Research Article

Molecular Docking - An Overview

Narender Boggula*, Sruthi Katta, Mahaboobi, Subhash Megavath, Jayanti Mukherjee, Rama Rao Tadikonda

CMR College of Pharmacy, Kandlakoya, Medchal, Hyderabad, Telangana, India.
*Corresponding author. Tel: +91 9666552211. E-mail: drnarenderboggula@gmail.com

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. Molecular docking has been widely employed as a fast and inexpensive technique in the past decades, both in academic and industrial settings. Although this discipline has now had enough time to consolidate, many aspects remain challenging and there is still not a straightforward and accurate route to readily pinpoint true ligands among a set of molecules, nor to identify with precision the correct ligand conformation within the binding pocket of a given target molecule. Nevertheless, new approaches continue to be developed and the volume of published works grows at a rapid pace. These recent developments incrementally contribute to an increase in accuracy and are expected, given time, and together with advances in computing power and hardware capability, to eventually accomplish the full potential of this area. Applications of molecular docking in drug development have evolved significantly since it was first created to aid in the study of molecular recognition processes between small and large compounds.

Keywords: Molecular docking, virtual screening, drug repositioning, simulation approach.

1. Introduction

In the field of molecular docking, Docking is the method which predicts the preferred orientation of one molecule to another when they are bonded together to create a stable complex. Using for instance, scoring functions knowledge of the preferred orientation can be utilized to forecast the strength of association or binding affinity between two molecules. The molecular docking method can be used to stimulate the atomic-level interaction between a tiny chemical and a protein. Which enable us to understand basic biochemical processes as well as the behavior of tiny compounds in the binding site of target proteins. The preliminary prediction of the binding characteristics of the medicines and nucleic acids is greatly aided by molecular docking. To determine if a substance or medication will interact with protein or DNA, medicinal chemists are undertaking computer simulation observations.

Virtual screening methods used in academic and commercial drug development and screening processes have made molecular docking an essential tool and methodology. It also plays a significant role in the effectiveness of these protocols. The development of highly accurate, efficient, and practical computational screening methods has been facilitated by recent advancements in high performance computing, optimized software and environmental platforms, and enriched publicly available compound libraries.

A tiny ligand's interaction with a target molecule can be determined using molecular docking techniques, as can whether or not they could act together as the binding site for two or more constituent molecules with a specific

structure. With the mentioned values, a pool of strong candidates can be determined by comparing docking molecules for proteins, other drug-like molecules, or even fragments from the original molecule. This technique is interesting because it allows for the exploration of a wide range of molecular binding interactions, including lipid-protein, lipid-lipid, enzyme-substrate, drug-enzyme, drug-nucleic acid, protein-nucleic acid, nucleic acid- nucleic acid, protein-drug, and protein-protein potential affinities, which play important roles in each molecular biological stage as well as structuring coupling.

Since 1975, advances in nuclear magnetic resonance spectroscopy, high throughout protein purification, and X-ray crystallography have mainly contributed to a better understanding of the structural features of macromolecules and complexes with ligands. While molecular docking has become increasingly popular and simpler to use in the field of drug development, it is not wholly dependent on molecular structure databases, unlike many other in silico technologies. Working with compounds that are not in databases is not impossible because they can be modeled using one or more related structures to create a novel chimeric output that can resemble the original molecule. To evaluate the functionality of the medication molecule, additional parameter adjustments might be made throughout the docking procedure.

Commercial docking programs are quite effective and have been very helpful to the academic community and pharmaceutical sector. Their costs are rising quickly, despite the fact that they are not considerably more accurate than programs that are freely available. Therefore, we considered using a consensus docking strategy, in which we can use freely accessible software and efficiently integrate their output, making the final anticipated solution much better than the individual programs and commercial programs.

Molecular docking is the process of fitting two or more molecules together. Molecular Docking = Target + Ligand. Docking has great promise for the screening of novel drugs and therapeutic targets as well as the elucidation of biomolecular interactions. Its widespread applications may be seen in open-source initiatives like OpenZika, which includes the screening of potential medications against structural models of the Zika virus. By examining the novel uses for existing well-known medications, the approach has lately attained conceptual mainstream status and is believed to be very helpful in accelerating drug development. With an emphasis on docking's applications in the fields of adverse response prediction or medication relocation, this research offers a basic understanding of the distinguishing qualities and principles of docking in order to develop a more logical and targeted treatment.

2. History

The development of docking algorithms in the 1980s led to rise in molecular docking as the most widely used approach among the numerous logical approaches currently being investigated for pharmaceutical research and development. This growth was also facilitated by improvements in techniques like nuclear magnetic resonance spectroscopy, X-ray crystallography, and protein rich filtration. Simulated docking procedures aim to use computational techniques to estimate the connection between specified structures (including such receptors or proteins) and at least each ligand, in order to find composites that illustrate energies for the dynamic site of the important objective particles.

This is done by putting different postures (binding conformations between the protein and the ligand) to the test, which are then scored using scoring formulas. In particular unchanging body docking and semi-adoptable ligand docking (where the ligand's inner bond revolutions is permitted but the receptors is maintained fixed or the receptor is regarded flexible but the ligand is regarded as a fixed molecule), the receptor as well as ligand guidelines and standards are fixed. And flexible docking (where the ligand is treated as a fixed molecule but the receptor is regarded as being flexible).

Similar to flexible docking, in which both the receptor and the ligand are treated as flexible molecules (both molecules are treated as flexible). The great majority of docking programs employ rigid docking. Searching the field of docked confirmations uses a lot less computational power. Contrarily, flexible docking requires more computation but yields better results since it is more accurate in its predictions of ligand binding geometries than rigid-receptor docking.

Developmental coding, quick fourier transform, genetic programming, guided differential evolution, incremental constructions, fragment-based approaches, simulated annealing, multiple copy methods, matching algorithms,

molecular mechanics, Monte Carlo simulations, and Tabu search are just a few of the computational methods and tools that computational biologists use in their docking studies. Each method has a unique set of advantages for conducting docking investigations.

In the article, we describe a variety of docking tools features and shortcomings so that a user can select the optimal strategy for their research. Vitality sceneries are frequently used to discuss protein structures. It becomes extremely challenging to find global minima when two molecules interact. Conventions nowadays are based on theories from material science (steric complimentary) as well as methods from software engineering and other designing disciplines that include design acknowledgement, improvement, AI, and other related concepts.

Knowledge-based docking strategies draw their methods from comparative modelling systems. These methods can be based solely on structures because, by definition, the configurations of the protein to be based on sequence comparison and alignment, sequences and structures (such as threading), or both. According to a 2012 study, docking accommodations can be found for structures speaking to nearly all identified protein-protein associations, provided that these segments have a recognized construction or can be Homology-manufactured, despite the limited number of protein compounds in the Protein Data Bank.

In 2005, the TM-adjust method which combines the TM-score revolution network with Dynamic programming to establish a foundation for layout-based docking was introduced to find the ideal fundamental configuration between protein matches. There are numerous binding strategies between the ligand and protein molecules due to the flexibility of translation, rotation, and conformation. As a result, many sampling strategies have been employed to get around the difficulty of calculating an attainable conformation. These algorithms' development and validation are supported by affinity and structural data found in databases as Protein Records Bank, ZINC, PubChem, Drug Bank, PDBBIND, Chem DB, PLD, and CREDO.

3. Theory

By simulating the ideal conformation based on complementarity and pre-organization, Molecular docking can forecast and determine the binding affinity and interaction mode between ligands and receptors. Figure 1A shows the first proposed "lock-and-key" model, which refers to the rigid docking of receptors and ligands to determine the key's proper orientation in order to unlock the lock. The significance of geometric complementarity is emphasized by this model. Receptors and ligands must alter their conformation to match each other well, though, because the genuine docking mechanism is so flexible. So, we create an "Induced fit model". Based on geometric complementarity, energy complementarity, and pre-organization, it is guaranteed that receptors and ligands will achieve the most stable structure while reducing free energy.
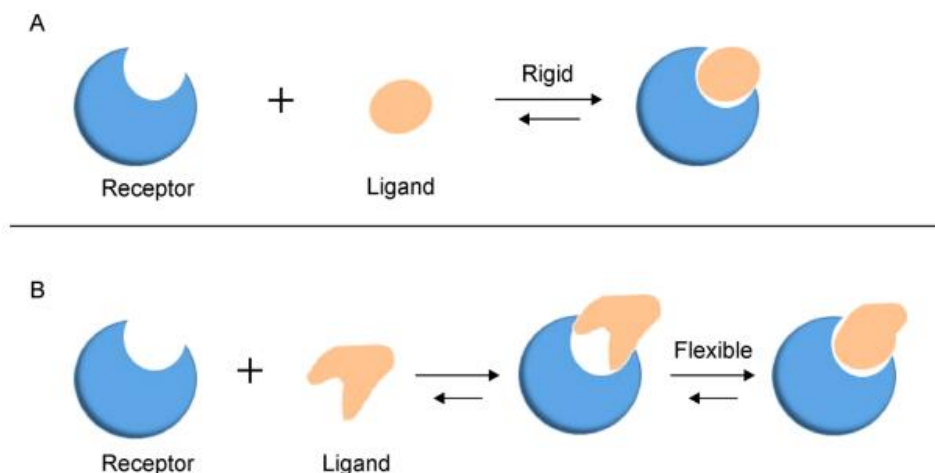

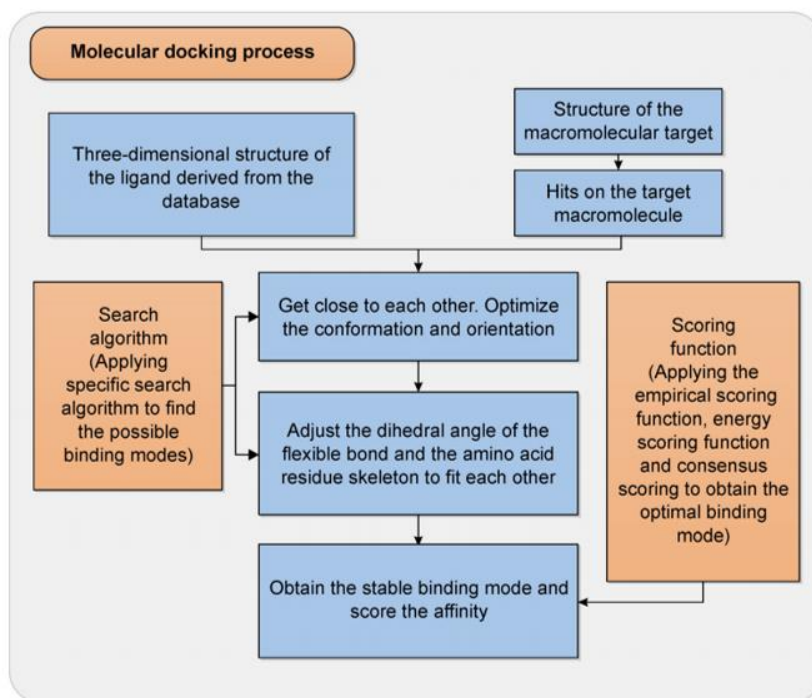*Figure 1: Two models of molecular docking, (A) A lock-and-key model, (B) Induced fit model*

Figure 2: Molecular docking process

As shown in Figure 2, By using a specialized algorithm and the molecular docking program, we can determine the best conformation and orientation for each molecule in terms of complementarity and pre-organization. Next, we can apply a scoring function to forecast the building affinity and evaluate the interactive mode. Figure 3 shows the protein-DNA docking with Autodock Vina displayed in PyMOL.
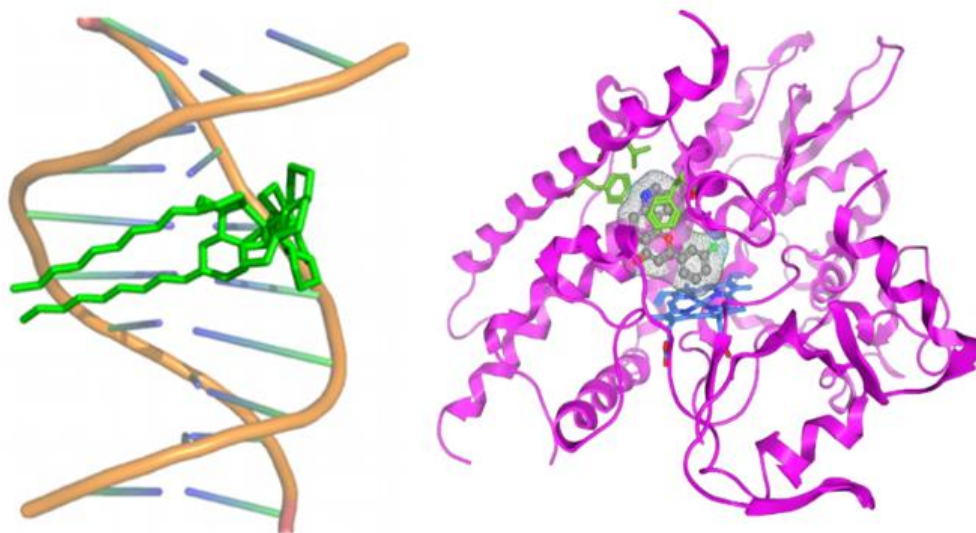

Figure 3: The protein-DNA docking

3.1 Approaches of molecular docking

There are mainly two sorts of techniques utilized for molecular docking. One method makes use of computer simulations to determine the energy profiling for ligand-target docked conformers. While the second strategy makes use of a method that determines the surface complementarity between the ligand and the target.

Chemistry Research Journal

Simulation approach

In this approach, after a "Definite time of moves" in the target molecules conformational space and a physical separation of the ligand and target molecules, the ligand is permitted to bind into the groove or pocket of the target molecule. The movements involve internal (torsional angle rotations) or external (rigid body transformations such rotations and translations) changes to the ligand structure. The "total energy of the system" is computed as the amount of energy produced by each movement within the conformational limitations of the ligand. As it is easier to incorporate ligand flexibility into a molecular modeling tool, this technique is preferable to the shape complementarity one. This method also makes it easier to evaluate the chemical recognition between the ligand and the target molecule. However, because extensive energy landscapes must be generated for each pose, molecular docking using this method takes longer to evaluate the best docked conformer. However, quick optimization techniques and grid-based tools have significantly improved this flaw to make computer simulation methods more approachable.

Shape complementarity approach

This method uses the surface structure characteristics of the ligand and target to aid in the molecular docking process. The molecular surface of the target is described in terms of its solvent-accessible surface area in order to achieve molecular docking. In contrast, the molecular surface of the ligand is described in terms of a matching surface illustration, which aids in the search for a complementary groove or pocket for ligand docking on the target molecular surface. In specifically, the number of turns in the main-chain atoms is used to predict hydrophobicity for protein target molecules. The relatively quick and reliable shape complementarity approach scans thousands of ligands in a matter of seconds to determine their potential binding characteristics to the target molecular surface.

Monte Carlo approach

It produces a randomized translation, rotation, and conformation of the ligand. In a busy place. It gives the configuration a starting value. It then creates a new setup and scores it. Using the metropolitan criterion, it decides whether to maintain the new configuration. (Metropolis criteria: If a new strategy outperforms the old one, it is immediately accepted. If the arrangement is not novel, a Boltzmann's law- focused likelihood study is used. If the solution passes the probability function test, the arrangement is accepted; otherwise, it is rejected.

Matching approach

This technique places an emphasis on redundancy, determines the best place for the ligand atom to be in the site, and produces a ligand-receptor arrangement that might also benefit from improvement.

Ligand fit approach

A rapid and accurate method for docking small molecule ligands into protein active sites while taking form complementarity into consideration is referred to as "Ligand fit".

Point complimentarily approach

These methods concentrate on contrasting the physical or chemical characteristics of several substances. Blind docking is a method for screening the whole interface of target molecules to find probable peptide ligand binding sites and mechanisms of action.

Fragment-based method

Fragment-based methods can be defined as breaking the ligand down into individual photons or particles, attaching the fragments, and then joining the fragments.

Blind docking

It was developed to scan the full surface of protein targets for potential peptide ligand binding sites and modes.

Inverse docking

- ✓ In this case, a computer approach is used to make decisions on a tiny molecule's protein targets for toxicity and side effects.
- ✓ Understanding these targets and proteomics pharmacokinetic profile can make it easier to evaluate any potential toxic side effects of medication candidates.
- ✓ In order to conduct docking experiments on a certain ligand, one of these procedures is used.

Metropolis criterion

If a new answer receives a higher score than the previous one, it is approved right away. A prospect function based on Boltzmann is helpful if the configuration is not brand-new. The solution is established if it satisfies the possibility function test; else, the configuration is undesirable.

Types of docking

The first molecular docking algorithm was created in 1982, using by calculation of the released binding energy. Docking analyses are carried out to control the ligand-target interaction profile and look for the best ligand conformation within the complex. Additionally, empirical scoring functions that convert docking score into binding energy are investigated. To create 3D ligand and target interaction profiles, a variety of free online programs are available, including Biovia DSV, Pymol, Chimera, Rasmol, SwissPDB viewer, etc. Three general categories of docking are outlined below:

Flexible docking

The side chains of the protein and ligand are kept flexible during flexible docking. The induced-fit concept put forward by Daniel Koshland in 1958 serves as the foundation for the general notion of flexible docking. It is also referred to as "induced-fit docking" as a result. Wherein the binding energies of the suggested ligand's different conformations are estimated at protein or receptor pockets. The target chain should also be adaptable enough to cooperate with receptor and ligand conformational alterations. The most widely used and accurate method allows for the prediction of a wide range of potential changes in the ligand's structure, but it is also time-consuming and expensive.

Semi-flexible docking

This strategy makes the protein the only hard component and the ligand molecules the only flexible ones. The conformational degrees of freedom of the ligand are also examined in addition to the six translational and rotational degrees of freedom. These methods rely on the fixed conformation of a protein being able to recognize the ligands that need to dock. This assumption isn't always true, as was previously mentioned.

Rigid docking

As a result, lock and key docking is what it is, which causes a number of issues. The Rigid docking preserves and freezes the primary geometry of the target and ligand during docking analysis.

The 'Lock and Key' concept, put forth by Emil Fischer in 1894, serves as the foundation for this kind of docking research. As a result, it is for the purpose of detecting drug-target interaction, the analysis of ligand-target docking is highly important, but there is an issue when the ligand docks at the pocket site of a receptor protein. Due to both entities' stiff structures, it is difficult to see interactions and to get the best confirmation of a ligand.
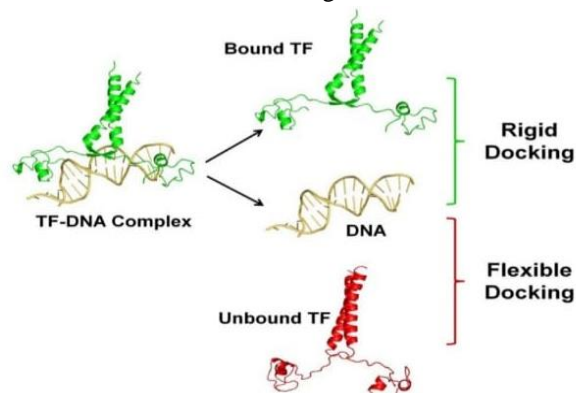

Figure 4: Structural image of rigid and flexible docking

Docking models

The lock and key theory

Emil Fischer first suggested a paradigm known as the "lock and key model" in 1890. It explains the operation of biological systems. Similarly, how a key fits into a lock, the substrate slips into the big molecule's active site. The function of biological locks depends on certain stereochemical characteristics.
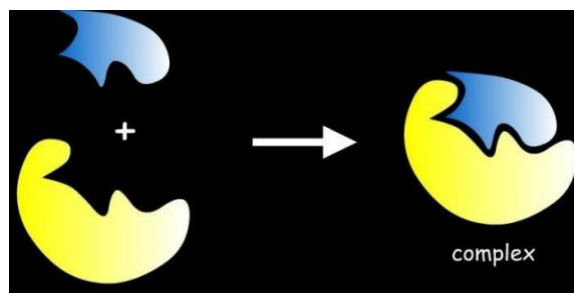
Chemistry Research Journal

Figure 5: Lock and key theory

The induced fit theory

Daniel Koshland put forth the induced fit idea in 1958. The fundamental tenet is that during recognition, the ligand and target adapt to one another through gradual conformational changes until the best alignment is reached.
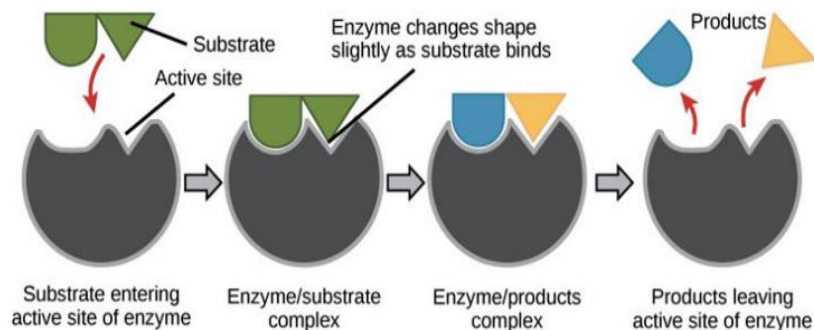

Figure 6: Induced fit theory

The conformational ensemble models

Conformational Ensemble model was proposed in 2003 by BUYONG Ma et al. Proteins have been found to be capable of significantly bigger conformational changes in addition to the tiny induced fit adaptation. According to the hypothesis, proteins already exist in a variety of conformational states. A protein's ability to transition from one state to another is known as Ductility.
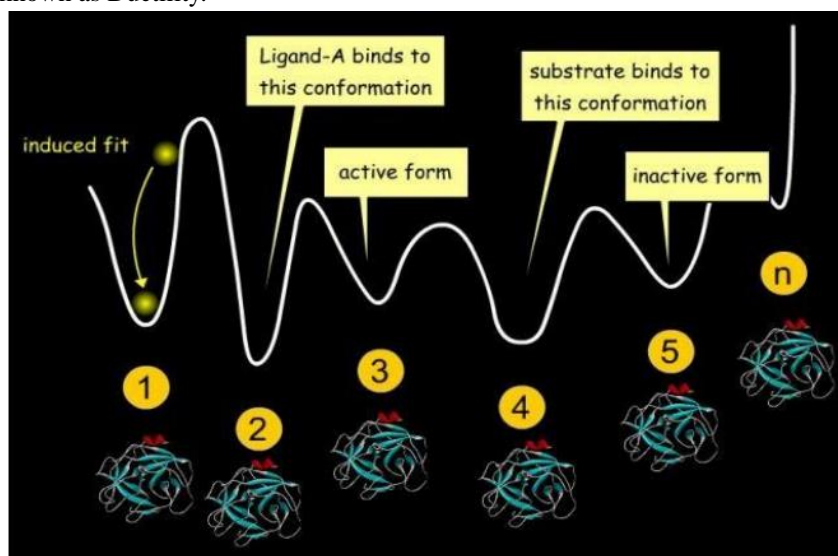

Figure 7: Conformational ensemble model

From the lock and key to the ensemble model

The induced-fit, conformation ensemble, and lock and key models do not conflict. Every single one of them focuses on a distinct facet of the recognition process. The ensemble model demonstrates the structural complexity of proteins, the induced-fit model illustrates how complementarity is achieved, and the lock-and-key model introduces the concept of 3D complementarity.
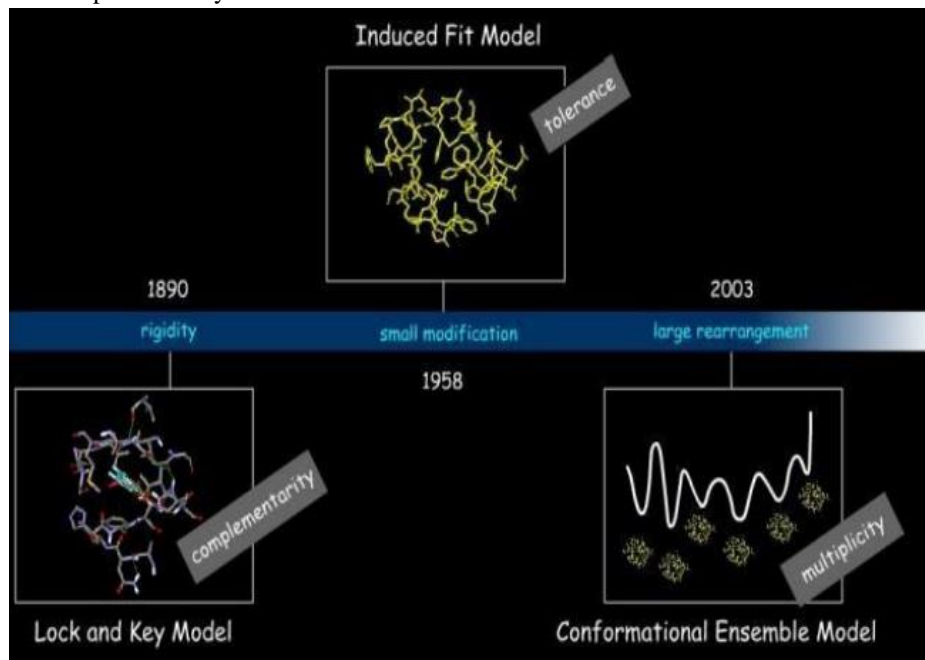

Figure 8: From the lock and key to the ensemble model

Mechanism of docking

- The arrangement of the attention protein is the first requirement for creating a docking screen. Typically, a biophysical technique like as X-ray crystallography or, less frequently, NMR spectroscopy has been used to maintain the structure. A docking agenda uses this protein organization and a folder of ligands as input.
- The search algorithm and scoring function are two methods that determine whether a docking program will be successful. The protein and ligand pairs potential orientations and conformations make up the study space. With current computing capabilities, it is impossible to fully identify the research domain that would list every possible molecule distortion as well as every potential translational and rotational orientation of the ligand with respect to the protein at a given level of granularity.
- The majority of docking programs in use take flexible ligands into account, and many are working to mimic a flexible protein receptor.
- The process known as molecular docking was used to examine the intermolecular communication between two molecules in silica. The macromolecule in this improvement is the protein receptor. The ligand is the tiny particle.
- A molecule has inhibitory properties.

Major steps involved in mechanics of molecular docking

Consequently, the docking procedure involves the following steps:

Step-1: Protein preparation

The Protein Data Bank (PDB) must be used to retrieve the three-dimensional structure of the protein; the structure must then be pre-processed. According to the provided parameters, this should permit amputation of the water molecules from the cavity, stabilize the charges, substantial the missing residue, produce the side chains, etc.

Step-2: Prediction of the active site

The active site of the protein must be predicted following protein production. There are several active sites on the receptor, but just the one that is of concern should be selected. When present, hetero atoms and water molecules are often unimportant.

Step-3: Making the ligand

Chemistry Research Journal

Ligand can be found in many databases, including ZINC and PUBCHEM, or it can be sketched using the chem. Sketch tool. The LIPINSKY'S RULE OF 5 should be applied while choosing the ligand. The Lipinski rule of five helps to distinguish between drug-like and non-drug-like behaviors. The CADDD approach (Computer Aided Drug Design and Detection). Due to drug similarity for molecules surviving with two or more of the following requirements, it offers great potential for success or failure.

Step-4: Docking

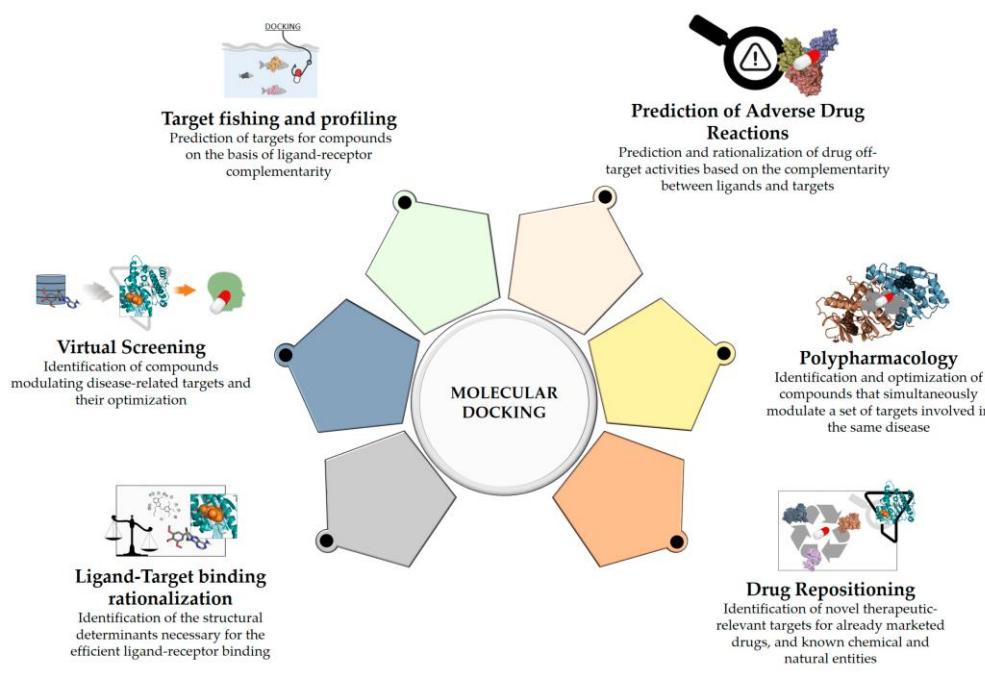The protein and ligand are docked, and the interactions are examined.

Figure 9: Applications of molecular docking in current drug discovery. Molecular docking is currently employed to help rationalizing ligands activity towards a target of interest and to perform structure-based virtual screening campaigns, similarly to as when it was first developed. Besides these applications, it can also be used to identify series of targets for which the ligands present good complementarity (target fishing and profiling), some of them being potentially responsible for unexpected drug adverse reactions (off-targets prediction). Moreover, docking is also currently employed for the identification of ligands that simultaneously bind to a pool of selected targets of interest (polypharmacology) and for identifying novel uses for chemical compounds with already optimized safety profiles (drug repositioning).

4. Conclusion

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.

Chemistry Research Journal

Molecular docking has been established as a pivotal technique among the computational tools for structure-based drug discovery. Here we addressed key aspects of the methodology and discussed recent trends in the literature for advancing and employing the technique for successful drug design. Benchmarking sets and the various metrics available are crucial for validating performance gains achieved by new docking software but must be carefully chosen since no single one can be regarded as the absolute best for molecular docking. This is expected to provide novel valuable opportunities in future drug discovery and development and, in particular, in the design of challenging and innovative drugs (i.e., multi-target ligands), as well as in assisting ligand profiling and repositioning.

Author contributions

The manuscript was written through the contributions of all authors. All of them approved the final version of the manuscript.

Funding

This research received no external funding.

Disclosure

The authors declare no conflict of interest.

Ethical approval

Not required.

References

[1]. Lengauer T, Rarey M. Computational methods for biomolecular docking. Curr. Opin. Struct. Biol. 1996; 6(3):402-406.

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

[3]. Zev, S.; Raz, K.; Schwartz, R.; Tarabeh, R.; Gupta, P. K.; Major, D. T. Benchmarking the Ability of Common Docking Programs to Correctly Reproduce and Score Binding Modes in SARS-CoV-2 Protease Mpro. J. Chem. Inf. Model. 2021; 61:2957-2966.

[4]. Alberg DG, Schreiber SL. Structure-based design of a cyclophilin-calcineurin bridging ligand. Science. 1993; 262(5131):248-250.

[5]. Zhao H, Caflisch A. Discovery of ZAP70 inhibitors by high-throughput docking into a conformation of its kinase domain generated by molecular dynamics. Bioorg. Med. Chem. Lett. 2013; 23:5721-5726.

[6]. Ayaz Mahmood Dar, Shafia Mir. Molecular docking: approaches, types, application and basic challenges. J Anal Bioanal Tech. 2017; 8(2):1-3.

[7]. B. Mukesh, K. Rakesh. Molecular docking: a review. International Journal of Research in Ayurveda and Pharmacy. 2011; 2(6):1746-1751.

[8]. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A geometric approach to macromolecule-ligand interactions. Journal of Molecular Biology. 1982; 161(2):269-288.

[9]. Ahmed A, Mam B, Sowdhamini R. DEELIG: A deep learning approach to predict protein-ligand binding affinity. Bioinformatics and Biology Insights. 2021; 15:1-9.

[10]. Guedes IA, Pereira FSS, Dardenne LE. Empirical scoring functions for structure-based virtual screening: Applications, critical aspects, and challenges. Frontiers in Pharmacology. 2018; 9:1089.

[11]. Koshland DE. The key–lock theory and the induced fit theory. Angew Chemie. Int Ed English. 1995; 33(23–24):2375-2378.

[12]. Lexa KW, Carlson HA. Protein flexibility in docking and surface mapping. Quarterly Reviews of Biophysics. 2012; 45(3):301-343.

[13]. Andrusier N, Mashiach E, Nussinov R, Wolfson HJ. Principles of flexible protein-protein docking. Proteins: Structure, Function and Genetics. 2008; 73(2):271-289.

[14]. Anderson AC, O'Neil RH, Surti TS, Stroud RM. Approaches to solving the rigid receptor problem by identifying a minimal set of flexible residues during ligand docking. Chemistry & Biology. 2001; 8(5):445-457.

Chemistry Research Journal

[15]. Meng X-Y, Zhang H-X, Mezei M, Cui M. Molecular docking: A powerful approach for structure-based drug discovery. Current Computer-Aided Drug Design. 2011; 7(2):146-157.

[16]. Du X, Li Y, Xia YL, Ai SM, Liang J, Sang P, et al. Insights into protein–ligand interactions: Mechanisms, models, and methods. International Journal of Molecular Sciences. 2016; 17(2):144-177.

[17]. Cramer F. Biochemical correctness: Emil Fischer's lock and key hypothesis, a hundred years after-An essay. Pharmaceutica Acta Helvetiae. 1995; 69(4):193-203.

[18]. Salmaso V, Moro S. Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: An overview. Frontiers in Pharmacology. 2018; 9:923.

[19]. Meng X-Y, Zhang H-X, Mezei M, Cui M. Molecular docking: A powerful approach for structure-based drug discovery. Curr Comput Aided-Drug Des. 2012; 7(2):146-157.

[20]. T. Supriya, M. Shankar, S. Kavya Lalitha, J. Dastgir, M. Niranjan Babu. American Journal of Biological and Pharmaceutical Research, 2016; 3(2):83-89.

[21]. Avinash R, Veerabhadra Rao A, et al. A review on molecular docking, Novel tool in drug design and analysis. Journal of Harmonized Research in Pharmacy. 2013; 2(4):215-218.

[22]. Kitchen D, De cornez H, Furr J, Bajorath J, et al. Docking and scoring in virtual screening for drug discovery: methods and applications. International Journal of Pharma and Bioscience. 2004; 3(11):95-97.

[23]. McMartin C, Bohacek RS, et al. QXP: Powerful, Rapid Computer Algorithms for Structure-based Drug Design. J Comput Aid. Mol. Des. 1997; 11:333-344.

[24]. Schnecke V, Kuhn LA, et al. Virtual Screening with Solvation and Ligand induced Complementarity, Perspect. Drug Discov. 2000; 20:171-190.

[25]. Jain AN, et al. Surflex: Fully Automatic Flexible Molecular Docking Using a Molecular Similarity-Based Search Engine. J Med Chem. 2003; 46:499-511.

[26]. Pedro H. M. Torres, Ana C. R. Sodero, Paula Jofily and Floriano P. Silva-Jr. Int. J. Mol. Sci. 2019; 20:4574.

[27]. Guedes IA, De Magalhães CS, Dardenne LE. Receptor-ligand molecular docking. Biophys Rev. 2014; 6(1):75-87.

[28]. Ahmad S, Singh S, Srivastava MR, Shukla S, Rai S, Shamsi AS. Molecular Docking Simplified: Literature Review. Adv. Med. Dental Health Sci. 2021; 4(4):37-44.

[29]. Claussen H, Buning C, Rarey M, Lengauer T. FlexE. Efficient molecular docking considering protein structure variations. J Mol Biol. 2001; 308(2):377-395.

[30]. Lamb ML, Jorgensen WL. Computational approaches to molecular recognition. Curr Opin Chem Biol. 1997; 1(4):449-457.