Journal of Ethnopharmacology and Toxicology
Review Article

The impact of molecular docking on drug discovery and development: a review of recent advances and applications

Sudhakar Pachiappan, Akshaya Sivakumar*, Gayathiri Muthusamy, Parkavi Gopalakrishnan

Department of Pharmacology, Swamy Vivekanandha College of Pharmacy, Elayampalayam, Tiruchengode - 637205, Namakkal, Tamil Nadu, India.

***Correspondence**
Akshaya Sivakumar
akshayasivakumar2111@gmail.com

Volume: 3, Issue: 2, Pages: 1-13
DOI: https://doi.org/10.37446/jet/ra/3.2.2025.1-13
Received: 20 March 2025 / Accepted: 15 August 2025 / Published: 31 December 2025

Molecular docking is a structure-based computational approach that predicts the binding location and affinity of ligands for their targets, aiding in drug discovery and development. determined by platform compatibility, processing resources, and research requirements. While it speeds up and lowers the cost of drug development, issues including protein flexibility, scoring function constraints, and prediction accuracy persist. Hybrid modeling, molecular dynamics simulations, and consensus scoring are all examples of advances that improve dependability. As computational tools and structural biology advances, molecular docking is projected to become more accurate, allowing for quicker and more efficient drug creation. Cloud computing, GPU acceleration, and high-throughput virtual screening now allow the analysis of vast chemical libraries in reduced time frames. Docking has been successfully applied in the repurposing of approved drugs, the design of covalent inhibitors, and the exploration of allosteric modulators. However, persistent challenges such as limitations in scoring functions, treatment of solvent effects, and conformational variability necessitate further methodological refinement. This review summarizes the principles, applications, and recent technological advancements in molecular docking, highlighting its evolving role in bridging computational predictions with experimental validation and accelerating the drug development pipeline.

Keywords: *AI and machine learning, drug development, drug discovery, lead optimization, molecular docking*

Introduction

Molecular docking is a computational method to identify the structure of compounds with multiple molecules. Docking studies try to predict the desired three-dimensional structures. Docking uses maths to forecast the best orientation of molecules when they join to form a stable complex. Scoring functions can evaluate the binding strength between two molecules based on their preferred orientation. Molecular docking has a big impact on the unified design of drugs. Scientists often use several docking methods to create an ideal shape for both the protein and the ligand, as well as the relative position between them, to minimize the total system's free energy. *De-novo* drug design involves creating new compounds based on a receptor's 3D structure.

Molecular modeling techniques help to determine the structural features of lead target complexes and lead changes (Surana et al., 2021). Emil Fisher compared the drug-receptor interaction to the interaction between keys and locks in the early 1890s. He believed that the drug and receptor interacted as solid bodies that did not change their conformations. Daniel Koshland recently proposed that both molecules undertake conformational changes throughout interaction and choose the most appropriate conformation to connect. X-ray structures and in silico simulations have repeatedly confirmed this theory (Doytchinova, 2022).

Types of Molecular Docking

Docking is a method of organizing molecules to interact optimally with receptors. Docking occurs when molecules form a stable complex within a cell. Molecular modeling is used to create, describe, and manipulate the three-dimensional structures and interactions of substances (Raval & Ganatra, 2022).

Rigid docking

The target protein and selected ligand conformations are designed to be stiff during the docking process. There is no flexibility in the molecular bond angle, bond length, or torsion angle; they must all be constant. This method of docking does not allow for such significant molecular changes (Halperin et al., 2002).

a. Methods

Preparation: Water molecules are eliminated, hydrogen atoms are added, partial charges are assigned, and, if required, the geometry is optimized to create the receptor and ligand structures. Search algorithm: To discover the ideal binding posture, docking algorithms use a variety of search techniques to examine the ligand's conformational space. Geometric hashing, Monte Carlo techniques, genetic algorithms, and grid-based search are examples of common search algorithms (Halperin et al., 2002). Scoring function: Candidate ligand poses are evaluated and ranked based on expected binding affinity. Typically, scoring functions take into account elements including desolation energy, van der Waals forces, electrostatic interactions, and geometric complementarity. Post-processing and analysis involve refining predicted binding postures and analyzing intermolecular interactions. Visualization techniques let researchers discover important binding residues and interactions in docked complexes (Sahu et al., 2024).

b. Applications

Drug discovery: In computer-aided drug design (CADD), rigid docking is frequently used to screen sizable chemical libraries and find possible drug candidates that have a high affinity and specificity for binding to a target protein. Rigid docking can anticipate protein-protein interactions, allowing for the creation of inhibitors and modulators. Enzyme mechanisms: Docking small compounds into enzyme active sites helps researchers understand enzyme-substrate interactions and catalysis mechanisms, facilitating rational drug design and enzyme engineering. Rigid docking plays a crucial role in virtual screening procedures, which prioritize compounds for experimental testing based on projected binding affinity to a target protein (Sahu et al., 2024).

Flexible docking

This molecular docking allows complete freedom to both the target and the ligand being processed. They account for all conformational changes at the molecular level (Sahu et al., 2024).

a. Methods

Flexible docking algorithms use conformational sampling and classic search methodologies to predict the best ligand binding posture inside the receptor's flexible binding region. Algorithms can employ scoring functions to assess the compatibility of ligand conformations with receptors, guiding the search for optimal binding poses. Scoring function: In flexible docking, scoring functions are critical for determining the energetics of ligand-receptor interactions and rating projected binding poses. Usually, these scoring systems take into account elements like conformational strain, van der Waals forces, electrostatic interactions, geometric complementarity, and solvation effects (Sahu et al., 2024).

b. Applications

Flexible docking is a popular method for virtual screening and lead optimization in drug development. It predicts the binding affinity and selectivity of small molecules to target proteins. Flexible docking can accurately predict binding affinity and effectiveness by identifying compounds that rigid docking approaches may miss due to receptor flexibility. Flexible docking can help understand the molecular mechanics of protein-ligand interactions, particularly induced-fit effects during binding. Flexible docking models receptor conformational changes to better understand ligand recognition and binding selectivity (Cavasotto et al, 2004). Flexible docking is a technique used in enzymology to anticipate enzyme inhibitor binding modalities and processes. Flexible docking allows for the creation of strong, targeted enzyme inhibitors with potential for therapeutic use by taking into account the flexibility of both the ligand and the receptor. Virtual screening

methods rely on flexible docking to uncover potential drug candidates from huge chemical databases. Flexible docking improves virtual screening predictions and lead optimization by taking into account receptor flexibility (Sahu et al., 2024).

Induced-fit docking

Induced fit docking includes elements of rigid and flexible docking. The process starts with docking stiff structures, then refines the complex using flexible side chains or backbone motions to account for induced fit effects. This computational method in molecular modeling predicts binding mechanisms and affinity between ligands and receptors, taking into account conformational changes in both structures during interaction. In contrast to rigid docking, which assumes rigid conformations for both the ligand and receptor during binding, induced fit docking takes into consideration the dynamic nature of biomolecular interactions and the induced-fit phenomena found in receptor-ligand binding.

a. Methods

Conformational sampling involves sampling both the ligand and receptor structures to create an ensemble of conformations that reflect their flexibility. This might entail methods like normal mode analysis, systematic conformational search tools, or molecular dynamics simulations. Induced fit docking techniques use conformational sampling and classical search methodologies to estimate the best ligand binding posture inside the receptor's flexible binding region. Algorithms employ scoring systems to match ligands to receptors and find the best binding postures (Chaudhury et al., 2008). In induced fit docking, scoring functions evaluate the energetics of ligand-receptor interactions and provide a ranking of the anticipated binding positions according to many criteria, including conformational strain, van der Waals forces, electrostatic interactions, geometric complementarity, and solvation effects. These scoring systems help identify the most stable and physiologically relevant binding postures among docked complexes. Analysis and post-processing: Following the creation of potential ligand positions, post-processing methods are used to improve the predictions and examine the interactions between molecules. Through the use of visualization tools, scientists may analyze the docked complexes and pinpoint important binding residues, conformational shifts, and induced-fit effects (Sahu et al., 2024).

a. Applications

In lead optimization, induced fit docking predicts the affinity and selectivity of small-molecule ligands for target proteins. Through the explicit consideration of receptor flexibility, induced fit docking makes it possible to identify ligands that rigid docking approaches would miss, resulting in more precise estimates of binding effectiveness and affinity. Fragment-based drug design: Using induced fit docking, tiny molecular fragments are screened to predict their binding patterns to target proteins. Induced fit docking identifies fragment hits that can lead to high-affinity molecules by taking receptor flexibility into account (Xu et al., 2018). Induced fit docking is a crucial part of virtual screening procedures, which use computer screening to find possible drug candidates from huge databases of chemical compounds. Induced fit docking improves virtual screening predictions, hit detection, and lead optimization by taking into account receptor flexibility. Induced fit docking helps understand the molecular mechanics of ligand-receptor interactions, particularly induced-fit effects during binding. Induced fit docking models receptor conformational changes to better understand ligand recognition and binding selectivity (Sahu et al., 2024).

Ligand-based docking

Lipid-based docking determines a ligand's binding mechanism and affinity to a target receptor by comparing its similarities to existing ligands with empirically confirmed activities. Ligand-based docking techniques use the ideas of pharmacophore mapping and molecular similarity to pinpoint the crucial structural elements and functional groups required for ligand binding and biological activity. Liquid-based docking predicts the binding affinity and selectivity of novel ligands to target receptors by analyzing their molecular characteristics and spatial arrangement (Tang et al., 2025).

a. Methods

Ligand selection and preparation of the dataset: The first step in ligand-based docking is the selection of a dataset that includes ligands with known activity against the target receptor and a variety of structural types. These ligands are used as reference compounds for creating pharmacophore models and doing similarity searches. Usually, virtual screening libraries or experimental databases are used to curate the dataset.

Pharmacophore generation: Pharmacophore models are created by aligning essential functional groups and molecular characteristics that influence ligand binding and biological activity. Hydrophobic areas, aromatic rings, acceptors, and

donors of hydrogen bonds are typical pharmacophore characteristics. Pharmacophore creation techniques uncover shared properties across active ligands and their spatial correlations. Virtual screening and similarity search: After the pharmacophore model is created, it is utilized to look for compounds with comparable spatial arrangements and chemical characteristics in virtual compound libraries. The pharmacophore model is matched with potential ligands using ligand-based similarity search techniques such as shape-based similarity, 3D pharmacophore overlays, and 2D fingerprints. Virtual screening compares prospective ligands to known active molecules and predicts their affinity for the target receptor. After virtual screening, ligands are scored and validated to determine their binding affinity and selectivity. Scoring functions use pharmacophore fit, chemical characteristics, and energy estimates to rank compounds for experimental testing (Sahu et al., 2024).

b. Applications

Ligand-based docking identifies compounds with potential action against certain receptors or enzymes. Using virtual compound libraries, ligand-based docking finds molecules that share chemical characteristics and biological activity with known ligands, making it easier to find new lead compounds for additional experimental testing. Virtual screening uses ligand-based docking to rank molecules for experimental testing from massive databases of chemical compounds. Ligand-based docking enables the rapid identification of potential drug candidates with desirable pharmacological properties by leveraging molecular similarity and pharmacophore mapping (Ajjarapu et al., 2022). Lead optimization: To direct the creation of novel compounds with enhanced binding affinity and selectivity, ligand-based docking is utilized. Ligand-based docking aids in the optimization of chemical scaffolds and functional groups to improve ligand potency and drug-like qualities by examining the structure-activity connections of existing ligands and finding important pharmacophore traits. Ligand-based docking makes bioisosteric substitution easier by finding structurally related compounds that can replace important chemical moieties or functional groups in existing ligands. By exploring the chemical space and discovering analogs with comparable pharmacological characteristics, ligand-based docking assists in the discovery of new ligands with increased efficacy and fewer off-target effects (Sahu et al, 2024).

Protein-protein docking

Protein-protein docking predicts binding mode and affinity by searching conformational space and discovering energetically advantageous configurations. Protein-protein interactions occur through complementary surfaces of protein molecules, where certain amino acid residues establish intermolecular contacts such as hydrogen bonds, hydrophobic interactions, and electrostatic interactions. Protein-protein docking models interactions and predicts their spatial organization in bound complexes (Grassmann et al, 2024).

a. Methods

Protein-protein docking begins with preparing protein structures, which involves removing water molecules, adding missing atoms, assigning partial charges, and optimizing hydrogen bond networks. Protein structures are commonly obtained by experimental approaches like X-ray crystallography and NMR spectroscopy. Protein-protein docking algorithms use diverse search methodologies to investigate the interacting proteins' conformational space and locate probable binding sites. Genetic algorithms, stochastic optimization strategies, Monte Carlo methods, and geometric hashing are examples of popular search algorithms. These algorithms take samples of the proteins' various orientations and conformations and assess their compatibility using intermolecular interactions and geometric complementarity (Holmes et al., 2023). A scoring function ranks potential binding positions based on projected affinity and stability. Typically, scoring functions take into account variables including interface area, van der Waals forces, electrostatic interactions, form complementarity, and desolvation energy. The scoring algorithm is designed to discriminate between native-like binding postures and nonspecific interactions. Post-processing and Analysis: Post-processing methods are used to improve predicted binding postures and evaluate intermolecular interactions. Researchers can use visualization tools to analyze docked complexes and find important binding residues and interfaces. Molecular dynamics simulations or energy reduction techniques can optimize docked structures and analyze their dynamic behavior (Sahu et al, 2024).

a. Applications

Protein-protein docking reveals the three-dimensional structure of protein complexes and the molecular underpinnings of their interactions. Protein-protein docking predicts the binding mode and interface residues of interacting proteins, offering insights into complex formation and function. Drug Discovery: Protein-protein docking is critical for structure-based drug design since it identifies potential inhibitors or modulators of protein-protein interactions. Protein-protein docking allows for the rational creation of tiny compounds or peptides to disrupt disease processes by targeting particular protein-protein interactions (Kaczor et al., 2018). Protein-protein docking predicts protein interactions and binding modes,

allowing for functional annotation. Protein-protein docking identifies complexes engaged in cellular processes and signaling cascades, shedding light on protein function and regulation. Virtual screening methods rely on protein-protein docking to find putative inhibitors of protein-protein interactions. Protein-protein docking prioritizes candidate drugs for experimental testing based on expected binding affinity and interaction specificity by docking them into target protein binding sites (Sahu et al, 2024).

Blind docking

Blind docking involves searching the receptor's full surface for possible binding sites and predicting ligand binding mechanism and affinity. Blind docking is ideal for targets with unknown or flexible binding sites, as it doesn't require previous knowledge of their location or structure. Blind docking methods sample a grid or mesh of locations on the receptor surface and evaluate probable binding poses using intermolecular interactions and scoring functions.

a. Methods

Grid creation involves representing the receptor structure as a three-dimensional grid or mesh that covers the protein's whole surface. Grid spacing determines search resolution, with smaller spacing resulting in finer sampling but higher computing cost. Grid creation may include masking or eliminating portions of the receptor surface that are unsuitable for ligand binding, such as solvent-exposed or conformationally flexible regions. Ligand sampling involves creating a library of ligand conformations and docking them into the receptor grid. Ligand sampling methods might include random, systematic, or fragment-based approaches. Ligand conformations are often created by exploring their structural space or sampling from a library of existing structures (Ritchie et al., 2008). Blind docking systems use diverse search tactics, including geometric hashing, Monte Carlo methods, evolutionary algorithms, and stochastic optimization, to identify probable ligand binding poses inside the receptor grid. The grid is used to put ligand conformations, which are then analyzed for potential binding poses based on geometric complementarity and intermolecular interactions with the receptor. A scoring function ranks potential binding positions based on projected affinity and stability. The scoring function evaluates geometric complementarity, electrostatic interactions, van der Waals forces, desolvation energy, and interface area. The scoring system differentiates between native-like binding postures and non-specific interactions (Sahu et al., 2024).

b. Applications

Blind docking identifies possible binding sites on a protein and predicts ligand affinities and binding modes. Blind docking allows for the discovery of new binding sites on the receptor's full surface, which may not be visible in experimental structures or homology models. Virtual screening methods rely on blind docking to find potential drug candidates from chemical compound libraries. Blind docking allows for the discovery of ligands with varied chemical structures and binding behaviors on the receptor surface, resulting in the identification of new lead compounds for testing. Blind docking is a technique used in fragment-based drug discovery to assess and predict the binding patterns of tiny molecule fragments to target proteins. Blind docking identifies fragment hits on the receptor surface that may be improved into high-affinity lead compounds using fragment linking or growth techniques (Kumar et.al, 2012). Blind docking can predict protein function by identifying putative ligand binding sites and analyzing their binding modalities. Blind docking explores the receptor's full surface, revealing the structural basis of protein-ligand interactions and aiding in annotation of function and control (Sahu et.al, 2024).

Molecular docking methods and approaches

This review study categorizes Vina algorithms into four macro groups: Modification can be based on target, ligand, scoring function, or computational system **(Table 1)**.

Table 1. Different methods and approaches for molecular docking

Tools	Description	Reference
VinaLC	Identification of Antagonistic Action of Pyrrolizidine Alkaloids in Muscarinic Acetylcholine Receptor M1 by Computational Target Prediction Analysis	(Abdalfattah et al., 2024)
QuickVina	Efficient analysis of toxicity and mechanism of food contaminants using network toxicology and molecular docking strategy: A Case Study of Aflatoxin B1	(Chu & Zi, 2024)

AutoDockFR	SwissDock 2024: major enhancements for small-molecule docking with Attracting Cavities and AutoDock Vina	(Bugnon et al., 2024)
AutoDock VinaXB	A Combination of Structure-based Virtual Screening and Experimental Strategies to Identify the Potency of Caffeic Acid Ester Derivatives as SARS-CoV-2 3CLpro Inhibitors from an In-house Database	(Pojtanadithee et al., 2024)

Principle of molecular docking

Docking is a process that examines how proposed molecules fit into cavities and interact with residues. Computational drug discovery typically involves docking between small compounds and macromolecules, similar to protein-ligand docking. This method of docking is referred to as molecular docking. Molecular mechanics, which applies classical physics to polyatomic systems, serves as the foundation for most docking programs. Experimental parameters help align experimental findings with molecular mechanics. To overcome experimental restrictions, quantum mechanics, semiempirical, and ab initio computations are used to turn mathematical equations into parameters. These equations use several characteristics to create molecular force field systems, including potential energy, torsional properties, bond geometry, electrostatic terms, and the Lennard-Jones potential. AMBER, CHARMM, GROMOS, OPLS-AA, and UFF are examples of force fields (Adelusi et al., 2022).

Guidelines for Molecular Docking

Hardware and software requirements for molecular docking

Docking computations do not require an intense processing unit (CPU) due to their quick completion time. Small docking efforts (500-1000 compounds) can now be run efficiently on most personal computers (PCs). Virtual screening of public databases employing docking-based approaches can result in a quick increase in the number of molecules (106 compounds). To complete the operation in a reasonable amount of time, additional data processors are required. GPU data processing is more efficient and appealing for heavy tasks than CPU-based computations (Murugan et al., 2022).

Program selection for molecular docking

There are various docking methods and approaches. For novices, accessible academic or free software is preferred. Some docking programs were not designed to run on Windows. Beginners can use Linux to overcome such issues. Additionally, Windows-friendly apps like AutoDock, Vina, and LeDock may be preferred (Bender et al., 2021).

Steps for molecular docking

The molecular docking process involves identifying and preparing the target protein, preparing the ligand, selecting the appropriate docking scoring function, and validating the results.

Target protein identification and preparation

The qualities of the selected protein structure are critical for the result of docking. The use of technology, such as X-ray crystallography, NMR, and cryo-EM, has increased to a large extent the number of proteins with known three-dimensional (3D) structures, which may now be stored in a public database, such as the Protein Data Bank. The first step in docking is to obtain the 3D structure of the protein preferably associated with a ligand, but also preferred from the PDB. 3D structures should be highly resolved or bound by high-affinity ligands. This may not be the case for some proteins. If a protein's 3D structure is unavailable in the PDB, homology modelling is used to generate it. Protocol preparation involves adding hydrogens, removing water, establishing charges, and minimizing energy. AutoDock and SwissDock implement an in-program force field, while MOE utilizes AMBER, and LeDock uses CHARMM charges and atomic species.

Ligand preparation

Ligand structures are generated by applications such as ChemDraw or can be found in databases such as PubChem and ZINC. Preclinical stress, before docking to a Ligand structure, is recommended to reduce the energy level. The grid represents binding space, which is then used to calculate the binding energetics. A box of known size is composed of small squares with probe atoms on each of them, specifying the boundary of possible interactions. The resolution of the grid and the size of the grid affect the docking results (Aguiar & Camps et al., 2024).

Docking type determination

Different types of docking are selected depending on the research needs of the researcher. Flexible docking algorithms are ideal for docking multiple compounds to a protein binding site with low pH, no water, and low solubility, but when inspecting thousands of compounds from databases, it may not be possible without the high processor power and fast computer, so the researcher can choose various docking algorithms based on computer work capacity and the target characteristics (Mursal et al., 2024).

Choose the most effective docking scoring function

The stability of the ligand-protein complex is a major consideration in selecting the optimal docking scoring function. Choosing a scoring function that accurately detects binding patterns and ligand-binding candidates may be challenging, and in theory, ligand protein complexes with lower binding free energy (G) are more stable. Another scoring function is known as empirical scoring, which considers the free energy of binding from hydrogen bonding and Van der Waals contacts, electrostatics, hydrophobic interactions, and the conformational free energy produced by ligand binding. Force field computing computes energy using molecular mechanics, analogous to CHARMM and AMBER, that includes internal energies, coulombic interactions (including Van der Waals), and hydrogen bonding (Meli et al., 2022).

Docking validation

As with any other technique, validation should be performed before the start and end of the docking, as well as when the same ligand/target is docked again. For highly complex target-ligand docking evaluation you should use molecular dynamics simulations to validate your targets, improve your selection criteria and then fix those complexes that would not come out in your docked results, calculate binding free energy, make sure you accounted for solvent effects in the ligand sequence and finally know what binding poses, residues and energy you obtained from the docking (Stanzione et al., 2021).

Application of molecular docking in drug design and discovery (Figure 1)

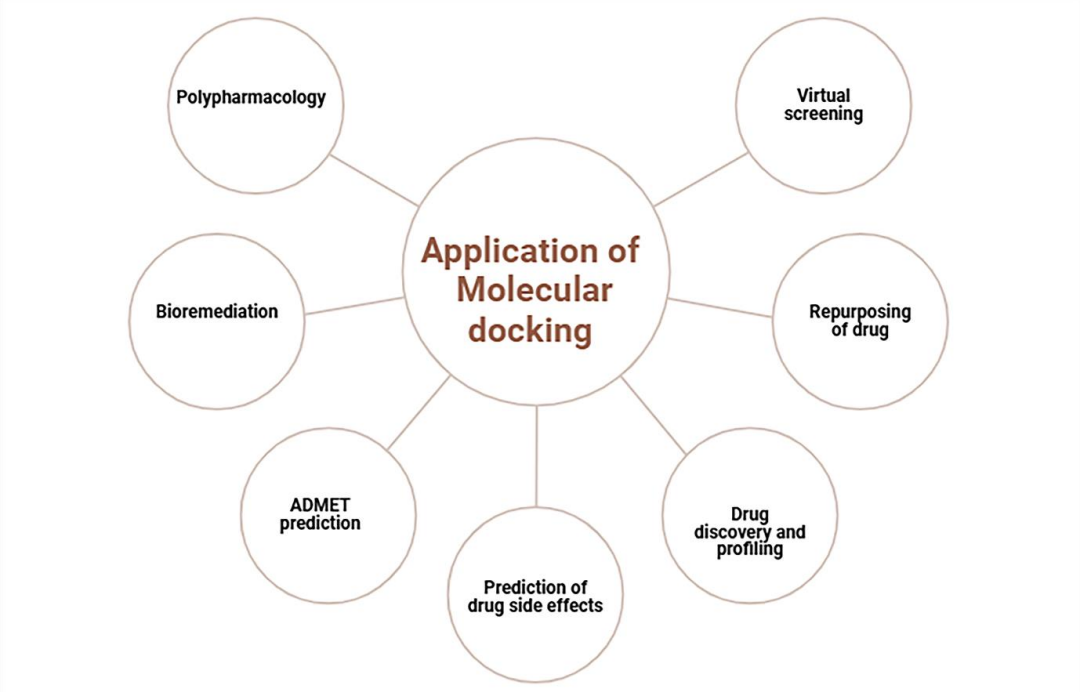

Figure 1. Application of molecular docking

Virtual screening

In drug discovery, molecular docking is often used to identify hits. Through the prediction of the binding affinity of small molecules to a protein or receptor of interest, molecular docking can aid in identifying novel therapeutic candidates. A huge database of small molecules can be screened by docking to identify compounds that have a high affinity for binding to a target protein (Zhu et al., 2022).

Lead optimization

Molecular docking can be used to optimize the structure of the target molecule to increase its affinity and selectivity after finding a hit compound. By modelling the binding patterns of modified structures, the process can also be used to make new analogs (Budipramana & Sangande, 2022).

Bioremediation

In bioremediation, molecular docking can be used to model how well small molecules will bind to enzymes that digest environmental contaminants and to design inhibitors or activators of these enzymes to increase bioremediation efficiency (Balan et al., 2021).

ADMET prediction

It is possible to predict the Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) properties of small molecules with the use of docking. Early in the drug development process, molecules with undesirable properties can be filtered out using the predicted ADMET properties (Ibrahim et al., 2023). AutoDock Vina, GOLD (Genetic Optimization for Ligand Docking), Glide, and the Schrödinger Suite are among the programs that offer efficient algorithms and computational methods for ligand-receptor docking simulations, allowing for the identification of potential therapeutic candidates and elucidating binding affinity. Moreover, they provide ADMET prediction modules that enable to investigation of the drug's actions in terms of its absorption, distribution, metabolism, excretion, and possible toxicity (Kumar et al., 2021).

Molecular dynamics simulation

Differential dynamics of protein–ligand complexes can be studied using molecular docking as well as molecular dynamics simulations. The stability and stability of the complex and conformational changes seen during ligand binding can be better investigated with these simulations. Molecular docking and molecular dynamics simulation are combined in some software tools. These include the most commonly used software such as Glide, AutoDock, Vina, and GOLD, which allow molecular dynamics simulation in addition to molecular docking, allowing for the investigation of protein–ligand interactions at long-term scales as well as the dynamic behavior of proteins (Vidal-Limon et al., 2022).

Recent developments and innovations in molecular docking

Molecular docking has lately made substantial advances, thanks to the integration of artificial intelligence (AI) and machine learning (ML) techniques, as well as improvements in computer capacity and algorithms (Gupta et al., 2021).

Integrating AI and machine learning

DIFDOCK, which works with high precision to achieve comprehensive prosperity and propagation of pose, is one of the remarkable AI-based docking approaches. Deep learning models, such as deep neural networks, can process large datasets and capture complex patterns that can be ignored by traditional methods (Dhudum et al., 2024; Guo et al., 2024). AI has transformed drug discovery, development, and precision medicine through the use of machine learning, deep learning (e.g., convolutional neural networks), and natural language processing is transforming drug discovery by leveraging large amounts of data and computer power to identify and validate targets. Machine learning and deep learning algorithms evaluate various datasets, including genomes, proteomics, and clinical data, to identify potential therapeutic targets (Wen et al., 2020). The many methodologies utilized in AI-assisted target identification and validation are as follows: (1) Statistical analysis uses omics data, such as genome-wide association studies (GWAS) and summary data-based Mendelian randomization (SMR), to identify disease-related target genes. (2) Network-based techniques can uncover complex biological linkages. Gene co-expression and miRNA-disease networks can uncover disease-associated gene sets and pathways. Target identification uses knowledge graphs, which include entities, connections, and semantic information, to represent and evaluate data. (3) Machine learning-based techniques. ML approaches, such as classifiers (e.g., random forest, support vector machine, Neural Net) and regression models, are used to determine a gene's potential therapeutic target. AI algorithms can verify new targets by predicting druggability and therapeutic appropriateness. This technique allows for research of hitherto undiscovered targets while reducing reliance on experimentally proven theories (Tiwari et al., 2023).

Improved computational techniques

In computational approaches, advances have greatly increased the speed and accuracy of molecular docking simulations. EXSCALATE4COV and other projects have used high-scale, high-performance computing to perform large-scale virtual screening, accelerating the discovery of new medicinal chemicals. In addition, technologies such as Rosetta@home use distributed computing to predict protein structures and docking interactions, allowing more fuller understanding of molecular interactions (Gayathiri et al., 2023). Diffusion models (DM) are a crucial methodology that drives these advancements. Diffusion models emerged in computer vision and quickly became popular for picture production due to their high quality and performance. These models were then adapted for application in other fields, such as computational structural biology. DMs can effectively model high-dimensional geometric data and leverage deep learning capabilities. In structural biology, they have produced cutting-edge achievements for protein 3D structure synthesis and small molecule docking (Yim et al., 2024). With the recent development of chemical libraries exceeding one billion molecules, more effective virtual screening technologies are required. The Deep Docking (DD) technology accelerates structure-based virtual screening by docking just a portion of a chemical library repeatedly while predicting the remaining docking scores using ligands. This technique allows the screening of chemical libraries the size of billions of molecules without the need for enormous computer resources since it produces virtual hit enrichment hundreds to thousands of times without significantly sacrificing possible therapeutic candidates (Gentile et al., 2022).

Approaches: fragmentation and consensus

Recent developments have also focused on piece-based methods, where small chemical pieces are used for the manufacture of potential ligands, and unified approaches that add several scoring functions to improve the docking accuracy. The purpose of these strategies is to increase the reliability of docking predictions and is an integral part of modern drug search processes (Wilson III et al., 2021). Fragment-based drug discovery identifies potential hit compounds by targeting binding hotspots. Combining pieces enables exploration of huge chemical spaces. This approach effectively identifies lead compounds. Growing, merging, and connecting are the three ideas used to combine pieces into a compound. Artificial intelligence (AI) has enhanced molecular design accuracy and efficiency (Yoo et al., 2025).

Advanced tools in molecular docking

Based models are developing as effective tools for designing molecules and property prediction. However, the absence of explainability in these models remains a major concern. ABIET (Attention-Based Importance Estimation Tool) is an explainable Transformer model that identifies key functional groups (FGs) for drug-target interactions in physiologically active compounds. Functional groups are essential for regulating biological interactions and chemical activity. This method uses focus weights via Transformer-encoder architectures based on SMILES representations to determine the relevance of molecular subregions. Using a particular approach to analyze attention ratings, taking into account layer-based extraction, bidirectional interactions, and activation transformations (Pereira et al., 2025).

Limitations and challenges in molecular docking

Molecular docking, despite its advances, has significant constraints and obstacles that affect its accuracy and dependability in drug discovery. These issues can be divided into three categories: computational, methodological, and biological (Agu et al., 2023).

Limitations in computation

Large-scale docking experiments require adequate computer power, which can be expensive and time-consuming. Most of the methods of docking use scoring functions that are for the approximate binding efficiency, but they often fail to capture complex molecular interactions such as water-mediated hydrogen bonding and entropy contributions. Binding (Gentile et al., 2022).

Challenges with methodology

Despite the progress, docking algorithms sometimes produce incorrect ligand orientation within the binding site. Money docking programs employ simplified molecular representations that ignore solvent effects, induced fit, and oligomer binding sites.(Ugurlu, 2024).

Scoring functions: desolvation and entropic effects

Scoring functions are critical in virtual screening for predicting binding affinity, and different techniques (e.g., target-specific scoring such as GOLD for kinase inhibition, empirical scoring, and consensus scoring) can assist in improving predictions. Accounting for protein flexibility is one of the most difficult tasks. Scoring functions must be both rapid and accurate. While several have been designed with these aims in mind, they have inherent flaws, particularly those related to solvent (desolvation) and entropy effects. For solvent effects, force-field-based scoring algorithms have difficulty adequately modeling the action of water. Solutions include employing a distance-dependent dielectric constant (e.g., in DOCK) or more advanced but computationally expensive approaches such as Free Energy Perturbation (FEP), Thermodynamic Integration (TI), OWFEG, and LIEPROFEC. Empirical scoring functions can help to mitigate entropy effects. For example, PHOENIX adds entropic effects based on experimental data from isothermal titration calorimetry. Knowledge-based scoring functions also face challenges in defining an ideal reference state (as per the inverse Boltzmann relation), which is very hard for complex systems like proteins.

Crystallographic water

A key problem in structure-based drug design is the existence of tightly bonded crystallographic water molecules within the receptor's binding pocket, which frequently mediate hydrogen bonding between the ligand and the protein. During docking, these water molecules might be added or removed from the binding site. Excluding them is often entropically advantageous, although it may result in a loss of enthalpy, which can be compensated for if the ligand is tailored to replicate the hydrogen bond network. Alternatively, the water molecule can be kept intact to establish hydrogen bonds with the binding site. To determine whether to include or omit such water molecules, free energy perturbation (FEP) computations using Monte Carlo simulations can be performed to predict binding energies.

Conclusion

Molecular docking, a computationally powerful technique, is now indispensable for accurate and rapid prediction of ligand-target interactions in biomolecular research and discovery. In addition to various docking methods like semi-rigid, flexible docking, and the most complex algorithm, which is the key component of molecular docking virtual screening, lead optimization, and structure-based drug design. Molecular docking, although advantageous, has drawbacks. What can affect the accuracy of docking results includes protein flexibility, weighting functions, and computational limitations. More importantly, however, recent in time advancements in docking were artificial intelligence, machine learning integration, and hybrid modeling techniques provide us with better accuracy and speed to dock. We can improve its predictability with effort in molecular docking & further advancement of consensus scoring technology, and high-throughput virtual screening. Objectives will need integration of experimental and computational validation strategies to bridge the current interoperability barriers, as well as to maximize this focused pharmaceutical search. At last, molecular docking is used in current medical research since it helps to decode new treatments and molecular interactions.

Acknowledgments

The authors would like to acknowledge the management of Swamy Vivekanandha College of Pharmacy, Tiruchengode, Namakkal, Tamil Nadu, India, for access to the scientific journals and for providing all the facilities.

Author contributions

All Authors have contributed equally and substantially.

Funding

No funding.

Conflict of interest

The authors declare no conflict of interest. The manuscript has not been submitted for publication in any other journal.

Ethics approval

Not applicable

Consent to participate

Not applicable

Consent to Publish

Not applicable

AI tool declaration

The authors have not used AI and it's related to tools to write this manuscript.

References

Abdalfattah, S., Knorz, C., Ayoobi, A., Omer, E. A., Rosellini, M., Riedl, M., ... & Efferth, T. (2024). Identification of antagonistic action of pyrrolizidine alkaloids in muscarinic acetylcholine receptor M1 by computational target prediction analysis. *Pharmaceuticals*, *17*(1), 80.

Adelusi, T. I., Oyedele, A. Q. K., Boyenle, I. D., Ogunlana, A. T., Adeyemi, R. O., Ukachi, C. D., ... & Abdul-Hammed, M. (2022). Molecular modeling in drug discovery. *Informatics in Medicine Unlocked*, *29*, 100880.

Agu, P. C., Afiukwa, C. A., Orji, O. U., Ezeh, E. M., Ofoke, I. H., Ogbu, C. O., ... & Aja, P. M. (2023). Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management. *Scientific reports*, *13*(1), 13398.

Aguiar, C., & Camps, I. (2024). Molecular Docking in Drug Discovery: Techniques, Applications, and Advancements.

Ajjarapu, S. M., Tiwari, A., Ramteke, P. W., Singh, D. B., & Kumar, S. (2022). Ligand-based drug designing. In *Bioinformatics* (pp. 233-252). Academic Press.

Balan, B., Dhaulaniya, A. S., Varma, D. A., Sodhi, K. K., Kumar, M., Tiwari, M., & Singh, D. K. (2021). Microbial biofilm ecology, in silico study of quorum sensing receptor-ligand interactions and biofilm mediated bioremediation. *Archives of Microbiology*, *203*(1), 13-30.

Bender, B. J., Gahbauer, S., Luttens, A., Lyu, J., Webb, C. M., Stein, R. M., ... & Shoichet, B. K. (2021). A practical guide to large-scale docking. *Nature protocols*, *16*(10), 4799-4832.

Budipramana, K., & Sangande, F. (2022). Molecular docking-based virtual screening: Challenges in hits identification for Anti-SARS-CoV-2 activity. *Pharmacia, 69*(4), 1047-1056.

Bugnon, M., Röhrig, U. F., Goullieux, M., Perez, M. A., Daina, A., Michielin, O., & Zoete, V. (2024). SwissDock 2024: major enhancements for small-molecule docking with Attracting Cavities and AutoDock Vina. *Nucleic acids research*, *52*(W1), W324-W332.

Cavasotto, C. N., & Abagyan, R. A. (2004). Protein flexibility in ligand docking and virtual screening to protein kinases. *Journal of molecular biology*, *337*(1), 209-225.

Chaudhry, H., Schleip, R., Ji, Z., Bukiet, B., Maney, M., & Findley, T. (2008). Three-dimensional mathematical model for deformation of human fasciae in manual therapy. *Journal of Osteopathic Medicine*, *108*(8), 379-390.

Chu, Z.-Y., & Zi, X.-J. (2024). Network pharmacology, single gene survival analysis and molecular docking to study the mechanism of Sotetsuflavone in the treatment of pancreatic cancer. *Pharmacological Research-Reports, 2*, 100010.

Dhudum, R., Ganeshpurkar, A., & Pawar, A. (2024). Revolutionizing drug discovery: a comprehensive review of AI applications. *Drugs and Drug Candidates*, *3*(1), 148-171.

Doytchinova, I. (2022). Drug Design—Past, Present, Future. *Molecules, 27*(5), 1496. https://www.mdpi.com/1420-3049/27/5/1496.

Gayathiri, E., Prakash, P., Kumaravel, P., Jayaprakash, J., Ragunathan, M. G., Sankar, S., ... & Govindasamy, R. (2023). Computational approaches for modeling and structural design of biological systems: A comprehensive review. *Progress in Biophysics and Molecular Biology*, *185*, 17-32.

Gentile, F., Yaacoub, J. C., Gleave, J., Fernandez, M., Ton, A.-T., Ban, F., ... & Cherkasov, A. (2022). Artificial intelligence–enabled virtual screening of ultra-large chemical libraries with deep docking. *Nature Protocols*, *17*(3), 672-697.

Grassmann, F., Mälarstig, A., Dahl, L., Bendes, A., Dale, M., Thomas, C. E., ... & Czene, K. (2024). The impact of circulating protein levels identified by affinity proteomics on short-term, overall breast cancer risk. *British Journal of Cancer*, *130*(4), 620-627.

Guo, Z., Liu, J., Wang, Y., Chen, M., Wang, D., Xu, D., & Cheng, J. (2024). Diffusion models in bioinformatics and computational biology. *Nature reviews bioengineering, 2*(2), 136-154.

Gupta, R., Srivastava, D., Sahu, M., Tiwari, S., Ambasta, R. K., & Kumar, P. (2021). Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Molecular diversity*, *25*(3), 1315-1360.

Halperin, I., Ma, B., Wolfson, H., & Nussinov, R. (2002). Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins: Structure, Function, and Bioinformatics*, *47*(4), 409-443.

Holmes, J., Liu, Z., Zhang, L., Ding, Y., Sio, T. T., McGee, L. A., ... & Liu, W. (2023). Evaluating large language models on a highly-specialized topic, radiation oncology physics. *Frontiers in Oncology*, *13*, 1219326.

Ibrahim, H., Liu, F., Asim, R., Battu, B., Benabderrahmane, S., Alhafni, B., ... & Zaki, Y. (2023). Perception, performance, and detectability of conversational artificial intelligence across 32 university courses. *Scientific reports*, *13*(1), 12187.

Kaczor, K., Gerula-Szymańska, A., Smektała, T., Safranow, K., Lewusz, K., & Nowicka, A. (2018). Effects of different etching modes on the nanoleakage of universal adhesives: A systematic review and meta-analysis. *Journal of Esthetic and Restorative Dentistry*, *30*(4), 287-298.

Kumar, A., Kini, S. G., & Rathi, E. (2021). A recent appraisal of artificial intelligence and in silico ADMET prediction in the early stages of drug discovery. *Mini Reviews in Medicinal Chemistry, 21*(18), 2788-2800.

Kumar, S., Flinn, I., Richardson, P. G., Hari, P., Callander, N., Noga, S. J., ... & Rajkumar, S. V. (2012). Randomized, multicenter, phase 2 study (EVOLUTION) of combinations of bortezomib, dexamethasone, cyclophosphamide, and lenalidomide in previously untreated multiple myeloma. *Blood, The Journal of the American Society of Hematology*, *119*(19), 4375-4382.

Meli, R., Morris, G. M., & Biggin, P. C. (2022). Scoring functions for protein-ligand binding affinity prediction using structure-based deep learning: a review. *Frontiers in bioinformatics, 2*, 885983.

Mursal, M., Ahmad, M., Hussain, S., & Khan, M. F. (2024). Navigating the computational seas: a comprehensive overview of molecular docking software in drug discovery. *Unravelling Molecular Docking-From Theory to Practice*.

Murugan, N. A., Podobas, A., Gadioli, D., Vitali, E., Palermo, G., & Markidis, S. (2022). A review on parallel virtual screening softwares for high-performance computers. *Pharmaceuticals, 15*(1), 63.

Pereira, J. F. S., Bessa, C., Matos, P. & Jordan, P. (2025). *Correction: Pereira et al. Pro-Inflammatory Cytokines Trigger the Overexpression of Tumour-Related Splice Variant RAC1B in Polarized Colorectal Cells. Cancers, 17*(10), 1642.

Pojtanadithee, P., Isswanich, K., Buaban, K., Chamni, S., Wilasluck, P., Deetanya, P., ... & Rungrotmongkol, T. (2024). A combination of structure-based virtual screening and experimental strategies to identify the potency of caffeic acid ester derivatives as SARS-CoV-2 3CLpro inhibitor from an in-house database. *Biophysical Chemistry*, *304*, 107125.

Raval, K., & Ganatra, T. (2022). Basics, types and applications of molecular docking: A review. *IP International Journal of Comprehensive and Advanced Pharmacology, 7*(1), 12-16.

Ritchie, A., Robertson, S., & Teufel, S. (2008). Comparing citation contexts for information retrieval. In *Proceedings of the 17th ACM conference on Information and knowledge management* (pp. 213-222).

Sahu, D., Rathor, L. S., Dwivedi, S. D., Shah, K., Chauhan, N. S., Singh, M. R., & Singh, D. (2024). A Review on Molecular Docking As an Interpretative Tool for Molecular Targets in Disease Management. *ASSAY and Drug Development Technologies, 22*(1), 40-50. https://doi.org/10.1089/adt.2023.060.

Stanzione, F., Giangreco, I., & Cole, J. C. (2021). Use of molecular docking computational tools in drug discovery. *Progress in medicinal chemistry, 60*, 273-343.

Surana, K. R., Ahire, E. D., Sonawane, V. N., & Talele, S. G. (2021). Biomolecular and molecular docking: A modern tool in drug discovery and virtual screening of natural products. In *Applied Pharmaceutical Practice and Nutraceuticals* (pp. 209-223). Apple Academic Press.

Tang, Z., Bai, Y., Fang, Q., Yuan, Y., Zeng, Q., Chen, S., Xu, T., Chen, J., Tan, L., Wang, C., Li, Q., Lin, J., Yang, Z., Wu, X., Shi, G., Wang, J., Yin, C., Guo, J., Liu, S., Peng, S., & Kuang, M. (2025). Spatial transcriptomics reveals tryptophan metabolism restricting maturation of intratumoral tertiary lymphoid structures. *Cancer Cell, 43*(6), 1025–1044.e14.

Tiwari, A., Adhikari, S., Kaya, D., Islam, M. A., Malla, B., Sherchan, S. P., ... & Pitkänen, T. (2023). Monkeypox outbreak: Wastewater and environmental surveillance perspective. *Science of the total environment*, *856*, 159166.

Ugurlu, S. Y. (2024). Computational Methods in Drug Discovery and Development.

Vidal-Limon, A., Aguilar-Toalá, J. E., & Liceaga, A. M. (2022). Integration of molecular docking analysis and molecular dynamics simulations for studying food proteins and bioactive peptides. *Journal of Agricultural and Food Chemistry, 70*(4), 934-943.

Wen, W., Su, W., Tang, H., Le, W., Zhang, X., Zheng, Y., ... & Wang, H. (2020). Immune cell profiling of COVID-19 patients in the recovery stage by single-cell sequencing. *Cell discovery*, *6*(1), 31.

Wilson III, D. M., Deacon, A. M., Duncton, M. A., Pellicena, P., Georgiadis, M. M., Yeh, A. P., Arvai, A. S., Moiani, D., Tainer, J. A., & Das, D. (2021). Fragment-and structure-based drug discovery for developing therapeutic agents targeting the DNA Damage Response. *Progress in Biophysics and Molecular Biology, 163*, 130-142.

Xu, L. D., Xu, E. L., & Li, L. (2018). Industry 4.0: state of the art and future trends. *International journal of production research*, *56*(8), 2941-2962.

Yim, W. W., Abacha, A. B., Fu, Y., Sun, Z., Xia, F., Yetisgen-Yildiz, M., & Krallinger, M. (2024). Overview of the MEDIQA-M3G 2024 Shared Task on Multilingual Multimodal Medical Answer Generation. In *Proceedings of the 6th Clinical Natural Language Processing Workshop* (pp. 581-589).

Yoo, C., Zhou, Y., & Weng, Q. (2025). Mapping 10-m industrial lands across 1000+ global large cities, 2017–2023. *Scientific Data*, *12*(1), 278.

Zhu, H., Zhang, Y., Li, W., & Huang, N. (2022). A comprehensive survey of prospective structure-based virtual screening for early drug discovery in the past fifteen years. *International Journal of Molecular Sciences, 23*(24), 15961.