Computational Methods & Tools for Drug Design & Discovery

Bhavesh B. Amrute[1], Nayan R. Chordiya[2], Yash S. Deshmukh[3], Ronak P. Jain[4]
Chandrashekhar D. Upasani[5]
[1]*Assistant Professor, Dept. Ph. Chem, SNJBs Shriman Sureshdada Jain College of Pharmacy, Neminagar, Chandwad, Nashik, Maharashtra, India.*
[2,3,4,5] *SNJBs Shriman Sureshdada Jain College of Pharmacy, Neminagar, Chandwad, Nashik, Maharashtra, India.*

***Abstract*—Chemical biology and computational drug design methods are used in drug development to effectively identify and improve lead molecules. The main applications of chemical biology are to clarify the biological function of a target and the mode of action of a chemical modulator. In contrast, computer-aided drug design uses structural information about the target (structure-based) or information about existing ligands with bioactivity (ligand-based) to make it easier to identify prospective candidate medications.[1] Pharmaceutical corporations and academic research teams are now using a variety of virtual screening techniques to shorten the time and expense involved in finding a potent medicine.[2,3,8 & 9]**

The laborious process of finding new drugs can be expedited with the use of computer-aided drug discovery (CADD) systems, potentially lowering the cost of R&D. Computational approaches can now be used successfully at different stages of the drug discovery and development pipeline because to advances in biological structural knowledge and computer capacity.[4,5,6,7] Here, we provide an overview of computational approaches employed in many aspects of drug development and highlight a few cutting-edge methodologies and tools.[11, 12]

***Index Terms*—CADD, LBDD, SBDD, Drug discovery, Drug development, Docking.**

1. INTRODUCTION

1.1. Computer Aided Drug Design (CADD)

The process of locating, creating, and analyzing drugs and related biological active substances using computer techniques is known as computer-aided drug design (CADD).[2,12]

The application of CADD techniques facilitates and expedites drug discovery while accelerating the early stages of chemical development. [14, 16]

Structure-based drug design and ligand-based drug design approaches are known as very effective and potent strategies in drug discovery and development. [19, 20] Various CADD approaches are rated as promising methodologies based on their necessity. Both of these approaches can be used in conjunction with molecular docking for virtual lead optimization and identification. [21-25]

In order to increase the efficiency and effectiveness of drug discovery and development, computational techniques have become increasingly popular in the pharmaceutical industry and in research sectors pipeline. [27-30]

1.2. Steps Involved in Computer Aided Drug Design (CADD)

Step 1: Target Identification:
There is a designated therapeutic target for which a medication must be created. It is possible to utilize a structure-based approach or a ligand-based approach, depending on the availability of structure information. Multiple lead compounds will be able to be identified by a successful CADD campaign. [26, 28]

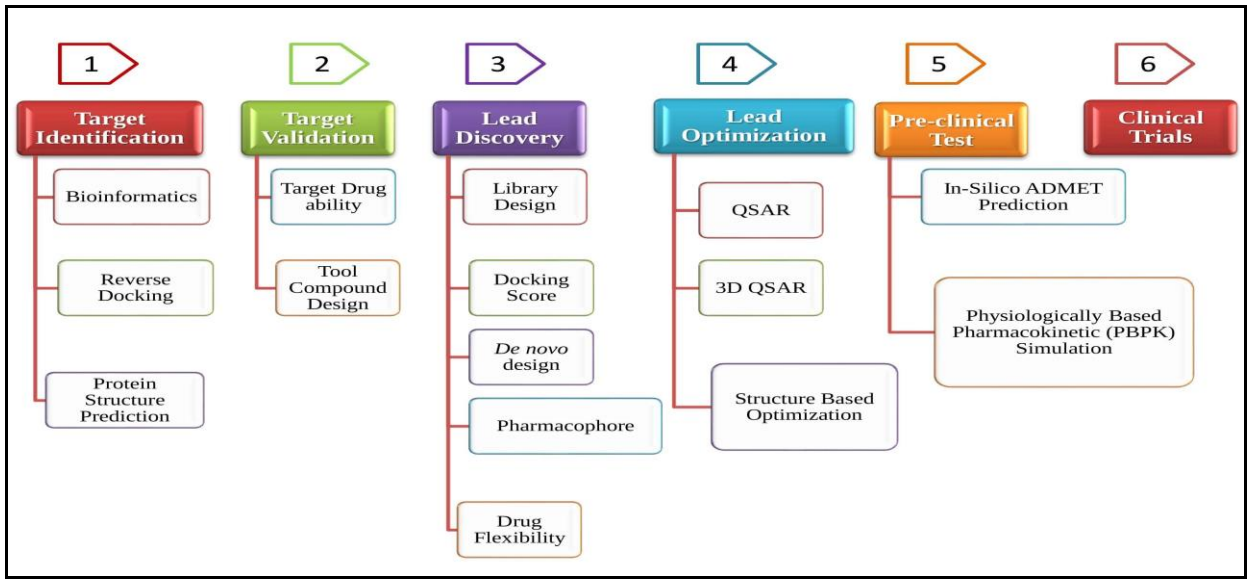

Figure No.1: Steps Involved in Computer Aided Drug Design (CADD)

1.1 Bioinformatics:

Techniques in bioinformatics are also used to determine whether a target is "drugable." It is conceivable to lower the probability of project failure later in the discovery process by performing such assessments in the early stages of drug research. [31]

1.2 Reverse Docking:

Drug repositioning and drug rescue are becoming more and more effective with the use of reverse or inverse docking. It entails docking a small-molecule drug or ligand into a variety of clinically significant macromolecular targets' potential binding cavities. [32-35]

1.3 Protein Structure Prediction:

Various biological processes depend heavily on proteins, which are significant molecules.

Sequence and structural homology serve as the main foundation for the prediction of protein structure. [36] Since a protein's activity primarily depends on its three-dimensional structure, protein structure prediction or modeling is crucial. Similar to this, a protein's amino acid makeup determines its 3D structure. [37, 38] A slight change in the protein's sequence can result in significant structural changes in the protein's native structure. Although it is critical to have a thorough understanding of protein 3D structure, it can be challenging to determine a protein's native structure when it exists in the physiological environment of the body. [39]

II. TARGET VALIDATION

Target validation is a complex stage of drug research that is essential to a medicine's success.

To make sure that the medicine is actively interacting with the target to have the desired therapeutic effect validated targets, biomarkers, and assessment methods are required. [40]

2.1 Target Drug Ability:

The analysis of ligands known to interact with a target of interest is a component of the ligand-based computer-aided drug discovery (LB-CADD) method. [40]

2.2 Tool Compound Design

III. LEAD DISCOVERY

A thorough search is conducted during lead discovery to identify a biological or small molecule therapeutics that is similar to a drug and will advance into preclinical and, if successful, clinical development (Figure 2) and eventually become a marketed drug. These candidates are referred to as development candidates. [34,35]

3.1 Library Design:

The term "compound library" refers to a collection of substances that can be used in high-throughput screening and other drug development procedures.

At various stages of drug discovery, the careful design of molecular libraries can speed up SAR

exploration and reveal "breakthrough" lead structures. [35]

3.2 Docking Score:

Molecules' conformation and orientation—collectively referred to as their "pose"—when they enter the binding site of a macromolecular target are examined using a technique called molecular docking. Poses that might be taken are generated by search algorithms, and scoring functions rank them. Over the past few decades, a variety of software has been created, including some well-known examples like AutoDock, AutoDock Vina, DockThor, GOLD [6, 7], FlexX [8], and Molegro Virtual Docker. [32, 33, 34]

3.3 *De novo* design:

De novo drug design (DNDD) is the process of creating novel chemical entities using computational growth algorithms to conform to a set of constraints.

De novo, which means "from the beginning," implies that this technique can be used to create novel molecular entities without the use of a starting template. [39]

3.4 Pharmacophore:

The collection of steric and electronic properties known as a pharmacophore is required to ensure the best supramolecular interactions with a particular biological target structure and to activate (or inhibit) its biological response. [48]

3.5 Drug Flexibility:

Flexibility affects oral bioavailability and is frequently measured by NRot in the early stages of drug discovery. The appeal of large, flexible molecules that are not Ro5 compliant as drug candidates is growing. [40]

IV. LEAD OPTIMIZATION

Following the identification of a primary lead compound, a drug candidate is created via the lead optimization procedure.

In order to get the required pharmacological profile, such as affinity, safety, pharmacokinetics, and ADME, lead optimization often necessitates the synthesis of hundreds of compounds over the course of several years (absorption, distribution, metabolism, elimination, toxicology). [39,40]

4.1 QSAR:

A ligand-based drug design technique called quantitative structure-activity relationship (QSAR) analysis was created more than 50 years ago by Hansch and Fujita (1964). [38]

Since that time and up to the present, QSAR has remained an effective technique for creating mathematical models. It uses regression and classification techniques, respectively, to find a statistically significant correlation between the chemical structure and a continuous (pIC50, pEC50, Ki, etc.) or categorical/binary (active, inactive, toxic, nontoxic, etc.) biological or toxicological property. [30]

The dimensionality of molecular descriptors (from 1D to nD) and various techniques for determining a correlation between chemical structures and biological property have all changed significantly during the past few decades.

4.2 3D QSAR:

The test set ligands' binding affinities were predicted using the 3D-QSAR based pharmacophore model, which allowed for the early prediction of those ligands' pharmacological activity. [21]

With PharmQSAR, 3D QSAR/QSPR models may be generated automatically utilising a library of substances with known activities or characteristics as well as their molecular fields. An external data collection of molecules is used to validate the produced model. [18, 19]

4.3 Structure Based Optimization:

Compared to the conventional approach, structure-based drug design is quickly becoming a crucial tool for lead discovery that is both quicker and more cost-effective.

Numerous novel targets and potential for future drug discovery have been made possible by genomic, proteomic, and structural investigations. [29]

V. PRECLINICAL TEST

Phases of testing must be conducted on the molecule in question during the pre-clinical testing phase of developing a substance into medicine. To treat an illness, a possible target first needs to be found. Multiple compounds are then screened, and those that show promise as a treatment for the disease must go through toxicity testing beforehand in order to minimize the risk of adverse effects. [31-34]

5.1 *In-Sillico* ADMET Prediction:

The precise prediction of the *in-vivo* pharmacokinetics of a potential drug molecule in a

human being, while it only exists as a virtual structure, is the ultimate goal of the in-silico prediction of ADME properties. [35]

5.2 Physiologically based Pharmacokinetic (PBPK) Simulation:

An analytical method for predicting the absorption, distribution, metabolism, and excretion (ADME) of synthetic or natural chemical compounds in people and other animal species is known as physiologically based pharmacokinetic (PBPK) modelling. [14]

PBPK modelling is employed in pharmaceutical research and drug development, as well as in the evaluation of the health risks associated with chemicals or cosmetics.

VI. CLINICAL TRIALS

A type of research called clinical trials examines novel diagnostic and therapeutic approaches and assesses how they affect patient outcomes.

Step 1: Discovery and Development
Step 2: Preclinical Research
Step 3: Clinical Research
Step 4: FDA Drug Review
Step 5: FDA Post-Market Drug Safety Monitoring

Drug Discovery Strategies:

Phenotypic screening, target-based screening, and medication repurposing have been the three main sequential approaches utilized in drug discovery. In a systems-based assay, phenotypic screening entails examining a large number of randomly chosen substances. In target-based screening, a chosen enzyme or receptor is modified to elicit the desired therapeutic response. In general, phenotypic or target-based screening requires a minimum of 10 years to complete the drug discovery cycle from concept to market, whereas drug repurposing only needs three. [29]

Structure-based drug discovery (SBDD) and ligand-based drug discovery (LBDD), which can be used in all of the aforementioned drug discovery methodologies, are the two basic approaches in computer-aided drug discovery (CADD).

In SBDD, a protein's three-dimensional (3-D) structure is examined to find probable binding sites and significant interactions that result in the corresponding pharmacological activity. Attempts are undertaken to find innovative medications with high potency and selectivity using this information. [36]

LBDD focuses on the chemistry of bioactive ligands to establish a structure-activity relationship between physiochemical qualities and bioactivities as the target protein's structure is unknown.

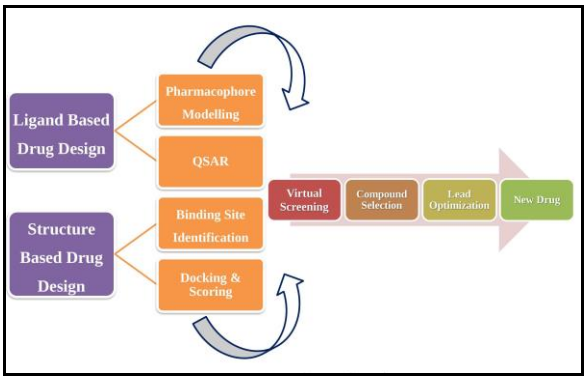

Figure No. 2: Drug Discovery Strategies

Structure-based Drug Discovery (SBDD)

If the target protein structure and the ligand are both known, this strategy is applied, in which the ligand inhibits the protein's activity by a method of competitive binding. It is a combination of experimental and computational methodologies rather than a single tool or technique. [44] It is the recommended CADD methodology since it has the best success rate. Ligand-based Drug Discovery (LBDD)

A single compound or a group of compounds that is well-known to be effective against a target serve as the starting point for ligand-based drug discovery. Based on an understanding of structure-activity relationships (SAR), the potency and other key properties are then improved by creating suitable analogs. [32]

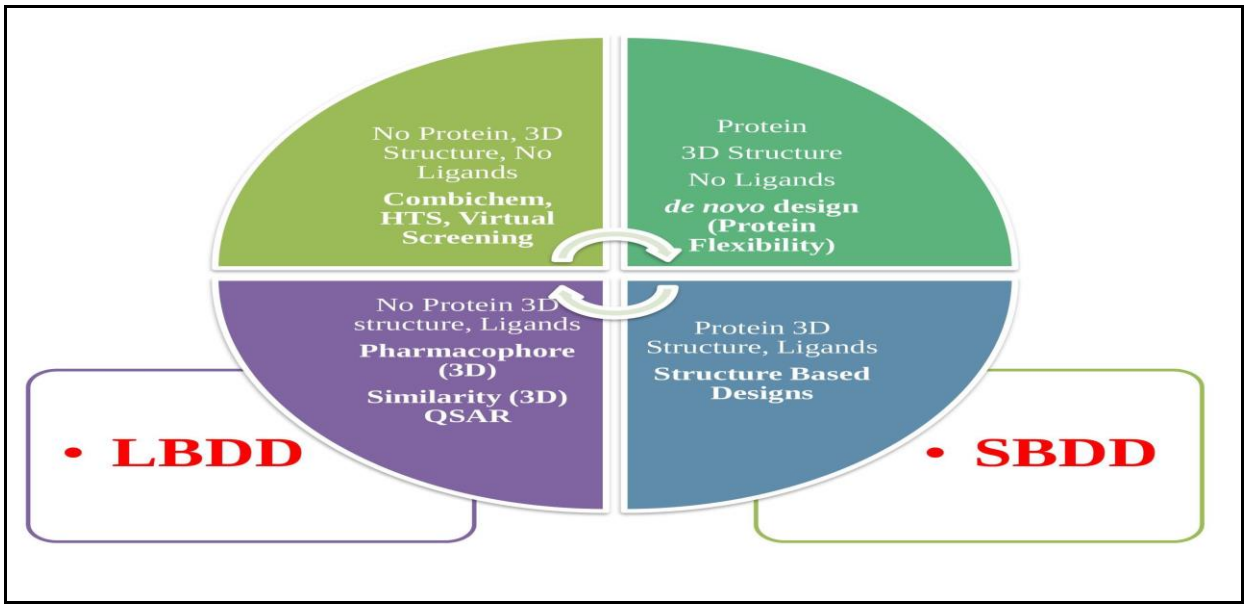

Figure No. 3: Strategies in Drug Designing

Table 1: Computer-aided techniques used in drug design and discovery

Technique	Roles in drug design and discovery
Docking	Determine a compound's approximate binding energy and binding mode to a target.
Structure-based virtual screening	Using docking methods choose active chemicals from a chemical library for a particular target.
Pharmacophore modeling	Understand and describe the molecular characteristics required for a biological macromolecule to recognize a ligand on a molecular level.
Ligand-based virtual screening	Using pharmacophore modelling approaches choose from a chemical library the compounds that are active against a certain target.
Homology modeling	Create a 3D structure based on related protein 3D structures for a target for which a crystal structure is not yet available.
Molecular dynamics	Molecular mechanics-based simulation to determine a compound's binding affinity to a target, examine the flexibility of the drug target for structure-based drug design, and/or comprehend the dynamic behaviour of proteins or other biological macromolecules
2D quantitative structure–activity relationship	Constructing a model that can be used to predict a property from a compound's molecular structure
3D quantitative structure–activity relationship	A method for statistically predicting how a molecule will interact with a target's active site; this method makes use of 3D conformation-derived data.
Quantum mechanics	A first-principles-based electron-orbital technique to optimize ligand and protein-ligand complex structures, increase docking precision, and determine things like free-binding energy
Absorption, distribution, metabolism, elimination, and toxicity prediction	Prediction of chemical substance toxicity, distribution, metabolism, absorption, and elimination in the human body to prevent expensive drug development failures at later stages.

Table 2: Techniques and their software

Docking	
Autodock4	It is a computational docking tool built using a quick Lamarckian genetic algorithm search technique and an empirical free energy force field. It is efficient for virtual screening and general-purpose docking of ligands to biomolecular targets, and it has specialized functions for covalent ligand complex prediction, ligands with flexible rings, explicit hydration, and metalloprotein target prediction.
AutoDockVina	A quick gradient-optimization conformational search and a straightforward scoring function serve as the foundation of our turnkey computational docking tool. For docking drug-like ligands to protein targets, it is quick and efficient.
AutoDockFR	It is a computer programme for docking proteins that has adaptable targets for side chain motion and induced fit.
AutoDockCrankPep	It is a program for computational docking of peptides to protein targets.
DOCK	Docking molecules
LIGPLOT	Program for automatically plotting protein-ligand interactions
SPROUT	de novo ligand design
STALK	A Molecular Docking System

Structure-based virtual screening	
DeepFrag	An Open-Source Browser App for Deep-Learning Lead Optimization (online)
e-Drug3D	Offers a facility to explore FDA approved drugs by for instance docking with PLANTS (online)
LINK - DecoyFinder	A graphical tool which helps finding sets of decoy molecules for a given group of active ligands - standalone
LigBuilder	It may automatically create ligand molecules within the binding pocket based on the three-dimensional structure of the target protein. - Solitary, standalone compound searching, ligand-based
Simtk	Simbios, the National NIH Center for Biomedical Computing, which focuses on Physics-based Simulation of Biological Structures, is the source of the software framework that was created there. There are numerous techniques for chemical docking, including ROCS-style tools, ADME Tox, Cytochrome P450 site of metabolism, etc.
RDock	A very fast open source small molecule protein docking programme that was initially developed by Vernalis - standalone
Autogrow	Ligand design using fragment-based growing, docking, and evolutionary techniques - standalone

Molecular dynamics

Molecular Modelling Toolkit	
3D-DOCK Suite	Includes – FT Dock which performs rigid-body docking between biomolecules; RPS core which uses a pair potentials to screen output from FT Dock; and Multi Dock which performs multiple copy side-chain refinement

Absorption, distribution, metabolism, elimination, and toxicity prediction	
PreADMET ADMET Prediction	Predict permeability for Caco-2 cell, MDCK cell, blood-brain barrier, human intestinal absorption, skin permeability, and plasma protein permeability predictions binding
PreADMET Toxicity	Prediction predict toxicological properties from chemical structures, such as Mutagenicity and carcinogenicity
Molinspiration	Calculation of Molecular Properties and Drug-likeness

2D quantitative structure–activity relationship
QSAR-Co, CORAL software
3D quantitative structure–activity relationship
3D-QSAR, Simplify Drug Design
Quantum mechanics
CADPAC, DMol3, Q-Chem
Homology modeling
RaptorX, Biskit, ESyPred3D, Molecular Operating Environment, SWISS-MODEL Local similarity/fragment assembly
Ligand-based virtual screening
LiSiCA, VSFlow
Pharmacophore modeling
HipHop, HypoGen, Pharmer, PHASE, GASP, PharmaGist, PharmMapper, MOE, Ligand Scout

VII. CONCLUSION

In the field of drug discovery and development, computer-aided drug design is a useful technology that allows for the rapid and cost-effective identification of the most promising therapeutic candidates.

With the advancements made to date by developing number of softwares for various techniques used in CADD process, computer-aided drug design has a bright future in helping to find many more cures.

VIII. ACKNOWLEDGMENT

All authors are very much thankful to Dr. C. D. Upasani, Principal of SNJBs Shriman Sureshdada Jain College of Pharmacy, Chandwad, Nashik for constant guidance and motivation.

Conflict of Interest:

There is no any conflict of interest.

REFERENCES

[1] Veselovsk and Ivanov Strategy of computer aided drug design. Russian academy of medical science. 2003; 3(1): 33-40.

[2] Yun T, Weiliang Z, Kaixian C, Hualiang J. New technologies in computer-aided drug design: Toward target identification and new chemical entity discovery. Drug discovery today technologies. 2006; 3(3): 307-313.

[3] Pranita PK, Madhavi MM, Rishikesh VA, Rajesh JO, Sandip SK. Computer-aided drug design: an innovative tool for modeling. Journal of medicinal chemistry. 2012; 2: 139-148.

[4] Pugazhendhi D, Umamaheswari TS; In silico methods in drug discovery–a review; International journal of advanced research in computer science and software engineering; 2013; 3(5); 680-683.

[5] Choudhary L.K., Shukla A., Zade S., Charde R; B(C.A.D.D.) - a new modern software-based approach in drug design and discovery; International journal of pharmaceutical chemistry; 2011; 1(1): 10-20.

[6] Vyas VK, Ukawala RD, Ghate M, Chintha C. Homology modeling a fast tool for drug discovery: current perspectives; Indian journal of pharmaceutical sciences; 2012; 74(1): 1–17.

[7] Roman AL, Mark BS; LigPlot+: Multiple ligandprotein interaction diagrams for drug discovery; Journal of chemical information and modeling; 2011; 51; 2778-2786.

[8] Alejandra HS, Aldo YT, Victor A, Héctor VC, Claudia MB; Protein-protein and protein-ligand docking; Intech open science and open minds; 2013; Chapter 3; 64-81.

[9] Gaba M, Gaba P, Singh S, Gupta GD; An overview on molecular docking; International Journal of Drug Development and Research; 2010; 2(2): 219-231.

[10] Udugade BV, Gawade SP; 3D QSAR and pharmacophore modeling on substituted cyanopyrrolidines as type ii anti-diabetic agent's potential dipeptidyl peptidase-iv inhibitors; Pharmacophore; 2016; 7(5): 342-348.

[11] David RK, Carlos JC; ZINC Pharmer: pharmacophore search of the ZINC database; Nucleic acids research; 2012; 40: 409-414.

[12] Reddy AS, Pati SP, Kumar PP, Pradeep HN, Sastry GN; Virtual screening in drug discovery-A computational perspective; Current protein and peptide science; 2007; 8:329-351.

[13] Chanin N, Chartchalerm INA, Thanakorn N, Virapong PA; Practical overview of quantitative structure-activity relationship; EXCLI Journal; 2009: 74-88.

[14] Denis F, Eugene M, Alexander T; Trust but verify: On the importance of chemical structure curation in cheminformatics and QSAR modeling research; Journal of Chemical Information and Modeling; 2010; 50(7): 1189–1204.

[15] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ; Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings; Advanced drug delivery reviews; 2001; 46(1-3): 3–26

[16] Lipinski CA; Lead- and drug-like compounds: the rule-of-five revolution; Drug discovery today: Technologies; 2004; 1(4): 337–341.

[17] Leo A, Hansch C, Elkins D; Partition coefficients and their uses; Chemical reviews; 1971; 71(6): 525–616.

[18] Ghose AK, Viswanadhan VN, Wendoloski JJ. A knowledge-based approach in designing Combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases; Journal of combinatorial chemistry; 1999; 1(1); 55–68.

[19] Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD; Molecular properties that influence the oral bioavailability of drug candidates; Journal of medicinal chemistry; 2002; 45(12): 2615–2623.

[20] Congreve M, Carr R, Murray C, Jhoti H; A 'rule of three' for fragment-based lead discovery? Drug discovery today; 2003; 8(19); 876–877.

[21] Oprea TI, Davis AM, Teague SJ, Leeson PD; Is there a difference between leads and drugs? A historical perspective; Journal of chemical information and modeling; 2001; 41(5); 1308–1315

[22] Leeson PD, Springthorpe B; The influence of drug-like concepts on decision-making in Medicinal chemistry; Nature reviews drug discovery; 2007; 6(11); 881–890.

[23] RamaKrishna K, Rao CVK, Rao RS; ChemoInformatics Part I: Molecular descriptors in omnimetrics research; Journal of applicable chemistry; 2015; 4(4); 1024-1144.

[24] Marialke J, Korner R, Tietze S, Apostolakis J (2007) Graph-based molecular alignment (GMA). Journal of chemical information and modeling 47 (2), 591-601.

[25] Vainio MJ, Johnson MS (2007) Generating conformer ensembles using a multi objective genetic algorithm. Journal of chemical information and modeling 47 (6), 2462-74.

[26] Shelley JC, Cholleti A, Frye LL, Greenwood JR, Timlin MR, Uchimaya M (2007) Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. J Comput Aided Mol Des 21 (12), 681-91.

[27] Stroganov OV, Novikov FN, Stroylov VS, Kulkov V, Chilov GG (2008) Lead finder: an approach to improve accuracy of protein-ligand docking, binding energy estimation, and virtual screening. Journal of chemical information and modeling 48 (12), 2371-85.

[28] Marialke J, Tietze S, Apostolakis J (2008) Similarity based docking. Journal of chemical information and modeling 48 (1), 186-96.

[29] Gorelik B, Goldblum A (2008) High quality binding modes in docking ligands to proteins. Proteins 71 (3), 1373-86.

[30] Goto J, Kataoka R, Muta H, Hirayama N (2008) ASEDock-docking based on alpha spheres and excluded volumes. Journal of chemical information and modeling 48 (3), 583-90.

[31] Singh S (2008) India takes an open-source approach to drug discovery. Cell 133 (2), 201-3.

[32] Amrute, B.B.; Amrutkar, R.D.; Tambe, S.R. Design and molecular docking studies of some 2, 3 di-substituted quinazolin-4-one analogues against Staphylococcus aureus UDG. Curr. Comput-Aid. Drug Des., 2019, 15, 1.

[33] Brylinski M, Skolnick J (2008) Q-Dock: Low-resolution flexible ligand docking with pocket-specific threading restraints. Journal of computational chemistry 29 (10), 1574-1588.

[34] Taminau, J.; Thijs, G.; De Winter, H., Pharao: pharmacophore alignment and optimization. Journal of molecular graphics & modelling 2008, 27 (2), 161-9.

[35] Irwin JJ, Shoichet BK, Mysinger MM, Huang N, Colizzi F, Wassam P, Cao Y (2009) Automated docking screens: a feasibility study. J Med Chem 52 (18), 5712-20.

[36] Bikadi Z, Hazai E (2009) Application of the PM6 semi-empirical method to modeling proteins enhances docking accuracy of AutoDock. Journal of cheminformatics 1, 15.

[37] Liu Y., Zhang Y., Zhong H., Jiang Y., Li Z., Zeng G., Chen M., Shao B., Liu Z., Liu Y. Application of molecular docking for the degradation of organic pollutants in the environmental remediation: A review. Chemosphere. 2018; 203:139–150. 52.

[38] Morris G.M., Trott O., Olson A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 2009; 28:455–461. 54..

[39] De Magalhães C.S., Almeida D.M., Barbosa H.J.C., Dardenne. De Magalhães C.S., Barbosa H.J.C., Dardenne L.E. Selection-Insertion Schemes in Genetic Algorithms for the Flexible Ligand Docking Problem. Lect. Notes Comput. Sci. 2004; 3102:368–379.

[40] Jones G., Willett P., Glen R.C., Leach A.R., Taylor R., Uk K.B.R. Development and Validation of a Genetic Algorithm for Flexible Docking. J. Mol. Biol. 1997; 267:727–748.