMETERS IN DOCKING STUDIES
LIPINSKI five rule [16]
- ✓ To evaluate drug likeness or determine if a chemical compound has properties that would make it a likely orally active drug in humans derived.
- ✓ Because of the realization, that HTS (high through put screening) is identifying large numbers of hit compounds and many of which did not possess 'drug-like' properties.
- ✓ By Christopher Lipinski in 1997, most orally administered drugs are relatively small and moderately lipophilic (A molecular mass less than 500 Daltons).
- ✓ RO5 identifies molecular properties important for a drug's pharmacokinetics in the human body: absorption, distribution, metabolism, and excretion ("ADME").
  - ➢ No more than 5 hydrogen bond donors (the total number of nitrogen–hydrogen and oxygen–hydrogen bonds).
  - ➢ Not more than 10 hydrogen bond acceptors (all nitrogen or oxygen atoms).

- An octanol-water (O/W) partition coefficient log P not greater than 5.

✓ However, the rule does not predict when a pharmacologically active lead structure is optimized to increase the activity and selectivity.

Other rules to define drug-like properties

- Majority of compounds with good oral bioavailability in rats had less than 10 rotatable bonds (ROTB) and polar surface area (PSA) less than 140 $Å^2$ [17].
- Compounds with log P less than three and PSA greater than 75 $Å^2$ were six times less likely to exhibit adverse events in in-vivo tolerance studies [18].
- Number of aromatic rings greater than three significantly increases the risk of compound attrition [19].
- "Flatness" of compounds as defined by the fraction of carbons that are SP^3 hybridized, guarantees a success in clinical development [20].
- Used as application in design of drug and treatment in major diseases and including cancer [21].

How this rule benefits?

The rule describes molecular properties important for a drug's pharmacokinetics in the human body, including their absorption, distribution, metabolism, and excretion ("ADME"). However, the rule does not predict if a compound is pharmacologically active. This rule helps Pharmaceutics/Industrial Pharmacy students in proper selection of the drug and knowing whether the drug is suitable for oral formulations[22]. For Medicinal chemistry students involved in drug designing, CADD (Computer Aided Drug Designing), understanding this rule will help you a lot in designing suitable homologues of rugs and fine tuning your drug with suitable modifications[23]. Comprehensively utilized docking tools employ search algorithms such as genetic algorithm, fragment-based algorithms, Monte Carlo algorithms[24] and molecular dynamics algorithms. Besides this, there are some tools [10] such as DOCK, GOLD, Flex-X and ICM which are mainly used for high throughput docking simulations. There are various kinds of molecular docking procedures involving either ligand/target flexible or rigid based upon the objectives of docking simulations like flexible ligand docking (target as rigid molecule), rigid body docking (both the target and ligand as rigid molecules) and flexible docking (both interacting molecules as flexible)[25]. On the basis of the results for the top scored poses, the performance of the academic programs conform to the following order[26]: LeDock (57.4%) > rDock (50.3%) ~ Auto Dock Vina (49.0%) > Auto Dock (PSO) (47.3%) > UCSF DOCK (44.0%) > Auto Dock (LGA) (37.4%), and that of the commercial programs confirm to the following order: GOLD (59.8%) > Glide (XP) (57.8%) > Glide (SP) (53.8%) > Surflex-Dock (53.2%) > Ligand Fit (46.1%) > MOE-Dock (45.6%).

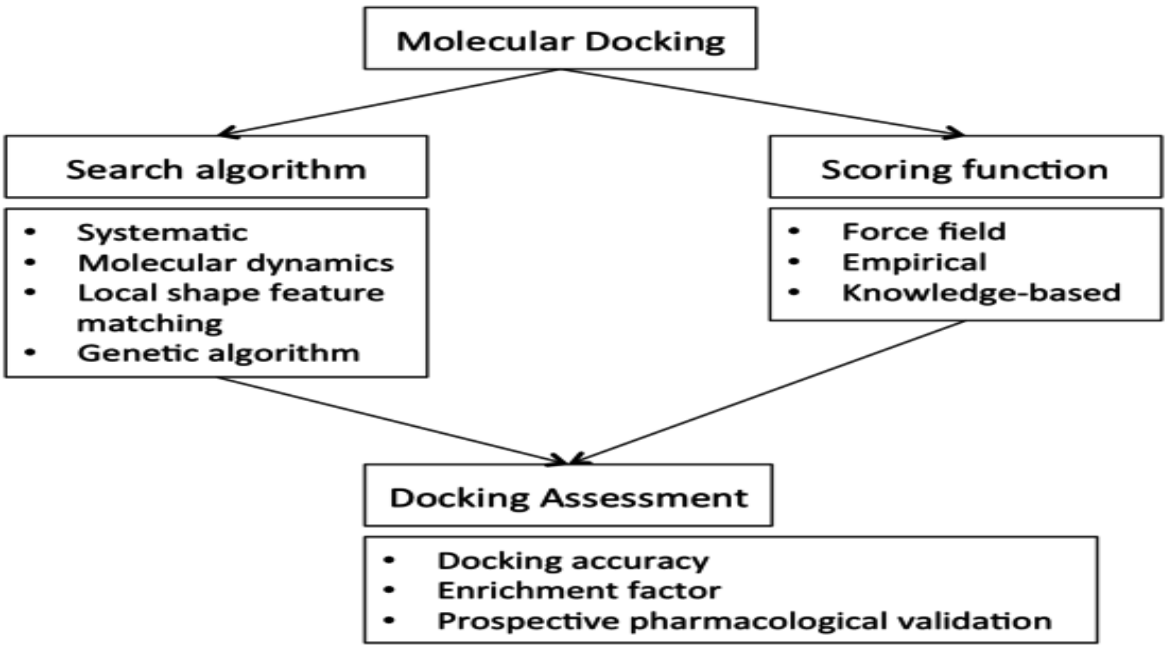

Fig. 1: Flow Chart

IS AN ACCURATE? IF YES HOW CAN I CONSIDER IN INVOLVING RESEARCH?

Molecular docking is an invaluable tool the field of molecular biology, computational structural biology, computer aided drug designing and pharmacogenomics [27]. It is an important common component of the drug discovery toolbox and its relative low-cost implications and perceived simplicity of use has stimulated an ever increasing popularity within academic communities [28]. Molecular docking is a study of how two or more molecular structures, for instance, drug and catalyst or macromolecule receptor, match along to be a perfect fit [29]. Binding orientation of small-molecule drug candidates to their macromolecular targets predicts the affinity and activity of a given small molecule [30].

Advantages of docking [31]

The application of docking in a targeted drug-delivery system is a huge benefit. One can study the size, shape, charge distribution, polarity, hydrogen bonding, and hydrophobic interactions of both ligand (**drug**) and receptor (**target site**).

Molecular docking helps in the identification of target sites of the ligand and the receptor molecule.

Docking also helps in understanding of different enzymes and their mechanism of action.

The "scoring" feature in docking helps in selecting the best fit or the best drug from an array of options.

Not everything can be proved experimentally as traditional experimental methods for drug discovery take a long time. Molecular docking helps in moving the process of computer-aided drug designing faster and also provides every conformation possible based on the receptor and ligand molecule.

Docking has a huge advantage when it comes to the study of protein interactions.

There are millions of compounds, ligands, drugs, and receptors, the 3D structure of which has been crystallized. Virtual screening of these compounds can be made.

Limitations of docking[32, 33]

In protein–small-molecule docking, there can be problems in the receptor structure. A reliable resolution value for small-molecule docking is below 1.2 Å while most crystallographic structures have a resolution between 1.5 and 2.5 Å.

Increasing the use of homology models in docking should be looked at with care as they have even poorer resolution [34]. Most applications accept and yield good results for structures below 2.2 Å. All the same, care should be taken while picking a structure.

The scoring functions used in docking, almost all of them, do not take into account the role played by covalently bound inhibitors or ions [35].

The methodology and research in protein–protein docking have to be greatly increased as the success in this field is greatly hampered by many false positives and false negatives [36].

DISCUSSION

Each docking program operates slightly differently; they share common features that involve ligand and receptor, sampling and scoring. Sampling entails conformational and orientational location of the ligand within the constraints of the receptor-site binding. A scoring function selects the best ligand conformation, orientation, and translation (referred to as poses), and classifies ligands in rank order. A successful docking exercise must accurately predict either or both ligand structure (pose prediction) and its binding propensity (affinity prediction). Available docking programs differ essentially in ligand placement in the "combining" site, exploration of conformational space and scoring or binding estimate[37]. Possible pitfalls in the docking studies are discussed and hints are provided to resolve commonly occurring problems increase in computing power, how to improve the accuracy is the future [38]. There three important aspects of protein-ligand docking: protein flexibility, ligand sampling, and scoring functions. Rapid advances in the last two decades have almost solved the ligand sampling issue. Speed and accuracy are the two important characteristics of a scoring function [39]. Because of the rapid direction for scoring function development, the computational methods for protein flexibility is still in its infancy and thereby remain one of the major future directions in protein-ligand docking.

CONCLUSION

Molecular Docking is a valuable and knowledgeable tool for *in silico* screening. It is playing an important and ever increasing role in rational drug design Docking is a computational procedure of searching for an appropriate ligand that fits both energetically and geometrically the protein's binding site and has been proved very efficient tool for novel drug discovery for targeting protein. Among different types of docking, protein-ligand docking is of special interest, because of its application in medicine industry**.** Protein-ligand docking refers to search for the accurate ligand conformations within a targeted protein when the structure of

proteins is known**.** Our goal of this study was to explore the feasibility of four different docking approaches: (AutoDock/Vina, GOLD, FRED and FlexX) for our target ASMT and to find out the lead compound. We compared the predictive power of each docking and scoring function. Based on review of suggest that all docking programs studied here do a reasonable job in docking and should aid significantly the drug discovery process.

REFERENCES

1. Ankur Dhanik and Lydia E Kavraki. Protein-ligand interactions: Computational Docking. Citable Reviews in Life Sciences. 2012.
2. Gschwend DA, Good AC and Kuntz ID. Molecular docking towards drug discovery. J Mol Recognit. 1996;9:175-186.
3. Ferreira LG, dos Santos RN, Oliva G and Andricopulo AD. Molecular docking and structure-based drug design strategies. Molecules. 2015;20:13384-13421.
4. Mehrotra R, Jangir DK, Agarwal S, Ray B and Singh P. Interaction studies of anticancer drug lomustine with calf thymus DNA using surface enhanced Raman spectroscopy. MAPAN. 2013;28:273-277.
5. Holt PA, Chaires JB and Trent JO. Molecular docking of intercalators and groove-binders to nucleic acids using Autodock and Surflex., J Chem Inf Model. 2008;48:1602-1615.
6. http://sbcb.bioch.ox.ac.uk/users/greg/teaching/docking-2012.html (8/12/2017)
7. Daniel Seeliger and Bert L de Groot. Ligand docking and binding site analysis with PyMOL and Autodock/Vina, J. Computer Aided Mol Des. 2010;24(5):417–422.
8. http://www.bio.iitb.ac.in/~sanjeeva/virtual_lab/Virtual_Laboratory/exp-12/procedure.pdf (21/06/18).
9. http://ncbr.muni.cz/~martinp/C3210/StructBioinf9.pdf (21/06/18).
10. Berman HM. The Protein Data Bank: A historical perspective. Acta Crystallogr A. 2008;64: 88–95.
11. Berman H, Henrick K and Nakamura H. Is announcing the worldwide, Protein Data Bank. Nat Struct Biol. 2003;10(12):980.
12. Maria Jesus Garcia-Godoy, Jose Garcia-Nieto, Antonio J Nebro and Jose F Aldana-Montes. A New Multi-objective Approach for Molecular Docking Based on RMSD and Binding Energy, Springer link. 2016; 65-77.
13. Ayaz Mahmood Dar and Shafia Mir. Molecular Docking: Approaches, Types, Applications and Basic Challenges. J Anal Bioanal Tech. 2017;8(2):1-3.
14. Seeliger D and De Groot BL. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. J Comput Aided Mol Des. 2010;24:417-422.
15. Shoichet BK, McGovern SL, Wei B and Irwin JJ. Lead discovery using molecular docking. Curr Opin Chem Biol. 2002;6:439-446.
16. Junfeng Gu, Xu Yang, Ling Kang, Jinying Wu and Xicheng Wang. MoDock. A multi-objective strategy improves the accuracy for molecular docking. Algorithms for Molecular Biology. 2015;10(8): 1-9.
17. Michael P and Pollastri. Overview on the Rule of Five. Curr Protoc Phamracol. 2010;49(9):1-9.
18. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW and Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem. 2002;45(12):2615-2623.
19. Hughes JD, Blagg J, Price DA, Bailey S, Decrescenzo GA, Devraj RV, Ellsworth E, Fobian YM, Gibbs ME, Gilles RW, Greene N, Huang E, Krieger-Burke T, Loesel J, Wager T, Whiteley L and Zhang Y. Physiochemical drug properties associated with in vivo toxicological outcomes. Bioorg Med Chem Lett. 2008;18(17):4872-4875.
20. Ritchie TJ and Macdonald SJ. The impact of aromatic ring count on compound developability are too many aromatic rings a liability in drug design. Drug Discovery Today. 2009;(21-22):1011-1020.
21. Frank Lovering, Jack Bikker and Christine Humblet. Escape from Flatland. Increasing Saturation as an Approach to Improving Clinical Success. J Med Chem. 2009;52(21):6752–6756.
22. Ziad Omran and Cyril Rauch Acid-mediated Lipinski's second rule: application to drug design and targeting in cancer. Eur Biophys J. 2014;43(4-5):199–206.
23. http://pharmawiki.in/lipinskis-rule-of-five-an-important-rule-for-drug-selection/ (6/12/2017).
24. Badry D Bursulaya and Maxim Totrov. Comparative study of several algorithms for flexible ligand docking Journal of Computer-Aided Molecular Design. 2003;17:755–763.
25. Shweta Agarwal and Ranjana Mehrotra. An overview of Molecular Docking. JSM Chemistry. 2016; 4(2):1024.
26. Nataraj S Pagadala. Software for molecular docking: a review. Biophys Rev. 2017;9(2):91–102.
27. Jahan B Ghasemi and Azezeh Abdolmaleki. Molecular docking challenges and limitations (chapter-3). Applied Case Studies and Solutions in Molecular Docking-Based Drug Design. Google text books. 2016;61-62.

28. https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/docking-molecular (6/12/2017).
29. Michael Berry and Junaid Gamieldien. Practical Considerations in Virtual Screening and Molecular Docking., Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology. 2015; 27:487–502.
30. Gane PJ and Dean PM. Recent advances in structure based rational drug design, current opinion in structural biology. 2010;10:401-404.
31. Hakes L, Lovell SC, Oliver SG and Robertson DL. Specificity in protein interactions and its relationship with sequence diversity and co-evolution. Proc Natl Acad Sci. 2007;104(19):7999-8004.
32. C. George Priya Doss and Thirumal Kumar D. Investigating the Inhibitory Effect of Wortmannin in the Hotspot Mutation at Codon 1047 of PIK3CA Kinase Domain: A Molecular Docking and Molecular Dynamics Approach, Advances in Protein Chemistry and Structural Biology. 2016;102:267-297.
33. Gohlke H and Klebe G. Approches to the description and prediction of binding affinity of small molecule ligands to macromolecular receptors. Angew Chemie Int Ed Endl. 2002;41:2644-2676.
34. Mihasan M. Basic protein structure prediction for the biologists. A review. Archives of biological sciences. 2010;62:857-871.
35. Mihasan M. What in silico molecular docking can do for the bench working biologists, journal of biosciences. 2012;37(1):1089-1095.
36. Moreira IS, Fernandes PA and Ramos MJ. Protein-protein docking dealing with the unknown. J Comput Chem. 2010;31(2):317-42.
37. Kunal Roy and Rudra Narayan Das. Understanding the basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment, Elsevier, 1st edition. 2015;151-189.
38. Aldo Jongejan, Chris de Graaf, Nico PE Vermeulen, Rob Leurs and Iwan JP de Esch. The Role and Application of In Silico Docking in Chemical Genomics Research Chapter-5. Chemical Genomics: Reviews and Protocols. 2005;63-85.
39. Sheng-You Huang and Xiaoqin Zou, Advances and Challenges in Protein-Ligand Docking. Int J Mol Sci. 2010;11(8):3016–3034.