Int. J. Drug Res. Tech. 2012, Vol. 2 (2), 189-197 ISSN 2277 - 1506

International Journal of Drug Research and Technology

Available online at http://www.ijdrt.com/

Review Article

REVIEW ON INTRODUCTION TO MOLECULAR DOCKING SOFTWARE TECHNIQUE IN MEDICINAL CHEMISTRY

Kharb Manju* Jat R. K., Parjapati Gunjan and Gupta Anju

Department of Pharmacy,
Suresh Gyan Vihar University, Rajasthan,
Jaipur, India

ABSTRACT

Molecular docking is a study of receptor of protein fit together. The problem is like solving a 3 dimensional puzzle. For example, the action of a harmful protein in human body may be prohibited by finding an inhibitor, which binds to that particular protein. Molecular modeling encompasses all theoretical methods and computational techniques used to model or mimic the behaviour of molecules. Molecular docking software's are mainly used in drug development. The most important application of docking software is virtual screening. In virtual screening the most interesting and promising molecules are selected from an existing database for further research. This review has basic Information about molecular modeling, molecular docking, basic concepts of docking, docking approaches, and mechanics of docking, docking software's for further development in this field.

Keywords: Molecular docking, Molecular modelling, Virtual screening, Ligand, Computational chemistry.

INTRODUCTION

Molecular Modeling is a tool for doing chemistry. Models are central for understanding of chemistry. Molecular modeling allows us to do and teach chemistry better by providing better tools for investigating, interpreting, explaining and discovering new phenomena. Like experimental chemistry, it is a skill-demanding science and must be learnt by doing and not just reading. Molecular modeling is easy to perform with currently available software, but the difficulty lies in getting the right model and proper interpretation. Molecular modeling is the general term used to describe the use of computers to construct molecules and perform a variety of calculations on these molecules in order to predict their chemical characteristics and behaviour. The term molecular modeling is often used synonymously with the term computational chemistry. Computational chemistry is a broader term, referring to any use of computers to study chemical system.[1] The benefit of molecular modeling is that it reduces the complexity of the system, allowing many more particles (atoms) to be considered during simulations.[2] In recent years the search for novel drugs has evolved from a process of trial and error into a sophisticated procedure including several computer- based approaches. In structure-based design the structures of known target proteins are used to discover new compounds of therapeutically relevance. The approaches can be classified roughly into two categories: de novo design and docking.[3]

BASIC CONCEPT

Docking is the formation of non dent protein-ligand complexes. Given the structures of a ligand and a protein, the task is to predict the structure of the resulting complex. This is the so-called docking problem. Because the native geometry of the complex can generally be assumed to reflect the global minimum of the binding free energy, docking is actually an energy-optimization problem. The development of docking methods is therefore also concerned with making the right assumptions and finding acceptable simplifications and that still provide a sufficiently accurate and predictive model for protein-ligand interactions.[4] The former method designs new ligands to fit the protein target, whereas the latter is used to decide whether existing compounds possess a good steric and chemical complementarily to the given protein. [3]

MOLECULAR DOCKING

In the field of molecular modelling ,docking is a method which predicts the preferred orientation of one molecule to a second when bound to each other to form a stable complex. Knowledge of the preferred orientation in turn may be used to predict the strength of association or binding affinity between two molecules using for example scoring functions. The associations between biologically relevant molecules such as proteins, nucleic acids, carbohydrates, and lipids play a central role in signal transduction. Furthermore, the relative orientation of the two interacting partners may affect the type of signal produced. Therefore docking is useful for predicting both the strength and type of signal produced. Docking is frequently used to predict the binding orientation of small molecule drug candidates to their protein targets in order to in turn predict the affinity and activity of the small molecule. Hence docking plays an important role in the rational design of drugs. Given the biological and pharmaceutical significance of molecular docking, considerable efforts have been directed towards improving the methods used to predict docking.[3] The number of accessible states grows exponentially with the degrees of freedom of the docking molecules. Energy calculations in condensed phases must subtract large numbers to arrive at small differences, almost guaranteeing inaccuracy.[5] Molecular docking can be thought of as a problem of "lock-and-key", where one is interested in finding the correct relative orientation of the "key" which will open up the "lock" (where on the surface of the lock is the key hole, which direction to turn the key after it is inserted, etc.). Here, the protein can be thought of as the "lock" and the ligand can be thought of as a "key". Molecular docking may be defined as an optimization problem, which would describe the "best-fit" orientation of a ligand that binds to a particular protein of interest. "Key". Molecular docking may be defined as an optimization problem, which would describe the "best-fit" orientation of a ligand that binds to a particular protein of interest. During the course of the process, the ligand and the protein adjust their conformation to achieve an overall "best-fit" and this kind of conformational adjustments resulting in the overall binding is referred to as "induced-fit".[6] The focus of molecular docking is to computationally simulate the molecular recognition process. The aim of molecular docking is to achieve an optimized conformation for both the protein and ligand and relative orientation between protein and ligand such that the free energy of the overall system is minimized.[7] The molecular docking tool has been developed to obtain a preferred geometry of interaction of ligand - receptor complexes having minimum interaction energy based on different scoring functions viz. only electrostatics, sum of steric and electrostatic (parameters from MMFF force field) and Dock Score. This utility allows one to screen a set of compounds for lead optimization.

- Rigid docking where a suitable position for the ligand in receptor environment is obtained while maintaining its rigidity
- Flexible docking where a favoured geometry for receptor-ligand interaction is obtained by changing internal torsions of ligand into the active site while receptor remains fixed
- Full flexible docking where the ligand is flexed via its torsion angles as well as the side chain of active site residues (selected active site residues within a user specified radius around the ligand) are flexed.

The challenge of the lead-generation phase of the receptor-ligand docking approach is to quickly screen millions of possible compounds that fit a particular receptor and to specifically select those that show a high affinity. The set of ligands thus selected can then be screened further by either more involved computational technique, such as free-energy perturbation theory, or directly in assays. Many techniques have been proposed that address specific parts of this challenge. Among the first were methods that simply evaluate whether a particular ligand can fit into the receptor pocket under the assumption of both rigid ligands and a rigid protein. This problem allows an enumerative approach; because there are only six degrees of freedom that completely specify the relative position of the ligand with respect to the receptor. Such techniques are reasonably fast. Possible geometries can be scored by force field, empirical or knowledge-based methods. (MDS allows user to select different intermolecular interactions viz. steric, electrostatic). In addition, a flexible ligand docking includes molecules internal degree of freedom along with values of translation and rotation in search of its suitable bound conformation that makes it computationally more expensive than rigid ligand docking. The Dock score or XC score as it is called compute binding affinity of a given protein ligand complex with known 3-D structure. Dock/XC score scoring function include terms for Vander Walls interaction, hydrogen bonding, deformation penalty, hydrophobic effects. Cavity points are found and the centre of mass of the ligand is moved to each cavity point. All rotations of ligand are scanned at each cavity point where ligand is placed. For each rotation a pose of the ligand is generated and the corresponding bumps are checked for each pose of ligand. The X-C score is calculated for each valid pose (determined by the cut off criteria fed by user in terms of max no of allowed bumps) and the pose of the ligand with the best score is given as output to user. Though this method is for one ligand for a given receptor, it can also be applied to a set of ligands/their conformers in a batch grid docking mode. MDS also incorporates the Piecewise Linear Pair wise Potential (PLP) function in PLP docking (rigid docking) method that includes ligand-receptor interactions of hydrogen bonding (donor-acceptor), repulsions (donor-donor, acceptor-acceptor) and dispersion (involving non-polar group interactions) types.[8]

DOCKING APPROACHES

Two approaches are particularly popular within the molecular docking community. One approach uses a matching technique that describes the protein and the ligand as complementary surfaces.[9,10]The second approach simulates the actual docking process in which the ligand-protein pair wise interaction energies are calculated Bo.[11] The approaches have significant advantages as well as some limitations. These are outlined below. Shape Complementarity Geometric matching/shape complementarily methods describe the protein and ligand as a set of features that make them dock able.[12] These features may include molecular surface/ complementary surface descriptors. In this case, the receptor's molecular surface is described in terms of its solvent-accessible surface area and the ligand's

molecular surface is described in terms of its matching surface description. The complementarily between the two surfaces amounts to the shape matching description that may help finding the complementary pose of docking the target and the ligand molecules. Another approach is to describe the hydrophobic features of the protein using turns in the main-chain atoms. Yet another approach is to use a Fourier shape descriptor technique.[13,14,15]

SIMULATION

The simulation of the docking process as such is a much more complicated process. In this approach, the protein and the ligand are separated by some physical distance, and the ligand finds its position into the protein's active site after a certain number of "moves" in its conformational space. The moves incorporate rigid body transformations such as translations and rotations, as well as internal changes to the ligand's structure including torsion angle rotations. Each of these moves in the conformation space of the ligand induces a total energetic cost of the system, and hence after every move the total energy of the system is calculated. The obvious advantage of the method is that it is more amenable to incorporate ligand flexibility into its modeling whereas shape complementarily techniques have to use some ingenious methods to incorporate flexibility in ligands. Another advantage is that the process is physically closer to what happens in reality, when the protein and ligand approach each other after molecular recognition. A clear disadvantage of this technique is that it takes longer time to evaluate the optimal pose of binding since they have to explore a rather large energy landscape. However grid-based techniques as well as fast optimization methods have significantly ameliorated these problems.[5]

Docking can be between.

- Protein / small ligand
- Protein / peptide
- Protein / protein
- Protein / nucleotide

MECHANICS OF DOCKING

To perform a docking screen, the first requirement is a structure of the protein of interest .Usually the structure has been determined using a biophysical technique such as x-ray crystallography, or less often, NMR spectroscopy. This protein structure and a database of potential ligands serve as inputs to a docking program. The success of a docking program depends on two components: the search algorithm and the scoring function. [5] The search space in theory consists of all possible orientations and conformations of the protein paired with the ligand. However in practice with current computational resources ,it is impossible to exhaustively explore the search space for this would involve enumerating all possible distortions of each molecule (molecules are dynamic and exist in an ensemble of conformational states) and all possible rotational and translational orientations of the ligand relative to the protein at a given level of granularity. Most docking programs in use account for a flexible ligand, and several attempt to model a flexible protein receptor. Each "snapshot" of the pair is referred to as a pose. A rigorous search algorithm would exhaustively elucidate all possible binding modes between the ligand and receptor. For a simple system comprising a ligand with four rotatable bonds and six rigid-body alignment parameters, the search space has been estimated as follows. The alignment parameters are used to position the ligand relative to the protein in a cubic active site measuring 103 Å3.If the angles are considered in 10 degree increments and translational parameters on a 0.5 Å grid there are approximately 4×108 rigid body degrees of freedom to sample, corresponding to 6×1014 configurations (including the four rotatable torsions) to be searched. This would require approximately 2000000 years of

computational time at a rate of 10 configurations per second. As a consequence only a small amount of the total conformational space can be sampled, and so a balance must be reached between the computational expense and the amount of the search space examined.[16]

A variety of conformational search strategies have been applied to the ligand and to the receptor. These include:

- Systematic or stochastictorsional searches about rotatable bonds
- Molecular dynamics simulations
- Genetic algorithms to "evolve" new low energy conformations.

LIGAND FLEXIBILITY

Conformations of the ligand may be generated in the absence of the receptor and subsequently docked.[17] Or conformations may be generated on-the-fly in the presence of the receptor binding cavity.[18] Force field energy evaluation are most often used to select energetically reasonable conformations.[19] But knowledge-based methods have also been used.[20]

RECEPTOR FLEXIBILITY

Computational capacity has increased dramatically over the last decade making possible the use of more sophisticated and computationally intensive methods in computer-assisted drug design. However, dealing with receptor flexibility in docking methodologies is still a thorny issue. The main reason behind this difficulty is the large number of degrees of freedom that have to be considered in this kind of calculations.5A single, fixed conformation, even the average provided by a crystal structure, may not be an adequate representation of the protein, unless the system is very rigid. Instead, even under standard equilibrium conditions, the native folded state of a protein is best characterized by a collection or ensemble of energetically nearly equivalent conformations. If the conditions are changed, the local minima and the population of these states may shift, eventually resulting in an observable change of the average structure. Also, the introduction of a ligand corresponds to a change of the environment that may lead to similar effects.[4]

DOCKING AND DE NOVO DESIGN METHODS

For the purpose of this review, a broad distinction is made between docking algorithms and de novo design methods. This is arguably subjective and in many cases significant overlap in methodology occurs between the two strategies. Examples of de novo design tools are BUILDER, CONCEPTS, CONCERTS, DLD/MCSS, Gens tar, Group-Build, Grow, HOOK, Ligand, LUDI, MCDNLG, SMOG and SPROUT. LUDI is given as an example of a de novo design tool applied to the docking problem.[16]

SCORING FUNCTION

Scoring functions are fast approximate mathematical methods used to predict the strength of the non-covalent interaction between two molecules after they have been docked. Most commonly one of the molecules is a small organic compound such as a drug and the second is the drug's biological target such as a protein receptor.[21] Scoring functions have also been developed to predict the strength of other types of intermolecular interactions for example between two proteins.[22] Or between protein and DNA.[23,24]

Scoring is actually composed of three different aspects relevant to docking and design:

- Ranking of the configurations generated by the docking search for one ligand interacting with a given protein; this aspect is essential to detect the binding mode best approximating the experimentally observed situation.
- Ranking different ligands with respect to the binding to one protein, that is,

prioritizing ligands according to their affinity; this aspect is essential in virtual screening.

- Ranking one or different ligands with respect to their binding affinity to different proteins; this aspect is essential for the consideration of selectivity and specificity.
- Scoring methods can range from molecular mechanics force fields such as AMBER, OPLS or CHARMM through to empirical free energy scoring functions or knowledge based functions. The currently available docking methods utilize the scoring functions in one of two ways. The first approach uses the full scoring function to rank a protein ligand conformation. The system is then modified by the search algorithm, and the same scoring function is again applied to rank the new structure.[16]

VIRTUAL SCREENING

Virtual screening (VS) is a computational technique used in drug discovery research. It involves the rapid in silico assessment of large libraries of chemical structures in order to identify those structures which are most likely to bind to a drug target, typically a protein receptor or enzyme. [25,26] Virtual screening has become an integral part of the drug discovery process. Related to the more general and long pursued concept of database searching, the term "virtual screening" is relatively new. Walters, et al. define virtual screening as "automatically evaluating very large libraries of compounds" using computer programs.[27]

DOCKING SOFTWARES

Auto Dock

Auto Dock uses Monte Carlo simulated annealing and Lamarckian genetic algorithm to create a set of possible conformations.LGA is used as a global optimizer and energy minimization as a local search method.

Dock

DOCK is one of the oldest and best known ligand-protein docking programs. The initial version used rigid ligands; exibility was later incorporated via incremental construction of the ligand in the binding pocket. As said DOCK is a fragment-based method using shape and chemical complementary methods for creating possible orientations for the ligand. These orientations can be scored using three different scoring functions; however none of them contain explicit hydrogen-bonding terms, solvation/desolvation terms, or hydrophobicity terms thus limiting serious use. DOCK seems to handle well a polar binding site and is useful for fast docking, but it is not the most accurate software available.

Gold

Gold has won a lot of new users during the last few years because of its good results in impartial tests. It has a good hit rate overall, however it somewhat when dealing with hydrophobic binding pockets. Gold uses genetic algorithm to provide docking of exible ligand and a protein with exible hydroxyl groups. The development of GOLD is currently focused on improving the computational algorithm and adding a support for parallel processing. GOLD has one of the most comprehensive validation test sets and is also available for use at CSC.[29]

V Life MDS

V life has provides following functions: Building polypeptides using V Life MDS, Molecular Docking using V Life MDS, Homology modeling using Biopredicta, Homology modeling using Biopredicta, Protein complex optimization using V Life MDS, Using alignment method in V Life MDS, Building molecules using V Life MDS, Conformational search using V Life MDS, Optimizing Molecules using V Life MDS, Using

miscellaneous utilities in V Life MDS, QSAR using V Life MDS.

Flex X

Flex X is another fragment based method using exible ligands and rigid proteins. It uses MIMUMBA torsion angle database for the creation of conformers. The MIMUMBA is an interaction geometry database used to exactly describe intermolecular interaction pattern

SIGNIFICANCE

A binding interaction between a small molecule ligand and an enzyme protein may result in activation or inhibition of the enzyme. If the protein is a receptor, ligand binding may result in agonism or antagonism. Docking may be applied to:

- Hit Identification – docking combined with a scoring function can be used to quickly screen large databases of potential drugs in silico to identify molecules that are likely to bind to protein target of interest (see virtual screening).
- Lead Optimization – docking can be used to predict in where and in which relative orientation a ligand binds to a protein (also referred to as the binding mode or pose).This information may in turn be used to design more potent and selective analogs.
- Bioremediation– Protein ligand docking can also be used to predict pollutants that can be degraded by enzymes. Estimating the binding affinity
- Searching for lead structures for protein targets
    - Comparing a set of inhibitors
    - Estimating the influence of modifications in lead structures
    - De Novo Ligand Design
- Design of targeted combinatorial libraries predicting the molecule complex
    - Understanding the binding mode / principle
    - Optimizing lead structures5

CONCLUSION

Molecular Docking is safe and easy to use tool helps in investigating, interpreting, explaining and identification of molecular properties using three-dimensional structures. Molecular docking tries to predict the structure of the intermolecular complex formed between two or more constituent molecules. The techniques are used in the fields of computational chemistry, computational biology and materials science for studying molecular systems ranging from small chemical systems to large biological molecules and material assemblies Most of the docking programs presently being used simulate the binding of a flexible ligand to a rigid biological receptor. This model does not reflect the actual physical process of binding and limits or in some cases even prevents the correct identification of potential drug candidates.

REFERENCES

1. Nadendla Rama, Rao (2004), “Molecular Modeling: A Powerful Tool for Drug Design and Molecular Docking”, ***Resonance***, 51-60.
2. (2011), Molecular Modelling from Wikipedia the free encyclopedia Available from: ***http://en.wikipedia.org/wiki/Molecular_modelling***.
3. Clauuen, Holger; Buning, Christian; Rarey, Matthias and Lengauer, Thomas (2001), “FLEXE: Efficient Molecular Docking Considering Protein Structure variations”, ***J. Mol. Biol.***, 308, 377-395.
4. Abraham, Donald and J, Burger’s (1998), “Medicinal Chemistry and Drug

Discovery", ***Docking and Scoring Functions/Virtual Screening.***, 1(06), 281-330.

5. Molecular Docking from Wikipedia, the free encyclopedia (2011), Available from: ***http://en.wikipedia.org/wiki/Docking_(molecular).***
6. Jorgensen, WL (1991), "Rusting of the lock and key model for protein-ligand binding", ***Science***, 254 (5034), 954-5.
7. Wei, BQ; Weaver, LH; Ferrari, AM; Matthews, BW and Shoichet, BK (2004), "Testing a flexible-receptor docking algorithm in a model binding site", ***J. Mol. Biol***., 337 (5), 1161-82.
8. (2011), Molecular Docking using VLifeMDS: VLife Product Documentation, Tutorial: ***BioPredicta***, VLife Sciences Technologies Pvt. Ltd.; 1-17.
9. Meng, EC; Shoichet, BK; and Kuntz, ID (2004), "Automated docking with grid-based energy evaluation", ***Journal of Computational Chemistry***, 13 (4), 505-524.
10. Morris, GM; Goodsell, DS; Halliday, RS; Huey, R; Hart, WE; Belew, RK and Olson, AJ (1998), "***Journal of Computational Chemistry***", 19 (14), 1639-1662.
11. Feig, M; Onufriev, A; Lee, MS; Im, W; Case, DA and Brooks CL (2004), " Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures", ***Journal of Computational Chemistry***, 25 (2), 265-84.
12. Shoichet, BK; Kuntz, ID and Bodian, DL (2004), "Molecular docking using shape descriptors", ***Journal of Computational Chemistry***, 13 (3), 380-397.
13. Cai, W; Shao, X and Maigret, B (2002), "Protein-ligand recognition using spherical harmonic molecular surfaces: towards a fast and efficient filter for large virtual throughput screening", ***J. Mol. Graph. Model***, 20 (4), 313-28.
14. Morris, RJ; Najmanovich, RJ; Kahraman, A and Thornton, JM (2005), "Real spherical harmonic expansion coefficients as 3D shape descriptors for protein binding pocket and ligand comparisons", ***Bioinformatics***, 21 (10), 2347-55.
15. Kahraman, A; Morris, RJ; Laskowski, RA and Thornton, JM (2007), "Shape variation in protein binding pockets and their ligands", ***J. Mol. Biol.***,368 (1), 283-301.
16. Taylor, RD; PJ, Jewsbury and JW, Essex1 (2002), "A review of protein-small molecule docking methods", ***Journal of Computer-Aided Molecular Design***, 16, 151-166.
17. Kearsley, SK; Underwood, DJ; Sheridan, RP and Miller, MD (1994), "Flexibases: a way to enhance the use of molecular docking methods", ***J. Comput. Aided Mol. Des.***, 8 (5), 565-82.
18. Friesner, RA; Banks, JL; Murphy, RB; Halgren, TA; Klicic, JJ; Mainz, DT; Repasky, MP; Knoll, EH; Shelley, M; Perry, JK; Shaw, DE and Francis, *et al*.(2004), "a new approach for rapid, accurate docking and scoring, Method and assessment of docking accuracy", ***J. Med. Chem.***, 47 (7),1739-49.
19. Wang, Q; Pang, YP and Romesberg (2007), "Preference of small molecules for local minimum conformations when binding to proteins", ***PLoS ONE***, 2 (9), 820.
20. Klebe, G and Mietzner, T (1994), "A fast and efficient method to generate biologically relevant conformations", ***J. Comput. Aided Mol. Des.***, 8 (5), 583-606.
21. Jain, AN (2006), "Scoring functions for protein-ligand docking", ***Curr. Protein Pept. Sci.***, 7 (5), 407-20.

22. Lensink, MF; Méndez, R and Wodak, SJ (2007), "Docking and scoring protein complexes:CAPRI 3rd Edition", ***Proteins Structure Function and Bioinformatics***, 69 (4), 704.
23. Robertson, TA and Varani, G (2007), "An all-atom, distance-dependent scoring function for the prediction of protein-DNA interactions from structure", ***Proteins.***, 66 (2), 359-74.
24. (2011), Scoring Function for Docking from Wikipedia, the free encyclopedia, Availablefrom: ***http://en.wikipedia.org/wiki/Scoring_functions_for_docking.***
25. Rester, U (2008), "From virtuality to reality-Virtual screening in lead discovery and lead optimization:A medicinal chemistry perspective",***Curr Opin Drug Discov level***., 11 (4), 559-68.
26. Rollinger, JM; Stuppner, H; and Langer, T (2008), "Virtual screening for the discovery of bioactive natural products", ***Prog Drug Res.*** 65 (211), 213-49.
27. Walters, WP; Stahl, MT and Murcko, MA (1998), "Virtual screening – an overview", ***Drug Discov. Today***, 3 (4), 160-178.
28. Virtual Screening from Wikipedia, the free encyclopedia (2011), Available from: ***http://en.wikipedia.org/wiki/Virtual_screening***.
29. Kaapro, Aatu and Ojanen, Janne (2002), "***Protein Docking***".