Molecular Docking in Drug Discovery: Techniques, Applications, and Advancements

C. Aguiar[1] and I. Camps[1,*]

[1]Laboratório de Modelagem Computacional - *La*Model, Instituto de Ciências Exatas - ICEx. Universidade Federal de Alfenas - UNIFAL-MG, Alfenas, Minas Gerais, Brazil

*icamps@unifal-mg.edu.br

Abstract

Objective: The primary objective of this study is to conduct a comprehensive review of the significance of molecular docking in the field of drug discovery. This includes an examination of the various approaches and methods used in molecular docking, as well as an exploration of the techniques used for interpreting and validating docking results.

Methods: To gather relevant data, a systematic search was conducted using Web of Science, PubMed, and Google Scholar. The search focused on articles related to molecular docking methodologies and their applications in drug discovery. Additionally, alternative techniques that can be used for more precise simulations of ligand-protein interactions were also considered.

Results: Molecular docking has proven to be an incredibly rich and valuable process in the field of drug discovery. Its flexibility allows for the incorporation of advanced computational techniques, thereby enhancing the reliability and efficiency of drug discovery processes. The results of the study highlight the significant strides made in the field of molecular docking, demonstrating its potential to revolutionize drug discovery.

Conclusions: Molecular docking continues to evolve, with new advancements being made regularly. Despite the challenges faced, these advancements have significantly contributed to the enhancement of molecular docking, solidifying its position as a crucial tool in the field of drug discovery.

Keywords: molecular docking; scoring functions; validation; machine learning; ligand-protein

INTRODUCTION

Molecular docking stands as a cornerstone in the modern drug discovery process, providing a computational lens through which the interaction between small molecule ligands and their macromolecular targets can be scrutinized. At its core, molecular docking simulates the "lock-and-key" mechanism that underlies molecular recognition, which is pivotal for the identification and optimization of compounds with therapeutic potential [1–4]. The technique's allure lies in its ability to predict how a ligand binds to a protein, thereby offering insights into the binding affinity and biological activity of the ligand, which are crucial for the rational design of drugs [2,5].

Despite its widespread adoption and success, molecular docking is not without its challenges. One of the primary hurdles is the accuracy of scoring functions, which are mathematical models used to predict the binding affinity between the ligand and its target [6]. These functions must strike a delicate balance between computational efficiency and the ability to accurately replicate the complex physicochemical phenomena occurring at the molecular level [6,7]. The inherent limitations of scoring functions, such as their simplified treatment of solvent effects and protein flexibility, often necessitate the use of additional validation strategies to ensure the reliability of docking predictions [7,8].

The field of molecular docking is continuously evolving, with recent advances aimed at addressing these challenges. The integration of machine learning algorithms with docking simulations has shown promise in enhancing the predictive power of scoring functions [3,9]. Furthermore, the development of methods that account for the dynamic nature of proteins, such as ensemble docking and induced-fit models, has improved the representation of protein-ligand interactions [7,10]. These innovations, coupled with the increasing computational power and the availability of high-quality structural data, are propelling molecular docking towards more accurate and efficient drug discovery workflows [2,3,10].

Figure 1 shows the exponential evolution of the number of indexed papers related to molecular docking since year 2000. These data indicates that molecular docking remains a vital interpretative tool in drug discovery, enabling the identification of novel therapeutics and the management of various diseases. As computational methods become more sophisticated and integrated with experimental data, the potential of molecular docking to contribute to the discovery of next-generation drugs continues to expand [2,3,5].

APPROACHES AND TECHNIQUES

There are two main approaches within molecular docking: rigid docking and flexible docking.

Rigid docking is a computational approach that treats both the ligand and the receptor as rigid bodies, meaning that their conformations do not change during the docking process. This method simplifies the interaction between the two molecules by considering only their translational and rotational degrees of freedom without accounting for any flexibility in their structures. It is particularly useful for initial screening in virtual screening campaigns where the goal is to quickly identify potential binders from large libraries of compounds. It can also be used when the binding site is well-defined and when the ligand and receptor are known to

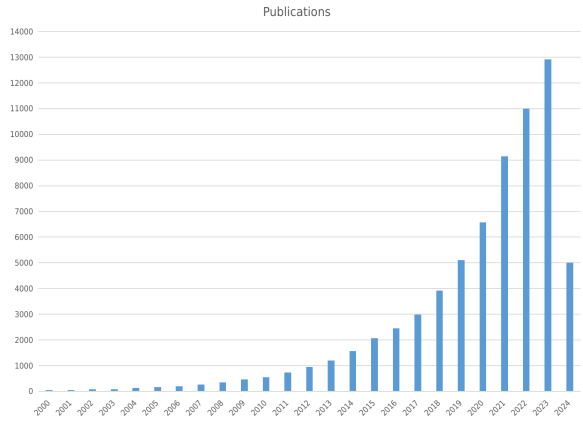

Figure 1 Evolution of number of papers with "molecular docking" keyword. Data included herein are derived from Clarivate™ (Web of Science™). © Clarivate 2024. All rights reserved.

undergo minimal conformational changes upon binding. Because it does not account for conformational changes, rigid docking can be much faster than flexible docking approaches. It is often used to examine large systems, such as protein-protein interactions, where the number of potential conformations can be prohibitively large for flexible docking methods [11].

Flexible docking is a sophisticated approach in molecular docking that accounts for the flexibility of the ligand and, in some cases, the receptor (protein) during the docking process. This method is crucial for accurately predicting how a molecule binds to a target, as it considers the dynamic nature of molecular interactions, which can significantly influence binding affinity and specificity. It allows the ligand to adopt multiple conformations during the docking process. This is essential because ligands can change their shape to fit more snugly into the binding site of the receptor. Techniques such as incremental construction from rigid parts, global energy optimization, and multi-conformer docking are used to explore the conformational space of the ligand [12]. While more computationally intensive, some flexible docking approaches also account for receptor flexibility. This can involve modeling key side chains of the receptor as flexible, or using molecular dynamics simulations to generate multiple receptor conformations that can be used in docking [13–15].

High-throughput docking (HTD)

High-throughput docking (HTD) is a computational technique used in drug discovery to rapidly screen vast libraries of small molecules (ligands) to identify those that are most likely to bind to a target protein with high affinity. This process is crucial for identifying potential drug candidates in the early stages of drug development. HTD leverages the principles of molecular docking, which predicts the preferred orientation of a ligand to a protein (receptor) to form a stable complex, thereby estimating the binding affinity between the two molecules. HTD is primarily used in structure-based drug design to enrich a sub-library with potential ligands from a larger chemical library. This enriched sub-library is then prioritized for further experimental evaluation. The main goal is to efficiently and effectively identify molecules that are likely to interact with the target protein, thereby reducing the time and cost associated with experimental high-throughput screening (HTS) [16].

Ensemble docking

Ensemble docking is a computational technique used in the field of molecular docking to address the challenge of protein flexibility when predicting the binding of ligands to their protein targets. Traditional docking methods often treat the protein as a rigid structure, which can lead to inaccuracies because proteins are dynamic and can adopt multiple conformations. Ensemble docking overcomes this limitation by considering multiple conformations of the protein target during the docking process. It uses a set of different protein structures, or an ensemble to represent the various conformations that a protein can adopt. These structures can be sourced from experimental data, such as X-ray crystallography or NMR spectroscopy, or generated computationally through techniques like molecular dynamics simulations [17,18]. By accounting for protein flexibility, ensemble docking can provide a more accurate prediction of the ligand's binding mode and affinity. It allows the ligand to interact with different conformations of the binding site, which can lead to the identification of the most favorable binding pose across the ensemble [17,19]. Ensemble docking is computationally more intensive than traditional docking because it involves multiple docking runs, one for each protein conformation. However, advancements in computational power and parallel processing have made it more feasible to perform ensemble docking on a large scale [20,21].

Incorporation of solvent

The solvent can be considered as implicit or explicit. The difference between implicit and explicit solvent models in molecular docking lies primarily in how they simulate the solvent environment's effect on the molecular interactions between the ligand and the receptor.

Implicit solvent models simplify the solvent environment by treating it as a continuous medium rather than simulating individual solvent molecules. This approach averages the solvent's effects on the solute, focusing on the macroscopic properties of the solvent, such as its dielectric constant. Because they do not require the simulation of individual solvent molecules, implicit models are computationally less demanding. This allows for faster simulations and the ability to study larger systems or longer time scales with reduced computational resources. While implicit models can speed up simulations and are useful for exploring conformational space, they may lack the detailed representation of specific solvent-solute interactions, such as hydrogen bonding or the precise effect of solvent molecules on the solvation shell of the ligand and receptor [22,23].

The incorporation of explicit solvent in molecular docking is a critical aspect that significantly influences the accuracy of docking predictions. This approach involves explicitly including water molecules or other solvent molecules in the docking simulations to more accurately model the real-life conditions under which molecular interactions occur. The presence of solvent can affect the conformation of the ligand and the receptor, as well as the interaction between them, making the consideration of solvent effects essential for realistic docking studies. Water molecules often play a crucial role in ligand recognition and complex stabilization for both nucleic acids and proteins. The polarization effects caused by the solvent can significantly impact the docking process, especially for nucleic acids where the phosphate groups and counter ions like Mg_2^+ or K^+ influence the water molecules and functional groups of drugs. Despite the recognized importance of explicit solvent in docking, incorporating water molecules into the docking process presents significant challenges. The large, solvent-accessible

interface of macromolecules makes it extremely challenging to accurately model the effects of water molecules within a reasonable computational timeframe [24]. It is worth noting that even the water model used can affect the calculated parameters in the simulations [25,26].

Machine learning and artificial intelligence

The integration of machine learning (ML) and artificial intelligence (AI) into molecular docking represents a significant advancement in the field of drug discovery and computational chemistry. These technologies are being used to enhance the accuracy, efficiency, and predictive power of molecular docking simulations. One of the primary applications of ML in molecular docking is the development of improved scoring functions. Traditional scoring functions often struggle to accurately predict binding affinities due to their reliance on simplified physical models. ML models, trained on large datasets of known ligand-receptor interactions, can learn complex patterns and interactions, leading to more accurate predictions of binding affinities. ML algorithms can significantly speed up the virtual screening process by learning to predict which molecules are likely to bind to a given target. This can be particularly useful in high-throughput screening, where ML models can prioritize potentially active compounds for further investigation, thereby reducing the need for extensive experimental assays. AI techniques, particularly deep learning models, can be trained to predict the likely binding poses of ligands within a receptor's active site. These models can learn from large datasets of known ligand-receptor complexes, capturing subtle patterns that may not be apparent through traditional docking methods. AI models can also be used to address the challenge of protein flexibility in docking simulations. By training on dynamic data from molecular dynamics simulations, ML models can predict how protein conformational changes affect ligand binding, which is a significant challenge for traditional docking methods that often consider the protein as a rigid body [27,28].

Deep learning architectures like convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are particularly suited for modeling the spatial and temporal complexities of molecular interactions. These models have been used to develop new docking algorithms that can automatically learn from structural data. Reinforcement learning (RL) has been applied to optimize the docking process itself. In this approach, the docking algorithm, modeled as an agent, learns to improve its strategy for exploring the conformational space of the ligand and receptor to maximize the predicted binding affinity. Transfer learning involves applying knowledge gained from one problem domain to a different but related problem. In the context of docking, models trained on large datasets of molecular interactions can be fine-tuned with smaller, specific datasets to improve performance on particular types of targets or ligands [3].

Hybrid methods

Hybrid methods in molecular docking are integrated computational approaches that combine different techniques to enhance the accuracy and effectiveness of docking simulations. These methods leverage the strengths and mitigate the drawbacks of individual virtual screening (VS) strategies, thereby improving the outcomes of computer-aided drug design. Hybrid methods can incorporate a variety of computational techniques, including but not limited to, molecular docking, molecular dynamics simulations, quantum mechanics, machine learning, pharmacophore modeling, and ligand-based drug design. The integration of these methods allows for a more comprehensive analysis of ligand-receptor interactions by considering factors such as protein flexibility, solvent effects, and the dynamic nature of molecular interactions. Here are some examples of hybrid methods used in molecular docking.

Combination of Scoring Functions: A hybrid method may involve the use of multiple scoring functions to evaluate the binding affinity of ligand-receptor complexes. This can include a combination of empirical, knowledge-based, and physics-based scoring functions to provide a more comprehensive assessment of the docking results. *Integration of Ligand-Based and Structure-Based Docking*: Hybrid methods can integrate ligand-based pharmacophore modeling with structure-based docking to enhance the virtual screening process. This approach can improve the identification of bioactive molecules by considering both the chemical properties of known ligands and the structural details of the target protein. *Use of Molecular Dynamics Simulations*: Molecular dynamics (MD) simulations can be integrated with docking to account for protein flexibility and solvent effects. In this hybrid approach, MD simulations are used to generate multiple conformations of the protein, which are then used in docking simulations to predict how ligands bind to different protein states. *Machine Learning-Enhanced Docking*: Machine learning algorithms can be used to refine docking predictions. For example, ML models can be trained on large datasets of known interactions to improve the prediction of binding poses and affinities or to re-score docking results to better correlate with experimental data. *Quantum Mechanics/Molecular Mechanics (QM/MM)*: This hybrid approach combines the accuracy of quantum mechanics (QM) for modeling the electronic interactions within the docking site with the efficiency of molecular mechanics (MM) for the rest of the system. QM/MM methods are particularly useful for capturing the electronic effects that are critical for understanding enzyme mechanisms, reaction energetics, and the influence of metal ions in docking simulations. *Pharmacophore-based Docking*: Pharmacophore models represent the spatial arrangement of features that are necessary for a molecule to interact with a specific biological target. Hybrid methods that incorporate pharmacophore models into docking simulations can improve the identification of bioactive molecules by considering both the chemical properties of known ligands and the structural details of the target protein. This approach can implicitly account for the flexibility of the receptor, which is often a limitation in traditional docking methods. *High-Throughput Docking Using Quantum Mechanical Scoring*: High-throughput docking (HTD) is a computational technique used to rapidly screen large libraries of compounds. A hybrid method that uses quantum mechanical scoring within HTD can provide a more accurate description of protein-ligand interactions, capturing the underlying physics of the molecular system and accounting for all contributions to the energy, including those effects usually missing in classical force fields [16,27,29–36].

Fragment-based docking

Fragment-based drug discovery (FBDD) involves screening a library of small, low-molecular-weight compounds against a target protein to identify those that bind with high efficiency and specificity. These fragments are then used as starting points for the development of more potent and selective drug candidates. Compared to traditional high-throughput screening (HTS), FBDD can efficiently cover chemical space with a higher hit rate. It requires fewer compounds to be screened and can identify binders with high ligand efficiency, which is a measure of the binding energy per atom of the ligand. In silico methods, including molecular docking, play a crucial role in FBDD. Docking is used to predict the ligand-

receptor interaction modes and to identify hits by structure-based virtual screening. Computational strategies can also guide the growth of fragment hits into potent leads. Fragments are particularly sensitive to scoring problems because they are weak ligands that form few interactions with the protein. Therefore, the correct and incorrect poses can be difficult to distinguish. However, incorporating binding mode information can improve fragment docking, as it helps to sort poses by similarity to the crystallographic structure of the protein in complex with known binders. FBDD is applicable to a wide range of targets, including those considered "undruggable" by traditional methods. The versatility of fragment-based docking allows it to be used in conjunction with other biophysical methods to guide fragment growth and confirm molecular interactions between a target and a ligand [37–40].

Virtual screening and drug repurposing

Virtual screening and drug repurposing are two interconnected strategies in the field of molecular docking that leverage computational methods to expedite the drug discovery process. These approaches are particularly valuable for identifying new therapeutic applications for existing drugs or compounds, thereby reducing the time and cost associated with traditional drug development. Virtual screening (VS) is a computational technique used to evaluate a large library of compounds against a biological target to identify those most likely to bind to and modulate the target's activity. It can be broadly categorized into ligand-based and structure-based (receptor-based) virtual screening. *Ligand-Based Virtual Screening (LBVS)* relies on the knowledge of known active ligands to search for new compounds with similar chemical properties or structural features [11,41]. *Structure-Based Virtual Screening (SBVS)* requires the three-dimensional structure of the target receptor. Molecular docking is a key component of SBVS, where compounds are computationally "docked" into the binding site of the target protein to predict their binding mode and affinity [11,41,42]. Virtual screening has become increasingly important with the advancement of technologies for protein structural science, including cryo-electron microscopy and X-ray crystallography, making receptor-based virtual screening a crucial step in lead compound discovery [11].

DOCKING FRAMEWORK

The steps required for molecular docking can be broadly categorized into several key stages, from preparation of the molecules involved to the analysis of the docking results. Here is a detailed breakdown of these steps:

1. Preparation of ligands and proteins.
2. Identification of the binding site.
3. Setup of docking calculations.
4. Running the docking.
5. Evaluation of results.
6. Validation of results.

Preparation of ligands and proteins

The preparation of ligands and proteins is a crucial initial step in the molecular docking process, ensuring that the molecular structures are suitable for accurate and efficient docking simulations. Ligands can be sourced from databases such as PubChem [43] and ZINC [44] or drawn using chemical drawing tools like ChemDraw, ChemSketch or Maestro if they are not available in databases. Once the ligand structure is obtained, it is important to optimize its geometry and calculate partial charges. This can be done using computational chemistry tools that perform energy minimization and charge distribution calculations. Ligands may need to be prepared in multiple conformations, especially if they are flexible. This involves generating different spatial arrangements of the ligand to explore various possible interactions with the target protein. The prepared ligand structures are then converted into a format suitable for the docking software, such as PDBQT for AutoDock, which includes information on atom types and charges [31,45,46].

Protein structures are typically retrieved from databases like the Protein Data Bank (PDB) [47]. If the structure of the target protein is not available, homology modeling may be used to predict the structure based on similar proteins. The retrieved or modeled protein structure often requires cleaning, which includes the removal of water molecules, ligands, or ions that are not relevant to the docking study. Additionally, missing atoms or residues in the protein structure need to be added. The protonation states of ionizable residues are adjusted, and the protein structure is subjected to energy minimization to relieve any steric clashes and to stabilize the structure [45,48].

Identification of the binding site

The identification of the binding site in molecular docking is of paramount importance because it is where the ligand interacts with the protein, and this interaction is central to the ligand's biological activity. Accurately identifying the binding site is crucial for several reasons [2,31,45]:

Efficiency: Knowing the location of the binding site before docking processes significantly increases the efficiency of the docking. It allows the docking algorithms to focus computational resources on the regions of the protein most likely to interact with the ligand, rather than searching the entire protein surface.

Accuracy: Correctly identifying the binding site ensures that the ligand is docked in the correct location, which is essential for predicting the binding affinity and the biological activity of the ligand accurately.

Mechanistic Insights: Analyzing the interactions within the binding site can provide insights into the mechanism of action of the ligand. Understanding these interactions can guide the optimization of the ligand's structure to improve its efficacy and specificity.

Drug Discovery and Development: In drug discovery, the identification of the binding site is a critical step for structure-based drug design. It enables the discovery of novel compounds with therapeutic potential and assists in the optimization of existing drugs.

Drug Repurposing: Identifying binding sites can also contribute to drug repurposing efforts by revealing alternative targets or pathways that an existing drug might modulate.

Target Validation: Accurate binding site identification helps in validating the target protein for therapeutic intervention, which is a crucial step in the early stages of drug development.

Flexibility and Dynamics: Understanding the binding site's flexibility and dynamics can inform the design of drugs that can accommodate or induce changes in the protein's conformation, which is important for the induced-fit model of enzyme activity.

Setup of docking calculations

Before starting the docking calculations, ensure that the docking software is properly installed on your system. You may also need additional software for preparing the inputs and analyzing the results. The setup of docking calculations involves several critical

steps to ensure accurate and efficient simulation of ligand-receptor interactions. In general, each software has a proper set of parameters. One step involves defining the search space within the protein, usually by setting up a grid around the binding site. This can be based on known ligand binding data or predicted using software tools that identify potential binding pockets. The binding site can be defined using coordinates or by selecting specific residues. Parameters such as grid size and spacing are crucial because they influence the accuracy and speed of the docking process. Some programs allow you to adjust the scaling of van der Waals radii and specify the ligand size to be docked and to define constraints to bias the docking if the docked poses do not match experimentally validated poses.

Running the docking

Running molecular docking involves various computational strategies depending on the resources available and the scale of the docking project. For small-scale docking tasks, molecular docking can be performed on a single CPU. This involves installing docking software, preparing the input files, and running the docking process using command-line or graphical interface. This method is straightforward but can be time-consuming if the ligand library is large. Almost all docking software can utilize multiple CPU cores to parallelize the docking process, significantly reducing the computation time. This is achieved by setting the number of CPUs in the docking command, which allows the software to run multiple docking calculations simultaneously. For larger datasets, high-performance computing (HPC) clusters can be used. These systems involve multiple CPUs working in parallel, managed by job scheduling systems like PBS or SLURM. Docking tasks are distributed among the CPUs, greatly enhancing the throughput of virtual screening campaigns. Some docking tools have been optimized to run on Graphics Processing Units (GPUs), which can perform calculations faster than CPUs for certain types of computational tasks.

Evaluation of results

Evaluating the results of molecular docking is a critical step in the drug discovery process, as it helps determine the potential efficacy and specificity of ligands based on their predicted interactions with the target protein.

Analysis of docking poses *Pose Selection*: Review the top-ranked poses based on the docking scores provided by the docking software. These scores typically reflect the predicted binding affinity of the ligand to the receptor. They are mathematical models used to predict the binding affinity of a ligand to a protein. They assess the quality of the ligand poses generated during the docking process by calculating a score that typically reflects the free energy of binding. These functions can be categorized into several types, including force-field-based, empirical, and knowledge-based scoring functions. Each type uses different parameters and methodologies to estimate the interaction energies. During the docking process, multiple poses (conformations) of a ligand are generated [49,50].

Scoring functions evaluate each pose based on its potential energy, interaction with the protein, and other physicochemical properties. The poses are then ranked based on their scores, with lower scores generally indicating more favorable interactions (higher binding affinity). The pose with the best (lowest) score is typically considered the most likely binding conformation. However, selecting the top-scoring pose as the best solution is not always reliable. It is essential to consider the distribution of scores and the differences between the top poses. In some cases, additional criteria such as the presence of key interactions (e.g., hydrogen bonds with crucial residues) or the fit within the binding pocket are also considered to validate the selected pose. Scoring functions are not perfect and often face challenges such as accurately capturing solvent effects, protein flexibility, and entropic changes. Due to these limitations, the top-scoring pose might not always represent the experimentally observed binding mode. Therefore, it is recommended to analyze multiple top-ranking poses and consider additional validation methods such as molecular dynamics simulations or experimental data. To improve the reliability of pose selection, consensus scoring can be used. This approach combines scores from multiple scoring functions to balance out their individual biases and errors, potentially leading to a more accurate selection of the best pose [50–53].

Recent developments include the integration of machine learning techniques with traditional scoring functions to enhance their predictive accuracy. These advanced scoring functions are trained on large datasets of known protein-ligand interactions to better generalize across different systems. While scoring functions provide a quantitative measure to evaluate and rank ligand poses, their limitations necessitate careful analysis and often the use of additional validation strategies to ensure the selection of the most biologically relevant pose [3,54].

Visual Inspection: Visual inspection of docking poses complements the quantitative assessments provided by scoring functions. This qualitative analysis involves manually examining the docked conformations of a ligand within the binding site of a protein to ensure that the interactions are plausible and consistent with known biochemical and pharmacological data. While scoring functions provide a numerical estimate of binding affinity, they can sometimes be misleading due to their inherent limitations, such as the inability to accurately model solvent effects or protein flexibility. Visual inspection helps to confirm that the top-scoring poses are chemically and physically reasonable [55–57].

By visually inspecting the docking poses, researchers can identify critical interactions such as hydrogen bonds, ionic interactions, and hydrophobic contacts that are known to be important for binding but might not be adequately captured by the scoring function. It can also reveal artifacts or unrealistic binding modes that might occur due to algorithmic limitations or errors in the input data (e.g., incorrect protonation states or unusual torsional angles). Molecular visualization tools such as PyMOL [58], UCSF Chimera [59], or VMD [60] are commonly used for inspecting docking poses. These tools provide 3D visualizations of the protein-ligand complex, allowing for detailed examination of the interaction interface.

2D interaction diagrams are also a valuable tool as they provide a simplified, yet informative, visualization of the interactions between a ligand and its protein target. These diagrams help quickly understand the key interactions that contribute to binding, such as hydrogen bonds, hydrophobic contacts, salt bridges, $\pi - \pi$, and ionic interactions. This type of diagrams simplify the visualization of complex 3D molecular interactions, making it easier to understand and analyze the interactions between ligands and proteins. This simplification is crucial for quickly assessing the nature and strength of these interactions without the need to manipulate 3D structures. They allow for the comparative visualization of multiple ligand-protein complexes, particularly useful in drug discovery projects where multiple ligands are being assessed for their interaction with the same protein. 2D interaction diagrams are available from tools like PLIP [61], LigPlot+ [62], LeView [63], PoseEdit [64] among many others.

Validation of results

Validating docking results is essential to ensure that the predicted interactions between the ligand and the protein target are accurate and reliable. Validation can be approached through both experimental and computational methods.

From the experimental point of view, the most direct way to validate a docking pose prediction is by determining the experimental protein structure in complex with the ligand, typically through X-ray crystallography or NMR spectroscopy. Also, assays such as surface plasmon resonance (SPR), isothermal titration calorimetry (ITC), and enzyme inhibition studies can provide experimental evidence for the binding affinity and kinetics of the ligand-protein interaction, which can be compared to docking predictions. Another experimental technique is site-directed mutagenesis that can be used to alter specific amino acids within the binding site. The effects of these mutations on ligand binding can validate the importance of predicted interactions [4, 65–67].

From the computational point of view, several methods can be useful validating the docking results. *Re-docking and Cross-docking*: Re-docking involves docking the co-crystallized ligand back into the binding site of the protein to see if the docking method can reproduce the experimentally observed binding mode. Cross-docking tests the docking protocol on a set of proteins with known ligands to assess its robustness. *Scoring Function Evaluation*: Comparing different scoring functions and evaluating their performance on benchmark datasets or against experimental data can help validate the docking results. The use of consensus scoring, where multiple scoring functions are applied, can also improve validation. *Ensemble Docking*: Docking against multiple conformations of the protein, generated through techniques like molecular dynamics simulations, can account for receptor flexibility and provide a more comprehensive validation of the docking results. *Decoy Sets*: Docking a set of inactive compounds (decoys) seeded with known active compounds can help assess the ability of the docking protocol to distinguish between binders and non-binders, providing a measure of its predictive accuracy. *Molecular Dynamics Simulations*: Post-docking molecular dynamics simulations can be used to assess the stability of the ligand-receptor complex over time, providing further validation of the docking results. *Area Under Curve (AUC) and Enrichment Factor (E.F)*: The AUC value from receiver operating characteristic (ROC) curves and the enrichment factor at certain percentages can be used to validate the docking tool's ability to rank active compounds higher than decoys. *Ligand Efficiency Metrics*: Metrics such as ligand efficiency (LE) and binding efficiency index (BEI) can be used to evaluate the quality of the docking poses in terms of binding affinity relative to the size of the ligand [4, 65–72].

OTHER METHODS TO STUDY THE LIGAND-PROTEIN INTERACTION

MOZYME

MOZYME is a semiempirical quantum chemistry method that is particularly useful for studying large molecular systems, including protein-ligand interactions. it is based on the Localized Molecular Orbital (LMO) method, which allows for the efficient modeling of large systems by focusing on local interactions within the system. This method is particularly advantageous for studying proteins and sections of DNA, where the size of the system can make traditional quantum mechanical methods computationally prohibitive. The MOZYME method has been applied to study the interactions between ligands and proteins by enabling the semiempirical modeling of these large complexes. It uses a divide-and-conquer approach, dividing the system into smaller subsystems, which significantly speeds up the computational process. It can be integrated with other computational methods to enhance its capabilities. For example, it has been combined with Density Functional Theory (DFT) to study the pKa values of ionizable residues in proteins, providing a more comprehensive understanding of protein-ligand interactions. One of the key advantages of MOZYME is its linear-scaling SCF method, which allows systems of up to 15,000 atoms to be modeled. This makes it particularly suitable for studying protein-ligand interactions, where the size and complexity of the system can otherwise limit the applicability of quantum mechanical methods. The MOZYME method is implement in the MOPAC software [73–77].

Fragment Molecular Orbital

The Fragment Molecular Orbital (FMO) method is a computational approach used to study the interactions between a ligand and a protein by breaking down large molecular systems into smaller, more manageable fragments. This method allows for the application of quantum mechanical calculations to systems that would otherwise be too large and computationally demanding. The FMO method was developed to compute very large molecular systems by dividing them into fragments and performing ab initio or density functional theory (DFT) calculations on these smaller units. his approach enables the study of complex biomolecular interactions, including protein-ligand binding, by providing detailed insights into the chemical nature of each residue and water molecule's contribution to ligand binding. FMO offers faster computational speeds than traditional quantum-mechanical (QM) methods and provides accurate information for investigating the chemical nature and binding characteristics of protein-ligand interactions. It is particularly useful for revealing atomistic details about individual contributions toward ligand binding, which are difficult to detect without using QM methods. The FMO method can be applied to structure-based drug design (SBDD) processes, as it allows for the rationalization of affinity and selectivity in ligand binding. By analyzing crystal structures of receptors with their corresponding agonists and antagonists, FMO can provide valuable insights for computer-aided drug design (CADD). FMO-based Pair Interaction Energy Decomposition Analysis (PIEDA) provides a breakdown of interaction energies into components such as electrostatic, exchange, charge transfer, and dispersion contributions. This detailed analysis helps identify key residues and interactions critical for binding. The FMO method is implemented in software packages like GAMESS (US), ABINIT-MP, PAICS, and OpenFMO, which are distributed free of charge [78–83].

FINAL REMARKS

In conclusion, molecular docking has emerged as a pivotal method in the realm of drug discovery, significantly enhancing the efficiency and effectiveness of the process. By predicting the binding orientation of small molecule ligands to their protein targets, it provides valuable insights into the molecular interactions and mechanisms underpinning drug-receptor recognition. This computational approach not only reduces the time and cost associated with traditional experimental methods but also facilitates the exploration of a vast chemical space, enabling the identification of potential drug candidates with improved specificity and efficacy. However, it's important to note that molecular docking is not without its challenges, including the need for accurate protein and ligand structures, and the complexity of scoring functions. Despite

these, the continuous advancements in computational power, algorithms, and structural biology promise a promising future for molecular docking in drug discovery.

LIST OF ABBREVIATIONS

AI = Artificial Intelligence
AUC = Area Under Curve
BEI = Binding Efficiency Index
CADD = Computer Aided Drug Design
CNN = Convolutional Neural Network
CPU = Central Processor Unit
DFT = Density Functional Theory
DNA = Deoxyribo Nucleic acid
EE = Enrichment Factor
FBDD = Fragment Based Drug Discovery
FMO = Fragment Molecular Orbital
GPU = Graphics Processing Unit
HPC = High Performance Computing
HTD = High Throughput Docking
HTS = High Throughput Screening
ITC = Isothermal Titration Calorimetry
LBVS = Ligand-Based Virtual Screening
LE = Ligand Efficiency
LMO = Localized Molecular Orbital
MD = Molecular Dynamics
ML = Machine Learning
MM = Molecular Mechanics
NMR = Nuclear Magnetic Resonance
PBS = Portable Batch System
PDB = Protein Data Bank
PIEDA = Pair Interaction Energy Decomposition Analysis
QM = Quantum Mechanics
QM/MM = Quantum Mechanics/Molecular Mechanics
RNN = Recurrent Neural Network
RL = Reinforcement Learning
ROC = Receiver Operating Characteristic
SCF = Self Consistent Field
SLURM = Sophisticated Linux Utility for Resource Management
SBDD = Structure Based Drug Design
SBVS = Structure Based Virtual Screening
SPR = Surface Plasmon Resonance
VS = Virtual Screening

CONSENT FOR PUBLICATION

Not applicable.

FUNDING

The authors declare that no funds, grants, or other support was received during the preparation of this manuscript.

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT AUTHORSHIP CONTRIBUTION STATEMENT

C. Aguiar: Investigation, Formal analysis, Writing-original draft, Writing-review & editing. **I. Camps**: Conceptualization, Methodology, Software, Formal analysis, Resources, Writing-review & editing, Supervision, Project administration.

ACKNOWLEDGEMENTS

We would like to acknowledge financial support from the Brazilian agencies CNPq, CAPES and FAPEMIG.

Literature cited

[1] L. Ferreira, R. dos Santos, G. Oliva, A. Andricopulo, Molecular docking and structure-based drug design strategies, Molecules 20 (2015) 13384–13421 (2015). doi:10.3390/molecules200713384.

[2] L. Pinzi, G. Rastelli, Molecular docking: shifting paradigms in drug discovery, Int. J. Mol. Sci. 20 (2019) 4331 (2019). doi:10.3390/ijms20184331.

[3] A. T. McNutt, P. Francoeur, R. Aggarwal, T. Masuda, R. Meli, M. Ragoza, J. Sunseri, D. R. Koes, GNINA 1.0: molecular docking with deep learning, J. Cheminf. 13 (2021) 43 (2021). doi:10.1186/s13321-021-00522-2.

[4] B. J. Bender, S. Gahbauer, A. Luttens, J. Lyu, C. M. Webb, R. M. Stein, E. A. Fink, T. E. Balius, J. Carlsson, J. J. Irwin, B. K. Shoichet, A practical guide to large-scale docking, Nat. Protoc. 16 (2021) 4799–4832 (2021). doi:10.1038/s41596-021-00597-z.

[5] D. Sahu, L. S. Rathor, S. D. Dwivedi, K. Shah, N. S. Chauhan, M. R. Singh, D. Singh, A review on molecular docking as an interpretative tool for molecular targets in disease management, ASSAY and Drug Development Technologies 22 (2024) 40–50 (2024). doi:10.1089/adt.2023.060.

[6] I. A. Guedes, F. S. S. Pereira, L. E. Dardenne, Empirical scoring functions for structure-based virtual screening: Applications, critical aspects, and challenges, Front. Pharmacol. 9 (2018) 1089 (2018). doi:10.3389/fphar.2018.01089.

[7] M. L. Verdonk, R. D. Taylor, G. Chessari, C. W. Murray, Structure-Based Drug Discovery, Springer Netherlands, 2007, Ch. Illustration of current challenges in molecular docking, pp. 201–221 (2007). doi:10.1007/1-4020-4407-0_8.

[8] S. Saikia, M. Bordoloi, Molecular docking: challenges, advances and its use in drug discovery perspective, Curr. Drug Targets 20 (2019) 501–521 (2019). doi:10.2174/1389450119666181022153016.

[9] I. A. Guedes, A. M. S. Barreto, D. Marinho, E. Krempser, M. A. Kuenemann, O. Sperandio, L. E. Dardenne, M. A. Miteva, New machine learning and physics-based scoring functions for drug discovery, Sci. Rep. 11 (2021) 3198 (2021). doi:10.1038/s41598-021-82410-1.

[10] P. H. M. Torres, A. C. R. Sodero, P. Jofily, F. P. Silva-Jr, Key topics in molecular docking for drug design, Int. J. Mol. Sci. 20 (2019) 4574 (2019). doi:10.3390/ijms20184574.

[11] B. Zhang, H. Li, K. Yu, Z. Jin, Molecular docking-based computational platform for high-throughput virtual screening, CCF Trans High Perform Comput. 4 (2022) 63–74 (2022). doi:10.1007/s42514-021-00086-5.

[12] S.-Y. Huang, Comprehensive assessment of flexible-ligand docking algorithms: current effectiveness and challenges, Brief. Bioinform. 19 (2017) 982–994 (2017). doi:10.1093/bib/bbx030.

[13] A. M. Bonvin, Flexible protein-protein docking, Curr. Opin. Struct. Biol. 16 (2006) 194–200 (2006). doi:10.1016/j.sbi.2006.02.002.

[14] K. Schleinkofer, T. Wang, R. C. Wade, Encyclopedic reference of genomics and proteomics in molecular medicine, Springer Berlin Heidelberg, 2006, Ch. Molecular docking, pp. 1149–1153 (2006). doi:https://doi.org/10.1007/3-540-29623-9_3820.

[15] N. Andrusier, E. Mashiach, R. Nussinov, H. J. Wolfson, Principles of flexible protein-protein docking, Proteins 73 (2008) 271–289 (2008). doi:10.1002/prot.22170.

[16] C. N. Cavasotto, M. G. Aucar, High-throughput docking using quantum mechanical scoring, Front. Chem. 8 (2020) 246 (2020). doi:https://doi.org/10.3389/fchem.2020.00246.

[17] R. E. Amaro, J. Baudry, J. Chodera, O. Demir, J. A. McCammon, Y. Miao, J. C. Smith, Ensemble docking in drug discovery, Biophys. J. 114 (2018) 2271–2278 (2018). doi:10.1016/j.bpj.2018.02.038.

[18] K. McKay, N. B. Hamilton, J. M. Remington, S. T. Schneebeli, J. Li, Essential dynamics ensemble docking for structure-based GPCR drug discovery, Front. Mol. Biosci. 9 (2022) 879212 (2022). doi:10.3389/fmolb.2022.879212.

[19] S.-Y. Huang, X. Zou, Ensemble docking of multiple protein structures: Considering protein structural variations in molecular docking, Proteins 66 (2006) 399–421 (2006). doi:10.1002/prot.21214.

[20] W. E. Falcon, S. R. Ellingson, J. C. Smith, J. Baudry, Ensemble docking in drug discovery: How many protein configurations from molecular dynamics simulations are needed to reproduce known ligand binding?, J. Phys. Chem. B 123 (2019) 5189–5195 (2019). doi:10.1021/acs.jpcb.8b11491.

[21] K. Kapoor, S. Thangapandian, E. Tajkhorshid, Extended-ensemble docking to probe dynamic variation of ligand binding sites during large-scale structural changes of proteins, Chem. Sci. 13 (2022) 4150–4169 (2022). doi:10.1039/d2sc00841f.

[22] A. V. Onufriev, D. A. Case, Generalized Born implicit solvent models for biomolecules, Annu. Rev. Biophys. 48 (2019) 275–296 (2019). doi:10.1146/annurev-biophys-052118-115325.

[23] K. W. Wang, J. Lee, H. Zhang, D. Suh, W. Im, CHARMM-GUI implicit solvent modeler for various generalized born models in different simulation programs, J. Phys. Chem. B 126 (2022) 7354–7364 (2022). doi:10.1021/acs.jpcb.2c05294.

[24] B. Z. Zsidó, B. Bayarsaikhan, R. Börzsei, V. Szél, V. Mohos, C. Hetényi, The advances and limitations of the determination and applications of water structure in molecular engineering, Int. J. Mol. Sci. 24 (2023) 11784 (2023). doi:10.3390/ijms241411784.

[25] F. Gentile, J. C. Yaacoub, J. Gleave, M. Fernandez, A.-T. Ton, F. Ban, A. Stern, A. Cherkasov, Artificial intelligence-enabled virtual screening of ultra-large chemical libraries with deep docking, Nat. Protoc. 17 (2022) 672–697 (2022). doi:10.1038/s41596-021-00659-2.

[26] M. Marcisz, S. A. Samsonov, Solvent model benchmark for molecular dynamics of glycosaminoglycans, J. Chem. Inf. Model. 63 (2023) 2147–2157 (2023). doi:10.1021/acs.jcim.2c01472.

[27] A. Clyde, X. Liu, T. Brettin, H. Yoo, A. Partin, Y. Babuji, B. Blaiszik, J. Mohd-Yusof, A. Merzky, M. Turilli, S. Jha, A. Ramanathan, R. Stevens, AI-accelerated protein-ligand docking for SARS-CoV-2 is 100-fold faster with no significant change in detection, Sci. Rep. 13 (2023) 2105 (2023). doi:10.1038/s41598-023-28785-9.

[28] T. Sivula, L. Yetukuri, T. Kalliokoski, H. Käsnänen, A. Poso, I. Pöhner, Machine learning-boosted docking enables the efficient structure-based virtual screening of giga-scale enumerated chemical libraries, J. Chem. Inf. Model. 63 (2023) 5773–5783 (2023). doi:10.1021/acs.jcim.3c01239.

[29] Z. Chen, H.-l. Li, Q.-j. Zhang, X.-g. Bao, K.-q. Yu, X.-m. Luo, W.-l. Zhu, H.-l. Jiang, Pharmacophore-based virtual screening versus docking-based virtual screening: a benchmark comparison against eight targets, Acta Pharmacol. Sin. 30 (2009) 1694–1708 (2009). doi:10.1038/aps.2009.159.

[30] B. Hu, M. A. Lill, PharmDock: a pharmacophore-based docking program, Journal of Cheminformatics 6 (2014) 14 (2014). doi:10.1186/1758-2946-6-14.

[31] X.-Y. Meng, H.-X. Zhang, M. Mezei, M. Cui, Molecular docking: A powerful approach for structure-based drug discovery, Curr Comput Aided Drug Des. 7 (2011) 146–157 (2011). doi:10.2174/157340911795677602.

[32] G. Khodarahmi, H. Farrokhpour, P. Asadi, F. Hassanzadeh, L. Saghaei, Quantum mechanical/molecular mechanical and docking study of the novel analogues based on hybridization of common pharmacophores as potential anti-breast cancer agents, Res Pharm Sci. 12 (2017) 233 (2017). doi:10.4103/1735-5362.207204.

[33] D. Giordano, C. Biancaniello, M. A. Argenio, A. Facchiano, Drug design by pharmacophore and virtual screening approach, Pharmaceuticals 15 (2022) 646 (2022). doi:10.3390/ph15050646.

[34] M. Mohanty, P. S. Mohanty, Molecular docking in organic, inorganic, and hybrid systems: a tutorial review, Monatsh Chem. 154 (2023) 683–707 (2023). doi:10.1007/s00706-023-03076-1.

[35] C. Uma Maheswari, S. Muthu, T. Sundius, Combined quantum-mechanics/molecular-mechanics, molecular docking studies on hemorrhoid drug, Results in Chemistry 5 (2023) 100962 (2023). doi:10.1016/j.rechem.2023.100962.

[36] M. T. Muhammed, E. Aki-Yalcin, Molecular docking: principles, advances, and its applications in drug discovery, Lett. Drug Des. Discovery 21 (2024) 480–495 (2024). doi:10.2174/1570180819666220922103109.

[37] Y. Chen, B. K. Shoichet, Molecular docking and ligand specificity in fragment-based inhibitor discovery, Nat. Chem. Biol. 5 (2009) 358–364 (2009). doi:10.1038/nchembio.155.

[38] T. Wang, M.-B. Wu, Z.-J. Chen, H. Chen, J.-P. Lin, L.-R. Yang, Fragment-based drug discovery and molecular docking in drug design, Curr. Pharm. Biotechnol. 16 (2015) 11–25 (2015). doi:10.2174/1389201015666141122204532.

[39] C. Jacquemard, M. N. Drwal, J. Desaphy, E. Kellenberger, Binding mode information improves fragment docking, J. Cheminf. 11 (2019) 24 (2019). doi:10.1186/s13321-019-0346-7.

[40] Q. Li, Application of fragment-based drug discovery to versatile targets, Front. Mol. Biosci. 7 (2020) 180 (2020). doi:10.3389/fmolb.2020.00180.

[41] J.-h. Gan, J.-x. Liu, Y. Liu, S.-w. Chen, W.-t. Dai, Z.-X. Xiao, Y. Cao, DrugRep: an automatic virtual screening server for drug repurposing, Acta Pharmacol. Sin. 44 (2022) 888–896 (2022). doi:10.1038/s41401-022-00996-2.

[42] E. H. B. Maia, L. C. Assis, T. A. de Oliveira, A. M. da Silva, A. G. Taranto, Structure-based virtual screening: from classical to artificial intelligence, Front. Chem. 8 (2020) 343 (2020). doi:10.3389/fchem.2020.00343.

[43] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang, E. E. Bolton, PubChem 2023 update, Nucleic Acids Res. 51 (2022) D1373–D1380 (2022). doi:10.1093/nar/gkac956.

[44] J. J. Irwin, K. G. Tang, J. Young, C. Dandarchuluun, B. R. Wong, M. Khurelbaatar, Y. S. Moroz, J. Mayfield, R. A. Sayle, ZINC20-A free ultralarge-scale chemical database for ligand discovery, J. Chem. Inf. Model. 60 (2020) 6065–6073 (2020). doi:10.1021/acs.jcim.0c00675.

[45] P. C. Agu, C. A. Afiukwa, O. U. Orji, E. M. Ezeh, I. H. Ofoke, C. O. Ogbu, E. I. Ugwuja, P. M. Aja, Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management, Sci. Rep. 13 (2023) 13398 (2023). doi:10.1038/s41598-023-40160-2.

[46] I. Azad, Molecular Docking - Recent Advances, IntechOpen, 2023, Ch. Molecular docking in the study of ligand-protein recognition: An overview (2023). doi:10.5772/intechopen.106583.

[47] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, The Protein Data Bank, Nucleic Acids Res. 28 (2000) 235–242 (2000). doi:10.1093/nar/28.1.235.

[48] G. M. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments, J. Comput. Aided. Mol. Des. 27 (2013) 221–234 (2013). doi:10.1007/s10822-013-9644-8.

[49] D. B. Kitchen, H. Decornez, J. R. Furr, J. Bajorath, Docking and scoring in virtual screening for drug discovery: methods and applications, Nat. Rev. Drug Discovery 3 (2004) 935–949 (2004). doi:10.1038/nrd1549.

[50] P. Suriana, R. O. Dror, Machine learning in structural biology workshop, 2023, Ch. Enhancing ligand pose sampling for molecular docking (2023).

[51] D. Ramírez, J. Caballero, Is it reliable to take the molecular docking top scoring position as the best solution without considering available structural data?, Molecules 23 (2018) 1038 (2018). doi:10.3390/molecules23051038.

[52] S. Makeneni, D. F. Thieker, R. J. Woods, Applying pose clustering and MD simulations to eliminate false positives in molecular docking, J. Chem. Inf. Model. 58 (2018) 605–614 (2018). doi:10.1021/acs.jcim.7b00588.

[53] K. Palacio-Rodríguez, I. Lans, C. N. Cavasotto, P. Cossio, Exponential consensus ranking improves the outcome in docking and receptor ensemble docking, Sci. Rep. 9 (2019) 5142 (2019). doi:10.1038/s41598-019-41594-3.

[54] R. Meli, G. M. Morris, P. C. Biggin, Scoring functions for protein-ligand binding affinity prediction using structure-based deep learning: A review, Front. Bioinform. 2 (2022) 885983 (2022). doi:10.3389/fbinf.2022.885983.

[55] A. Fischer, M. Smieško, M. Sellner, M. A. Lill, Decision making in structure-based drug discovery: Visual inspection of docking results, J. Med. Chem. 64 (2021) 2489–2500 (2021). doi:10.1021/acs.jmedchem.0c02227.

[56] M. Xu, C. Shen, J. Yang, Q. Wang, N. Huang, Systematic investigation of docking failures in large-scale structure-based virtual screening, ACS Omega 7 (2022) 39417–39428 (2022). doi:10.1021/acsomega.2c05826.

[57] M. Schäfer, N. Brich, J. Byska, S. M. Marques, D. Bednář, P. Thiel, B. Kozlíková, M. Krone, InVADo: interactive visual analysis of molecular docking data, IEEE Trans. Vis. Comput. Graphics 30 (2024) 1984–1997 (2024). doi:10.1109/TVCG.2023.3337642.

[58] Schrödinger, LLC, The PyMOL molecular graphics system, version 1.8 (2015).

[59] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera-A visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612 (2004). doi:10.1002/jcc.20084.

[60] W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38 (1996). doi:10.1016/0263-7855(96)00018-5.

[61] M. F. Adasme, K. L. Linnemann, S. N. Bolz, F. Kaiser, S. Salentin, V. Haupt, M. Schroeder, PLIP 2021: expanding the scope of the protein-ligand interaction profiler to DNA and RNA, Nucleic Acids Res. 49 (2021) W530–W534 (2021). doi:10.1093/nar/gkab294.

[62] R. Laskowski, M. Swindells, LigPlot+: multiple ligand-protein interaction diagrams for drug discovery, J. Chem. Inf. Model. 51 (2011) 2778–2786 (2011). doi:10.1021/ci200227u.

[63] S. Caboche, LeView: automatic and interactive generation of 2D diagrams for biomacromolecule/ligand interactions, J. Cheminf. 5 (2013) 40 (2013). doi:10.1186/1758-2946-5-40.

[64] K. Diedrich, B. Krause, O. Berg, M. Rarey, PoseEdit: enhanced ligand binding mode communication by interactive 2D diagrams, J. Comput.-Aided Mol. Des. 37 (2023) 491–503 (2023). doi:10.1007/s10822-023-00522-4.

[65] R. Wahlström, Validation of docking performance in the context of a structural water molecule using model system, Master's thesis, University of Dundee (2009).

[66] C. O. Nwonuma, V. C. Nwatu, G. Mostafa-Hedeab, O. S. Adeyemi, O. O. Alejolowo, O. A. Ojo, S. A. Adah, O. J. Awakan, C. E. Okolie, N. T. Asogwa, I. A. Udofia, G. O. Egharevba, N. H. Aljarba, S. Alkahtani, G. E.-S. Batiha, Experimental validation and molecular docking to explore the active components of cannabis in testicular function and sperm quality modulations in rats, BMC Complement Med. Ther. 22 (2022) 227 (2022). doi:10.1186/s12906-022-03704-z.

[67] A. S. Adah, J. O. Ayo, D. A. Adah, C. O. Nwonuma, T. A. Lawal, Molecular docking and experimental validation of the effect of ergothioneine on heat shock protein-70 following endurance exercise by Arabian stallions, BMC Vet. Res. 19 (2023) 27 (2023). doi:10.1186/s12917-023-03584-6.

[68] K. E. Hevener, W. Zhao, D. M. Ball, K. Babaoglu, J. Qi, S. W. White, R. E. Lee, Validation of molecular docking programs for virtual screening against dihydropteroate synthase, J. Chem. Inf. Model. 49 (2009) 444–460 (2009). doi:10.1021/ci800293n.

[69] S. Sokalingam, G. Munussami, J.-R. Kim, S.-G. Lee, Validation on the molecular docking efficiency of lipocalin family of proteins, J. Ind. Eng. Chem. 67 (2018) 293–300 (2018). doi:10.1016/j.jiec.2018.06.041.

[70] P. Agrawal, H. Singh, H. K. Srivastava, S. Singh, G. Kishore, G. P. S. Raghava, Benchmarking of different molecular docking methods for protein-peptide docking, BMC Bioinf. 19 (2019) 426 (2019). doi:10.1186/s12859-018-2449-y.

[71] C. Shivanika, S. D. Kumar, R. Venkataraghavan, T. Pawan, A. Sumitha, P. B. Devi, Molecular docking, validation, dynamics simulations, and pharmacokinetic prediction of natural compounds against the SARS-CoV-2 main-protease, J. Biomol. Struct. Dyn. 40 (2020) 585–611 (2020). doi:10.1080/07391102.2020.1815584.

[72] E. M. Terefe, A. Ghosh, Molecular docking, validation, dynamics simulations, and pharmacokinetic prediction of phytochemicals isolated from croton dichogamus against the HIV-1 reverse transcriptase, Bioinf. Biol. Insights 16 (2022) 1 (2022). doi:10.1177/11779322221125605.

[73] J. J. P. Stewart, Application of localized molecular orbitals to the solution of semiempirical self-consistent field equations, Int. J. Quantum Chem. 58 (1996) 133–146 (1996). doi:10.1002/(SICI)1097-461X(1996)58:2<133::AID-QUA2>3.0.CO;2-Z.

[74] K. Ohno, N. Kamiya, N. Asakawa, Y. Inoue, M. Sakurai, Application of an integrated MOZYME+DFT method to pKa calculations for proteins, Chem. Phys. Lett. 341 (2001) 387–392 (2001). doi:10.1016/S0009-2614(01)00499-7.

[75] J. J. P. Stewart, Application of the PM6 method to modeling proteins, J. Mol. Model. 15 (2009) 765–805 (2009). doi:10.1007/s00894-008-0420-y.

[76] A. Sulimov, D. Kutov, I. Ilin, V. Sulimov, Quantum-chemical quasi-docking for molecular dynamics calculations, Nanomaterials 12 (2022) 274 (2022). doi:10.3390/nano12020274.

[77] J. E. Moussa, J. J. P. Stewart, Mopac (2024). doi:10.5281/zenodo.6511958.
URL https://github.com/openmopac/mopac

[78] M. Schmidt, K. Baldridge, J. Boatz, S. Elbert, M. Gordon, J. Jensen, S.Koseki, N.Matsunaga, K. Nguyen, S. Su, T. Windus, M. Dupuis, J. Montgomery, General atomic and molecular electronic structure system, J. Comput. Chem. 14 (1993) 1347–1363 (1993). doi:10.1002/jcc.540141112.

[79] Y. Alexeev, M. P. Mazanetz, O. Ichihara, D. G. Fedorov, GAMESS as a free quantum-mechanical platform for drug research, Curr. Top. Med. Chem. 12 (2012) 2013–2033 (2012). doi:10.2174/156802612804910269.

[80] E. I. Chudyk, L. Sarrat, M. Aldeghi, D. G. Fedorov, M. J. Bodkin, T. James, M. Southey, R. Robinson, I. Morao, A. Heifetz, Exploring GPCR-ligand interactions with the fragment molecular orbital (FMO) method, Methods Mol Biol. 1705 (2017) 179–195 (2017). doi:10.1007/978-1-4939-7465-8_8.

[81] J. Choi, H.-J. Kim, X. Jin, H. Lim, S. Kim, I.-S. Roh, H.-E. Kang, K. T. No, H.-J. Sohn, Application of the fragment molecular orbital method to discover novel natural products for prion disease, Sci. Rep. 8 (2018) 13063 (2018). doi:10.1038/s41598-018-31080-7.

[82] M. Gordon, M. Schmidt, Theory and applications of computational chemistry, the first forty years, Elsevier, Amsterdam, 2005 (2005).

[83] H. Lim, D. H. Kang, J. Kim, A. Pellow-Jarman, S. McFarthing, R. Pellow-Jarman, H.-N. Jeon, B. Oh, J.-K. K. Rhee, K. T. No, Fragment molecular orbital-based variational quantum eigensolver for quantum chemistry in the age of quantum computing, Sci. Rep. 14 (2024) 2422 (2024). doi:10.1038/s41598-024-52926-3.