IJPSR (2018), Volume 9, Issue 10 (Review Article)

INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES AND RESEARCH

Received on 16 January, 2018; received in revised form, 09 May, 2018; accepted, 13 May, 2018; published 01 October, 2018

# COMPUTATIONAL METHODS IN DRUG DISCOVERY

Shruti Gupta, Aparajita Basu, Anamika Jaiswal and Nidhi Mishra *

Department of Applied Sciences, Indian Institute of Information Technology Allahabad, Devghat, Jhalwa, Allahabad - 211015, Uttar Pradesh, India.

**Keywords:**

Drug discovery, Computational methods, High-throughput screening

**Correspondence to Author:**
**Dr. Nidhi Mishra**

Assistant Professor, Department of Applied Sciences, Indian Institute of Information Technology Allahabad, Devghat, Jhalwa, Allahabad - 211015, Uttar Pradesh, India.

**E-mail:** nidhimishra@iiita.ac.in

**ABSTRACT:** The efficiency of drug discovery and designing process can be increased by effective strategies given by computational methods. Drug designing is an intense, time consuming and an interdisciplinary venture. Traditionally, drugs were discovered by synthesizing compounds in a long-drawn-out and multi-step process. Recent studies have shown an increase in day by day demand of novel and more efficient drugs. Due to availability of low-cost computer power, the use of such computers has become a leading topic in medicinal chemistry. In search of potent drugs computational techniques like docking, homology modelling, pharmacophore modelling, are employed by researchers around the globe in order to achieve the goal. The main purpose of this document is to give a summary of drug design process and specifically the role of computational modelling techniques. These involve some techniques from binding sites prediction to high throughput screening of large compound libraries.

**INTRODUCTION:** Drugs are crucial for the prevention and alleviation of disease. On October 5, 1981 the cover article - "Next Industrial Revolution: Designing drugs by computer at Merck" by Van Drie, 2007 was published by Fortune magazine. It's acknowledged for its potential in computer aided drug design. Although advancements were continuing in CADD, the capability for high-throughput screening (HTS) as a way for finding novel therapeutics, was gaining its epitome [1]. Drug design, is the technique of finding new medicines with the help of natural target. The entire process of discovering novel drug designing is said to be an arduous, risky and an exorbitant task.

QUICK RESPONSE CODE

DOI: 10.13040/IJPSR.0975-8232.9(10).4102-08

Article can be accessed online on: www.ijpsr.com

DOI link: http://dx.doi.org/10.13040/IJPSR.0975-8232.9(10).4102-08

The novel drug discovery and development cycle, from concept to production, takes around 14 years. Therefore, to reduce the risk of failure, and to shorten the research cycle for drug discovery, various approaches have been developed. CADD or computer-aided drug design is that one powerful technique used for reaching these goals. Computational methods now replace time consuming process of drug discovery and designing via traditional methods. CADD is generally helpful in three major aspects:

1. Filtering large libraries of compounds into smaller more active sets of compounds.
2. Oversee lead optimization of compounds by checking ADMET - Absorption, distribution, metabolism, excretion, potential for toxicity [2].
3. Designing new compounds.

To identify drug targets *in-silico*, bioinformatics tools are used. These tools are also used to study the related structures of target molecules for possible binding or active sites, then further

produce potential candidate molecules, check for ADMET properties and drug likeness, and carry out docking with target molecules and rank molecules according to their binding energy. One advantage of computational tools is that it efficiently delivers new drug candidates at a higher speed with lower cost. Now-a-days, in the review of nature of interactions like drug-nucleic, drug-protein, and enzyme-substrate interactions a specific potent lead molecule against any particular disease can be designed. With the involvement of different fields of science like, chemistry, pharmacology, molecular biology, computational methods have now become an interdisciplinary science. Some noteworthy methodologies based on computational drug designing have been developed which involve target recognition [3], Virtual high throughput screening (VHTS), QSAR, fragment based screening, virtual library design. This review provides succinct overview of some of these methodologies.

**Docking:** Docking is the computational method to determine the binding affinity between molecules (mainly protein structure and ligand). Process of docking is shown in **Fig. 1**, in which the ligand is the functional group binding to a biomolecule (target) to form a complex structure.

CADD can be divided into receptor based and ligand based. The former uses the structure of the target protein while the latter uses the knowledge of known inhibitors.

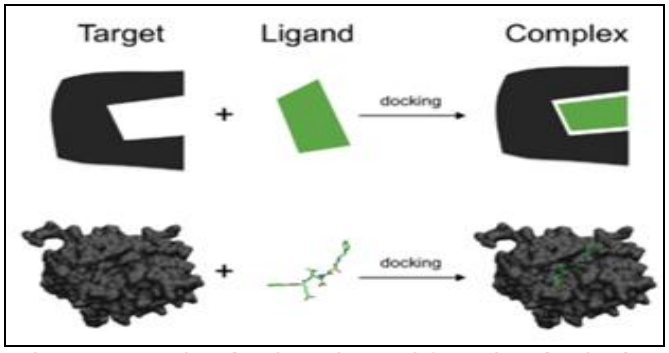


**FIG. 1: ILLUSTRATION OF DOCKING: A SMALL MOLECULE LIGAND (GREEN) TO A PROTEIN TARGET (BLACK) PRODUCING A STABLE COMPLEX [22]**

**A. Structure Based:** Structure-based computer-aided drug design or SB-CADD depends on the knowledge of the targeted protein structure and inculcate the calculated interaction energies of the tested compounds. It is based on the technique of using 3D structure of target receptors to search for the probable candidate compounds which can further alter their target function. The concept behind this approach is that a molecule's tendency to interact with a specific protein and express a desired biological effect is modified based on its ability to interact with a particular binding site of the protein [3].

Molecules that have common favourable interactions tend to exploit similar biological effects. The compounds are then ranked using an appropriate scoring function. Therefore, novel compounds can be explained by carefully analysing the protein's binding site.

**B. Ligand Based:** The LBCADD ligand based computer-aided drug discovery approach exploits the analysis of ligands that have previously interacted with a target of interest. These techniques use various reference structures that were collected from compounds which interact with the target of interest and analyse their 2-dimensional or 3-dimensional structures. The aim is to present these compounds in such a way that the properties, mainly physiochemical, which are crucial for the required interactions are retained while the irrelevant information is discarded. The techniques employed in LB-CADD use various methods to describe attributes of small molecules with the help of computational algorithms, which balances information content and efficiency. The optimal descriptor set relies on the biological functions predicted as well as on the LB-CADD technique; therefore, several algorithms for deriving chemical information have been developed [1].

**Virtual High-Throughput Screening:** Virtual high-throughput screening or vHTS is a technique done with the help of computations, in which the *in-silico* compound libraries are screened so that the binding affinity of the target receptors with the library compounds is checked and analysed, explained in **Fig. 2**. It allows the researchers to focus on the compounds which have high binding affinity and this is done by the screening of virtual compound libraries. These screening techniques rejects the compounds which have low affinity, therefore limiting the expenditure of time and other resources.

Two categories have been described for vHTS techniques:

- **Ligand Based:** Ligand-based technique is basically based on pharmacophore modelling where the model of the receptor is derived from the information contained in the ligands. Another technique which is used is - 2D chemical similarity analysis method which is used to scan the database against one or two ligands [4].

- **Structure Based:** Structure-based virtual high-throughput screening, SB-vHTS, is an *in-silico* method which stabilises on a comparison between the 3-Dimensional structure of the molecule with the repetitive binding pocket. This helps in selection for the ligand, a particular binding site. To simplify things, SB-vHTS takes into account the narrow conformational sampling of ligand and protein with an approximation of binding energy, and that can be computed. The problems which arise due theses approximations can be removed by clustering of ligand poses and iterative docking. Preparation of SB-vHTS is done by the following steps [1].

1. Preparing the target protein and compound library for the purpose of docking.
2. Determination of the most favourable binding pose for each compound.
3. Sorting the docked structure according to their ranks.

Although vHTS techniques are cost effective but they require precise compound and target library, optimal parameters and carefully analysed results [4].

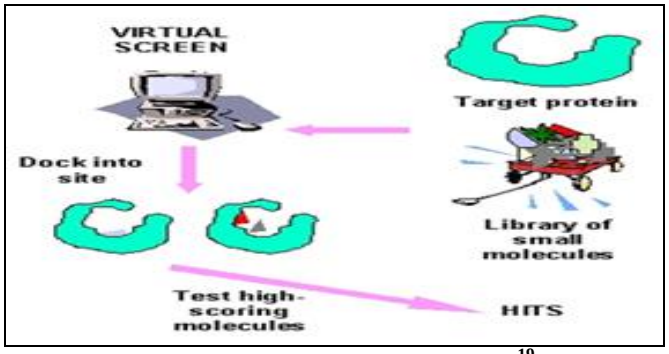


**FIG. 2: VIRTUAL SCREENING [19]**

**QSAR (Quantitative Structure Activity Relationships):** As mentioned, it is critical to know about the geometrical structure of the target protein and the ligand in order to implement molecular docking techniques. QSAR is a method which can be used even if the structure is unknown.

QSAR predicts by experimentation of how the protein interacts with the tested molecules. It may also be known about how the protein undergoing the study reflects positive activity with one group of compounds and negative with another [5].

Metaphorically, we have no idea how the lock appears to be. To build a QSAR model a set of descriptors are chosen, they sway the binding success of the compound with the target. The classical descriptors are variables such as - molecular volume, molecular weight, thermo dynamical and electrical properties. The QSAR model is mainly utilized for virtual screening of compounds to explore the listed drug candidate's descriptors for the target.

**3D QSAR:** As shown in **Fig. 3**, it is a natural extension to the classical approach of Hansch and Free Wilson methodologies, which uses 3D properties of the ligands to speculate their biological activities *via* tough chemometric techniques, for example - ANN, PLS or G/PLS. It has been used as a predictive tool for designing of agrochemicals and pharmaceuticals [6].

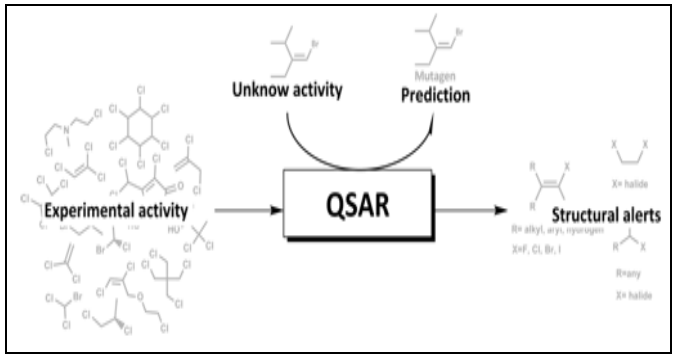


**FIG. 3: PREDICTION STUDIES OF BIOLOGICAL ACTIVITY OF CHEMICAL SUBSTANCES BY QSAR METHODS [21]**

CoMFA - comparative field molecular analysis is a 3 dimensional QSAR technique that places molecules and extricates aligned features which can be used to figure out the biological activity. The technique focuses on the alignment of molecular interaction fields as a whole instead of each individual atomic feature.

**Pharmacophore Mapping:** 'Pharmacophore' was first presented by Paul Ehrlich in 1909, who defined the term pharmacophore as 'a molecular framework that carries - phoros, the vital feature required for biological activity of a drug - pharmacon [7]. Peter Gund modified this definition in 1977 to a set of structural features in a molecule that is responsible for that molecule's biological activity.[15] Today, in computational chemistry, a pharmacophore is defined as the important feature of one or more molecules having same biological activity [8].

Pharmacophore modelling is a useful technique which classifies a group of ligands or molecules into active or inactive compounds. It is extensively used to identify new compounds when compared to drug targets [9]. Commercially available software catalyst has a module named HipHop which is used to generate and develop pharmacophores. The objective of pharmacophore mapping is to determine bioactive conformations of the ligand and in order to superimpose the mapping; one needs structural activity relationships of confor-mationally informative and structurally diverse molecules.

Similar to QSAR models, pharmacophores too can be constructed with no prior knowledge of the structure of target by extracting potentially wanted characteristics from the compounds which interact with the target. After the development of pharmacophore model, compound libraries can be searched for potential drug candidates [10].

**Structure Based Pharmacophore Modelling:** It works directly with the 3 dimensional structure of macromolecular target. The method includes analysis of the spatial relationships and the complementary chemical properties of the active site. This method is further classified in two categories

1. Macromolecule ligand - complex based and
2. Macromolecule (without ligand) based

The only drawback of this approach is that it can't be applied to cases when the compounds targeting the binding site of interest are not known. But macromolecule-based approach overcomes this disadvantage.

**Ligand-Based Pharmacophore Modelling:** When there is no macromolecular target structure present, Ligand - based pharmacophore modelling has become an important computational strategy for aiding drug designing. In this method, from a set of 3D structures of known ligands, common chemical features are procured. Pharmacophore development from several ligands involves two main steps:

- ✓ To indicate conformational flexibility of ligands, creating conformational space for each ligand.
- ✓ Arranging the multiple ligands in a training set and finding common chemical attributes to develop pharmacophore models [15].

**Pharmacophore Mapping Software:**

**Ligand Scout:** Ligand scout functions by taking a macromolecular structure, comprising of a bound ligand and identifies the primary features on the ligand which interact with points on the protein. The defined shape of the active site is naturally detected by the software and is combined into the pharmacophore. The software allows alignment of either a set of pharmacophores or a set of ligand molecules. Pharmacophores can also be created in the absence of a protein structure from the ligands.

**Discovery Studio Visualizer Accelrys:** It is a software company with its headquarters in the United States, with representation in Japan and Europe. It provides a platform for chemical research, especially in the areas of material science and drug discovery. DS ActiveX Control is a free plug-in that provides 3D visualization of small molecules, proteins, nucleic acid crystal structures and pharmacophore models.

**Fragment Based Screening:** Fragment based screening emerged as an effective and promising method in identification of best suitable hit which have higher binding affinity for target of interest, **Fig. 4** [11]. Fragment based screening gives out low molecular weight fragments and therefore it binds with weaker affinity with the target molecule. With good drug like properties, fragment has to be large and hydrophobic and follow Lipinski's rule. Hit to lead optimization would be performed based on the higher affinity interaction of ligand with the target molecule. The fragments which have weaker affinity for binding, their target need to be performed biophysical techniques like X-ray

crystallography, SPR - surface plasmon resonance or NMR to identify hits [16].

Fragment atoms should have elevated binding energy per unit molecular mass so that it can bind with higher efficacy. Though high throughput screening fragments are large but because of its lower molecular weight spaces, they are less potent and are less efficient binders. This gives fragment based screening obvious advantages over high throughput screening [12].

For lead optimization, fragment should have lower molecular mass, lower lipophilicity (cLogP), fewer hydrogen-bond acceptors and fewer aromatic rings [14]. For lead identification if two fragments are binding to separate binding sites then sum of free energies of binding molecules must be equal to the free binding energy of each fragments [11, 17].

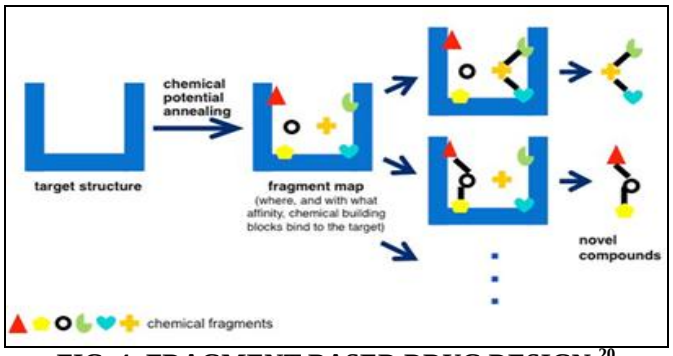


**FIG. 4: FRAGMENT BASED DRUG DESIGN [20]**

**Prediction and Optimization of Pharmaco-kinetics Properties of Drug (ADMET):** In the lead discovery and drug development, pharmacokinetic properties of drug like absorption, bioavailability, distribution to tissues, metabolism, excretion and toxicity is tested for lead optimization of best suited drug used as therapeutics [12].

Early computational approach for ADMET profile is by filtering database of compound using Lipinski's rule of five:

1. No more than 5 H-hydrogen bond donors.
2. No more than 10 H-hydrogen bond accepting sites (N, nitrogen or O, oxygen atoms).
3. The molecular mass should be less than 500 Daltons.
4. An octanol-water partition coefficient log P (ClogP) should not be more than 5.

Before any kind of screening is performed, this is a simple way of ADMET filtering. But this criterion is not hard set filter. There are examples of successful drugs that fail one or more criteria of Lipinski's rule.

Another estimation of drug-likeness is based on scoring of different molecular descriptor which includes molecular weight, H-bond acceptors and donors, rotatable bonds, log P, aromatic rings and the number of structural alerts.

**A. Profiling of Lead Compound:** The same computational tool is used to predict activity for ADMET profile which includes solubility, metabolism, membrane permeability, interaction with transcription proteins, and different aspects of toxicity [1].

**B. Lead optimization for Metabolism:** Metabolic profile of drug is important for the bioavailability, half-life and detection of harmful metabolites of drugs. Although computational tools are helpful in predicting the changes in the lead structure based on the target affinity, they can be used in parallel for interaction of target with metabolizing enzymes.

Two important phases of drug Metabolism are as follows:

**Phase I:** It includes hydrolases, oxidases, esterases and cytochrome p450 enzymes.

**Phase II:** It includes sulfation, methylation, acetylation and glutathione conjugation.

These reactions at the site of binding affinity of drug with lead compounds sometimes results in toxicity with release of toxic products like electrophile or free radicals [13].

Cyp450 enzyme is important metabolite because it is present in all domains of life as hemoprotein or membrane associated proteins in humans and it binds with drug for performing an enzymatic reaction. Docking studies predict the binding poses for specific sites of metabolism with the compound. Models of p450 binding sites orientation of amino acid side chain are developed by Monte Carlo and stochastic simulation or GOLD, FlexX, DOCK,

AutoDock are computational techniques which are used for predication of p450 structure.

SMARTCyp is another method used for determining the binding sites of ligand which are best suitable for CYP450 metabolism based on the Activation Energy calculated by the Quantum Mechanical Model.

**C. Lead Optimization for Absorption and Permeability:** Absorption of drug in the body requires 3 processes either through-

1. Transcellular - intramembrane or through cytoplasm or
2. Paracellular - *via* aqueous pore or small intestine.
3. Intestinal Cell lines - like Caco-2, MDCK or PAMPA.

Caco-2 cells resemble intestinal epithelial cell line in the form of a polarized monolayer, well-defined brush border on their apical surfaces and intercellular junctions. It promotes transport in direction from the basolateral to apical region and vice-versa. The P-glycoprotein (P-gp) inhibitor keeps check on active transport mediated by P-gp.

MDCK cells resemble epithelial cell line of canine kidney region. MDCK cells are generally transfected with transporter proteins to carry out drug efflux example being MDR1-MDCK cell line which is a useful model for the identification of P-gp inhibitors and substrates.

Other dimensions of a drug's ADMET profile that are speculated with computational tools includes permeability of membrane, which counts for a large part of bioavailability along with the distribution and penetration of the BBB and blood plasma protein-binding, implicated in drug's distribution and effective plasma concentrations. The development process of Computational modelling of ADMET prevents a probable blood pressure-lowering drug to lose in early state. However, the pharmacokinetic model cleared that this compound would actually have higher potency than the others that showed greater efficacy [18].

**CONCLUSION:** The computational techniques used in drug designing, are cost efficient, less time-consuming proving that, they are superior to the traditional methods of drug designing. It is a vast technology which uses *in-silico* models to increase the efficiency of the technique. During past years, there has been a significant progress in the development and application of novel method-logies due to advancement in genomic and molecular biology techniques. These methods go through many variability of drug discovery, starting from hit screening to lead optimization to drug target identification and ADMET assessment. Chemical leads are drug-like molecules, which modulate the function of a target protein and are optimized to act as a drug to fight against diseases.

**ACKNOWLEDGEMENT:** We thank our colleagues from Indian Institute of Information Technology, Allahabad who provided insight and expertise that greatly assisted in this article.

**CONFLICT OF INTEREST:** The authors have no conflicts of interest that are directly relevant to the content of this manuscript.

## REFERENCES:

1. Sliwoski G and Kothiwale SK: Computational Methods in Drug Discovery, Pharmacol Rev. 2014; 66(1): 334-395.
2. Fourches D, Williams AJ, Patlewicz G, Shah I, Grulke C, Wambaugh J, Ann R and Alexander T: Computational Tools for ADMET Profiling. 2018 John Wiley & Sons, Inc. January 2018.
3. Ekins Sean: Computational Toxicology: Risk Assessment for Chemicals. 2018 John Wiley & Sons, Inc. January 2018
4. Yuriev E and Holien J: Improvements, trends, and new ideas in molecular docking: 2012-2013 in review. Journal of Molecular Recognition 2015. doi.org/10.1002/jmr.2471.
5. Pugazhendhi D and Umamaheswari TS: *In silico* methods in drug discovery. International Journal of Advanced Research in Computer Science and Software Engineering. 2013; 3(5): 128X.
6. Virtual Screening 2018. Retrieved from
7. https://en.wikipedia.org/wiki/Virtual_screening
8. Subramaniam S, Mehrotra M and Gupta D: Virtual high throughput screening (vHTS) – A perspective. Bioinformation 2008; 3(1): 14-7.
9. Szostak JW: Introduction: Combinatorial Chemistry. Chemical Reviews 1997; 97(2): 347-348.
10. Verma J and Khedkar VM: 3D-QSAR in drug design. Curr Top Med Chem. 2010; 10(1): 95-115.
11. Katsila T, Georgios A and Spyroulias: Computational approaches in target identification and drug discovery. Computational and Structural Biotechnology Journal. 2016. doi.org/10.1016/j.csbj.2016.04.004.
12. Joseph-McCarthy D and Campbell AJ: Fragment-based lead discovery and design. J. Chem. Inf. Model 2014; 54(3): 693-704.
13. Yang SY: Pharmacophore modelling and applications in drug discovery: challenges and recent advances. Drug Discovery Today 2010; 15: 11-12.

14. Scotti T and Marcus: Theoretical Studies of the metabolism in drug discovery. Current Drug Metabolism, 2017; 18(6): 498-499. doi.org/10.2174/1389200218061707 21113611.
15. Wang, Tao, Wu and Bin M: Fragment-based drug discovery and molecular docking in drug design. Current Pharmaceutical Biotechnology, Volume 16, Number 1, January 2015; 15: 11-25.
16. Daisy P and Singh SK: A database for the predicted pharmacophoric features of medicinal compounds. Bioinformation 2011; 6(4): 167-168.
17. Kumar A and Voet A: Fragment based drug design: from experimental to computational approaches. Curr Med Chem. 2012; 19(30): 5128-47.
18. Rees DC and Congreve M: Fragment-based lead discovery. Nature Reviews 2004; 3(8): 660-72.
19. Tetko IV: ADMETox data, models and challenges. eADMET the reference in chemoinformatics. Kazan, Russia 2013.
20. Discovery of Novel Inhibitors by Docking: Shoichet Laboratory. University of California San Francisco.
21. Prediction studies of the biological activity of chemical substances by QSAR methods: Bioinformatics and High Performance Computing Research Group. BIO-HPC Website.
22. Pompliano D: Fragment-Based Drug Design: Delivering Drugs That Hit Multiple targets, Leveraging Insights From Systems Biology. Drug Development and Delivery 2012.
23. Lim H, Woong-Hee S, Lee MS and Chaok S: GalaxySite: ligand-binding-site prediction by using molecular docking. Nucleic Acids Research, NCBI. 2014 PMID: 24753427.

**How to cite this article:**
Gupta S, Basu A, Jaiswal A and Mishra N: Computational methods in drug discovery. Int J Pharm Sci & Res 2018; 9(10): 4102-08. doi: 10.13040/IJPSR.0975-8232.9(10).4102-08.