THE SCIENCE WORLD
A Monthly e Magazine
ISSN:2583-2212
Review Article
August, 2026 Vol.6(8), 2422-2442

Role of Virtual Screening In Drug Discovery And Development: An Overview

Dr Ravina Dhameliya
M.V.Sc scholar
ravinadhameliya8296@gmail.com
doi.org/10.5281/ScienceWorld.22236710

Abstract

Drug discovery is a highly complex and costly process, which demands integrated efforts in several relevant aspects involving innovation, knowledge, information, technologies, expertise, R&D investments and management skills. The shift from traditional to genomics- and proteomics-based drug research has fundamentally transformed key R&D strategies in the pharmaceutical industry addressed to the design of new chemical entities as drug candidates against a variety of biological targets. Therefore, drug discovery has moved toward more rational strategies based on our increasing understanding of the fundamental principles of protein-ligand interactions. The combination of available knowledge of several 3D protein structures with hundreds of thousands of small-molecules have attracted the attention of scientists from all over the world for the application of structure- and ligand-based drug design approaches. In this context, virtual screening technologies have largely enhanced the impact of computational methods applied to chemistry and biology and the goal of applying such methods is to reduce large compound databases and to select a limited number of promising candidates for drug design. This review provides a perspective of the utility of virtual screening in drug design and its integration with other important drug discovery technologies such as high-throughput screening (HTS) and QSAR, highlighting the present challenges, limitations, and future perspectives in medicinal chemistry.

Key Words: Introduction, Virtual Screening, Methods of VS, Advantages-Disadvantages, Applications, Conclusions, Future prospects

INTRODUCTION

The identification of novel leads is the essential factor and initiation point for all

new drug discovery initiatives. Recent developments in molecular biology, robotics, and microelectronics, particularly the thorough investigation of the human genome, will profoundly influence modern drug discovery and its effects on human diseases. Human genomics aims to deliver the sequences of all genes that encode proteins which constitute the biology of a specific organism. Thus, the count of possible targets for medication treatment could rise considerably. It is projected that the quantity of possible drug targets could rise from roughly 500 currently to approximately 5000-10000 within the next few years (Debouck & Metcalf *et al*., 2000).

The early phases of drug discovery rely heavily on the identification of viable hit molecules and their progression into strong lead compounds. Progress in medicinal chemistry, especially where chemical and biological sciences converge, has provided a solid basis for the development of novel drug candidates with enhanced pharmacodynamic and pharmacokinetic characteristics. Nevertheless, in spite of these advances, drug discovery continues to become more expensive and time-consuming (Hajduk & Greer *et al*., 2007; Hopkins & Groom *et al*., 2002). To address these challenges, advanced in silico approaches are increasingly employed for discovering new drugs, offering a promising solution (Lavecchia & Giovanni, 2013).

The extensive application of combinatorial chemistry and high-throughput screening (HTS) in lead discovery has driven a growing need for small organic molecules that can interact with specific therapeutic targets. These approaches emphasize the creation of vast molecular libraries alongside the biological evaluation of large numbers of samples. However, increasing demands to shorten development timelines and lower costs have prompted a shift away from random compound screening toward more rational, targeted strategies. This transition aims to enhance the success rate of new chemical entity (NCE) discovery and boost overall productivity in pharmaceutical research and development. The characterization of chemical and biological space has renewed screening models and highlighted the value of integrating traditional and modern methodologies in drug discovery. Within this framework, computational techniques have become an essential component of contemporary medicinal chemistry programs (Andricopulo & Montanari *et al*., 2005; Overington *et al*., 2006). Contemporary drug discovery relies on innovation and scientific insight achieved by combining experimental techniques with computational strategies. One of the key challenges for the pharmaceutical sector is the discovery of novel new chemical entities (NCEs) from an immense space of existing and hypothetical compounds.

Official Website
www.thescienceworld.net
thescienceworldmagazine@gmail.com
Published 30/08/2026

Computational methods can systematically improve several stages of the discovery process, such as hit finding, lead refinement, and the assessment of pharmacokinetic behavior. In recent years, progress in high-performance computing, algorithms, methodological approaches, and specialized expertise has greatly strengthened structure-based drug design (SBDD), making it a highly impactful approach in modern drug development (Grüneberg *et al*., 2002; Tame, 1999).

In this review, we focus on the principles and applications of VS in the SBDD framework, starting from the initial stages of the process that include receptor and library preprocessing, to docking, scoring, and post-processing of top scoring hits. We also highlight several successful studies and protocols that led to nM leads, discuss novel applications of Structure-Based VS (SBVS) such as substrate identification for the discovery of novel metabolic pathways, and provide recent trends in library design. Limitations of SBVS are also examined. Finally, we present two developed VS protocols that aim to enhance inhibitor selectivity for the target protein structure.

VIRTUAL SCREENING

Virtual screening generally involves using computational methods to evaluate virtual compound libraries—acting as a type of in-silico laboratory against target receptors, with the aim of speeding up the drug discovery process (Shen *et al*., 2003). This approach is designed to computationally analyze large hypothetical chemical databases or virtual libraries, allowing researchers to identify a smaller subset of candidate molecules that are most likely to be active against a particular biological target (Waszkowycz *et al*., 2001).

Official Website
www.thescienceworld.net
thescienceworldmagazine@gmail.com
Published 30/08/2026

Virtual Screening (VS) was introduced in the late 1990s as a computational approach that uses algorithms and modeling techniques to select compounds from virtual libraries with the highest likelihood of interacting with specific biological targets. Since then, VS has developed into a powerful tool for identifying novel drug-like

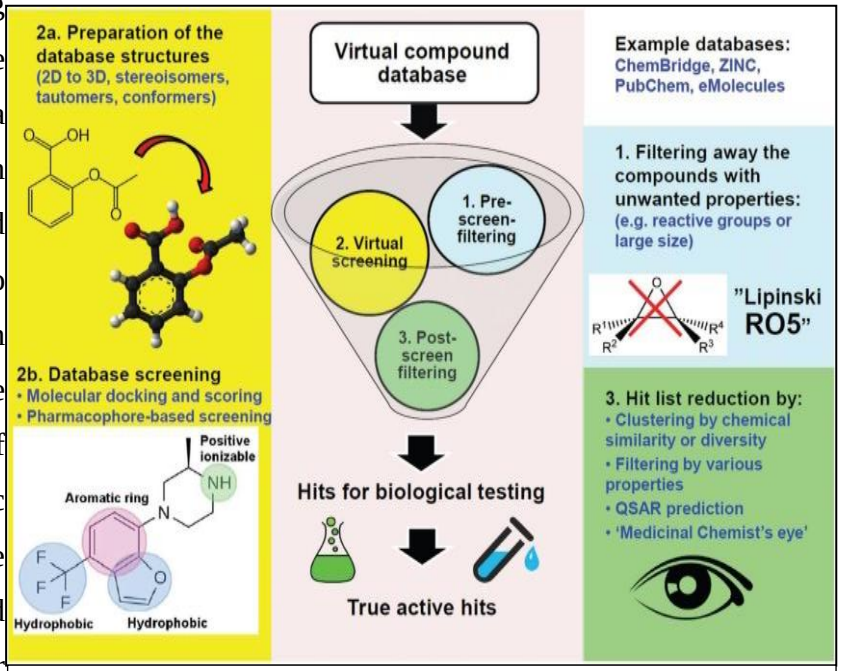

Figure 1: General workflow of virtual screening (Sahlgren *et al*., 2017)

2004). Its adoption has grown rapidly, making it a key component of modern drug research (Schneider, 2010). By facilitating the discovery of compounds with specific bioactivity, virtual screening contributes to lowering costs, reducing infrastructure demands, and accelerating the overall process of identifying new chemical entities (Macalino et al., 2015).

The central concept underlying all computational strategies in early drug discovery is to explore, in silico, large or moderately sized compound databases and identify a small set of candidates likely to exhibit the desired biological activity. The term “data mining” was introduced in 1996 to describe this process (Morrell J, 1996) and was succinctly defined by Gasteiger and colleagues as “the extraction of knowledge from a large dataset to predict new events” (Gasteiger *et al*., 2003).

DATABASES FOR VIRTUAL SCREENING

The efficiency of virtual screening depends on availability of quality and quantity of data and predictability of the underlying algorithms. The mining of data from sea of drug database and chemical libraries is a pre-requisite for successful screening. The currently available database and chemical libraries for this purpose includes ZINC (free database of 3.3 million commercially available compounds), Available Chemicals Directory (ACD; 4 million entries, not free), National Cancer Institute Compound Database (NCI; free database of 400,000 entries), and MDDR (MDL; Drug Data Report 4147,000 entries). Many online portals (Ambinter, Bioinfo- DB, ChemSpider, e-molecules, ZINC) provide access to chemical

Official Website
www.thescienceworld.net
thescienceworldmagazine@gmail.com
Published 30/08/2026

libraries. Theses portals are used in searching the structures and properties of molecules.

Table 1: Chemical databases available for virtual screening

Library	Description	No. of Compounds or Targets	Website
ZINC (ZINC15)	Free for commercially available compounds	980 million	http://zinc15.docking.org/
Pubchem	Freely available collection of chemical information resource at the U.S. (NCBI).	248 million	https://pubchem.ncbi.nlm.nih.gov/
PDB	A pool of 3D structures of proteins available	150,000 targets	http://www.rcsb.org/
LigandBox	3D structures of compounds available	4 million	http://ligandbox.protein.osaka-u.ac.jp/ligandbox/
Chem-Bridge	Commercial database of diverse and target-focused screening compounds	1.3 million	https://www.chembridge.com/

(da Silva *et al*., 2019)

The assessing pharmacokinetic properties and drug-likeness of chemical compounds is a crucial step in discovering molecules with potential biological activity (Mignani *et al*., 2018). Accordingly, evaluating a compound's drug-like characteristics is an essential part of virtual screening approaches. High-throughput screening (HTS) is commonly employed to identify hits or lead compounds, allowing large libraries of hundreds of thousands of molecules to be tested for biological activity. This modern method enables the daily screening of thousands of samples against a specific target-whether known or unknown-by providing a straightforward activity measurement for each compound (Schapira *et al*., 2003).

Over time, various cheminformatics or molecular filters have been created as valuable tools for screening compounds with favorable pharmacokinetic and pharmacodynamic properties, minimal toxicity, and reduced promiscuity or reactivity in

Official Website
www.thescienceworld.net
thescienceworldmagazine@gmail.com
Published 30/08/2026

inhibition assays. These filters help guide decision-making in the discovery of new chemical entities for applications in pharmaceuticals, cosmetics, agrochemicals, and biotechnology (Huggins *et al*., 2011). The most frequently applied filters aim to eliminate compounds with poor cell membrane permeability or limited distribution. Notable examples of widely used cheminformatics filters include those developed by Lipinski (Lipinski *et al*., 1997) and Veber (Veber *et al*., 2002).

Some filters function by choosing compounds according to defined ranges of physicochemical and structural properties that correspond to specific pharmacokinetic features (e.g., gastrointestinal absorption or blood-brain barrier permeability) and pharmacodynamic attributes (e.g., target specificity or promiscuity). These properties are determined using statistical thresholds, such as the 90th percentile for each molecular descriptor, to highlight the important characteristics of the compounds being evaluated (Daina & Zoete, 2016).

Table 2: Structural and physicochemical properties present in some molecular filters applied in virtual screening.

Molecular filter	MW (Da)	PSA (A2)	HBA	HBD	clogP	RTB	NAR
Lipinski's rule (RO5)	500	-	0-10	0-5	<5	-	-
Ghose's rule	160-480	-	-	-	- 0.4 to +5.6	-	20-70
Oprea's drug-like rule	-	-	2-9	0-2	2-8	-	-
Walters	200–500	<120	0-10	0–5	-	0-8	-
Veber's rule	-	<140	-	-	-	0-10	-
Beyond rule of five (bRO5	<1000	<250	<15	<6	−2- to 10	<20	-
Congreve's rule (RO3)	<300	<6	<3	<3	-	-	-

Official Website
www.thescienceworld.net
thescienceworldmagazine@gmail.com
Published 30/08/2026

(MW, molecular weight; PSA, polar surface area; HBA, Hydrogen bond acceptor; HBD, Hydrogen bond donor; RTB, Rotatable bond; NAR, Number of aromatic rings; clogP, Lipophilicity of Compound; clogD, Distribution coefficient)

METHODS OF VIRTUAL SCREENING (Santana *et al.*, 2021)

Two main computational strategies are used in virtual screening of molecules: (1) ligand-based virtual screening (LBVS) and (2) structure-based virtual screening (SBVS). Often, these two approaches are combined to achieve a more effective and comprehensive screening process and identify novel bioactive compounds against a specific molecular target or a biological system (Wang *et al.*, 2020). Currently, virtual screening methods are an integral part of the design and discovery of new bioactive compounds, and their applications have become popular in the academia and industry (Kar & Roy, 2013).

The LBVS approach include prefiltering methods such as structural and three- dimensional (2D and 3D) shape similarity searches modeling (Khan *et al.*, 2019; Xu *et al.*, 2019; Oum *et al.*, 2020), as well as ligand-based pharmacophore modeling (Lu *et al.*, 2016; Kaur & Silakari, 2017; Liang *et al.*, 2020). Screening is conducted based on the inherent features of the compound, including its electronic, topological, physicochemical, and structural properties, which are linked to its molecular activity (Berenger *et al.*, 2017; Garcia-Hernandez *et al.*, 2019).

SBVS is the preferred method when the 3D structure of the molecular target has been experimentally characterized and tries to predict the likelihood of coupling between candidate ligands and target protein, considering the binding strength of the complex. The most used SBVS technique is molecular anchoring due to the low computational cost and the satisfactory results (Meng *et al.*, 2011).

These two approaches may be used separately or together, depending on the information available about the molecules or targets. If only a single active molecule is known, similarity searching can be employed. Machine learning techniques are applied in virtual screening when a common pharmacophore cannot be identified and there are enough active and inactive compounds; they help derive structure–activity relationships from known actives to predict biological activity. Docking studies are utilized when the 3D structure of the target is known, as they predict how individual molecules bind and determine the orientation that best aligns with the observed structure (Guedes *et al.*, 2014; Yan *et al.*, 2016).

Official Website
www.thescienceworld.net
thescienceworldmagazine@gmail.com
Published 30/08/2026

3d-Shape Similarity Search Method

In similarity-based approaches, algorithms assume that a molecule with a similar structure to another whose biological activity is known may share similar activity (Willett, 2006). Similarity-based methods typically require only a small amount of reference information to develop predictive models and their simplicity and effectiveness have led to widespread adoption within the scientific community. Moreover, these algorithms can be integrated with other approaches, such as machine learning, meta-heuristics, and more, to enhance their performance (Hönig *et al*., 2023). Shape-similarity methods can screen vast compound libraries against a reference ligand with known bioactivity (Koes & Camacho, 2014).

Different methods have been used in the representation of the 3D molecular shape of the ligands, such as Gaussian overlay-based methods (Cai et al., 2013), atomic distance-based methods (Bonanno & Ebejer, 2020), and surface-based methods (Cleves *et al*., 2019).

Molecular shapes that are identified are transformed into 3D molecular fingerprints, which are subsequently compared using similarity or distance measures, including Tanimoto, Dice, and Tversky coefficients (Shin *et al*., 2015). Despite the structural complexity of natural products making target identification challenging-even with computational tools-3D shape- based similarity search methods have emerged as a powerful approach for predicting the macromolecular targets of these compounds (Chen *et al*., 2020).

Web-based platforms that utilize shape-similarity search techniques include SHAFTS (Liu *et al*., 2011) and USR-VS (Li *et al*., 2016), while installable open-source tools include Shapeit (Grant *et al*., 1996), gWEGA (Yan *et al*., 2014), and OptiPharm (Puertas-Martín *et al*., 2019). The principle behind 3D shape similarity is that molecules with similar shapes are likely to fit into the same binding site, potentially displaying comparable biological activity. These 3D shape-similarity search methods are used to identify compounds in chemical libraries that share similar molecular shapes, even if their structures differ (Santana *et al*., 2021).

Machine Learning

Machine learning (ML) involves the use of intelligent algorithms to analyze data, learn from it, and make decisions to solve specific problems. Advances in artificial intelligence have significantly accelerated the research and development of novel bioactive natural compounds with potential industrial applications. ML has been widely employed at

Official Website
www.thescienceworld.net
thescienceworldmagazine@gmail.com
Published 30/08/2026

various stages of virtual screening, such as predicting pharmacokinetic properties (Wei *et al*., 2017), evaluating compound penetration across the blood-brain barrier (Dai *et al*., 2021) and cell membranes (Wolfe *et al*., 2018), assessing side effects and toxicity (Pu *et al*., 2019), and identifying pan assay interference compounds (PAINS)—highly reactive and promiscuous molecules that often yield false positives in high-throughput screening assays (Jasial *et al*., 2018).

In certain instances, ML algorithms have demonstrated higher efficiency and are considered more effective than traditional QSAR methods for predicting hit compounds from chemical libraries (Tsou *et al*., 2020). Machine learning (ML) involves computer programs designed to mimic human-like cognitive abilities, such as learning, problem-solving, and pattern recognition. These algorithms are trained on large datasets, which serve as benchmarks for solving specific computational problems. In the context of virtual screening, the primary goal of an ML framework is to generalize knowledge from the training dataset to accurately evaluate a test dataset and support decision-making (Vamathevan *et al*., 2019).

In the LBVS approach, ML algorithms are used to predict the bioactivity or pharmacodynamic and pharmacokinetic properties of compounds by analyzing their similarity to known active molecules. These algorithms are often trained using chemo-structural and bioactivity data available in public databases, along with experimental results (Martínez- Treviño *et al*., 2020).

Molecular descriptors for evaluating similarity between compounds include physicochemical properties (e.g., cLogP, topological polar surface area (tPSA), molecular weight), structural characteristics (such as rotatable bonds and aromatic rings) (Lo *et al*., 2018), as well as functional groups, molecular shape, and pharmacophores (Bonanno & Ebejer, 2020). The selection of molecular representation and descriptor type plays a crucial role in determining both the accuracy and interpretability of results generated by ML algorithms (Jiménez-Luna *et al*., 2020). Open-source platforms like scikit-learn (Pedregosa *et al*., 2011) and SciPy (Virtanen *et al*., 2020) have been widely used to build machine learning models.

ML algorithms are generally classified into supervised and unsupervised learning. Supervised methods require retrospective validation with datasets containing both active and inactive compounds to identify the most appropriate algorithms for differentiating bioactive molecules (Sieg *et al*., 2019).

Official Website
www.thescienceworld.net
thescienceworldmagazine@gmail.com
Published 30/08/2026

Supervised learning methods are generally divided into two categories: (1) regression analysis and (2) classification techniques. Regression analysis includes methods such as decision trees, artificial neural networks, support vector machines, and random forests. In contrast, unsupervised algorithms detect patterns within datasets without requiring inactive compounds, arranging the data in an organized manner. These methods are commonly used for exploratory analyses, such as clustering (Patel *et al*., 2020). Examples of unsupervised approaches include clustering techniques like hidden Markov models, hierarchical clustering, and k-means (Santana *et al*., 2021).

Pharmacophore Modelling

A pharmacophore model defines a set of chemical features arranged in a specific three- dimensional (3D) configuration that are critical for biological activity against a specific molecular target. Important features within a pharmacophore include hydrogen bond donors and acceptors, hydrophobic regions, positively or negatively charged groups, and metal ion coordination (Schaller *et al*., 2020).

Ligand binding sites are constrained by both physicochemical and spatial factors, such as amino acid composition, cavity shape, and volume, which restrict non-specific interactions. These constraints dictate the ligand's binding orientation, allowing structurally different molecules to interact with the same target if they share the same pharmacophore characteristics (Vuorinen & Schuster, 2015).

A variety of software tools are currently available to generate pharmacophore models, varying in how they handle ligand conformational flexibility and structural alignment. Commercial programs for pharmacophore modeling include LigandScout (Wolber & Langer, 2005) and Molecular Operating Environment (MOE) (Molecular Operating Environment, 2019), both supporting ligand- and structure-based approaches and compatible with major operating systems.

Several open-source options for ligand-based pharmacophore modeling exist, such as Pharmer (https://sourceforge.net/projects/pharmer/) (Koes & Camacho, 2011) and Align-it (formerly Pharao; compatible with OS X) (Taminau et al., 2008). In addition, free web servers for structure-based pharmacophore screening have been developed, including Pharmit (http://pharmit.csb.pitt.edu/) (Sunseri & Koes, 2016) and PharmMapper (http://www.lilab- ecust.cn/pharmmapper/) (Liu et al., 2010).

Ligand-based pharmacophore modelling

Ligand-based pharmacophore virtual screening consists of several essential steps:

Official Website
www.thescienceworld.net
thescienceworldmagazine@gmail.com
Published 30/08/2026

(1) selecting experimentally confirmed active compounds, (2) generating 3D conformations of the ligands and performing their structural alignment, (3) identifying the structural features and functional groups involved in molecular recognition, (4) constructing and validating the pharmacophore model using a testing dataset from a compound library, and (5) screening the target library. In this process, the pharmacophore model is developed by aligning the 3D conformers of a set of bioactive compounds (training dataset) and subsequently validated using active compounds (true positives or hits) and inactive compounds (false positives or decoys) as a testing set (Pal *et al*., 2019).

It should be noted that highly restrictive pharmacophore models may select compounds with stronger activity against the target but can reduce the structural diversity of natural products. On the other hand, less restrictive models may capture a higher number of false positives (Schaller *et al*., 2020).

Pharmacophore modelling methods use two main types of scoring functions to assess how well compounds match the predicted model: root mean square deviation (RMSD)-based and overlay-based scoring. RMSD-based scoring evaluates the fitness of a compound by calculating the distances between its functional groups and the centres of the pharmacophore features. In contrast, overlay-based scoring estimates functional similarity by taking into account the radii of the functional groups or atoms relative to the pharmacophore model (Vuorinen & Schuster, 2015).

Structure-based pharmacophore modelling

The structure-based approach, in contrast, utilizes spatial information of a ligand bound to its molecular target, including ligand conformations and poses. This method is applicable only when experimentally determined structures of the target, typically in complex with an active ligand, are available, such as those obtained via X-ray crystallography (Jiang *et al*., 2020).

A structure-based model provides insights into ligand–target interactions and requires the 3D structure of the target, either obtained experimentally through X-ray crystallography or NMR, or modeled based on homologous proteins. The Brookhaven Protein Data Bank (PDB) is a key repository that stores the 3D coordinates of experimentally solved protein structures (Berman *et al*., 2000).

The optimal starting point for constructing a structure-based 3D model is a crystal complex where a ligand is bound to the protein's active site. This setup offers accurate information on the ligand's bioactive conformation within the binding site.

Official Website
www.thescienceworld.net
thescienceworldmagazine@gmail.com
Published 30/08/2026

LIGANDSCOUT, a recently developed software tool, aids in generating chemical feature–based models by interpreting ligands and mining the PDB to identify critical interaction points between ligands and proteins (Wolber & Langer, 2005).

Structure-Based Virtual Screening Approaches

Molecular docking

Molecular docking is a computational approach used in structure-based drug design to predict the binding geometries and affinities between proteins and ligands (Cosconati *et al*., 2010).

In docking-based virtual screening, compounds from a database are assessed and ranked based on their steric and electrostatic compatibility with the target site, with the top-ranked molecules selected for further evaluation through biochemical assays (Singh *et al*., 2006). The docking algorithm generates an optimal set of ligand conformations and binding poses, ideally capturing the experimentally determined binding mode. Scoring functions commonly employed in docking include force-field-based, empirical, and knowledge-based methods (Guedes *et al*., 2018). Consequently, careful evaluation is necessary when choosing the best compounds from docking-based virtual screening results (Jacob *et al*., 2012).

A variety of programs are available for database virtual screening (DBVS), including AutoDock, which is widely cited in scientific literature and freely accessible for academic use (Morris *et al*., 1996), as well as GOLD (Jones *et al*., 1997), DOCK (Ewing *et al*., 2001), and Glide (Halgren *et al*., 2004).

When choosing docking software, several considerations are important, including the capability to adjust docking parameters or protocols interactively based on new data, support for additional scoring functions, availability of pre- and post-docking filters, inclusion of design features, validation study results, user learning curve, customer support, cost, computational efficiency, user interface, compatible input/output structural file formats, accessibility of the source code, and the possibility for future software updates (Ghosh *et al*., 2006).

Official Website
www.thescienceworld.net
thescienceworldmagazine@gmail.com
Published 30/08/2026

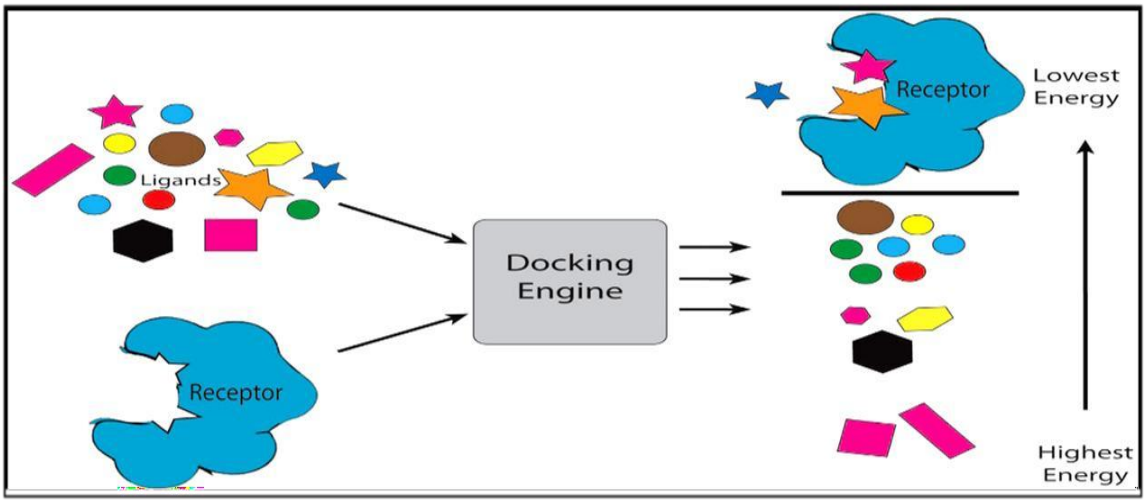

Figure : Demonstration of high-throughput virtual screening: Multiple ligands are docked to a receptor and ranked by energy estimate

ADVANTAGES OF VIRTUAL SCREENING

Reduces costs, infrastructure requirements, and time needed for discovering new chemical structures. Enables the identification of novel compounds with specific bioactivities by evaluating large structural libraries in silico. Filters compounds based on desirable physicochemical, pharmacokinetic, and pharmacodynamic properties, while eliminating those that do not meet the criteria. Allows in silico assessment of the "druggability" of new molecular targets. Virtual screening significantly impacts lead discovery; integrating chemical information with virtual screening is expected to play a key role in drug development in the post-genomic era. Achieves a substantially higher hit rate compared to traditional screening approaches, such as high-throughput screening (Shen *et al*., 2003; Macalino *et al*., 2015).

DISADVANTAGES OF VIRTUAL SCREENING

Dependent on the choice of scoring function. Heavily influenced by the target and the search algorithm used. Certain molecules often produce false-positive bioactivity results; including such molecules in the training or validation set can compromise the predictive accuracy of the model. Comparing different screening methods or models is essential to determine the most appropriate approach, yet evaluating and contrasting virtual screening protocols and models remains challenging (Kirchweger & Rollinger, 2018).

APPLICATIONS USED IN VIRTUAL SCREENING

Sr	Tool	Identified Molecules and	References

Official Website
www.thescienceworld.net
thescienceworldmagazine@gmail.com
Published 30/08/2026

No		other findings	
1	AutoDOCK Vina	Liquiritic acid, Terchenulin etc	Joshi *et al*., 2021
2	Molecular Docking	CID 25145634, CID 25145634, etc.	Karthick *et al*., 2013
3	Pharmacophore modeling	V12 and V14 compounds as a lead	Zhao *et al*., 2021
4	Molecular Docking	Rutin as a potential inhibitor	Xu *et al*., 2020
5	Molecular Docking	β hematin inhibitors	de Sousa *et al*., 2020
6	Molecular Docking	NSC-358078	Yazdani *et al*., 2023

CONCLUSIONS

The rational design of new drugs through the models employed in VS has proved effective and may save time, effort and money. Virtual screening will probably be crucial in speeding up the initial phase of drug discovery by effectively identifying lead compounds from nature. The VS, when compared to traditional in vitro screening, offers greater capacity and reduced experimental efforts for testing. In theory, it is possible to compute and forecast interactions between all identified products and all structurally characterized targets. The quality of hit compounds can be enhanced by implementing extra drug-like filters (ADME and drug- likeness), thereby reducing failures in the initial stages of drug development.

FUTURE PROSPECTS

The future of virtual-focused evaluation of databases by VS, utilizing well-validated pharmacophore models, will lead to a substantial enhancement in determining lead structures. Pharmaceutical history shows that a significant number of drugs have come from natural products. However, even for the drug targets currently in use, the natural products that are available have not been fully tested.

REFERENCES

Andricopulo, A. D., & Montanari, C. A. (2005). Structure-activity relationships for the design of small-molecule inhibitors. *Mini reviews in medicinal chemistry*, *5*(6), 585-593. https://doi.org/10.2174/1389557054023224

Berenger, F., Vu, O., & Meiler, J. (2017). Consensus queries in ligand-based virtual

screening experiments. *Journal of cheminformatics*, *9*(1), 60. https://doi.org/10.1186/s13321-017- 0248-5

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov I. N., & Bourne, P. E. (2000). The protein data bank. *Nucleic acids Research*, *28*(1), 235-242. https://doi.org/10.1093/nar/28.1.235

Bonanno, E., & Ebejer, J. P. (2020). Applying machine learning to ultrafast shape recognition in ligand-based virtual screening. *Frontiers in Pharmacology*, *10*, 1675. https://doi.org/10.3389/fphar.2019.01675

Cai, C., Gong, J., Liu, X., Gao, D., & Li, H. (2013). SimG: an alignment-based method for evaluating the similarity of small molecules and binding sites. *Journal of Chemical Information and Modeling*, *53*(8), 2103-2115. https://doi.org/10.1021/ci400139j

Chen, Y., Mathai, N., & Kirchmair, J. (2020). Scope of 3D shape-based approaches in predicting the macromolecular targets of structurally complex small molecules including natural products and macrocyclic ligands. *Journal of Chemical Information and Modeling*, *60*(6), 2858-2875. https://doi.org/10.1021/acs.jcim.0c00161

Cleves, A. E., Johnson, S. R., & Jain, A. N. (2019). Electrostatic-field and surface-shape similarity for virtual screening and pose prediction. *Journal of Computer-Aided Molecular Design*,*33*(10),865 886.https://doi.org/10.1007/s10822-019-00236-6

Cosconati, S., Forli, S., Perryman, A. L., Harris, R., Goodsell, D. S., & Olson, A. J. (2010). Virtual screening with AutoDock: theory and practice. *Expert Opinion on Drug Discovery*, *5*(6), 597-607. https://doi.org/10.1517/17460441.2010.484460

da Silva R., S. F., Olanda, C. G., Fokoue, H. H., & Santanna, C. M. (2019). Virtual screening techniques in drug discovery: review and recent applications. *Current Topics in Medicinal Chemistry*, *19*(19), 1751-1767. https://doi.org/10.2174/1568026619666190816101948

Dai, R., Zhang, W., Tang, W., Wynendaele, E., Zhu, Q., Bin, Y., De Spiegeleer, B., & Xia, J. (2021). BBPpred: sequence-based prediction of blood-brain barrier peptides with feature representation learning and logistic regression. *Journal of Chemical Information and Modeling*, *61*(1), 525-534. https://doi.org/10.1021/acs.jcim.0c01115

Daina, A., & Zoete, V. (2016). A boiled-egg to predict gastrointestinal absorption and brain penetration of small molecules. *ChemMedChem*, *11*(11), 1117-1121 https://doi.org/10.1002/cmdc.201600182

de Sousa, A. C. C., Combrinck, J. M., Maepa, K., & Egan, T. J. (2020). Virtual screening as a tool to discover new β-haematin inhibitors with activity against malaria parasites. Scientific Report 10: 3374. https://doi.org/10.1038/s41598-020-60221-0

Debouck, C., & Metcalf, B. (2000). The impact of genomics on drug discovery. *Annual Review of Pharmacology and Toxicology*, *40*(1), 193-208.

Ewing, T. J., Makino, S., Skillman, A. G., & Kuntz, I. D. (2001). DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *Journal of Computer- Aided Molecular Design*, *15*, 411-428. https://doi.org/10.1023/A:1011115820450

Garcia-Hernandez, C., Fernandez, A., & Serratosa, F. (2019). Ligand-based virtualscreening using graph edit distance as molecular similarity

measure.*Journal of Chemical Information and Modeling*, *59*(4), 1410-1421. https://doi.org/10.1021/acs.jcim.8b00820 Gasteiger, J., Teckentrup, A., Terfloth, L., & Spycher, S. (2003). Neural networks as data mining tools in drug design. *Journal of Physical Organic Chemistry*, *16*(4), 232-245.https://doi.org/10.1002/poc.597

Ghosh, S., Nie, A., An, J., & Huang, Z. (2006). Structure-based virtual screening of chemical libraries for drug discovery. *Current Opinion in Chemical Biology*, *10*(3), 194-202. https://doi.org/10.1016/j.cbpa.2006.04.002

Grant, J. A., Gallardo, M. A., & Pickup, B. T. (1996). A fast method of molecular shape comparison: A simple application of a Gaussian description of molecular shape. *Journal of Computational Chemistry*, *17*(14), 1653-1666. https://doi.org/10.1002/(SICI)1096- 987X(19961115)17:14%3C1653::AID-JCC7%3E3.0.CO;2-K

Grüneberg, S., Stubbs, M. T., & Klebe, G. (2002). Successful virtual screening for novel inhibitors of human carbonic anhydrase: strategy and experimental confirmation. *Journal of medicinal chemistry*, *45*(17), 3588-3602. https://pubs.acs.org/doi/10.1021/jm011112j

Guedes, I. A., Pereira, F. S., & Dardenne, L. E. (2018). Empirical scoring functions for structure-based virtual screening: applications, critical aspects, and challenges. *Frontiers in Pharmacology*, *9*, 1089. https://doi.org/10.3389/fphar.2018.01089

Guedes, I.A., de Magalhães, C.S., & Dardenne, L.E. (2014). Receptor–ligand molecular docking. *Biophysical Reviews*, *6*, 75–87. https://doi.org/10.1007/s12551-013-0130-2

Hajduk, P. J., & Greer, J. (2007). A decade of fragment-based drug design: strategic advances and lessons learned. Nature reviews Drug discovery, 6(3), 211-219. https://doi.org/10.1038/nrd2220

Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., & Banks, J. L. (2004). Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal of Medicinal Chemistry*, *47*(7), 1750- 1759. https://doi.org/10.1021/jm030644s

Hönig, S. M., Lemmen, C., & Rarey, M. (2023). Small molecule superposition: A comprehensive overview on pose scoring of the latest methods. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *13*(2), e1640. https://doi.org/10.1002/wcms.1640

Hopkins, A. L., & Groom, C. R. (2002). The druggable genome. *Nature reviews Drug discovery*, *1*(9), 727-730. https://doi.org/10.1038/nrd892

Huggins, D. J., Venkitaraman, A. R., & Spring, D. R. (2011). Rational methods for the selection of diverse screening compounds. *ACS Chemical Biology*, *6*(3), 208-217. https://doi.org/10.1021/cb100420r

Jacob, R. B., Andersen, T., & McDougal, O. M. (2012). Accessible high-throughput virtual screening molecular docking software for students and educators. *PLoS Computational Biology*, *8*(5), e1002499. https://doi.org/10.1371/journal.pcbi.1002499

Jasial, S., Gilberg, E., Blaschke, T., & Bajorath, J. (2018). Machine learning distinguishes with high accuracy between pan-assay interference compounds that are promiscuous or represent dark chemical matter. *Journal of Medicinal Chemistry*, *61*(22), 10255-10264. https://doi.org/10.1021/acs.jmedchem.8b01404

Jiang, S., Feher, M., Williams, C., Cole, B., & Shaw, D. E. (2020). Autoph4: An automated method for generating pharmacophore models from protein binding pockets. *Journal of Chemical Information and Modeling*, *60*(9), 4326-4338.https://doi.org/10.1021/acs.jcim.0c00121

Jiménez-Luna, J., Grisoni, F., & Schneider, G. (2020). Drug discovery with explainable artificial intelligence. *Nature Machine Intelligence*, *2*(10), 573-584.https://doi.org/10.1038/s42256-020-00236-4

Jones, G., Willett, P., Glen, R. C., Leach, A. R., & Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology*, *267*(3), 727-748. https://doi.org/10.1006/jmbi.1996.0897

Joshi, C., Chaudhari, A., Joshi, C., Joshi, M., & Bagatharia, S. (2021). Repurposing of the herbal formulations: molecular docking and molecular dynamics simulation studies to validate the efficacy of phytocompounds against SARS-CoV-2 proteins. *Journal of Biomolecular Structure and Dynamics*, *40*(18), 8405-8419.https://doi.org/10.1080/07391102.2021.1922095

Kar, S., & Roy, K. (2013). How far can virtual screening take us in drug discovery? *Expert Opinion on DrugDiscovery*, *8*(3 245-261.https://doi.org/10.1517/17460441.2013.761204

Karthick, V., Ramanathan, K., Shanthi, V., & Rajasekaran, R. (2013). Identification of potential inhibitors of H5N1 influenza A virus neuraminidase by ligand-based virtual screening approach. *Cell Biochemistry and Biophysics*, *66*, 657-669.https://doi.org/10.1007/s12013-012-9510-7

Kaur, M., & Silakari, O. (2017). Ligand-based and e-pharmacophore modeling, 3D-QSAR and hierarchical virtual screening to identify dual inhibitors of spleen tyrosine kinase (Syk) and janus kinase 3 (JAK3). *Journal of Biomolecular Structure and Dynamics*, *35*(14), 3043-3060. https://doi.org/10.1080/07391102.2016.1240108

Khan, S. U., Ahemad, N., Chuah, L. H., Naidu, R., & Htar, T. T. (2019). Sequential ligand- and structure-based virtual screening approach for the identification of potential G protein- coupled estrogen receptor-1 (GPER-1) modulators. *RSC advances*, *9*(5), 2525-2538. https://doi.org/10.1039/C8RA09318K

Kirchweger, B., & Rollinger, J. M. (2018). Virtual screening for the discovery of active principles from natural products. *Natural Products as Source of Molecules with Therapeutic Potential: Research & Development, Challenges and Perspectives*, 333-364. https://doi.org/10.1007/978-3-030-00545-0_9

Koes, D. R., & Camacho, C. J. (2011). Pharmer: efficient and exact pharmacophore search. *Journal of Chemical Information and Modeling*, *51*(6), 1307-1314.

Koes, D. R., & Camacho, C. J. (2014). Shape-based virtual screening with volumetric aligned molecular shapes. *Journal of Computational Chemistry*, *35*(25), 1824-1834. https://doi.org/10.1002/jcc.23690

Lavecchia, A., & Di Giovanni, C. (2013). Virtual screening strategies in drug discovery: a critical review. *Current Medicinal Chemistry*, *20*(23), 2839-2860.http://dx.doi.org/10.2174/09298673113209990001

Li, H., Leung, K. S., Wong, M. H., & Ballester, P. J. (2016). USR-VS: a web server for large- scale prospective virtual screening using ultrafast shape recognition

techniques. *Nucleic acids Research*, *44*(W1), W436-W441. https://doi.org/10.1093/nar/gkw320

Liang, J. W., Wang, M. Y., Wang, S., Li, S. L., Li, W. Q., & Meng, F. H. (2020). Identification of novel CDK2 inhibitors by a multistage virtual screening method based on SVM, pharmacophore and docking model. *Journal of enzyme inhibition and medicinal chemistry*, *35*(1), 235-244. https://doi.org/10.1080/14756366.2019.1693702

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, *23*(1-3), 3-25.https://doi.org/10.1016/S0169-409X(96)00423-1

Liu, X., Jiang, H., & Li, H. (2011). SHAFTS: a hybrid approach for 3D molecular similarity calculation. 1. Method and assessment of virtual screening. *Journal of Chemical Information and Modeling*, *51*(9), 2372-2385. https://doi.org/10.1021/ci200060s

Liu, X., Ouyang, S., Yu, B., Liu, Y., Huang, K., Gong, J., Zheng, S., Li, Z., Li, H., & Jiang, H. (2010). PharmMapper server: a web server for potential drug target identification using pharmacophore mapping approach. *Nucleic acids Research*, *38*(suppl_2), W609-W614. https://doi.org/10.1093/nar/gkq300

Lo, Y. C., Rensi, S. E., Torng, W., & Altman, R. B. (2018). Machine learning in chemoinformatics and drug discovery. *Drug Discovery Today*, *23*(8), 1538-1546.https://doi.org/10.1016/j.drudis.2018.05.010

Lu, F., Luo, G., Qiao, L., Jiang, L., Li, G., & Zhang, Y. (2016). Virtual screening for potential allosteric inhibitors of cyclin-dependent kinase 2 from traditional chinese medicine. *Molecules*, *21*(9), 1259. https://doi.org/10.3390/molecules21091259

Macalino, S. J. Y., Gosu, V., Hong, S., & Choi, S. (2015). Role of computer-aided drug design in modern drug discovery. *Archives of Pharmacal Research*, *38*, 1686-1701.https://doi.org/10.1007/s12272-015-0640-5

Martinez-Trevino, S. H., Uc-Cetina, V., Fernandez-Herrera, M. A., & Merino, G. (2020). Prediction of natural product classes using machine learning and 13C NMR spectroscopic data. *Journal of Chemical Information and Modeling*, *60*(7), 3376-3386. https://doi.org/10.1021/acs.jcim.0c00293

Meng, X. Y., Zhang, H. X., Mezei, M., & Cui, M. (2011). Molecular docking: a powerful approach for structure-based drug discovery. *Current Computer-Aided Drug Design*, *7*(2), 146–157. https://doi.org/10.2174/157340911795677602

Mignani, S., Rodrigues, J., Tomas, H., Jalal, R., Singh, P. P., Majoral, J. P., & Vishwakarma, R. A. (2018). Present drug-likeness filters in medicinal chemistry during the hit and lead optimization process: how far can they be simplified? *Drug Discovery Today*, *23*(3), 605- 615. https://doi.org/10.1016/j.drudis.2018.01.010

Molecular Operating Environment (2019). Chemical Computing Group ULC. Montreal, QC: Molecular Operating Environment, 202.

Morrell, J. (1996). Mining information from databases for drug discovery. *Book Abstr 21^{1th} ACS National Meeting* CINF-052.

Morris, G. M., Goodsell, D. S., Huey, R., & Olson, A. J. (1996). Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. *Journal of Computer-Aided Molecular Design*, *10*, 293-304.https://doi.org/10.1007/BF00124499

Oum, Y. H., Kell, S. A., Yoon, Y., Liang, Z., Burger, P., & Shim, H. (2020). Discovery of

novel aminopiperidinyl amide CXCR4 modulators through virtual screening and rational drug design. *European journal of medicinal chemistry*, *201*, 112479.https://doi.org/10.1016/j.ejmech.2020.112479

Overington, J. P., Al-Lazikani, B., & Hopkins, A. L. (2006). How many drug targets are there?. *Nature reviews Drug discovery*, *5*(12), 993-996. https://doi.org/10.1038/nrd2199

Pal, S., Kumar, V., Kundu, B., Bhattacharya, D., Preethy, N., Reddy, M. P., & Talukdar, A. (2019). Ligand-based pharmacophore modeling, virtual screening and molecular docking studies for discovery of potential topoisomerase I inhibitors. *Computational and Structural Biotechnology Journal*, *17*, 291-310. https://doi.org/10.1016/j.csbj.2019.02.006

Patel, L., Shukla, T., Huang, X., Ussery, D. W., & Wang, S. (2020). Machine learning methods in drug discovery. *Molecules*, *25*(22), 5277. https://doi.org/10.3390/molecules25225277 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D.,

Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, *12*, 2825-2830. https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf?ref=https:/

Pu, L., Naderi, M., Liu, T., Wu, H. C., Mukhopadhyay, S., & Brylinski, M. (2019). e toxpred: A machine learning-based approach to estimate the toxicity of drug candidates. *BMC Pharmacology and Toxicology*, *20*, 1-15. https://doi.org/10.1186/s40360-018-0282-6

Puertas-Martín, S., Redondo, J. L., Ortigosa, P. M., & Pérez-Sánchez, H. (2019). OptiPharm: an evolutionary algorithm to compare shape similarity. *Scientific Reports*, *9*(1), 1-24. https://doi.org/10.1038/s41598-018-37908-6

Sahlgren, C., Meinander, A., Zhang, H., Cheng, F., Preis, M., Xu, C., Salminen, T. A., Toivola, D., Abankwa, D., Rosling, A., Karaman, D. S., Salo-ahen, M. H., Osterbacka, R., Eriksson, J. E., Willfor, S., Petre, I., Peltonen, J., Leino, R., Johnson, M., Rosenholm, J., & Sandler, N. (2017). Tailored approaches in drug development and diagnostics: from molecular design to biological model systems. *Advanced Healthcare Materials*, *6*(21), 1700258. https://doi.org/10.1002/adhm.201700258

Santana, K., Do Nascimento, L. D., Lima e Lima, A., Damasceno, V., Nahum, C., Braga, R. C., & Lameira, J. (2021). Applications of virtual screening in bioprospecting: facts, shifts, and perspectives to explore the chemo-structural diversity of natural products. *Frontiers in Chemistry*, *9*, 662688. https://doi.org/10.3389/fchem.2021.662688

Schaller, D., Šribar, D., Noonan, T., Deng, L., Nguyen, T. N., Pach, S., Machalz, D., Bermudez, M., & Wolber, G. (2020). Next generation 3D pharmacophore modeling. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *10*(4), e1468. https://doi.org/10.1002/wcms.1468

Schapira, M., Raaka, B. M., Das, S., Fan, L., Totrov, M., Zhou, Z., Wilson, S. R., Abagyan, R., & Samuels, H. H. (2003). Discovery of diverse thyroid hormone receptor antagonists by high-throughput docking. *Proceedings of the National Academy of Sciences*, *100*(12), 7354-7359. https://doi.org/10.1073/pnas.1131854100

Schneider, G. (2010). Virtual screening: an endless staircase? *Nature Reviews Drug*

Discovery, *9*(4), 273-276. https://doi.org/10.1038/nrd3139

Seddon, M. P., Cosgrove, D. A., Packer, M. J., & Gillet, V. J. (2018). Alignment-free molecular shape comparison using spectral geometry: the framework. *Journal of Chemical Information and Modeling*, *59*(1), 98-116. https://doi.org/10.1021/acs.jcim.8b00676

Shen, J., Xu, X., Cheng, F., Liu, H., Luo, X., Shen, J., Chen, K., Zhao, W., Shen, X., & Jiang, H. (2003). Virtual screening on natural products for discovering active compounds and target information. *Current Medicinal Chemistry*, *10*(21), 2327 2342. https://doi.org/10.2174/0929867033456729

Shin, W. H., Zhu, X., Bures, M. G., & Kihara, D. (2015). Three-dimensional compound comparison methods and their application in drug discovery. *Molecules*, *20*(7), 12841- 12862. https://doi.org/10.3390/molecules200712841

Shoichet, B. K. (2004). Virtual screening of chemical libraries. *Nature*, *432*(7019), 862-865. https://doi.org/10.1038/nature03197

Sieg, J., Flachsenberg, F., & Rarey, M. (2019). In need of bias control: evaluating chemical data for machine learning in structure-based virtual screening. *Journal of Chemical Information and Modeling*, *59*(3), 947-961. https://doi.org/10.1021/acs.jcim.8b00712

Singh, S., Malik, B. K., Sharma, D. K. (2006). Molecular drug targets and structure-based drug design: A holistic approach. *Bioinformation*, *1*(8), 314-320. https://doi.org/10.6026%2F97320630001314

Sunseri, J., & Koes, D. R. (2016). Pharmit: interactive exploration of chemical space. *Nucleic acids Research*, *44*(W1), W442-W448. https://doi.org/10.1093/nar/gkw287

Tame, J. R. (1999). Scoring functions: a view from the bench. *Journal of Computer-Aided Molecular Design*, *13*(2), 99-108.

Taminau, J., Thijs, G., & De Winter, H. (2008). Pharao: pharmacophore alignment and optimization. *Journal of Molecular Graphics and Modelling*, *27*(2), 161-169.https://doi.org/10.1016/j.jmgm.2008.04.003

Tsou, L. K., Yeh, S. H., Ueng, S. H., Chang, C. P., Song, J. S., Wu, M. H., Chang, H. F., Chen, S. R., Shih, C., Chen, C. T., & Ke, Y. Y. (2020). Comparative study between deep learning and QSAR classifications for TNBC inhibitors and novel GPCR agonist discovery. *Scientific Reports*, *10*(1), 16771.https://doi.org/10.1038/s41598-020-73681-1

Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M., & Zhao, S. (2019). Applications of machine learning in drug discovery and development. *NatureReviews Drug Discovery*, *18*(6), 463-477.https://doi.org/10.1038/s41573-019-0024-5

Veber, D. F., Johnson, S. R., Cheng, H. Y., Smith, B. R., Ward, K. W., & Kopple, K. D. (2002). Molecular properties that influence the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry*, *45*(12), 2615-2623. https://doi.org/10.1021/jm020017n

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., Van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, Andrew R. J., Jones, E., Kern, R., Larson, E.,

Carey, C. J., Polat, I., Feng, Y., Moore, E. W., Vanderplas, J., Laxalde, D., perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M.,

Ribeiro, A. H., Pedregosa, F., & Van Mulbregt, P. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261-272.https://doi.org/10.1038/s41592-019-0686-2

Vuorinen, A., & Schuster, D. (2015). Methods for generating and applying pharmacophore models as virtual screening filters and for bioactivity profiling. *Methods*, *71*, 113-134. https://doi.org/10.1016/j.ymeth.2014.10.013

Wang, Z., Sun, H., Shen, C., Hu, X., Gao, J., Li, D., Cao, D., & Hou, T. (2020). Combined strategies in structure-based virtual screening. *Physical Chemistry Chemical Physics*, *22*(6), 3149-3159. https://doi.org/10.1039/C9CP06303J

Waszkowycz, B., Perkins, T. D. J., Sykes, R. A., & Li, J. (2001). Large-scale virtual screening for discovering leads in the postgenomic era. *IBM Systems Journal*, *40*(2), 360-376. https://doi.org/10.1147/sj.402.0360

Wei, L., Xing, P., Su, R., Shi, G., Ma, Z. S., & Zou, Q. (2017). CPPred-RF: a sequence-based predictor for identifying cell-penetrating peptides and their uptake efficiency. *Journal of Proteome Research*, *16*(5), 2044-2053. https://doi.org/10.1021/acs.jproteome.7b00019

Willett, P. (2006). Similarity-based virtual screening using 2D fingerprints. *Drug discovery today*, *11*(23-24), 1046-1053.

Wolber, G., & Langer, T. (2005). LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *Journal of Chemical Information and Modeling*, *45*(1), 160-169. https://doi.org/10.1021/ci049885e

Wolfe, J. M., Fadzen, C. M., Choo, Z. N., Holden, R. L., Yao, M., Hanson, G. J., & Pentelute,

B. L. (2018). Machine learning to predict cell-penetrating peptides for antisense delivery. *ACS CentralScience*, *4*(4), 512-520.https://doi.org/10.1021/acscentsci.8b00098

Xu, X., Ren, J., Ma, Y., Liu, H., Rong, Q., Feng, Y., Wang, Y., Cheng, Y., Ge, R., Li, Z., & Bian, J. (2019). Discovery of cyanopyridine scaffold as novel indoleamine-2, 3-dioxygenase 1 (IDO1) inhibitors through virtual screening and preliminary hit optimisation. *Journal of enzyme inhibition and medicinal chemistry*, *34*(1), 250-263.https://doi.org/10.1080/14756366.2018.1480614

Xu, Z., Yang, L., Zhang, X., Zhang, Q., Yang, Z., Liu, Y., Liu, Y., Wei, S. & Liu, W. (2020). Discovery of potential flavonoid inhibitors against COVID-19 3CL proteinase based on virtual screening strategy. *Frontiers in Molecular Biosciences*, *7*, 556481. https://doi.org/10.3389/fmolb.2020.556481

Yan, X., Li, J., Gu, Q., & Xu, J. (2014). gWEGA: GPU-accelerated WEGA for molecular superposition and shape comparison. *Journal of Computational Chemistry*, *35*(15), 1122- 1130. https://doi.org/10.1002/jcc.23603

Yan, X., Liao, C., Liu, Z., T Hagler, A., Gu, Q., & Xu, J. (2016). Chemical structure similarity search for ligand-based virtual screening: methods and computational resources. *Current Drug Targets*, *17*(14), 1580-1585.https://doi.org/10.2174/1389450116666151102095555

Yazdani, B., Sirous, H., Brogi, S., & Calderone, V. (2023). Structure-based high-throughput virtual screening and molecular dynamics simulation for the discovery of novel SARS- CoV-2 NSP3 Mac1 domain inhibitors. *Viruses*, *15*(12), 2291.https://doi.org/10.3390/v15122291

Zhao, S., Li, X., Peng, W., Wang, L., Ye, W., Zhao, Y., Yin, W., Chen, W. D., Li, W., & Wang, Y. D. (2021). Ligand-based pharmacophore modeling, virtual screening and biological evaluation to identify novel TGR5 agonists. *RSC Advances*, *11*(16), 9403-9409. https://doi.org/10.1039/D0RA10168K