



Molecular docking: A potential tool for Drug Designing, Docking Simulation and Redocking

¹Indu Bharkavi SK, ²Khadijah Mohideen, ³Chandrasekaran Krithika, ⁴Jayasri NP, ⁵Chandini Rajkumar, ⁶Saranya Ramsridhar, ⁷Murali Balasubramaniam A

^{1,4,5,6,7} Department of Oral Pathology and Microbiology, Sathyabama Dental College and Hospital, Sathyabama Institute of Science and Technology, Chennai – 600 119, Tamil Nadu, India.

² Department of Oral and Maxillofacial Surgery and Diagnostic Sciences, Faculty of Dentistry, Najran University, Najran - 66462, Kingdom of Saudi Arabia.

³ Meenakshi Academy of Higher Education and Research, West K.K. Nagar, Chennai-600 078, Tamil Nadu, India.

Corresponding Author:

Dr. Khadijah Mohideen, M.D.S, Ph.D., Assistant Professor, Department of Oral and Maxillofacial Surgery and Diagnostic Sciences, Faculty of Dentistry, Najran University, Najran - 66462, Kingdom of Saudi Arabia.

(Received: 16 March 2025

Revised: 20 April 2025

Accepted: 15 June 2025)

KEYWORDS

Molecular docking,
ligand,
protein,
docking simulations,
redocking

ABSTRACT:

Various infectious diseases are found to be prevalent globally, most of which are drug-resistant. Many new approaches and advancements are carried out in this scientific community to treat the same by discovering newer drugs. One such novel method is Molecular Docking. It has emerged as a powerful computational tool in drug design, enabling the prediction of binding interfaces between small molecules (ligands) and target proteins (receptors). Molecular docking contributes significantly to modern drug discovery, particularly in lead optimization, hit identification, and studying DNA-drug interactions. It offers a three-dimensional perspective, making it a cost-effective, safe, and user-friendly technology.

Additionally, one of its advantages lies in adjusting the flexibility of the system under study. This article comprehensively reviews molecular docking techniques, focusing on docking simulations and the crucial aspect of redocking for validation. Through redocking experiments, researchers can enhance the accuracy and reliability of docking methodologies, facilitating the discovery and optimization of potential drug candidates.

1. Introduction

In discovering novel therapeutic agents, molecular docking has revolutionized the drug design process by offering insights into ligand-receptor interactions at the atomic level [1]. It is an in-silico analysis, a computational technique for predicting the binding interactions between a ligand and receptor [2]. Molecular docking helps to assess the binding modalities of a ligand with a known three-dimensional protein structure, which is the primary goal and helpful in forecasting the signal's kind and strength, similar to which we can determine the drug's efficacy [3]. Docking simulations provide a cost-effective and time-efficient approach to evaluating potential drug candidates, enabling researchers to explore vast chemical spaces and identify molecules with the highest binding affinity for specific targets. In the

drug development process, the ability of the protein receptor to detect, interact with, and connect with molecular substrates and inhibitors is crucial, so scoring functions and docking programs are introduced [4]. The scoring functions and sample conformations of tiny molecules at protein binding sites are utilized in the computational process to determine which conformations best complement the protein binding sites [5]. For one of the targets, every docking program could generate ligand conformations that matched crystallographically confirmed protein/ligand composite structures [6]. Various docking software is available, including Hammerhead, ICM, MC Dock, GOLD, Gem Dock, Glide, and Yucca [5]. These docking programs employ a search algorithm that iteratively evaluates the



conformation of the ligand until it converges to the minimum energy state.

2. Basics of Molecular Docking

The docking methods can be categorized as rigid or flexible based on whether they consider changes in molecular geometry during docking.

Rigid Docking: Rigid docking refers to a docking approach where the molecules' bond angles and lengths remain unchanged throughout the docking process. In other words, the molecular structures are considered fixed and do not undergo any internal geometry adjustments during the docking simulation. Rigid docking is a straightforward method commonly used when the assumption of a static molecular structure is reasonable.

Flexible Docking: Flexible docking, on the other hand, allows for the simulation of changes in the internal geometry of the interacting molecules that can occur when they form a complex molecule. This method is more computationally demanding because it considers the molecules' flexibility and potential conformational changes as they interact. As computational processing capacity increases, flexible docking becomes a viable option for more accurate modeling of molecular interactions [4].

The docking methods can also be classified as conventional or induced fit based on how they treat the receptor's binding site.

Conventional Docking: In traditional docking, the receptor's binding site is considered static, and the ligand is docked into this fixed site.

Induced Fit Docking: Induced fit docking considers the binding site's dynamic nature. It allows the receptor's binding site to endure conformational modifications in response to the ligand's presence and better accommodate the ligand. This approach is more realistic but requires more computational resources [7,8].

The choice of which docking method to use depends on the specific research question, available computational resources, and the accuracy required for modeling molecular interactions [8]. By employing docking, scientists aim to predict the precise alignment of therapeutic small-molecule drugs with their target proteins, which plays a pivotal role in forecasting small-

molecule affinity and activity [3]. The utilisation of docking holds immense significance in the rational drug development process, driving numerous endeavours to enhance docking prediction algorithms.

At its core, molecular docking involves predicting the optimal spatial arrangement of one molecule relative to another when they join to form a stable complex, accomplished through mathematical techniques. Employing a scoring system, the strength of the interaction or binding affinity between these two molecules is quantifiable based on their preferred orientation.

This process is indispensable in signal transduction, facilitating interactions among diverse biological molecules such as lipids, proteins, and nucleic acids. Through docking, scientists can anticipate the nature and strength of the signals generated in these intricate physiological interactions, thus contributing significantly to the knowledge of biological processes [9].

3. Mechanism of Docking

Molecular docking is a fundamental molecular modeling technique employed to elucidate the binding interactions between enzymes or proteins and small molecules, often referred to as ligands. This methodology essentially operates on the principle of a lock and key mechanism. The primary prerequisite for running a docking screen is the protein sequence. The structure is identified via NMR spectroscopy or X-ray crystallography. The protein function and compound database serve as the docking tool. The search algorithm and grading procedures are crucial to docking success [2]. When exploring a protein's conformational space linked to a ligand, the search space includes all possible protein orientations and conformations. It is challenging to comb the space. So, most docking systems use flexible ligands [4].

Receptor flexibility

Managing stretchy proteins is a significant challenge in the docking process. A biomolecule or protein's shape changes based on the ligand it binds to, demonstrating that a single receptor conformation will result from docking with a stiff receptor. However, the ligands might need to attach to the receptor in multiple conformations when the docking is done using a flexible receptor. Different conformational states of proteins are typically the most overlooked component in molecular docking



research. Protein flexibility is significant because it increases affinity between a given medication and a target. Water molecules at the active site are another component of target flexibility. The water molecules must be adjusted to prevent artifacts during docking [10].

Ligand chemistry

Since the recognition of any ligand by a biomolecule depends on both electrostatic interaction and three-dimensional orientation, the formulation of the ligand has a significant impact on the docking findings, which indicates that the ligand's preparation and conformation are both crucial. Before this, there was still a significant difference between the approximate pKa values and the tautomeric and protomeric states of the molecules that needed to be docked, which was likely because the structure was optimized by removing or adding hydrogens. Since molecules are essentially ionized under physiological conditions, they are preserved in practically all databases in their neutral forms. For this reason, ionizing them before docking is required. However, achieving standard ionization is simple in other programs. The question of which tautomer to employ or whether to utilize them all remains concerning the tautomer [11].

Molecular Modeling

Molecular modeling is a versatile tool that enables the creation, characterization, and manipulation of compound configurations and their ensuing properties, all of which are intricately tied to the three-dimensional geometries of these molecules. This technique is crucial in exploring and understanding chemical structures and their behaviour [3].

Models of Molecular Docking

The lock and key theory: Emil Fischer developed the lock and fundamental model in 1890 to explain how biological processes work. Like how a key is inserted into a lock, a substrate is inserted into the active site of a macromolecule [3].

The induced fit theory: The induced fit theory was put forth by Daniel Koshland in 1958. The underlying principle is that both the ligand and target adapt to one another throughout character recognition by making subtle conformational changes until an optimum match is achieved [3].

The conformation ensemble model: Proteins have been found to undergo much higher conformational alterations compared to minor induced-fit adjustments [3]. A novel idea proposes that proteins comprise an ensemble of conformational states that already exist. The protein can switch between states because of its flexibility [2]. Multiple conformational selection and induced fit events occur in allosteric propagation [12]. Figure 1 explains the various types of docking mechanisms, such as the lock and key method, Induced fit and Conformational selection methods.

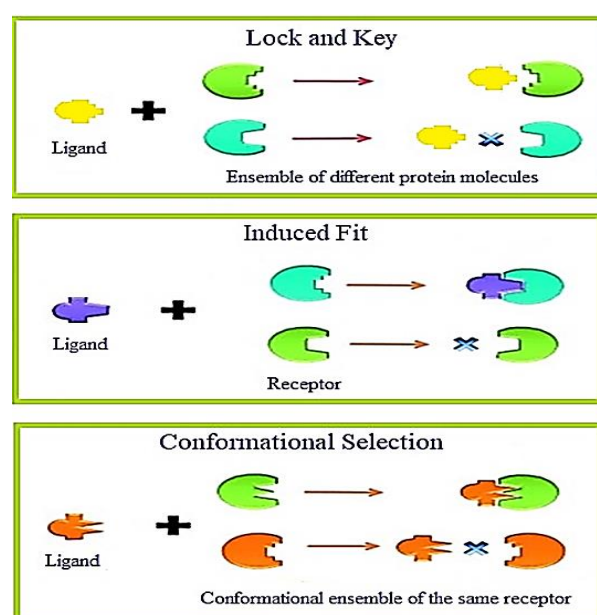


Figure 1: Various Types of Docking Mechanisms

4. Molecular Dynamic Simulations

Molecular dynamics (MD) simulations are computational techniques that offer insights into the temporal behaviour of individual atoms within a molecular system. These simulations rely on a comprehensive model of the physical principles that regulate interatomic interactions. In contrast, molecular docking primarily investigates the interaction between two biological molecules, often as a static representation of their binding, without accounting for dynamic changes. In contrast, MD simulations go beyond this static interaction view and provide a detailed and time dependent description of how these interacting molecules evolve and move over time [13].

Molecular docking simulations (MDS) predict the three-dimensional structure resulting from the interaction of



one or more molecules within an intermolecular complex. When it comes to protein-ligand docking, there are several dedicated software programs designed for this purpose. Autodock and Redock are two widely accepted programs within the MDS community. These tools utilize MDS techniques to ascertain the optimal spatial arrangement of two or more molecules, often focusing on locating the best fit between them [14].

Consequently, MDS delivers a set of predicted ligand conformations within the binding site of the target molecule. These poses often encompass various conformations and positions for the ligand. To refine and categorize this data and determine the optimal orientation through MDS, fitness functions, scoring systems, or energy functions come into play. These functions are crucial in evaluating binding free energy and intermolecular binding affinity [15].

The fundamental components of an MDS can be distilled into two primary stages. Firstly, a search method is essential to systematically explore the vast space of potential protein-ligand geometries or poses. Secondly, these generated poses require thorough analysis and ranking, achieved by applying scoring or energy functions. This process helps identify the most favourable binding orientations and key interaction sites within the complex [14].

5. Redocking

One of the future perspectives is Redocking. "Redocking" is a technique for determining a ligand's crystallographic location, and it is essentially a procedure to validate whether the molecular docking algorithm can recover the crystallographic position using MDS. Since hydrogen atoms' atomic locations are rarely detected in X-ray protein crystallography, RMSD calculations are typically performed for non-hydrogen atoms. High resolution, better than 1.5, X-ray diffraction data are required to determine the locations of hydrogen atoms [14]. Redocking is mainly used to validate the docking protocol [16].

6. Cross-docking

A method known as "cross-docking" can be utilized in addition to re-docking to extend the validity of a docking protocol. Cross-docking is an option when there are many crystallographic structures for the same protein target. In this process, several ligands from various

crystal structures of the same protein are docked to a single complex protein crystallographic conformation [17].

7. Docking Scores

Scoring functions serve as mathematical tools within computational chemistry and molecular modeling. Their primary role is to estimate the binding affinity between two molecules after the docking process. Typically, these two molecules comprise a drug, a compact chemical compound, and its corresponding biological target, often a protein receptor [18]. In predicting the strength of intermolecular interactions, scoring systems have been developed for interactions between two proteins [2] and interactions involving a protein and DNA [19]. These systems aid in gauging the intensity of associations formed between these biomolecules.

8. Utility of Scoring Function

Scoring functions are pivotal in various applications within drug discovery and molecular modeling. These applications encompass the exploration of virtual databases for screening small molecules with the potential to bind to specific protein targets, facilitating the discovery of novel compounds suitable for drug development [20,21]. Additionally, scoring functions aid in creating new compounds through de novo design, enabling them to attach to their intended protein targets effectively [22]. Furthermore, these functions are instrumental in optimizing screening hits, primarily enhancing their selectivity and affinity [23].

9. Classes of Scoring Function

A force field is utilized to evaluate the magnitudes of intermolecular interactions, specifically van der Waals and electrostatic forces, between every atom within two molecules as they come together to form a complex [24-26]. This computation plays a crucial role in determining their affinities for each other. Furthermore, it is a common practice to include the intramolecular energies of the binding partners, known as strain energies, in these calculations. In some instances, the desolvation energy of both the protein and the ligand is considered, and this is achieved using implicit solvation techniques like GBSA or PBSA, mainly when the binding process takes place in the presence of water [27].



Empirical scoring functions are built upon analyzing diverse interactions between the molecules involved in binding processes [22]. This analysis can encompass methods like quantifying the ligand and receptor atoms in direct contact or calculating changes in solvent-accessible surface area (SASA) when comparing the complexed state to the uncomplexed ligand and protein. Typically, multiple linear regression techniques are applied to establish the coefficients of the scoring function [28]. These coefficients encapsulate various factors, including the impact of unmet hydrogen bonds, which play a significant role in the enthalpy of binding [29]. The absence of a single hydrogen bond can substantially affect binding affinity, often altering it by one to two orders of magnitude. Other factors, such as the number of constrained rotatable bonds during complex formation (which contributes unfavourably to conformational entropy) and the number of hydrogen bonds (which contributes favourably to affinity, especially if shielded from the solvent, and has no contribution if solvent-exposed), are also considered.

On the other hand, knowledge-based scoring functions rely on statistical insights drawn from extensive 3D datasets, such as the Protein Data Bank or the Cambridge Structural Database. These datasets are employed to derive statistical "potentials of mean force." This methodology is grounded in the notion that close intermolecular interactions between specific atom types or functional groups, occurring more frequently than expected by random chance, are likely to confer an energetic advantage and thus positively impact binding affinity [30].

10. Docking Programs

Molecular docking software operates by employing a search process that iteratively assesses the conformation of the ligand until it reaches a state of minimal energy [31]. Over the past two decades, many docking tools and programs have been developed, catering to academic research and commercial applications. These tools encompass a diverse range and include well-known programs like DOCK, AutoDock, Flex, Glide, GOLD, Surflex, MOE-Dock, LeDock, AutoDock Vina, rDock, UCSF Dock, among numerous others [32-41]. AutoDock Vina, GOLD, and MOE-Dock were the three programs that predicted the top-ranked poses with the highest scores. LeDock and GOLD were successful in

determining the proper ligand binding positions. Glide (XP) and GOLD reliably forecast the positions with an accuracy of 90.0% [42].

Additionally, it was demonstrated that in a virtual screening trial against Factor Xa, GOLD produced more enrichment factors than Glide. In contrast, Glide outperformed GOLD against the same target in a comparable virtual screening study. A recent paper said these docking programs could forecast experimental poses with an average root-mean-squared deviation (RMSD) of 1.5 to 2 Å [43].

11. Application of Molecular Docking

Molecular docking serves multiple vital purposes. Molecular docking is instrumental in evaluating the specificity of a proposed drug for a homologous protein. It allows us to understand how well a drug candidate can bind to its intended target [44].

The lead Optimization technique

This technique can predict a ligand's most favourable orientation within its target protein's binding site. It also offers insights into various binding strategies, aiding in designing medications that are more effective and highly selective [9,44].

Hit Identification in Molecular Docking

The process of hit identification, also known as a hit finding, is employed in drug discovery to pinpoint compounds that selectively bind to a particular target protein and regulate its activity according to desired specifications, facilitating the drug discovery process [45].

Drug-DNA interaction

Molecular docking is a major factor in the initial prediction of a drug's ability to attach to a nucleic acid. This data demonstrates the relationship between the cytotoxicity of a medicine and its molecular structure. Considering this, medicinal chemists continuously work to clarify the molecular basis of pharmaceuticals' anti-cancer activity by examining how medications interact with nucleic acids in the presence of copper [46,47]. Predicting if a substance or medicine interacts with a protein or DNA is the primary goal of in silico observations by medicinal chemists. If the docking program correctly anticipates the mentioned interaction,



experimental methods are provided to determine the complex's actual binding mode. New anti-cancer drugs are developed because of this. Additionally, having this knowledge would help identify any structural changes in a medication that would cause it to bind to its target in a specific sequence or structure [48].

12. Recent Advancement in Molecular Docking Techniques

Machine learning scoring functions stand apart from traditional scoring functions by their unique approach to modeling the relationship between structural characteristics of protein-ligand complexes and binding affinity [49]. Unlike conventional scoring functions, machine learning scoring functions don't rely on predefined functional forms for this connection; instead, they derive the functional form directly from the available data.

It's worth noting that machine learning scoring functions have repeatedly demonstrated superior performance in predicting the binding affinities of diverse protein-ligand complexes compared to conventional scoring systems [18,50,51]. This advantage holds for target-specific complexes as well. However, the degree of improvement varies depending on the specific target and, to a significant extent, the amount of relevant data available [52,53].

In structure-based virtual screening, machine learning scoring functions typically outperform classical scoring systems, provided that appropriate precautions are taken. Notably, this performance gap widens further when target-specific data is incorporated into the modeling process [54]. It's essential to underscore that selecting suitable decoys for a specific target is a critical factor in the training and testing of any scoring function [55,56].

13. Virtual Screening

An in-silico method called VS is employed in the search for new drugs. Computational techniques routinely assess large databases of molecular structures during VS. The application of VS is anticipated to reveal compounds more likely to bind to the molecular target, which is usually an enzyme or protein receptor [57]. The flowchart below serves as a visual representation of the same process [Figure 2].

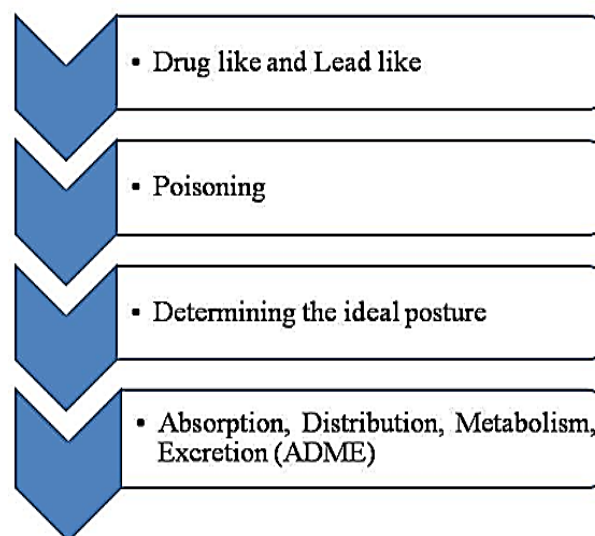


Figure 2: Virtual Screening Procedure

14. Challenges and Limitations

The success of the docking program depends upon the search algorithms and the scoring function. The search algorithm is defined by rules and parameters applied to predict the conformations. Two continuous methodological difficulties are needed for sampling ligand-receptor configurations and adequately assessing their complementarity. The treatment of ligand conformations is currently included in docking methods. However, the therapy of receptor flexibility is the biggest obstacle [58]. Meng et al. has noted that if receptor flexibility is introduced to receptor conformation sampled in high-energy and this is not accounted for in the docking calculation, then the docking calculation will be worse. Database docking scoring functions are still a topic of current research. However, experimental equipment is still needed to create realistic interactions between tiny compounds and receptors. Developing scoring functions with high accuracy and minimal computing expense may advance docking applications [59].

15. Conclusion

Molecular docking is a powerful technique to investigate the interconnections between small molecules and target proteins, shedding light on essential biochemical processes. Combining chemical information from natural products with virtual screening through docking promises to play a crucial role in drug discovery in the post-genomic era as new potential drug targets emerge



from functional genomic studies. Evaluating the molecular dynamic algorithm's ability to reproduce the crystallographic binding locations is supported by redocking and cross-docking, which represent prospects in this field. Furthermore, docking-based virtual screening has demonstrated a higher success rate than traditional screening methods. In conclusion, molecular docking stands as a cornerstone in the drug discovery process, facilitating the exploration of critical molecular interactions and the identification of potential drug candidates.

Conflicts of interest

The authors declare no conflicts associated with this work.

Abbreviations

MD - Molecular Docking

NMR - Spectroscopy-Nuclear magnetic resonance

DNA - Deoxyribonucleic acid

GBSA - Generalized born surface area

PBSA - Poisson Boltzmann surface area

SASA - Solvent accessible surface area

MDS - Molecular dynamic simulation

RMDS - Root mean square deviation

PAM - Pralidoxime

VS - Virtual screening

References

1. Sreepadmanabh M., Sahu A.K., Chande A. COVID-19: Advances in diagnostic tools, treatment strategies, and vaccine development. *J Biosci*, 2020. 45(1), 148.
2. Chen G., Seukep A.J., Guo M. Recent Advances in Molecular Docking for the Research and Discovery of Potential Marine Drugs. *Mar Drugs*, 2020. 18(11), 545.
3. Morris G.M., Lim-Wilby M. Molecular docking. *Methods Mol Biol*, 2008. 443, 365-382.
4. Raval K., Ganatra T. Basics, types and applications of molecular docking: A review. *IP Int J Compr Adv Pharmacol*, 2022. 7(1), 12-16.
5. Sousa S.F., Fernandes P.A., Ramos M.J. Protein-ligand docking current status and future challenges. *Proteins*, 2006. 65(1), 15-26.
6. Warren G.L., Do T.D., Kelley B.P., Nicholls A., Warren S.D. Essential considerations for using protein-ligand structures in drug discovery. *Drug Discov Today*, 2012. 17(23-24), 1270-1281.
7. Christoph S. Accounting for Induced-Fit Effects in Docking: What is Possible and What is Not? *Current topics in medicinal chemistry*, 2010. 11, 179-191.
8. Allegra M., Tutone M., Tesoriere L., Attanzio A., Culetta G., Almerico A.M. Evaluation of the IKK β Binding of Indicaxanthin by Induced-Fit Docking, Binding Pose Metadynamics, and Molecular Dynamics. *Front Pharmacol*, 2021. 12, 701568.
9. Shoichet B.K., McGovern S.L., Wei B., Irwin J.J. Lead discovery using molecular docking. *Curr Opin Chem Biol*, 2002. 6(4), 439-446.
10. Totrov M., Abagyan R. Flexible ligand docking to multiple receptor conformations: A practical alternative. *Curr Opin Struct Biol*, 2008. 18, 178-184.
11. Elokely K.M., Doerksen R.J. Docking Challenge: Protein Sampling and Molecular Docking Performance. *J Chem Inf Model*, 2013, 53, 1934-1945.
12. Nussinov R., Ma B., Tsai C.J. Multiple conformational selection and induced fit events take place in allosteric propagation. *Biophys Chem*, 2014. 186, 22-30.
13. Karplus M., McCammon J. Molecular dynamics simulations of biomolecules. *Nat Struct Mol Biol*, 2002. 9, 646-652.
14. Linus S.A., Fernanda P.M., Mariana M.X., Eduarda O.P., Bianca V., Jana A.F., et al. Recent Progress of Molecular Docking Simulations Applied to the Development of Drugs *Current Bioinformatics*, 2012. 7(4), 352-365.
15. De Azevedo Jr. W.F. MolDock Applied to Structure-Based Virtual Screening. *Curr Drug Targets*, 2010. 11, 327-334.
16. Kontoyanni M., McClellan L.M., Sokol G.S. Evaluation of Docking Performance: Comparative Data on Docking Algorithms. *J. Med. Chem*, 2004. 47, 558-565.
17. Thilagavathi R., Mancera R.L. Ligand-protein cross-docking with water molecules. *J Chem Inf Model*, 2010. 50, 415-421.
18. Jain A.N. Scoring functions for protein-ligand docking. *Curr Protein Pept Sci*, 2006. 7(5), 407-420.



19. Robertson T.A., Varani G. An all-atom, distance-dependent scoring function for the prediction of protein-DNA interactions from structure. *Proteins*, 2007. 66(2), 359-374.
20. Rajamani R., Good A.C. Ranking poses in structure-based lead discovery and optimization: current trends in scoring function development. *Current Opinion in Drug Discovery & Development*, 2007. 10(3), 308-315.
21. Seifert M.H., Kraus J., Kramer B. Virtual high-throughput screening of molecular databases". *Current Opinion in Drug Discovery & Development*, 2007. 10(3), 298-307.
22. Böhm H.J. Prediction of binding constants of protein ligands: a fast method for the prioritization of hits obtained from de novo design or 3D database search programs. *Journal of Computer-Aided Molecular Design*, 1998. 12(4), 309-323.
23. Joseph-McCarthy D., Baber J.C., Feyfant E., Thompson D.C., Humblet C. Lead optimization via high-throughput molecular docking. *Current Opinion in Drug Discovery & Development*, 2007. 10(3), 264-274.
24. Fenu L.A., Lewis R.A., Good A.C., Bodkin M., Essex J.W., Scoring Functions. In: *From Free Energies of Binding to Enrichment in Virtual Screening: Structure-Based Drug Discovery*, 1st ed., Dhоти H., Leach A.R., Eds., Dordrecht:Springer, 2007. pp. 223-246.
25. Sottriffer C., Matter H. Classes of Scoring Functions. In: *Principles, Challenges, and Practical Guidelines: Virtual Screening*, 1st ed., Mannhold R., Kubinyi H., Folkers G., Eds., John Wiley & Sons, 2011. Vol. 48, pp. 185-190.
26. Ain Q.U., Aleksandrova A., Roessler F.D., Ballester P.J. Machine-learning scoring functions to improve structure-based binding affinity prediction and virtual screening. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2015. 5(6), 405-424.
27. Genheden S., Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery*, 2015. 10(5), 449-461.
28. Schneider N., Lange G., Hindle S., Klein R., Rarey M. A consistent description of Hydrogen bond and Dehydration energies in protein-ligand complexes: methods behind the HYDE scoring function. *J Comput Aided Mol Des*, 2013, 27(1), 15-29.
29. Lange G., Lesuisse D., Deprez P., Schoot B., Loenze P., Bénard D., Marquette J.P., Broto P., Sarubbi E., Mandine E. Requirements for specific binding of low affinity inhibitor fragments to the SH2 domain of (pp60) Src are identical to those for high affinity binding of full-length inhibitors. *Journal of Medicinal Chemistry*, 2003. 46(24), 5184-5195.
30. Muegge I. PMF scoring revisited. *Journal of Medicinal Chemistry*, 2006. 49(20), 5895-5902.
31. Pagadala N.S., Syed K., Tuszynski J. Software for molecular docking: a review. *Biophys Rev*, 2017. 9(2), 91-102.
32. Venkatachalam C.M., Jiang X., Oldfield T., Waldman M. LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *J Mol Graph Model*, 2003. 21, 289-307.
33. Österberg F., Morris G.M., Sanner M.F., Olson A.J., Goodsell D.S. Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in Auto Dock. *Proteins*, 2002. 46, 34-40.
34. Friesner R.A., Banks J.L., Murphy R.B., Halgren T.A., Klicic J.J., Mainz D.T., Repasky M.P., Knoll E.H., Shelley M., Perry J.K., Shaw D.E., Francis P., Shenkin P.S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem*, 2004. 47(7), 1739-1749.
35. Jones G., Willett P., Glen R.C., Leach A.R., Taylor R. Development and validation of a genetic algorithm for flexible docking. *J Mol Biol*, 1997. 267, 727-748.
36. Jain A.N. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *J Med Chem*, 2003. 46(4), 499-511.
37. Corbeil C.R., Williams C.I., Labute P. Variability in docking success rates due to dataset preparation. *J Comput Aided Mol Des*, 2012. 26, 775-786.
38. 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.
39. Trott O., Olson A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring



- function, efficient optimization and multithreading. *J Comput Chem*, 2010. 31, 455-461.
40. Ruiz-Carmona S., Alvarez-Garcia D., Foloppe N., Garmendia-Doval A.B., Juhos S., Schmidtke P., et al. rDock: a fast, versatile and open-source program for docking ligands to proteins and nucleic acids. *PLoS Comput Biol*, 2014. 10(4), e1003571.
41. Allen W.J., Balias T.E., Mukherjee S., Brozell S.R., Moustakas D.T., Lang P.T., et al. DOCK 6: Impact of new features and current docking performance. *J Comput Chem*, 2015. 36(15), 1132-1156.
42. Wang Z., Sun H., Yao X., Li D., Xu L., Li Y., et al. Comprehensive evaluation of ten docking programs on a diverse set of protein-ligand complexes: the prediction accuracy of sampling power and scoring power. *Phys Chem Chem Phys*, 2016. 18(18), 12964-12975.
43. Bissantz C., Folkers G., Rognan D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *J Med Chem*, 2000. 43(25), 4759-4767.
44. Gschwend D.A., Good A.C., Kuntz I.D. Molecular docking towards drug discovery. *J Mol Recognit*, 1996. 9(2), 175-186.
45. Ferreira L.G., Dos Santos R.N., Oliva G., Andricopulo A.D. Molecular docking and structure-based drug design strategies. *Molecules*, 2015. 20(7), 13384-13421.
46. Agarwal S., Jangir D.K., Mehrotra R., Lohani N., Rajeswari M.R. A structural insight into major groove directed binding of nitrosourea derivative nimustine with DNA: a spectroscopic study. *PLoS One*, 2014. 9(8), e104115.
47. Mehrotra R., Jangir D.K., Agarwal S., Ray B., Singh P., Srivastava A.K. Interaction studies of anti-cancer drug lomustine with calf thymus DNA using surface enhanced Raman spectroscopy. *MAPAN*, 2013. 28, 273-277.
48. Holt P.A., Chaires J.B., Trent J.O. Molecular docking of intercalators and groove-binders to nucleic acids using Autodock and Surfex. *J Chem Inf Model*, 2008. 48(8), 1602-1615.
49. Ballester P.J., Mitchell J.B. A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking. *Bioinformatics*, 2010. 26(9), 1169-1175.
50. Li H., Leung K.S., Wong M.H., Ballester P.J. Improving AutoDock Vina Using Random Forest: The Growing Accuracy of Binding Affinity Prediction by the Effective Exploitation of Larger Data Sets. *Molecular Informatics*, 2015. 34 (2-3), 115-126.
51. Ashtawy H.M., Mahapatra N.R. A Comparative Assessment of Predictive Accuracies of Conventional and Machine Learning Scoring Functions for Protein-Ligand Binding Affinity Prediction. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2015. 12(2), 335-347.
52. Durrant J.D., Friedman A.J., Rogers K.E., McCammon J.A. Comparing neural-network scoring functions and the state of the art: applications to common library screening. *Journal of Chemical Information and Modeling*, 2013. 53(7), 1726-1735.
53. Wójcikowski M., Ballester P.J., Siedlecki P. Performance of machine-learning scoring functions in structure-based virtual screening. *Sci Rep*, 2017. 7, 46710.
54. Ragoza M., Hochuli J., Idrobo E., Sunseri J., Koes D.R. Protein-Ligand Scoring with Convolutional Neural Networks. *Journal of Chemical Information and Modeling*, 2017. 57(4), 942-957.
55. Imrie F., Bradley A.R., Deane C.M. Generating property-matched decoy molecules using deep learning. *Bioinformatics*, 2021. 37(15), 2134-2141.
56. Ballester P.J. Selecting machine-learning scoring functions for structure-based virtual screening. *Drug Discov Today Technol*, 2019. 32-33, 81-87.
57. Oliveira, T. A. d., Silva, M. P. d., Maia, E. H. B., Silva, A. M. d., & Taranto, A. G. Virtual Screening Algorithms in Drug Discovery: A Review Focused on Machine and Deep Learning Methods. *Drugs and Drug Candidates*, 2023. 2(2), 311-334.
58. Carlson H.A. Protein flexibility and drug design: how to hit a moving target. *Curr Opin Chem Biol*, 2002. 6(4):447-452.
59. 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*, 2011. 7(2), 146-157.