WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A REVIEW ON MOLECULAR DOCKING

Shaikh swaleha[1], Prof. Thange Trupti[2]

[1,2]Pratibhatai Pawar college of pharmacy, WadalaMahadev, Shrirampur, Maharashtra.
Prov. Thange Trupti, Pratibhatai Pawar college of pharmacy, WadalaMahadev, Shrirampur, Maharashtra.

Abstract: -

Molecular docking has become an increasingly important tool for drug discovery. In this review, we present a brief introduction of the available molecular docking methods, and their development and applications in drug discovery. The differences in and performance of available docking software are also discussed. What is molecular docking and its types are studied. Flexible receptor molecular docking approaches, the brief and indetailed process of the molecular docking is explained. Three application examples of molecular docking approaches for drug discovery are provided.

Keywords: - Flexible docking, Rigid docking, Molecular docking, Molecular docking Procedure, Software's, Application.

1. INTRODUCTION: -

Virtual screening (VS) is one such computational method that has become popular in the drug discovery process. VS optimizes drug discovery by highlighting molecules that fit the target. To achieve this, two distinct methods of VS are used, ligand-based and structure-based.[1]
To show how likely a specific ligand is to attach to a target protein, structure-based approaches use molecular docking. With this method, interactions between potential candidates and the target protein are modelled down to the atom level. This enables researchers to anticipate how the potential molecules will behave when they attach to proteins, enabling them to choose molecules with desired behaviour and discard those that exhibit undesired behaviour.[2]
Molecular docking is used in structure-based approaches to forecast the likelihood that a given ligand will attach to a certain target protein. With the use of this technology, interactions between potential candidates and the target protein may be modelled at the atom level. By predicting how potential molecules will behave when they bind to proteins, scientists are able to choose molecules that display desired behaviour and discard those that do not.[3]
The structure of the intermolecular complex produced between two or more molecules is predicted using molecular docking. The protein-ligand interaction is the most intriguing instance because of its medical uses. Small molecules known as ligands interact with the binding sites of proteins. Protein regions known as binding sites are involved in the synthesis of many substances. There are a variety of mutual conformations that could lead to binding. These are sometimes referred to as binding modes.[4]

2. History: -

A computer programme was typically used at the end of the modelling process to distinguish between the relatively few configurations that were left after all of the heuristic constraints had been imposed. Complex modelling in the 1970s centred around manually identifying features on the surfaces of the interactors and interpreting the consequences for binding, function, and activity. Computers were originally employed in a study on the interaction of haemoglobin in sickle-cell fibres. The work on the trypsin-BPTI complex came next in 1978. Using a scoring mechanism that favoured wide interface areas and pairs of molecules in touch but not sharing the same space, computers were able to distinguish between good and bad models. One interaction centre was employed in the computer's simplified model of the interacting proteins, one for each residue. Hand-searching was used to find advantageous electrostatic interactions, including hydrogen bonds. [5]
More complicated structure determinations were made in the early 1990s, and computational power was much more readily available. The attention shifted to creating generalised tools that could be used to study any number of complexes at a reasonable processing cost with the advent of bioinformatics. Any specific previous knowledge could still be added at the stage of choosing between the highest-ranking output models, or be provided as input if the algorithm catered for it. The new approaches were designed to work even in the absence of phylogeny or experimental hints. The correlation approach, an algorithm that utilised the fast Fourier transform to greatly enhance scalability for assessing coarse shape complementarity on rigid-body models, was published in 1992. In 1997, it was expanded to include electrostatics. [6]
Six research groups attempted to predict the complexed structure of TEM-1 Beta-lactamase with Beta-lactamase inhibitor protein (BLIP) in the first blind trial, the findings of which were published in 1996. The exercise highlighted the need to accommodate conformational change and the challenge of identifying conformers. Additionally, it served as a model for the 2001-debuting CAPRI evaluation series.[7]

3.Classification of Molecular Docking: -

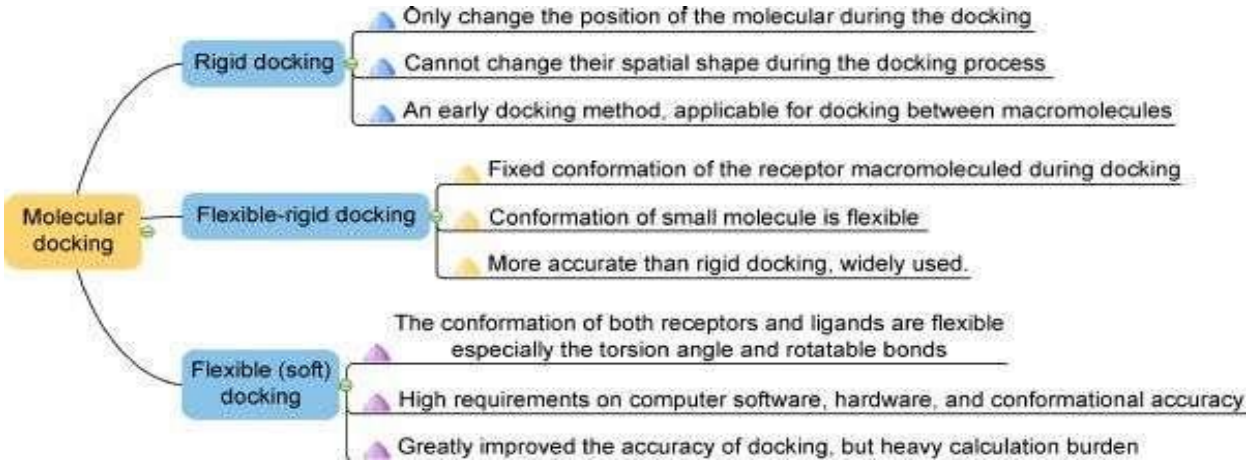

Fig 3: - classification of the molecular docking.

3.1. Rigid Docking: -

In the process of Rigid Docking calculation, the conformation of the ligands and receptors does not change, only the spatial position and posture of the two molecules change. In this kind of docking simulation, the spatial conformation of the ligand and the receptor is regarded as fixed. Namely, this docking method is the most convenient due to the simplest calculation difficulty and calculation amount. [8] Therefore, it is suitable for investigating the docking system with relatively large structures, such as the protein–protein and protein–nucleic acid complexes. In this field, Stoddard et al. treated the ligand and acceptor backbone structures with rigidity, and successfully implemented the docking simulation of maltose and protein by the binary docking method.[9]

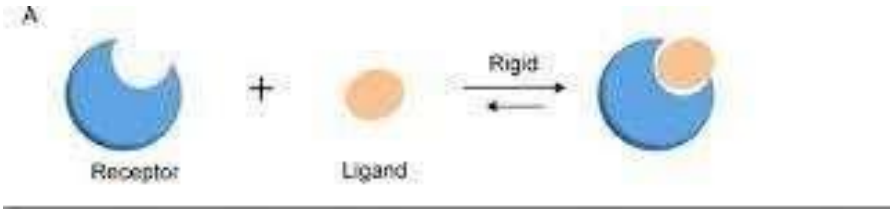

Fig3.1: – Rigid docking

3.2. Flexible Docking: -

During the flexible docking calculation, the conformation of ligand and receptor is allowed to change freely. Because this kind of docking simulation is of high accuracy, and closest to the real docking situation, it is often used to accurately investigate the recognition between two molecules. However, due to the geometric growth of variables with the number of atoms in the system, the flexible docking method is computationally intensive and time-consuming and requires high requirements on computer software and hardware systems. The most representative molecular docking software is FlexX, and Mangonietal. have used flexible ligands to dock with flexible receptors in this research area previously.[10]

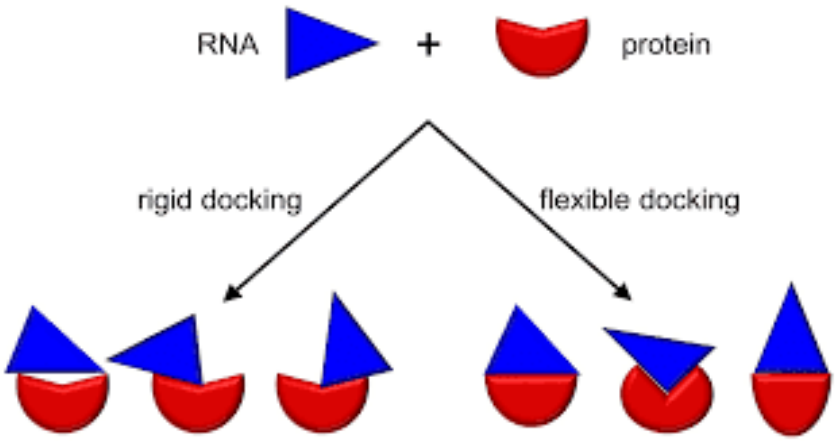

Fig3.2: - Flexible docking.

4. Molecular Docking : -

Molecular docking is a method which analyses the conformation and orientation (referred together as the "pose") of molecules into the binding site of a macromolecular target. Searching algorithms generate possible poses, which are ranked by scoring functions. [12]

What is molecular docking

Molecular docking is a key tool in structural molecular biology and computer-assisted drug design. The goal of ligand-protein docking is to predict the predominant binding mode(s) of a ligand with a protein of known three-dimensional structure. Successful docking methods search high-dimensional spaces effectively and use a scoring function that correctly ranks candidate dockings. Docking can be used to perform virtual screening on large libraries of compounds, rank the results, and propose structural hypotheses of how the ligands inhibit the target, which is invaluable in lead optimization. The setting up of the input structures for the docking is just as important as the docking itself, and analyzing the results of stochastic search methods can sometimes be unclear. This chapter discusses the background and theory of molecular docking software, and covers the usage of some of the most-cited docking software.[13]

5.Molecular Docking Procedure: -

5.1 Manual docking Molecular: -

Modelling can be used to dock, or fit, a molecule into a model of its binding site. If the binding groups on the ligand and the binding site are known, they can be defined by the operator such that each binding group in the ligand is paired with its complementary group in the binding site. The ideal bonding distance for each potential interaction is then defined and the docking procedure is started. The program then moves the molecule around within the binding site to try and get the best fit as defined by the operator. In essence, the procedure is similar to the overlay or fitting process, only this time the paired groups are not directly overlaid but fitted such that the groups are within preferred bonding distances of each other. Both the ligand and the protein remain in the same conformation throughout the process and so this is a rigid fit. Once a molecule has been docked successfully, fit optimization is carried out. This is essentially the same as energy minimization, but carried out on the ligand–target protein complex. Different conformations of the molecule can be docked in the same way the interaction energies measured to identify which conformation fits the best.[14]

5.2 Automatic docking: -

A variety of docking programs now exist that can automatically dock ligands into a binding site with the minimum of input from an operator.

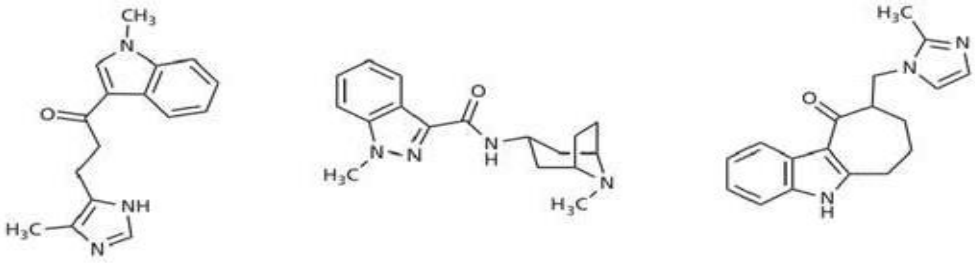

Fig5.2A: - Test structures

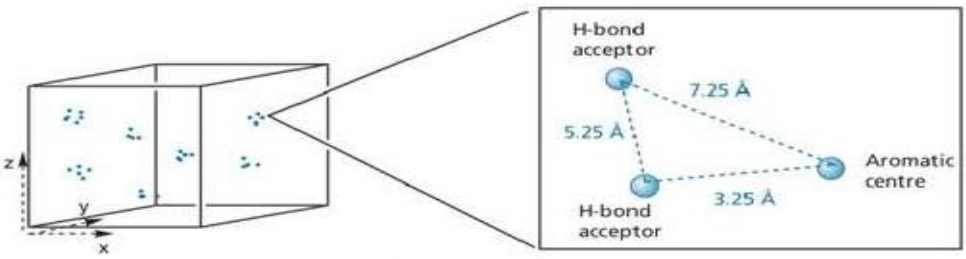

Fig 5.2B: - Pharmacophore Plot

They are also amenable to studying many diff erent molecules automatically. Indeed, an important application of automatic docking programs is to carry out virtual screening of hundreds of diff erent molecules with the aim of identifying new lead compounds that will interact with the target. Virtual screening can be seen as complementary to biological screening in that the former can identify the structures from a chemical 'library' that are most likely to bind to the target. Th ese can then be given priority for biological screening, making the latter more efficient. For virtual screening to be effective it has to use efficient algorithms which not only dock each molecule realistically, but also give an accurate 'score' of the relative binding energies of the

molecules concerned. Moreover, for each molecule studied, the docking program is likely to generate several different orientations or binding modes. It is necessary to score all of these in order to identify the most likely binding mode in terms of how well it fits the space available and how many intermolecular interactions it can form with the binding site. The calculations required for docking and scoring have to be rapid in order to process the number of molecules involved in a reasonable time period, but they also have to be accurate enough to give a good measure of relative binding energies. This is a difficult compromise to make as increasing the speed at which an algorithm operates involves assumptions or short cuts that inevitably reduce the accuracy of the calculation. As a result, this is an area of intense research interest in the development of new and improved docking programs. For reasons of space, it is not possible to go into the mathematical details of docking algorithms, and so this section focuses more on the general methods by which automatic docking can be carried out. The simplest approach to automatic docking is to treat the ligand and the macromolecular target as rigid bodies. This is acceptable if the active conformation of the ligand is known or if the ligand is a rigid cyclic structure. At the next level of complexity, the target is still considered as a rigid body, but the ligand is allowed to be flexible and can adopt diff erent conformations. Th e most complex situation is where both the target and the ligand are considered to be flexible. This last situation is extremely expensive in terms of computer time, and most docking studies are carried out by assuming a rigid target .

5.3 Defining the molecular surface of a binding site: -

In order to carry out docking calculations, it is necessary to know the structure of the protein target and the nature of the binding site. This can be obtained from an X-ray crystal structure of the protein which can be downloaded onto a computer. The amino acids lining the binding pocket can then be identified. The next step is to define the molecular surface of the binding site. One could do this by defining each atom within the binding site by its van der Waals radius, but this results in an extensive surface area, much of which would be inaccessible to a ligand (Fig. 5.3A). A simpler molecular surface can be defined by identifying the parts of the van der Waals surface that are accessible to a solvent molecule. In practice a probe sphere of radius 1.4–1.5 Å is used to represent a water molecule and this is 'rolled over' the surface of the binding site (Fig. 5.3A). Convex surfaces shown in dark blue are where the probe sphere makes contact with the van der Waals surfaces of a particular atom. Concave surfaces shown in light blue are known as re-entrants and represent how fast the probe atom can access the space between the atoms of the binding site. In this area, the probe is in contact with two or three atoms. This kind of molecular surficial so referred to as a Connolly surface. The surface is actually represented by a regular distribution of points or dots, and the crucial ones for docking are those on the convex surfaces. Each one of these has a vector associated with it which points into the binding site. The direction of the vector corresponds to the normal of the surface at that point and so it is a mathematical indication of curvature.[14]

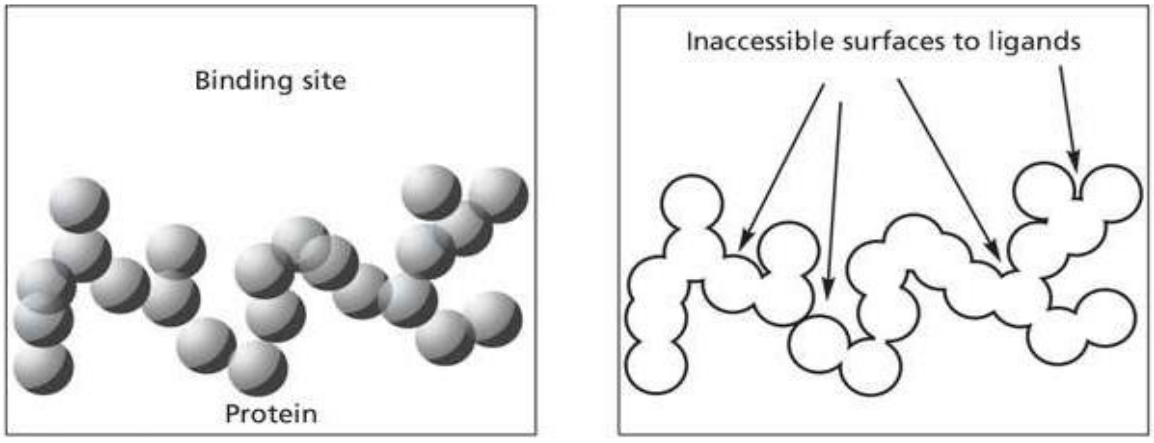

Fig5.3A: - Defining the surface of the binding site by atoms and van der waals surfaces.

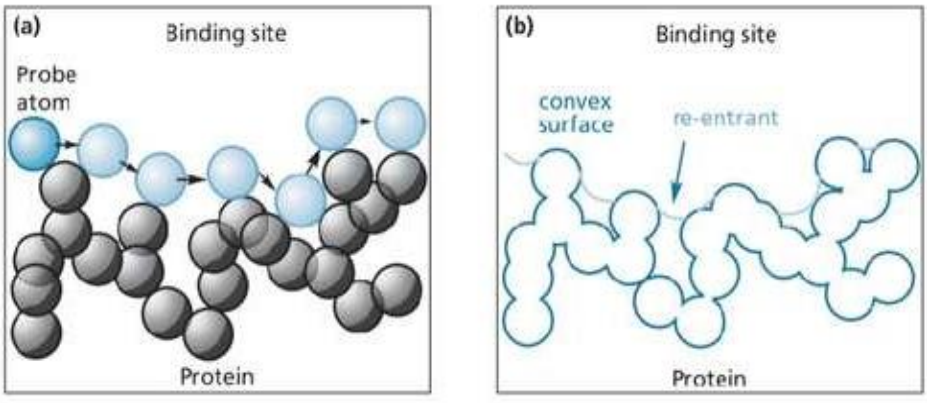

Fig5.3B: - Defining the Connolly surface of binding site with the probe atom, the Connolly surface shown in blue

5.4 Rigid docking by shape complementarity

The first problem with any docking program is how to position the ligand within the binding site. If you or I were handed real models of the target and the ligand, we would consider the space available in the binding site, eye up the ligand, and judge how we could place it into the binding site before we actually do it. In other words, humans have a spatial awareness, which includes the ability to assess the shape of an empty space. This does not come naturally to computers, and the empty space of a binding site has to be defined in a way that a computer program can understand before ligands can be inserted. The DOCK program was one of the earliest programs to tackle this problem. The Connolly surface is first defined, then the empty space of the binding site is defined by identifying a collection of differently sized spheres which will fill up the space available and give a 'negative image' of the binding site (Fig. 5.4A). This is achieved as follows.

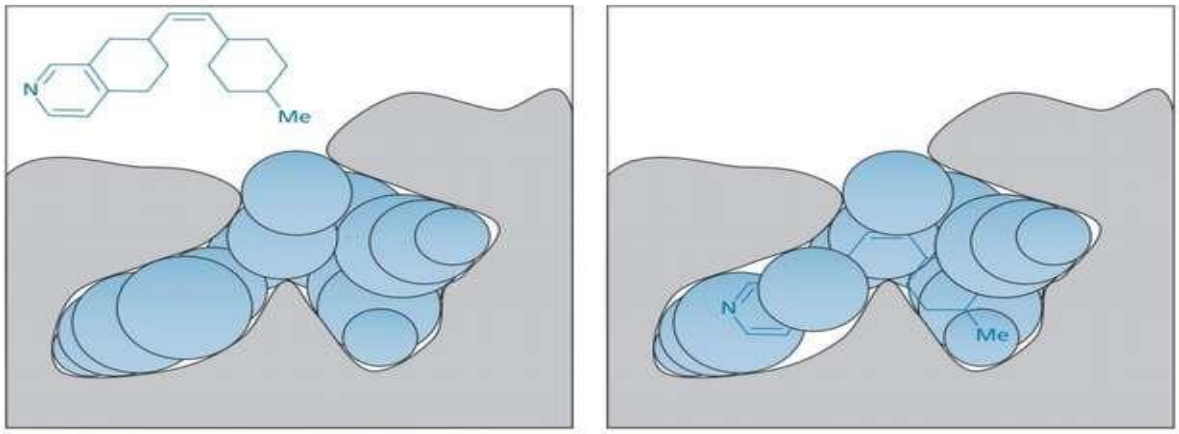

Fig5.4A: - The docking program.

For each dot representing the molecular surface, spheres are constructed that touch that dot plus one other dot on the molecular surface. Therefore, if there are n dots representing the molecular surface, n–1spheres will be created at each of the dots. This represents a massive number of spheres and so it is necessary to whittle these down. Th e number of spheres can be reduced significantly as follows:[14]

• For each dot on the molecular surface, the sphere of smallest radius touching it is chosen. This ensuring that none of the spheres chosen intersects the molecular surface;

• There are several dots associated with the surface of a particular atom and each of these now has one sphere associated with it. The next filtering process is to select the sphere with the largest radius. Once this has been completed, the number of spheres left is the same as the number of atoms lining the binding site. Spheres are allowed to overlap and the centre of each sphere accurately defines a unique position of 3D space within the binding site. Each sphere representing the binding site can be considered as a pseudoatomic and so it is now possible to carry out an overlay operation. where ligand atoms are matched with pseudoatoms then overlaid. However, how does the programmer decide which ligand atom and pseudoatom should be matched? One could try out every possible combination, but this would take up far too much computer time. Instead, a systematic matching operation takes place called distance matching or clique searching. Firstly, the distances between each of the ligand atoms are measured. This repeated for all of the pseudoatoms. These distances are then used to identify which ligand atoms and pseudoatoms can be matched up. The operation takes place as follows. A graph is prepared where each ligand atom (1, 2, 3, …) is matched to each of the receptor spheres (A, B, C, …) to give a list of paired atom/pseudoatoms (1A, 1B, 1C…, 2A, 2B, 2C….3A, 3B, 3C…, etc.). Th e next stage is to identify whether two of these pairs are compatible, for example is the pairing 1A possible at the same time as the pairing 2C? This is done by comparing the distance between the ligand atoms 1 and 2, with the distance between the receptor spheres A and C. If the distances are similar, then they are compatible. This process is now repeated for further pairings to see if they are compatible with those already identified. The minimum number of pairings required for an acceptable docking is four. The whole procedure is repeated systematically for each ligand atom to find a variety of matches which will eventually lead to different docking modes. As an example, consider a ligand represented by atoms 1–10 and a binding site represented by pseudoatoms A–G (Fig. 5.4B). The atom/pseudoatom pairs 1A, 6F, 7G, and 10E would be identified as compatible for the docking operation as the distances between the specified ligand atoms match the distances between the specified pseudoatoms. Once this procedure has been carried out, the actual docking process can take place. Docking then involves an overlay where ligand atoms are fitted onto their paired Pseudoatoms. For example, in Fig. 5.4B, ligand atoms 1, 6, 7, and 10 are matched to pseudoatoms A, F, G, and E respectively (Fig. 5.4B). This process is repeated for all the other possible matches to give a number of docking or binding modes. Note that this docking procedure is carried out purely in terms of steric complementarity (i.e. whether selected ligand atoms can match up with selected pseudoatoms). It takes no account of possible binding interactions, either favorable or unfavorable.[14]

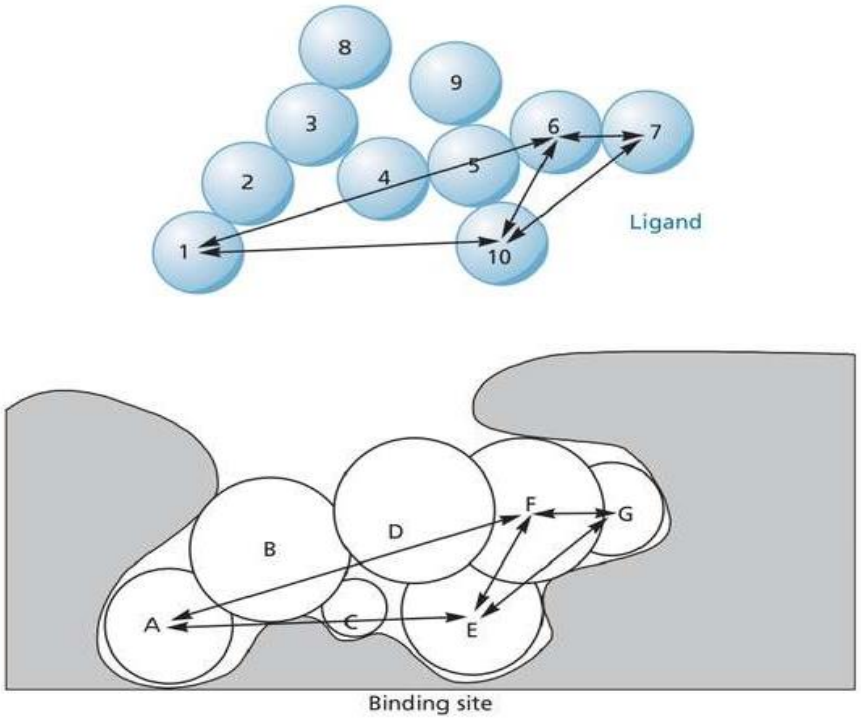

Fig5.4B: - Comparing the distance between the ligand atom and distance between Pseudoatoms.

Moreover, as selected ligand atoms are matched up with selected pseudoatoms, it is perfectly possibly that some of the binding modes obtained are impossible. For example, a ligand atom not used in the matching operation might be placed in the same space as an atom lining the binding site (Fig.5.4D). Therefore, a filtering process has to be included in the program to remove any such unacceptable binding modes. If the binding mode is acceptable, an optimization process is carried out which 'finetunes' the position of the ligand in the binding site. This minimizing unfavorable steric interactions and optimizes intermolecular interactions between the ligand and the binding site. The binding energy of the ligand is now measured and a score is given for that binding mode. This is repeated for all the possible matches and binding modes. Th e binding modes with the highest scores are then stored so that they can be analyzed further by the operator. In the original version of DOCK, this scoring operation took into account only steric interactions and hydrogen bond interactions, but many other factors have an impact on receptor–ligand binding, such as other types of intermolecular interactions, desolvation, the difference in energy between a ligand's diff erent conformations, and the decrease in entropy resulting from a flexible molecule being bound in a fixed conformation. Later versions of DOCK have tackled these issues, as have other docking programs.[14]

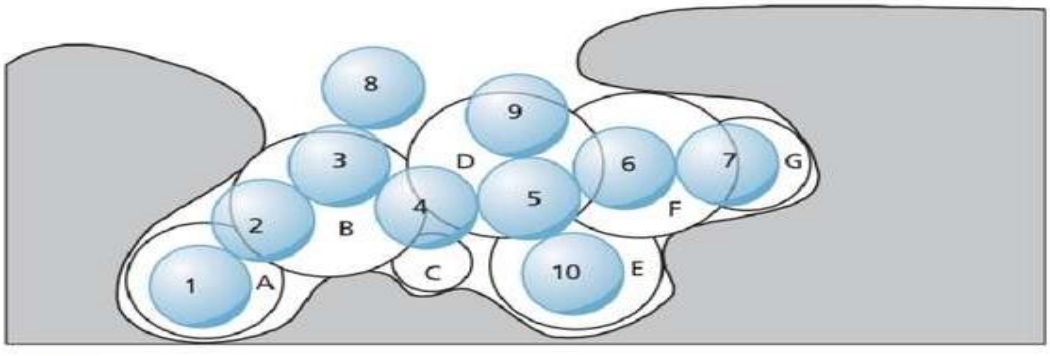

Fig5.4C: -Docking by overlaying the atom Pseudoatoms pairings of 1A,6F,7G and 10E.

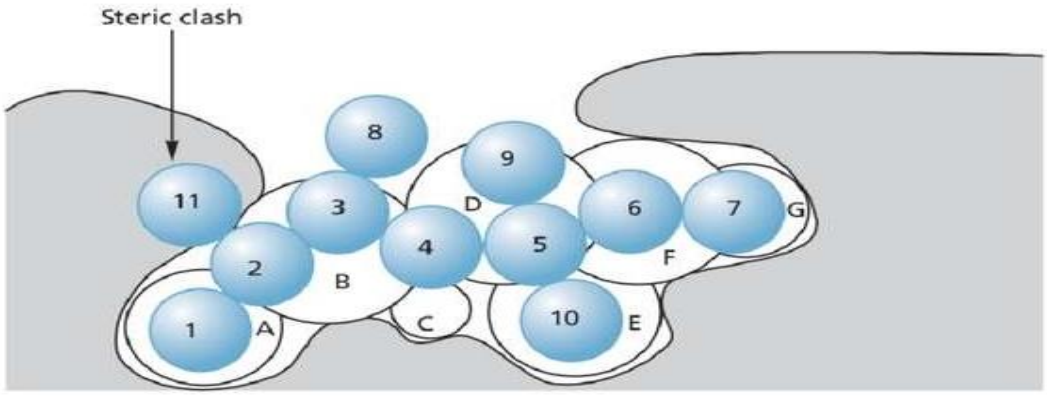

Fig5.4D: -An unacceptable binding mode due to steric clash

6.Molecular docking software's: -

Table 1

Some representative molecular docking programs, and their algorithm characteristics and applications.

Program Name	Algorithm Characteristics	Typical Applications
DOCK	**Step-by-step geometric matching strategy; AMBER force field experience-based scoring function. As a kind of commonly used molecular docking software, it can be used for docking between flexible small-molecule ligands and flexible proteins.**	**Protein–small molecule**
AutoDock	**Lamarck genetic algorithm and experience-based scoring function; the flexibilities of small molecules and some residue side chains can be fully taken into consideration.**	**Protein–small molecule**
AutoDock Vina	**The upgraded version of AutoDock; the success rate and calculation speed are greatly improved compared to AutoDock; simple parameter setting, easy to use, and parallel operation on multi-core machines for docking flexible ligands and flexible protein side chains.**	**Protein–small molecule**
MDock	**Using the knowledge-based atomic–atomic contact potential scoring function, the flexibilities of proteins and small molecules are considered by using the conformations of the multiple proteins and small molecules during the docking process.**	**Protein–small molecule**
FlexX	**The best conformation is selected according to the size of the docking free energy, which has the advantages of fast speed, high efficiency, and easy operation. It is the representative software of the flexible docking and can also be used for the virtual screening of small molecule database.**	**Protein–small molecule**
GOLD	**Based on the GA docking program, the ligand is completely flexible, the receptor binding position is partially flexible; the automatic docking program can be used for virtual screening of the database. Its accuracy and reliability are highly evaluated in the molecular docking simulation.**	**Protein–small molecule**
Surflex-Dock	**The Hammerhead scoring function is used; it combines a large number of conformations from the**	**Protein–small molecule**

	intact molecules through a crossover process to achieve flexible docking.		
eHiTS	**An accurate and fast molecular docking program, which can be used to study ligand and receptor interactions and perform high-throughput virtual screening.**	**Protein–small molecule**	
EADock	**Multi-objective evolutionary optimization algorithm for docking small molecules with the active sites of proteins.**	**Protein–small molecule**	**[15]**

7. Application: -

- Chemical mechanism studies.
- Protein-Protein Docking.
- Drug discovery (Lead optimization).
- Bioremediation
- Structure based drug design.
- Virtual screening.
- Complex solubility.
- Drug excipient-interaction.
- Solubility determination.
- Target fishing and profiling
- Prediction of adverse drug reaction.
- Drug repositioning. [Google Scholar]

8. Conclusion: -

Molecular docking also helps in rejuvenation of drug molecule for a different protean target, ultimately helping reuse the drug. It can be used to model the interaction between a small molecule and a protein at the atomic level. Software's are used for the docking procedure. Through out procedure and application are mention above.

9. Reference: -

1. Suleen Raad, Anu David, Wendy K Chung, Christophe Faure Molecular Docking: A powerful approach for structure-based drug discovery pub med Curr Comput Aided Drug Des. 2011 Jun 1; 7(2): 146–157.
2. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3151162/#:~:text=The%20molecular%20docking%20approach%20can,fundamental%20biochemical%20processes%20%5B19%5D.
3. Jorgensen WL. The many roles of computation in drug discovery. Science. 2004;303(5665):1813–1818. [PubMed] [Google Scholar]
4. Bajorath J. Integration of virtual and high-throughput screening. Nat Rev Drug Discov. 2002;1(11):882–894. [PubMed] [Google Scholar]
5. Walters WP, Stahl MT, Murcko MA. Virtual screening - an overview. Drug Discov. Today. 1998;3:160–178. [Google Scholar]
6. Langer T, Hoffmann RD. Virtual screening: an effective tool for lead structure discovery? Curr Pharm Des. 2001;7(7):509–527. [PubMed] [Google Scholar]
7. Adriano D. Andricopulo,Institute of physics of sao carlos, University of sao paulo Brazil. Front. Pharmacol., 22 August 2018 Sec. Experimental Pharmacology and Drug Discovery
8. https://www.frontiersin.org/articles/10.3389/fphar.2018.00923/full#:~:text=The%20first%20molecular%20docking%20algorithm,of%20spheres%20defining%20its%20volume.
9. Damm K.L., Carlson H.A. Exploring experimental sources of multiple protein conformations in structure-based drug design. J. Am. Chem. Soc. 2007;129:8225–8235. doi: 10.1021/ja0709728. [PubMed] [CrossRef] [Google Scholar]
10. Leach A.R., Shoichet B.K., Peishoff C.E. Prediction of protein-ligand interactions. Docking and scoring: successes and gaps. J. Med. Chem. 2006;49:5851–5855. doi: 10.1021/jm060999m. [PubMed] [CrossRef] [Google Scholar]
11. Key Topics in Molecular Docking for Drug DesignInt J Mol Sci. 2019 Sep; 20(18): 4574.Published online 2019 Sep 15.
12. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6769580/#:~:text=Molecular%20docking%20is%20a%20method,by%20scoring%20functions%20%5B1%5D.
13. A critical assessment of docking programs and scoring functions. Warren GL, Andrews CW, Capelli AM, Clarke B, LaLonde J, Lambert MH, Lindvall M, Nevins N, Semus SF, Senger S, Tedesco G, Wall ID, Woolven JM, Peishoff CE, Head MS.J Med Chem. 2006 Oct 5;49(20):5912-31. doi: 10.1021/jm050362n.PMID: 17004707
14. MEDICINAL CHEMISTRY Graham L. Patrick The moral rights of the author have been asserted Second Edition copyright 2001Third Edition copyright 2005Fourth Edition copyright 2009
15. https://en.wikipedia.org/wiki/List_of_protein-ligand_docking_software#cite_ref-1