PROTEIN LIGAND DOCKING: A MINI REVIEW WITH A CASE STUDY

Hani S.H Mohammad Ali

Department of Biological sciences, Faculty of science, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia

Received – February 18, 2019; Revision – April 24, 2019; Accepted – May 26, 2019
Available Online – June 10, 2019
DOI: http://dx.doi.org/10.18006/2019.7(3).243.248

ABSTRACT

Ligand protein docking is a molecular technique used to predict the structure to design a particular drug. It predicts the orientation and position of small molecule (ligand) when it is bound to a protein. It can be done by using different software's available. Molecular docking generates different structure ranked on scoring function. In this review, we discussed some general information regarding molecular docking with an example of NS1 protein from yellow fever virus.

KEYWORDS

Protein
Ligands
Docking
NS1

* Corresponding author
E-mail: hanoiolfat@hotmail.com; hanoiolfat@gmail.com (Hani S.H Mohammad Ali)

Peer review under responsibility of Journal of Experimental Biology and Agricultural Sciences.

All the articles published by Journal of Experimental Biology and Agricultural Sciences are licensed under a Creative Commons Attribution-NonCommercial 4.0 International License Based on a work at www.jebas.org.

Production and Hosting by Horizon Publisher India [HPI] (http://www.horizonpublisherindia.in/).
All rights reserved.
1 Introduction

The technique of molecular docking may be applied for modeling small molecule and protein interaction (atomic level), which permit the characterization and performance of these small molecules and binding location to proteins. This also helps in interpreting important biochemical processes. Docking process includes two major steps; first ligand conformational prediction along with location and pose while second step is assessment of binding strength. These steps mainly depend on sampling techniques and schemes scoring respectively (McConkey et al., 2002). The information of active sites can be obtained through comparative study of target protein with family of proteins having same function. If the information of binding sites is not available, few programs like cavity detection or online servers can be utilized for this purpose.

Understanding of drug bimolecular interactions, molecular editing is a striking framework, for designing normal drugs and break through along with interpretation of molecule manipulation (Ligand) into targeted preferred location of DNA/protein (receptor) primarily to produce stable complex of significant characteristics and with more specificity by non-covalent order (Rohs et al., 2005). The information based on docking approach can be integrated for free energy, binding energy and stability of complexes. Presently, the utilization of molecular docking approach to envisage the uncertain binding factors of ligand-receptor complex in advanced.

The major aim of molecular docking is to optimize conformation and with objective of processing less free energy binding to get ligand-receptor complex. The total free energy binding (ΔGbind) is exposed with the help of different factors, including hydrogen bond (ΔH bond), electrostatic (ΔGelec) torsional free energy (ΔGtor), dispersion and repulsion (ΔGvdw), desolvation (ΔGdesolv), total internal energy (ΔGtotal) and unbound system’s energy (ΔGunb). Hence, more evidences regarding the nature of different sorts of connections which trigger molecular docking is provided by good insight of the common principle that administer projected free energy binding (ΔGbind) (Guedes et al., 2014).

Data sources for the exploration of targeting with appropriate PDB format and a procedure to produce ligand as a PDB file are required by real usage of molecular docking. In order to perform these activities certain software like Discovery studiois are available where PDB format can be made for ligand. The ability to relate with available target proteins/DNA these tools deliver the organization of ligands based docking. Minor molecules docking to a target having a known sampling of probable conformation of ligand in a specific channel of target in a way to establish the adjust conformation of the complex. This may be possible by the use of scoring function of software (Agarwal et al., 2015). There are some approaches which include the infrared spectroscopy, X-ray crystallography and Nuclear Magnetic Resonance (NMR) spectroscopy, for the determination and creation of three-dimensional structure of every biomolecules or organic molecules. Therefore, it becomes possible due to homology modeling to investigate the possible proteins structure of unspecified structure with high order homology to defined structure.

Different databases can be used for offering information on minor ligand molecules. Many databases including CSD (Cambridge Structural Database), ACD (Available Chemical Directory), MDDR (MDL Drug Data Report) and NCI (National Cancer Institute Database) can assist in this regard. Various interacted conformers are produced and checked with each other, while during performing docking. In the process of rejection, new conformers are tested and still the search methodology continues till the acceptance of one conformation. According to the experimental binding tendencies of docked conformers and free energies binding appear to be more problematic than their binding orientation. Various scoring functions are engaged like consensus scoring to avoid this problem. Consensus scoring is appliance of number of score functions to the similar docked position to avoid false positive (Seeliger & Groot, 2010). For designing an effective docking protocol, lot of attempts were carried out. Certainly, success has been achieved in the computational prediction of docking methods. The current small review article is devoted to latest computational techniques, emphasis has been given to types and application of molecular docking to a specific case.

2 Approaches of molecular docking

Modeling the interaction of two molecules is an unsolved problem. Various major factors of forces are existed in the inter-molecular association, such as hydrophobic, van der walls or stacking interactions between the aromatic/ volatile amino acids, hydrogen bonding and electrostatic forces. Hence with the minimum information, ligand docking mimics the natural interaction of ligand with that of its receptor (Sousa et al., 2006). Easy methods are available for docking inflexible ligands with inflexible receptors and flexible ligands with inflexible receptors, but generally, methodologies of docking taking in to account conformational flexible ligands and sixty-four protein engineering-technology and there is problem in application receptors. A search algorithm combination and functions of the scoring is shown in figure 1.
The creation of search algorithm should have an optimal number of entries including binding modes determined experimentally. Though, amid two molecules, algorithm searching will show all probable binding modes, yet, the search of this algorithm would be unfeasible due to the amount of time taken and the size of the search space. Ultimately, a mere amount of total conformational space may be sampled, so an equilibrium must be attained between the amounts of the search spaced investigated and computational expense. Additionally, molecular dynamics, genetic algorithms, Tabu, and systematic searches, Monte Carlo methods, fragment-based, point complementary and distance geometry methods, are some examples of searching algorithm. Furthermore, on the other way a many mathematical approaches in scoring function applied to forecast nature of the non-covalent association known as affinity of binding. An important problem in all computational approaches is the creation of an energy scoring function responsible for sharply and precisely explain the association between ligand and protein (Jain, 2006). Three most significant applications of scoring functions in molecular docking are (i) determination of binding mode and location of ligand on a protein (ii) prediction of absolute binding affinity between ligand and protein, especially in lead optimization and (iii) most important application, is to point out a specific protein target in the potential drug by searching database of large ligand. In previous few years, various development has been made in discovery of different scoring functions which have various accuracies and efficiencies in computing. Force field, empirical, knowledge-based and consensus scoring are some of the common examples which enhance efficiencies of scoring. The procedure of protein docking may be divided specifically into rigid docking and flexible docking (Huang et al., 2010a).

2.1 Rigid Docking
Avoiding any sort of flexibility, this particular kind of docking deals the ligand and receptor both and recognizes only six degrees of translational and rotational freedom. Mostly, docking suites used rigid body docking procedure as a first step.
2.2 Flexible Docking

Flexible docking is used to model ligand flexibility of a rigid protein receptor. This model used conformational space of ligand to model its flexibility. However, in this regard some approaches have been reported about the flexibility of protein taking into account ideologically. Additionally, ligand flexibility: systematic, random or stochastic, and simulation approaches are the three common classes of algorithm for treating flexibility of ligand (Souse et al., 2006). Molecular docking of proteins may be more challenging due to multiple degrees of freedom and their flexibility. Huang et al. (2010b) have described protein flexibility groups which may include soft docking, side-chain flexibility, molecular relaxation and protein ensemble docking.

3 General procedure for docking

Many significant scientific reports are available regarding molecular docking approaches and many publications comparing different tools of molecular docking (Taylor et al., 2002; Souse et al., 2006; Huang et al., 2010a). Further, four step methodology will be described to execute the molecular docking.

3.1. Target selection

Preferably, X-ray crystallography or nuclear magnetic resonance could be used to investigate the target structure experimentally. This may be getting from PDB; nevertheless, docking is preferred over homology models or threading while performing comparison. Molprobity (Chen et al., 2010) validation software frequently used to evaluate the quality of model. Water molecules removal from the cavity, stabilization of charges, missing residues filling, and generation of side chains, in accordance to the available parameters, are carried out after selecting the model. The stability and biological activeness of receptor could be used at this point.

3.2. Preparation and Ligand selection

Selection of ligands type for docking is mainly depends on the aim of docking. Databases like ZINC or PubChem, or chemsketch tool might be useful for this purpose (Dias & Azvedo., 2008). Further, determination of various filters such as total charge, molecular weight, polar surface area, solubility, commercial availability, thresholds similarity of pharmacophores, synthetic accessibility, and absorption, distribution of molecules, metabolism, excretion, and toxicology characteristics, which required for the reduction of molecules are also required for the proper docking.

3.3. Docking

Docking is the final step, where receptor and the interactions of ligand which is docked upon are tested. The scoring function produces a score, which is dependent on the selection of best suitable ligand.

3.4. Results evaluation of docking

Normally, docking algorithm success in forecasting a ligand binding pose is evaluated in terms of root mean square deviation (RMSD) between heavy atom positions observed experimentally of the ligands and the one(s) forecasted in the logarithm. The elasticity of system is the main problem in searching of accurate pose and the searching efficiency is determined by the number of degrees of freedom present in the conformational search which is the main aspect (Souse et al., 2006).

3.5. Software of docking and description

For investigation and justifying ligand protein or protein-protein interaction there are many algorithms available. Obtaining successful docking technologies speed and accuracy are one of the important features in the process. Many algorithms show common methodologies along with new extensions concentrated on gaining quick and accurate method as higher as possible. For instance, AutoDock (Morris et al., 2009), DOCK (Kuntz et al., 1982), FlexX (Rarey et al., 1996), GOLD (Jones et al., 1997), ICM (Abagyan et al., 1994), ADAM (Mizutani et al., 1994), DARWIN (Taylor & Burnett., 2000), DIVALI (Clark., 1995), and DockVision (Hart & Read, 1992) are some of the most common docking programs.

3.6 Molecular docking application to a special case

Biochemical reaction as it goes on before experimental part of any examination molecular docking may determine the possibility of this reaction. Molecular docking has reformed the results in some area of research. Especially, regarding the interaction between protein target (may be an enzyme) and small molecules (ligands) which may show the forecast the activation or binding characteristics of drug’s to nucleic acid. This information determines the correlation between molecular structure and cytotoxicity of drug. Keeping in view this significance the researchers especially medicinal chemists are focusing to interpret the basic anticancer mechanism of drugs at molecular level by examining the nucleic acid and drugs interaction properties (Mehrotra et al., 2013; Agarwal et al., 2014). In-silico observations are being carried out by medicinal chemists to evaluate/predict their results whether the drug/compound is...
interacting with protein/DNA. Further, when the researchers become certain that the docking program is forecasting the so-called interaction, then the information is used to investigate the actual binding mode of complex that progress into drug discovery. Additionally, in the detection of these structural modifications in a drug would be instrumental information that may results in sequence/structure specific binding to their target (Holt et al., 2008).

3.7 Molecular docking of NS1 protein from yellow fever virus using SwissDock server

3.7.1 Yellow fever virus NS1 protein

With high fatality rate Yellow fever virus produces a non-structure protein (NS1) and is responsible for numerous epidemics of this fever (Beasley et al., 2015). This protein contains the virulence part that is responsible for causing yellow fever. In the beginning it exists in the form of dimer but later on it becomes hexamer after translocation into reticulum lumen. This form of NS1 is responsible for causing disease as it invade the immune system. In mature form this protein contains 352 amino acids and is used as potential marker for the identification of this virus (Rastogi et al., 2016). NS1 protein of flavivirus shows resembles with dengue and zika virus in its structure.

3.7.1 Methods

3.7.1.1 Ligand and Receptor

The C-terminal fragment structure of the NS1 protein from yellow fever virus was obtained from the Protein Data Bank [PDB ID: 5YXA] at 2.1 Å resolution. The crystal structure of protein HET atoms of water molecules was removed to clean and only chain A was selected for docking studies. The ligand molecules JJ3297, [ZINC: 02989649] were retrieved from ZINC databases

3.7.1.2 Molecular docking

SwissDock web server was used to perform molecular docking available at www.swissdock.ch/docking based on the docking software. This docking web server has interface to input selected structure of protein and ligand from databases, docking parameters change, and to visualize favorable clusters online. Results can also be downloaded and viewed in freely available UCSF Chimera package.

3.7.2 Results

In swissdock each run Gibbs free energy ($\Delta G$) and Full Fitness of the docking evaluated. On the basis of Full Fitness and cluster formation favorable binding modes were scored. The cluster ranking performed using Full Fitness value. Figure 2 shows the clustering results obtained from selected ligands docking into NS1 protein. For most favorable interaction, NS1 protein showed Full Fitness of -878.36 kcal/mol and estimated $\Delta G$ of -7.12 kcal/mol. Result represented by figure 2 shows the visualization of the favorable binding of the ligands into the protein NS1.

Conclusion

The protein-ligand interaction studies are important in the structure-based drug design. Influenza virus non-structural protein 1 (NS1) is demonstrated as potential therapeutic target for antiviral therapy by identification of specific small-molecule inhibitors. Based on the results of docking studies, it has been clearly expressed that ligand molecules JJ3297 showed favorable binding with NS1 protein. Present study, has given an insight to inhibit NS1. Further investigations are required to get more favorable interaction into NS1.

Conflict of Interest

The authors declare that there is no conflict of interest in this manuscript.
References

Abagyan R, Totrov M, Kuznetsov D (1994) ICM-A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. Journal of Computational Chemistry 15: 488-506.

Agarwal S, Chadha D, Mehrotra R (2015) Molecular modeling and spectroscopic studies of semustine binding with DNA and its comparison with lomustine–DNA adduct formation. Journal of Biomolecular Structure and Dynamics 33: 1653-1668.

Agarwal S, Jangir DK, Mehrotra R, Lohani N, Rajeswari M (2014) A Structural Insight into Major Groove Directed Binding of Nitrosourea Derivative Nimustine with DNA: A Spectroscopic Study. PLoS ONE 9: 104-115.

Beasley DW, McAuley AJ, Bente DA (2015) Yellow fever virus: genetic and phenotypic diversity and implications for detection, prevention and therapy. Antiviral Research 115:48-70.

Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010). MolProbity: all-atom structure validation for macromolecular crystallography. ActaCrystallographica Section D 1:12-21.

Clark KP (1995) Flexible ligand docking without parameter adjustment across four ligand-receptor complexes. Journal of Computational Chemistry 16: 1210-1226.

Dias R, de Azevedo WF Jr. (2008) Molecular docking algorithms. Current Drug Targets 9: 1040-1047.

Guedes IA, de Magalhães CS, Dardenne LE (2014) Receptor-ligand molecular docking. Biophysical Reviews 6: 75-87.

Hart TN, Read RJ (1992) A multiple-start Monte Carlo docking method. Proteins 13: 206-222.

Holt PA, Chaires JB, Trent JO (2008) Molecular docking of intercalators and groove-binders to nucleic acids using Autodock and Suflex. Journal of Chemical Information and Modeling 48: 1602-1615.

Huang SY, Grinster SZ, Zou X (2010a) Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions. Physical Chemistry Chemical Physics 12: 12899-908.

Huang SY, Zou X (2010b) Advances and challenges in protein-ligand docking. International Journal of Molecular Sciences 11: 3016-3034.

Jain AN (2006) Scoring functions for protein-ligand docking. Current Protein Peptide Science 7: 407-20.

Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. Journal of Molecular Biology 267: 727-748.

Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE (1982) A geometric approach to macromolecule-ligand interactions. Journal of Molecular Biology 161: 269-88.

McConkey BJ, Sobolev V, Edelman M (2002) The performance of current methods in ligand-protein docking. Current Science 83: 845–855.

Mehrotra R, Jangir DK, Agarwal S, Ray B, Singh P, Srivastava AK (2013) Interaction studies of anticancer drug lomustine with calf thymus DNA using surface enhanced Raman spectroscopy. MAPAN 28: 273-277.

Mizutani MY, Tomioka N, Itai A (1994) Rational automatic search method for stable docking models of protein and ligand. Journal of Molecular Biology 243: 310-326.

Morris GM, Huey R, Lindstrom W, Sanner M, Belew R, Goodsell DS, Olson A (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of Computational Chemistry 16: 2785.

Rastogi M, Sharma N, Singh SK (2016) Flavivirus NS1: a multifaceted enigmatic viral protein. Virology Journal 13:131.

Rohs R, Bloch I, Sklenar H, Shahked Z (2005) Molecular flexibility in ab-initio drug docking to DNA: binding-site and binding-mode transitions in all-atom Monte Carlo simulations. Nucleic Acids Research 33: 7048-7057.

Seeliger D, de Groot BL (2010) Ligand docking and binding site analysis with PyMOL and Autodock/Vina. Journal of Computer Aided Molecular Design 24: 417-422.

Sousa SF, Fernandes PA, Ramos MJ (2006) Protein-ligand docking: current status and future challenges. Proteins 65: 15-26.

Taylor JS, Burnett RM (2000) DARWIN: A program for docking flexible molecules. Proteins 41: 173-191.

Taylor RD, Jewsbury PJ, Essex JW (2002) A review of protein-small molecule docking methods. Journal of Computer Aided Molecular Design 16: 151-66.