



Insilico Methods in Drug Discovery - A Review

D. Pugazhendhi

Asst. Prof., PG & Research Dept. of Computer Science
Quaid-e-millath College, Chennai, India

T.S. Umamaheswari

Research Scholar, Bharathiyar University
Coimbatore, India

Abstract— Drug discovery and development is an intense, lengthy and an interdisciplinary endeavour. Drug discovery is mostly portrayed as a linear, consecutive process that starts with target and lead discovery, followed by lead optimization and pre-clinical *in vitro* and *in vivo* studies to determine if such compounds satisfy a number of pre-set criteria for initiating clinical development. Traditionally, drugs were discovered by synthesizing compounds in a time-consuming multi-step processes against a battery of *in vivo* biological screens and further investigating the promising candidates for their pharmacokinetic properties, metabolism and potential toxicity. Sophisticated insilico approaches has given a tremendous opportunity to pharmaceutical companies to identify new potential drug targets which in turn affect the success and time of performing clinical trials for discovering new drug targets. The main goal of this work is to review insilico methods for drug discovery process with emphasis on identifying drug targets.

Key words: Insilico, drug discovery, drug design, bioinformatics, insilico methods.

I. INTRODUCTION

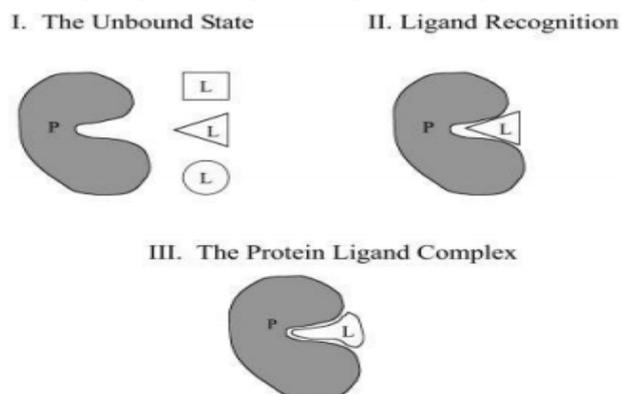
In silico methods can help in identifying drug targets via bioinformatics tools. They can also be used to analyze the target structures for possible binding/active sites, generate candidate molecules, check for their drug likeness, dock these molecules with the target, rank them according to their binding affinities, further optimize the molecules to improve binding characteristics. The use of computers and computational methods permeates all aspects of drug discovery today and forms the core of structure-based drug design. High-performance computing, data management software and internet are facilitating the access of huge amount of data generated and transforming the massive complex biological data into workable knowledge in modern day drug discovery process. The use of complementary experimental and informatics techniques increases the chance of success in many stages of the discovery process, from the identification of novel targets and elucidation of their functions to the discovery and development of lead compounds with desired properties [1]. Computational tools offer the advantage of delivering new drug candidates more quickly and at a lower cost. This review provides succinct overview of several methods in drug discovery by using insilico approaches on identifying drug targets where there are genes or proteins associated with specific diseases.

II. INSILICO METHODS IN DRUG DISCOVERY

There are five insilico methods in drug discovery. They are Molecular docking, Virtual High through put screening, QSAR (*Quantitative* structure-activity relationship), Pharmacophore mapping, Fragment based screening and which are discussed below.

A. Molecular docking

Docking is the computational determination of binding affinity between molecules (protein structure and ligand). Given a protein and a ligand find out the binding free energy of the complex formed by docking them. Figure 1 shows the stages of High throughput docking for protein –ligand complex binding.



Diller, D. J. and Merz, K. M. (2001), High throughput docking for library design and library prioritization. *Proteins*, 43: 113–124.

Fig. 1 A High throughput docking for Protein and Ligand complex

Docking or Computer aided drug designing can be broadly classified as “Receptor based methods” which make use of the structure of the target protein and “Ligand based methods” which is based on the known inhibitors

- 1) *Receptor based methods*: Uses the 3D structure of the target receptor to search for the potential candidate compounds that can modulate the target function. These involve molecular docking of each compound in the chemical database into the binding site of the target and predicting the electrostatic fit between them. The compounds are ranked using an appropriate scoring function such that the scores correlate with the binding affinity. Receptor based method has been successfully applied in many targets
- 2) *Ligand based methods*: In the absence of the structural information of the target, ligand based method make use of the information provided by known inhibitors for the target receptor. Structures similar to the known inhibitors are identified from chemical databases by variety of methods, some of the methods widely used are similarity and substructure searching, pharmacophore matching or 3D shape matching. Numerous successful applications of ligand based methods have been reported. Figure 2 shows the strategy is used for search the similar compounds

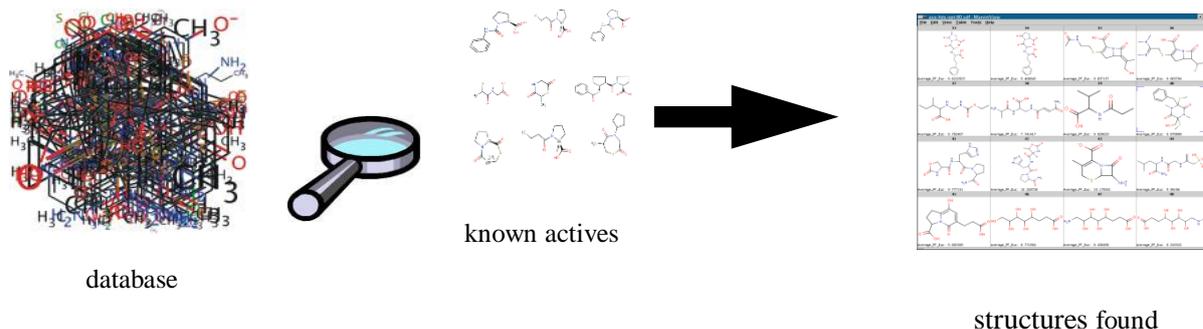


Fig.2 search for similar compounds

B. Virtual High Throughput Screening

Virtual screening is a computational method where large libraries of compounds are assessed for their potential to bind specific sites on target molecules such as proteins, and well-matched compounds tested. **Virtual screening (VS)** is a computational technique used in drug discovery research. By using computers, it deals with the quick search of large libraries of chemical structures in order to identify those structures which are most likely to bind to a drug target, typically a protein receptor or enzyme [2]. Virtual screening has become an integral part of the drug discovery process. Related to the more general and long pursued concept of database searching, the term "virtual screening" is relatively new. Walters, *et al.* define virtual screening as "automatically evaluating very large libraries of compounds" using computer program [3]. As this definition suggests, VS has largely been a numbers game focusing on questions like how can we filter down the enormous chemical space of over 10^{60} conceivable compounds [4] to a manageable number that can be synthesized, purchased, and tested. Although filtering the entire chemical universe might be a fascinating question, more practical VS scenarios focus on designing and optimizing targeted combinatorial libraries and enriching libraries of available compounds from in-house compound repositories or vendor offerings. It is less expensive than High Throughput Screening, Faster than conventional screening, scanning a large number of potential drugs like molecules in very less time. HTS itself is a trial and error approach but can be better complemented by virtual screening. Figure 3 shows the work flow of Virtual High Throughput Screening.

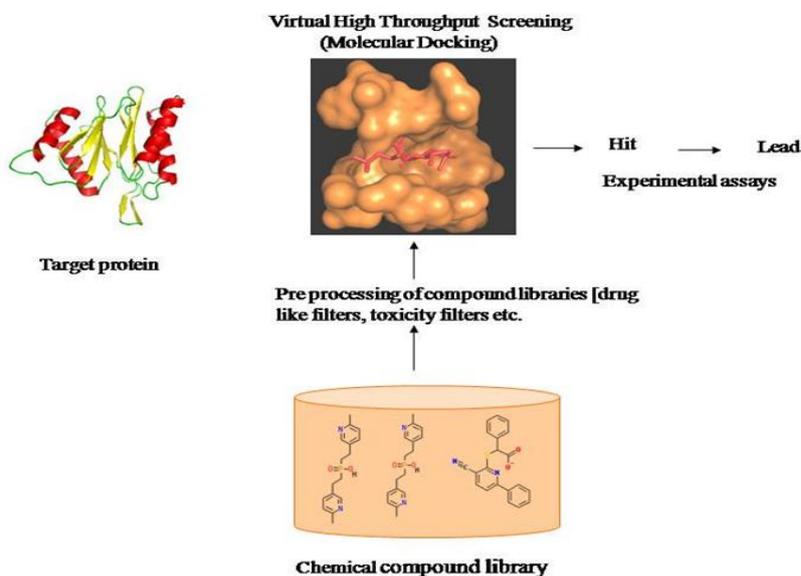


Fig 3. Virtual High Throughput Screening work flow

C. QSAR(Quantitative structure-activity relationship)

QSAR is statistical approach that attempts to relate physical and chemical properties of molecules to their biological activities. The aim of QSAR is the prediction of molecular properties from their structure without the need to perform the experiment using invitro or invivo. It saves times and resources [5]. Various descriptors like molecular weight, number of rotatable bonds LogP etc. are commonly used. Many QSAR approaches are in practice based on the data dimensions. It ranges from 1D QSAR to 6D QSAR. The methods called quantitative structure-activity relationship (QSAR) are based on the assumption that the activity of a certain chemical compound is related to its structure. More precisely, this approach says that the activity, or the property, for instance the toxic effect, is related to the chemical structure through a certain mathematical algorithm, or rule. It is considered the presence of a particular characteristic in the chemical compound, which is present or not in the structure. For instance, it is well known that if in the chemical compound there are certain groups, like an aromatic amine, or an epoxide, there is a higher probability that the chemical compound is genotoxic. Figure 4 shows the typical way to derive QSAR model.

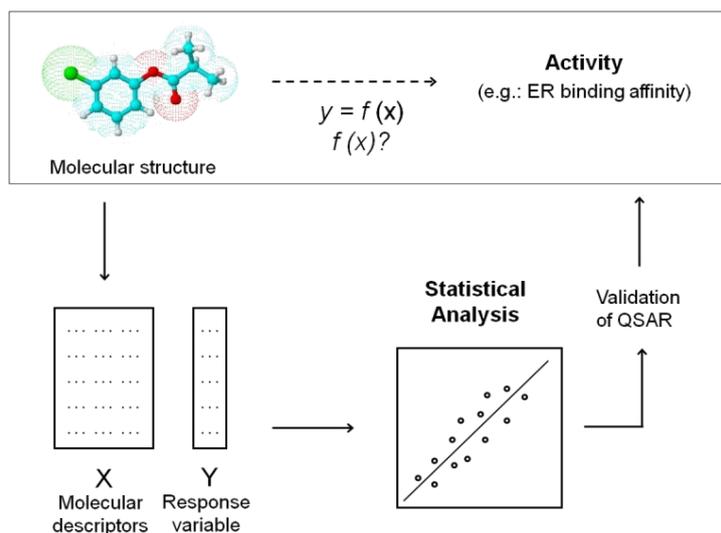


Fig 4. Derivation of QSAR model

The basic assumption is that there is a mathematical function of the chemical properties which is related to the effect. Thus, the effect called y is a function called f of the chemical properties, called x . Mathematically, $y = f(x)$. Typically, use a number of chemical compounds with known values of the toxic effect (y). For each chemical compound calculate a series of parameters, called chemical descriptors. Then to find an algorithm that provides a quite accurate value, similar to the real experimental value. The final step is to check if the so-obtained algorithm is capable to predict the property values for other chemicals, not used to build up the model. This last phase is called validation of the QSAR model. This last phase is very important. Indeed, it is very important to generate a model which is working not only for the chemical substances used within the training set, but also for other chemicals. The challenge is to define the correct statistical properties of the model [6]. Sometimes it is said that QSAR models represent a way for industry to spend less for toxicological research, or can be used to save animals to be used for experiments. The real challenge is not to identify the best method to protect human beings and environment. The challenge is to take advantage of all the contributions that each approach, in vivo, in vitro, and insilico offer.

D. Pharmacophore mapping

It is the process of deriving a 3D pharmacophore. A pharmacophore is a set of features together with their relative spatial orientation that are thought to be capable of interaction with a *particular* biological target such as Hydrogen bond donors and acceptors, positively and negatively charged groups, hydrophobic regions and aromatic rings. It depends on atomic properties rather than element types, it does not depend on specific chemical connectivity. Pharmacophore mapping is developed by specifying the nature of the key pharmacophoric features and the 3D distance map among all the key features. It has conformational flexibility and mapping the different combinations of pharmacophoric groups in the molecule. A Pharmacophore map can be generated by superposition of active compounds to identify their common features. Based on the pharmacophore map either de novo design or 3D database searching can be carried out. Frequently small molecules with very different 2D structures displace each other from a binding site on macromolecules. Even more often, mono modification of the structure of an active molecule renders it inactive. Such structure bioactivity relationships are an indirect probe of the 3D structure and chemical properties of the macromolecular recognition site for the ligands. The goal of pharmacophore mapping is to transform such 2D structure-activity information into the 3D requirements for binding to the target biomolecule. This allows one to search 3D databases for other molecules that match these 3D properties or to design new active molecules. A pharmacophore map identifies the bioactive confirmation of each active molecule and indicates how to superimpose, compare in 3D, the various active compounds. The map identifies which types of points match in what conformation of the compounds. The decisions as to the required points and the bioactive conformations are

interdependent. i.e. the choice of one affects the choices available for the other. A pharmacophore features include hydrogen bond acceptor atoms, hydrogen bond donor atoms, hydrogen bond donor site, hydrogen bond acceptor site, and hydrophobic centers. The Figure 5 shows the distance relationship between various pharmacophoric features present in the map[7].

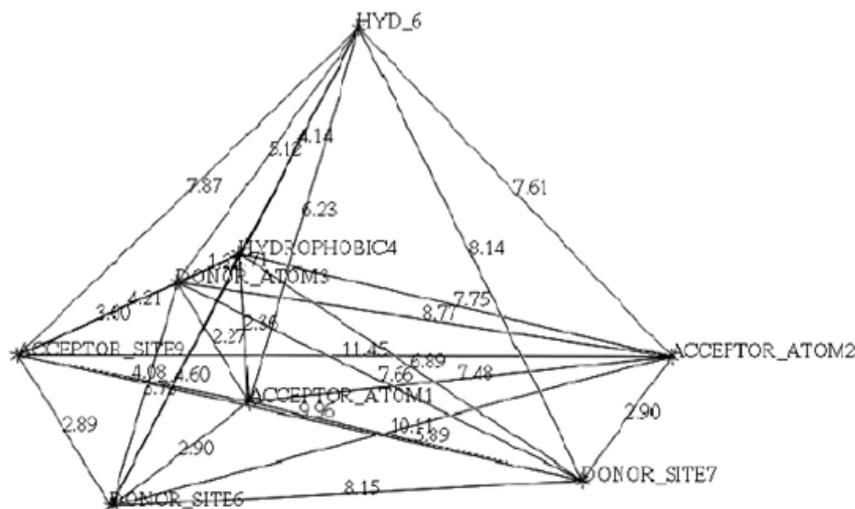


Fig 5. Relationship between various pharmacophoric features in 3D view

III. CONCLUSION

The drug discovery and development process is a long and expensive one. It starts from target identification, after that, validates the targets and identifies the drug candidates. Before any newly discovered drug is placed on the market, it must undergo extreme preclinical and clinical tests and get the FDA approval. Due to the limitation of throughput, accuracy and cost, experimental techniques cannot be applied widely, therefore, in recent times the drug discovery process has shifted to insilico approaches such as Molecular Docking, Protein-ligand interactions, QSAR, vHTS etc. Insilico approach has been of great importance to develop fast and accurate target identification and prediction method for the discovery.

REFERENCES

- [1] www.scfbio-iitd.res.in/tutorial/drugdiscovery.htm
- [2] http://en.wikipedia.org/wiki/Virtual_high_throughput_screening
- [3] Walters WP, Stahl MT, Murcko MA (1998). "Virtual screening – an overview". *Drug Discov. Today* 3 (4): 160–178. doi:10.1016/S1359-6446(97)01163-X
- [4] Bohacek RS, McMartin C, Guida WC (1996). "The art and practice of structure-based drug design: a molecular modeling perspective". *Med. Res. Rev.* 16: 3–50. doi:10.1002/(SICI)1098-1128(199601)16:1<3::AID-MED1>3.0.CO;2-6
- [5] <http://gepard.bioinformatik.uni-saarland.de/teaching/ws-2011-12/stl-bioinformatics-m added-ws1112/lectures/V05-QSAR.pdf>
- [6] <http://vega.marionegri.it/wordpress/qsar-intro/>
- [7] 28-01-10InsilicoDrugDesigningDGL7 (3)