Molecular Docking Technique

Bhawna Poudyal
Department of Pharmaceutics
“Kr College of Pharmacy
Bengaluru, Karnataka, India

Abstract
New Molecular docking software is a technique approach which is used mainly in drug design and development and also to perform various analysis process. Molecular docking is a process that involves computerized platforms. It involves combination of molecules where it involves one molecule bound other molecules to form a bound complex with all over minimum energy requirements. It not only involves target binding but also involves identification of correct conformation of ligand within target binding and identifies the strength between the target a ligand. Now a day’s variety of docking programs are available in i.e. (Auto Dock, Auto Dock vina, Le Dock, r Dock, UCSF Dock) which evaluates binding affinity estimation an visualize the different dimensional structure of molecule. In recent years molecular docking technique has good progress in various field it greatly improves the efficacy and reduce the cost. This technique greatly introduces the principles. The main application of molecular Docking is virtual screening. Molecular Docking in Insilco-based structure widely used in drug discovery. It involves identification of novel compounds of therapeutic interest. The applications of Molecular Docking have greatly changed it also side by side illustrate about the newer uses of Docking an as well as prediction of adverse effect, poly pharmacology, drug responses, drug profiling target fielding, a various newer drug discovery approach.

Keywords: molecular docking technique, drug designing, CADD

Introduction
Docking is a method which involves orientation and a best attempt to find a matching between two molecules it involves binding of one ligand to the active site of protein receptor to form a complex. The best way involved in explaining molecular docking is “Lock and Key system “the steps involve in this system is.

- finding the better orientation for the key which will go best in opening up the lock.
- on the surface the key lock is present.
- on which the direction to turn the lock is given.

Hence, the protein can be taken as the lock the ligand can be thought as a key.

The binding between the biologically relevant molecule like protein, nucleicacid, carbohydrates lipid play role in signal transduction. Therefore, Docking predicts the site for binding the ligand in order to estimate the activity an activity of small molecules. Hence, ligand -Protein are involved in various biological processes with pharmaceutical actions thus the community involved in scientific approaches of an investigation of binding phenomenon during years developing various theories.

The first discovery of explanation was introduced by Emil Fischer in the year (1894) with the very known “lock and key system “which is a model for explaining in year 1958 Koshland introduce “Induce fit” system according to this theory the enzyme-substrate interaction, the ligand induces confirmational changes optimizing ligand -target interactions.

In year 1965 described the landscape for energy which provides the work for the protein confirmations a ligand preferentially binds to them.

As we know identification is one of the most difficult steps in molecular docking operation a drug design discovery technique. Therefore, numerous methodology astategy was established in determining the identification of the binding targets amolecule. Among them high throughput screening (HTP) was largely used from year 1990swith which have been used in determination of high cost a lowhit rate but on the other hand virtual high – throughput screening (vHTS) got great importance in high performance A computing.

During past years varieties of programming approaches was developed in molecular Docking such as –

For Docking program, the most critical scoring function used for sampling algorithms are divided into following – Shape matching, systemic search, stochastic search, force field, empirical, knowledge based.

~ 8 ~
Most recently new scoring functions was developed on basis of quantum mechanical (QM) a semi empirical quantum mechanical (SEQM) to capture binding affinity an identification of target site amolecule. The problem of sampling was only partially overcome it still had a huge challenge for all scoring functions to predict binding affinities of small molecules.

The studies providing information on evaluation was established in year 2011. Previously in year 2013 Damm - Ganamet et al publish the document of structure activity resource for community. Even though various Docking programmes was organised in past two decades it still remains difficult to understand the Docking programs for specific targets.

<table>
<thead>
<tr>
<th>Program</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auto Dock</td>
<td>It is free for academic use maintained by molecular graphics laboratory, Scripps Research Institute LaJolla. It is LGA based docking technique software.</td>
</tr>
<tr>
<td>Auto Dock Vina</td>
<td>It is local search global optimizer. It is free for academic use. maintained by molecular graphics laboratory, Scripps Research Institute LaJolla.</td>
</tr>
<tr>
<td>Le Dock</td>
<td>It is based on annealing of ligand position taking bonds formation into considerations using a knowledge of function scoring. It is free for academic a maintained by Lephar Research.</td>
</tr>
<tr>
<td>r Dock</td>
<td>R Dock is usually a combination of stochastic a deterministic search technique which is used for low energy-based ligand posing. maintained by r Dock development team. it is free for academic use.</td>
</tr>
<tr>
<td>GOLD</td>
<td>It is GA based Docking technique. it is product of collaboration between university of Sheffield, Galaxo Smith Kline, a Cambridge crystallographic data centre.</td>
</tr>
<tr>
<td>Surflex - Dock</td>
<td>It is docking technique based on protocol that can be automatically generated.</td>
</tr>
<tr>
<td>GLIDE</td>
<td>It is exhaustive based docking technique. It has extra precision standard precision a high throughput virtual screening.</td>
</tr>
</tbody>
</table>

Types of docking –
- Rigid docking (lock and key); in this both receptor a ligand is rigid in geometry.
- Flexible docking; in this type of docking the rotation of molecules is performed in every rotation.

The features of evaluated docking programmes

- Electrostatic forces.
- Steric forces.
- Solvent related forces.

**Key stages in docking**
- Target selection a preparation
- ligand re a selection
- Docking-Evaluating docking result.

**Approaches of Molecular Docking**

For performing molecular docking two types of approaches are used Stimulation approach Here the target a ligand is being separated by a physical distance a then the ligand is allowed to bind by groove of target “after definite time of moves “in its conformations. In this the moves occurs as no of variations to the structure of ligand either internally or externally. the ligand in every conformation leads to release of energy as “total energy utilized”. this approach is advantageous in sense that is in more compatibility to accept ligand flexibility. Therefore, there is more recognition between the ligand a receptor. However, the approach takes longer duration to estimate no of docked conformers due to Large energy for each conformation. Recently fast optimization method has been developed which is user friendly.

**Shape Complementarity Approach**

The ligand employs ligand a target as surface structural feature that provides their molecular interaction. Here the surface target is shown with respect to solvent accessible surface area a ligand molecular surface is showed in terms of matching surface illustration. For example, in protein target molecules hydrophobicity estimated by no of turns in main chain atoms. It leads to scanning of no of ligands a finding possible binding property of ligand on target surface.

**Applications of Molecular Docking**

The binding of small molecule ligand an enzyme protein may result in activation an inhibition of enzymes. the main propose of ligand binding is antagonism or agonism. The main importance of docking technique is in drug design.
process most drug activity are organic molecule an can be applied for:

- Hit identification - docking when combined with scoring function results in potential of drug in silico to identify the molecules that are likable to bind protein of target interest.
- Lead optimization - docking can be used in predicting the orientation in which the ligand would bind the protein this is used in more proper implementation of drug design.
- Bioremediation - it’s used in prediction of pollutants that degrade enzymes produced.
- Docking is mainly used in determination of protein - protein docking.
- It is used in determination of side effect when used with another molecule.
- It is used in designing of drugs.
- It is used in study of geometry of particular complex.
- Drug – DNA interactions it establishes a correlation between drugs molecular structure and its cytotoxicity.
- Molecular docking is usually used in demonstrating the feasibility of any biochemical reaction as it is carried out before any experimental part investigation.
- Its main application is characterization of small molecules.

As compared to protein the nucleic acid has got less importance in drug targets. Drug which is known to interact with DNA include; all binders. The variables in DNA I less as compared to RNA they are more complex the those of protein but make RNA more attractive as targets. In years no of NMR an RNA based targets has been come into consideration in literature. The only difference appears in binding location. In case of protein the binding location is typically deep an interior is well separated by solvent whereas, in RNA target binding is usually located in surface a relatively exposed to solvents. Based on Dock screening are identified and capable of RNA duplex but no

Limitations

The major limitation of molecular docking technique is lack of confidence in the ability of scoring function to give binding an accurate energy required. The binding capability is hardly predicted between inter- molecules like solvation effect. In addition to some inter molecules interactions rarely considered in scoring function which have their significance for example - Transthyretin thyroxine complex - where energies function is failed that is transthyretin - thyroxine complex.

It’s still the unsolved problem to deal accurately with the water molecule in binding process during docking process, which is tough a lead to lots of attention in near future due to many reason:

- X -ray crystal structures lack of information of hydrogen, due to inefficient scattering by smaller atoms.
- Not knowing the exact position of hydrogen leads to in accuracies in identifying in water molecule which might be acting bridging molecule between the ligand and receptor.
- The theoretical approach is available to accurately predict how water molecules are affected by ligands a how strong the effect is.
- An on the other hand it is difficult to predict that how many water molecules in binding pocket would be replaced by potential ligands a how the hydrogen bonding network would be disturbed by ligand binding.
- One of the major challenges faced by docking field is rigid receptor.
- A protein adopts many confirmations depending upon the properties of ligand to bind as performed by rigid receptor corresponds to single receptor confirmations, which leads to negative results

Literature Review on Molecular docking Technique

Kolb et al (3) used molecular docking to screen library of close to 1 million commercially available lead like structures the field of computer aided drug design a discovery (CADD) in successive years (CADD) has gained success in growing area. many giant pharmaceutical companies adopt CADD for lead discovery. In the review the whole description on molecular docking is mentioned. To improve molecular docking modelling played a major role in multiple context of some reactions at current stage the technology does not fall under current scope of molecular docking, due to fact that the process is complex an it’s difficult to manage the interactions occur. In order to understand how chemistry works in nature more than two factor an methodology were predicted. The informatics market is estimated to develop by 1.5 billion in 2016 to 2.84 billion in yare 2022an may further keep on expanding

Kuntz [4]. The earliest work using structural shape contacts in which the fitting the outlines enables best possible configurations between two proteins identified. a little strategy was performed later undertaking algorithms by
Kuntz collaboration to continue searching for various configurations using geometry between receptor and ligand. The first discovery of explanation was introduced by Damm - Ganam et al. [4] previously in year 20113 publish the document of structure activity resource for community. Even though various Docking programmes was organised in past two decades it still remains difficult to understand the Docking programs for specific targets.

Emil Fischer in the year (1894) [5] The first discovery of explanation was introduced with the very known “lock and key system” which is a model for explaining In year 1958 Koshland introduce “Induce fit” system- according to this theory the enzyme-substrate interaction, the ligand induce confirmational changes optimizing ligand -target interactions. In year 1965 described the landscape for energy which provides the work for the protein confirmations a ligand preferentially binds to them the studies providing information on evaluation was established in year 2011.

(Morris et al., 2009) [6] performed with Auto Dock molecular docking. The protein molecules were processed by adding all hydrogen, merging non-polar hydrogen atoms using Auto Dock Tools. The charges were assigned using the Gasteiger method. The torsions were fixed for the ligand. The grid box of 60 × 60 × 60 with 3.75 Å was set around the active sites (ASP-127, TRP-179, ASN-234, GLU-235, TYR-313, GLU-340, TRP-381, ASN-396) where the ligand can interact. The rigid grid box was attained using Auto grid. It was followed by Auto Dock with Lamarckian Genetic Algorithm.

Some of the most popular docking programs are: - Auto Dock (Morris et al., 2009), Auto Dock Vina [7] (Trott and Olson, 2009), GOLD [8] (Jones et al., 1997), Quick Vina-W and Glide [9] (FlexX, Rarey et al., 1996), DOCK (Kuntz et al., 1982), ICM (Avagyan et al., 1994), (Ferreira et al. 2015; Guedes et al. 2014) [10]; Molecular docking is a prominent method for structure-based drug design, due to the prediction of the binding-conformation of molecular ligands to the target receptor binding site. Characterization of the active binding behaviour plays an important role in rational design of novel pesticides, herbicides, insecticides and fungicides After (Nataraja et al., 2017) [11], doing some research, have found that most appropriate software for our research would be Auto Dock Vina since it is newly designed and improved version of the Auto Dock program. This version adopted a new knowledge-based scoring function with a Monte Carlo sampling technique and the Brayden-Fletcher-Goldfarb-Shannon (BFGS) method for local optimization.

Computational methodology to study ligand -protein binding a molecular-Docking

Molecular docking occurs in three steps.
- Definition of the structure of the target
- Location of binding site
- Determination of the binding mode

Ways to bind two molecules together
- Need to quantify or rank solution
- Scoring function or force field
- Experimental structure may be amongst one of several predicted solution
- Selection of method
- In simple term for molecular docking, one must -
  1. Find a set of compounds to start with example from inspecting known ligands for a protein that is substrate in an enzyme.
  2. Compounds from a screening experiment of a combinatorial library.
  3. Compounds from other visualizing software is established.
  4. From varying another ligand.
  5. Virtual screening using a fast-docking algorithm.
  6. de novo design using fragments of compounds get several hundreds to thousands of ligaments to start with.

A. Genetic algorithms (GA) [12], are an interesting application of the stochastic search, which have been successfully used in molecular docking programs such as Auto Dock and Gold. The GA algorithm addresses the high computational cost associated with stochastic methods by applying concepts of the theory of evolution and natural selection. As a first step, the algorithm encodes all of the structural parameters of the initial structure in a chromosome, which is represented by a vector. Starting from this chromosome, the random search algorithm generates an initial population of chromosomes covering a
wide area of the energy landscape. This population is evaluated and the most adapted chromosomes (i.e., those with the lowest energy values) are selected as templates for the generation of the next population. This procedure decreases the average energy of the chromosome ensemble by transmitting favourable structural characteristics from one population to another, reducing therefore, the conformational space to be explored. The GA routine is recursively executed and, after a reasonable number of conformational search-and-evaluation cycles, converges to a conformation (chromosome) corresponding to the global energy minimum.

B. Scoring functions \[^{[13]}\] Are categorized in the three following groups: force-field-based, empirical, and knowledge-based functions. Force-field-based scoring functions estimate the binding energy by summing the contributions of bonded (bond stretching, angle bending, and dihedral variation) and non-bonded terms (electrostatic and van der Waals interactions) in a general master function. This type of scoring function applies an ab initio method to calculate the energy associated with each term of the function using the equations of classical mechanics. A major limitation of force-field-based methods is their inaccuracy in estimating entropic contributions. This shortcoming is due to the lack of a reasonable physical model to describe this phenomenon. Furthermore, the solvent is not explicitly considered, stopping the estimation of de-solvation energies.

Empirical scoring functions are another type of evaluation method. Each term of the function describes one type of physical event involved in the formation of the ligand-receptor complex. These include hydrogen-bonding, ionic and apolar interaction, as well as de-solvation and entropic effects, first step in the development of an empirical function, a series of protein-ligand complexes with known binding affinities is used as a training set to perform a multiple linear regression analysis. Then, the weight constants generated by the statistical model are used as coefficients that adjust the terms of the equation. A drawback of empirical scoring functions is their dependence on the accuracy of the data used to develop the model. However, because of the simplicity of the employed energy terms, empirical functions are faster than force-field-based methods. Surflex and Flexare broadly used molecular docking programs using empirical scoring functions. Every scoring function has its virtues and limitations. Therefore, the simultaneous use of different scoring methodologies has been increasingly employed as a way to obtain a consensus scoring. This can be very useful, as it combines the advantages and simultaneously attenuates the shortcomings of each method.

C. Molecular Dynamics: Flexibility of the target binding site is an essential but frequently overlooked aspect to be considered in molecular docking. Enzymes and receptors can undergo conformational changes during the molecular recognition process. In some cases, these structural rearrangements are small and the ligand fits in a binding site with little mobility. Otherwise, some proteins undertake significant conformational changes, which can involve elements of secondary and tertiary structure. Such flexibility issues can be handled by the use of techniques such as MD. Regardless of its usefulness MD has its limitations. Among them, we can stress the high computational cost demanded by the simulation of large systems, which usually consist of thousands of atoms when ligand-receptor complexes are under study. Some of the conformational changes undertaken by receptors during molecular recognition occur on time scales exceeding the available computational capacity. Despite its limitations, MD is able to deliver important contributions to SBDD, especially when combined with other molecular modelling methods, such as molecular docking.

A. Docking risk
The docked and native poses appeared to be correlated with the size and number of rotatable bonds. It has been widely acknowledged in the literature that re-docking accuracy decreases as the number of rotatable bonds increases, a phenomenon that is known to be independent of which docking program is used for the larger, dinucleotide structures in the re-docking test set which nucleotide fragment (i.e., nicotinamide or adenosine) where in the binding pocket, given the self-similarity present in the conjoined phosphor ribose groups.

Example of docking by binding to active site
- HIV -1 protease is target receptor
- Aspartyl is group of active site present for receptor to bind
- Generation of molecular surface for receptors
- Generate spheres to fill the active site of receptor. The sphere become potential location for ligand atoms.
- Sphere centers are matched with ligand atoms to determine the orientation for the ligand.
- Find the best ranking or score.

There are three scoring schemes
a. Shapescoring.
   b. Electrostatic scoring.
   c. Force field scoring.

Methods of Molecular Docking
Molecular docking can demonstrate the feasibility of any biochemical reaction as it is carried out before the experimental part of any investigation. There are some areas, where molecular docking has revolutionized the findings.
Following are the methods for performing molecular docking:

1. Protein-Protein Docking - involves the prediction of binding between two protein structures so as to form a protein complex using features such as steric and physicochemical complementarity at the protein-protein interface. This involves the prediction of conformational changes between unbound and bound structures. This is possible using structural data with a deeper understanding of the fundamental principles of protein interactions together with available advanced computational capabilities.

2. Protein-Ligand Docking - can be further classified into three types that are based on the structure of the Protein and Ligand:

   2.1. Rigid ligand and rigid receptor docking
   When the ligand and receptor are both treated as rigid bodies, the search space is very limited, considering only three translational and three rotational degrees of freedom. In this case, ligand flexibility could be addressed by using a pre-computed set of ligand conformations, or by allowing for a degree of atom-atom overlap between the protein and ligand.

   2.2. Flexible ligand and rigid receptor docking
   For systems whose behaviour follows the induced fit paradigm it is of vital importance to consider the flexibilities of both the ligand and receptor since in that case both the ligand and receptor change their conformations to form a minimum energy perfect-fit complex.

   2.3. Flexible ligand and flexible receptor docking
   The intrinsic mobility of proteins has been proved to be closely related to ligand binding behaviour and it has been reviewed by Teague. Incorporating receptor flexibility is a significant challenge in the field of docking. Ideally, using MD simulations could model all the degrees of freedom in the ligand-receptor complex.

3. Protein-Peptide Docking - the interest in peptide therapeutics triggered the rapid development of new techniques dedicated to protein-peptide docking which are being increasingly incorporated into the drug discovery and design process [14-16].

**Aim And Objective**

**Aim**

- The main aim of docking technique is to find the best way through which a ligand would bind an active site, to achieve optimized conformational for both receptor a ligand.
- To achieve orientation between protein a ligand that all the energy utilized is minimized.
- its main importance is in identification of its binding site that is correct binding site.
- Prediction of binding affinity.
- The importance in designing if drug.

**Objective of molecular docking**

A. Ligand-protein docking is an optimization problem based on predicting the position of a ligand energy.

B. Molecular docking problems are traditionally tackled with single-objective, as well as with multi-objective approaches, to minimize the binding with the lowest binding energy in the active site of the receptor.

C. A novel multi-objective formulation that considers: The Root Mean Square Deviation (RMSD) difference in the coordinates of ligands and the binding (intermolecular) energy, as two objectives to evaluate the quality of the ligand-protein interactions.

D. To determine a set of representatives that is multi-objective algorithms.

E. The performances have been assessed by applying two main quality indicators intended to measure convergence and diversity of the fronts.

F. In addition, a comparison with LGA, a reference single-objective evolutionary algorithm for molecular docking (Auto Dock) is carried out.

G. This new multi-objective approach shows an improvement over the ligand-protein docking predictions that could be promising in in silico docking studies to select new anticancer compounds for therapeutic targets that are multidrug resistant.

H. The significance of molecular docking is also understood in several environmental and industrial research, in order to untangle the interactions among macromolecules of non-medical interest.

I. This emphasizes the involvement of computational techniques in the aforementioned fields to expand our knowledge on macromolecular interacting mechanisms. Another knowledge-based approach is the use of three-dimensional similarity information from co-crystallized ligands as an additional Scoring schemes can also be improved by tailoring them to a specific target site, to designed sites or to multiple related receptors. For example, altered binding sites that emphasize distinct chemical features can be applied to specifically analyse or electrostatic contributions, hydrophobic interactions or salvation energies of scoring functions, as has been demonstrated in docking studies on structures of mutated T4 lysozyme active site.

**Conclusion**

When comes to developments its quite relevant that the stories highlighted above an many more other developments can find place in literature related to computerised aided drug design, that is silico approaches with biophysical data combination, toxicology clinical studies which leads to high importance in drug design a discovery. It assistscomputer-based ideas which gives new ideas to personnel in exploring an bringing theories together experimentally an bringing solutions to all problems occurred including cost. Therefore, many more challenges are still remaining like role of water molecules an solving the limitations aproblemaring, solvents effect, entropiceffects, receptor flexibility etc. This is a huge challenge a proof to huge discoveries which provides information about utilisation of computerised tools in drug design.

**References**


2. Holt PA, Chairs JB, Trent JO. Molecular docking of intercalators and groove-binders to nucleic acids using


12. Ferreira LG, Dos Santos RN, Oliva G, Andricopulo AD. Molecular docking and structure-based drug design 2015.


