Brief Introduction to In-Silico Drug Discovery Process and Virtual Screening Method; Ubiquitination Regulator in Cancer: A Review

[1*]Ms. Pooja S Murkute, [2]Mr. Roshan S. Raut, [3]Mr. Nakul P Kathar, [4]Dr. Gajanan S Sanap, [5]Mr. Radhesham G Gaikwad

*[1]Assistant professor, [1]Research Scholar, [2]Assistant professor, [2]Principal, [3]Research Scholar

*[1]Assistant professor Department of Pharmacognosy & Phytochemistry, LBYP College of D. Pharmacy (D. Pharm & B. Pharm, Aurangabad, India

***Abstract*: Virtual screening, especially the structure-based virtual screening, has emerged as a reliable, cost-effective and time-saving technique for the discovery of lead compounds. Here, the basic ideas and computational tools for virtual screening have been briefly introduced, and emphasis is placed on aspects of recent development of docking-based virtual screening, scoring functions in molecular docking and ADME/Tox-based virtual screening. This paper gives an overview of drug discovery process and emphasizes in the area of virtual screening. Virtual screening (VS) has emerged as an important tool in identifying bioactive compounds through computational means, by employing knowledge about the protein target or known bioactive ligands. In silico 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 in silico methods for drug discovery process with emphasis on identifying drug targets, where there are genes or proteins associated with specific diseases. This review provides a succinct overview of several recent approaches that employ bioinformatics for the systematic characterization of the targets of bioactive compounds.**

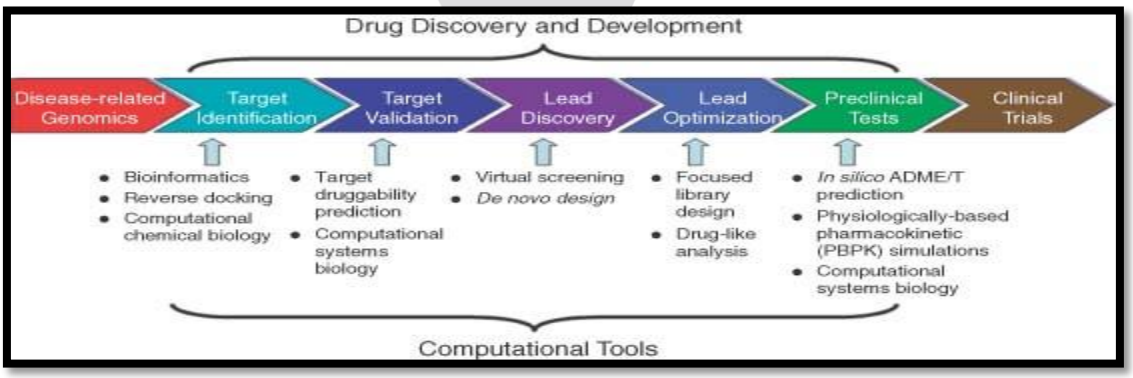
***Keywords*: Drug Discovery, In-Silico Method, Virtual Screening, Structural based virtual screening, Ligand based virtual screening Pharmacophore based virtual screening, Machine learning virtual screening, Structure Based Virtual Screening, ADME/Tox-Based Virtual Screening, Ubiquitination Regulator, Ubiquitination Regulator in Cancer.**

I. INTRODUCTION

Drug discovery and development is an intense, lengthy and on inter disciplinary [1], [2]. Drug discovery is multifaceted process which involves identification of a drug chemical therapeutically useful in treating and management of disease condition [3]. Drug discovery is lengthy process that take around 10-15 year and cost up to 2.558 billion USD for a drug to reach the market [4]. The process of drug discovery includes the identification of drug candidates, synthesis, characterization, screening and assay for therapeutic efficacy [3]. Drug discovery is a linear, consecutive process [1]. The process state that target and lead discovery followed by lead optimization and preclinical in vitro & in vivo studies [1]. The goal of drug discovery process is to search new drug molecules [5].

Drug Discovery is process

Drug Discovery is process involves the identification of candidate's synthesis characterization, screening and assays for therapeutic efficacy [3]. The drug discovery is a process for identifying compound therapeutically useful in treatment and curing the disease the drug discovery process and Development (pipeline) have a number of a distinct viral state steps. The drug discovery Pipeline Consist from the selecting a disease Continuation with target hypothesis, lead, optimization preclinical & clinical trials. Following fig. 1 Present the Representation of drug discovery process. (Pipeline) [1].



"Figure 1. Drug discovery process"

The Important Step involved in the process of drug discovery are
a) Lead Identification
b) Lead Optimization
c) Pre- Clinical Lead Development
d) Clinical Lead Development

a) Lead Identification
Lead identification is complex factor the search for lead structure. The natural products structure directed molecule design, modification of natural products, biochemical understanding of the disease process & broad screening of synthesis compound from which lead can be obtained [5].

b) Lead optimization
In this process a compound that displays an interesting biological action and ends with the identification of the best anlages. Molecules are chemically modified & susequitety characterized in order to obtain compound with suitable properties like a drug efficacy &potency in vitro and in vivo physical properties, pharmacokinetic properties and toxicological properties are characterizing by lead. Lead structures can have improved target affinity and selectivity. Docking techniques are applied to aid on structure-based absorption, distribution, metabolism and excretion (ADME) [6].

c) Preclinical development
Preclinical studies and testing programmer with and without the use of animal testing methods have the purpose of limiting risks. Whenever a new active substance is to be used as a medicinal product in humans. They should be designed in such a way as to achieve as early, risk-free, unproblematic, and economic a transition as possible from preclinical to clinical trials in medicinal products development (Glossary of Clinical Trial Terms, NIH Clinicaltrials.gov). Scientists carry out in vitro and in vivo tests. In vitro tests are experiments conducted in the lab, usually carried out in test tubes and beakers ("vitro" is "glass" in Latin) and in vivo studies are those in living cell cultures and animal models ("vivo" is "life" in Latin). Preclinical testing involves: pharmacology, toxicology, preformulation, formulation analytical and pharmacokinetics [6].

d) Clinical development

The NIH organizes clinical trials into 5 different types:

1. Treatment trials: test experimental treatments or a new combination of drugs.

2. Prevention trials: look for ways to prevent a disease or prevent it from returning.

3. Diagnostic trials: find better test or procedures for diagnosing a disease.

4. Screening trials: test methods of detecting diseases.

5. Quality of life trials: explore ways to improve comfort & quality of life for individuals with a chronic illness [3], [6].

II. PHARMACEUTICAL CLINICAL TRIALS ARE COMMONLY CLASSIFIED INTO 4 PHASES

Phase 0
A recent designation for exploratory, first in human trials designed to expedite the development of promising therapeutic agents by establishing early on whether the agent behaves in human subjects as was anticipated from preclinical studies.

Phase 1
A small group of healthy volunteers (20-80) are selected to assess the safety, tolerability, pharmacokinetics, & pharmacodynamics of a therapy. Normally include dose ranging studies so that doses for clinical use can be set/adjusted.

Phase 2
Performed on larger groups (20-300) & are designed to assess the activity of the therapy, & continue phase1 safety assessments.

Phase 3
Randomized controlled trials on large patient groups (hundreds to thousands) aimed at being the definitive assessment of the efficacy of the new therapy, in comparison with standard therapy. Side effects are also monitored. It is typically expected that there be at lest two successful phase 3 clinical trials to obtain approval from the FDA. Once a drug has proven acceptable, the trial results are manufacturing procedures, formulation details, shelf life, etc. This document is submitted to the FDA for review.

Phase 4
Post-launch safety monitoring & ongoing technical support of a drug may be mandated or initiated by the pharmaceutical company designed to detect rare or long term adverse effects over a large patient population & timescale than was possible during clinical trials [3].

III. IN SILICO DRUG DISCOVERY PROCESS

Latin in silico (i.e. performed using computer or via computer stimulation). Petro miramontes, a mathematician form National Autonomous University of Mexico (UNAM), presented the report "DNA and RNA physicochemical constrain cellular automata & molecular evolution. In his talk, Mira motes used the term "in silico" to characterize biological experiments carried out entirely in a computer [1]. The history of 'in silico' term is poorly defined, with several researchers claiming their role in its origination. However, some of the earliest published examples of the Word include the use by Sieburg (1990) and Danchin et al (1991). In a more recent book, Danchin (2002) provides aquotation that offers a concise and cogent depiction of the Potential of computational tools in chemistry, Biology and pharmacology [7]. In silico drug designing is a form of computer-based modeling whose technologies are applied in drug discovery processes [6]. In silico (or virtual) compound library design operates to reduce the number of compounds to be tested, and two basic applications can be distinguished: diversity and structure-based design. Diversity design aims to select a smaller sub-library from a larger compound library in such a way that the full range of chemical diversity is best represented13. When no structural information about the target and/or target ligands is available, diversity design is the method of choice [8].

IV. SIGNIFICANCE OF IN-SILICO DRUG DISCOVERY PROCESS

As structures of more and more protein targets become available through crystallography, NMR and bioinformatics methods, there is an increasing demand for computational tools that can identify and analyze active sites and suggest potential drug molecules that can bind to these sites specifically. Time and cost required for designing a new drug are immense and at an unacceptable level. According to some estimates it costs about $880 million and 14 years of research to develop a new drug before it is introduced in the market. Intervention of computers at some plausible steps is imperative to bring down the cost and time required in the drug discovery process [6].

V.NEED FOR MODERN IN SILICO TECHNIQUES

a) These techniques offer the advantage of delivering new drug candidates more quickly and at a lower cost.
b) They increase the chance of success in many stages of the discovery process.
c) They facilitate accessing huge amount of data generated.
d) They transform the massive complex biological data into workable knowledge

VI. COST OF INNOVATION

In 2001 Pharmaceutical research and manufacturers of America (PhRMA) estimated the cost at US$802 million over a period of 11 years from the initial research stage to the successful marketing of a new drug4. More recent estimates by DiMasi at the Tufts Center for Study of Drug Development (CSDD) that was published in 2003 put the average cost at US$802 million spread over 12 years, while the Boston Consulting Group estimates the cost as $880 million over 15 Boston Consulting Group estimates the cost as $880 million over 15 ranges from $800 million to $1.8 billion. These estimates are averages and there is significant variation in both time and cost averages and there is significant variation in both time and cost drug being developed and the nature and scope of the clinical trials required to gain regulatory approval [1].

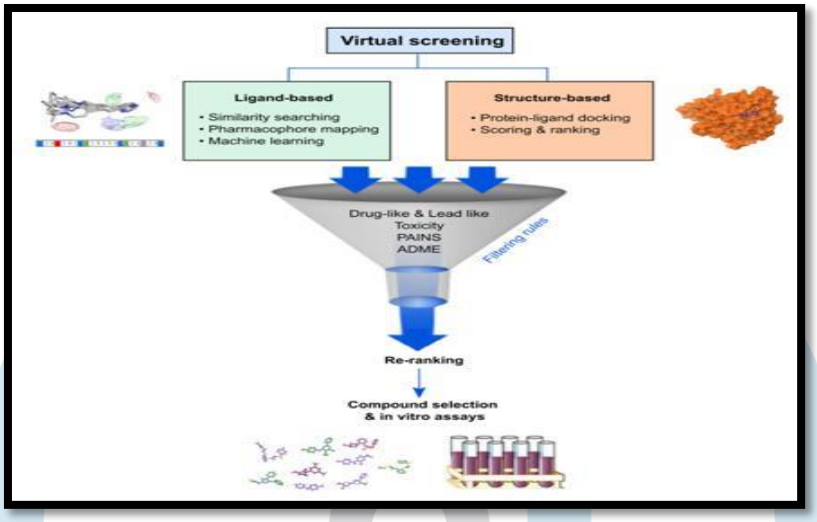
VII. CLASSIFICATION OF IN-SILLICO METHODS

a) QSAR (Qualitative structural activity relationship)
b) Computational method
c) Virtual screening method
d) Ligand screening

c) Virtual screening

Popular VS techniques originated in the 1980s, but the first publication about VS appeared in 1997. VS is an in silico technique used in the drug discovery process [9]. The challenges of in silico drug discovery include the evaluation of multiple binding modes, of accessible conformational states for both ligand and receptor, of affinity and selectivity versus efficacy, plasma protein binding, metabolic stability (site of reactivity and turn-over), absorption, distribution and excretion, as well as in vivo versus in vitro properties of model compounds, while seeking a favorable intellectual property position [10]. V.S is the computational method of in silico. The aim of VS is to score, rank and/or filter a set of chemical structures using one or more computational procedures in order to ensure those molecules with the largest prior probabilities of activity are assayed first in a "lead discovery programed" [11]. The application of computational screening is also called "virtual" or in silico screening. Recent advances in drug discovery have enabled a dramatic increase in the number of synthetic and naturally sourced molecules that are available for testing in vitro in biochemical and cellular assays [12]. Virtual screening is the ability to discover highly active, yet small, organic molecules which are used for treatment purposes [10]. Virtual screening (VS) is the use of high-performance computing to analyze large databases of chemical compounds to identify possible drug candidates, which is seen as a complementary approach to experimental HTS. Virtual screening is knowledge driven, which means that some information is available regarding either the nature of receptor binding pocket or the type of ligand that is expect to bind productively, or both. It should be noted that VS encompasses a variety of computational screens, from the simplistic to the sophisticated, and hence, can effectively exploit different types of information describing the receptor. we are witnessing an increased integration of in silico technologies that sift through enormous numbers of virtual chemicals based on 'soft modeling' quantitative structure–activity relationship (QSAR) techniques that estimate properties deemed important for orally available drug. The development of virtual screening methods extends the possibilities to molecules

that do not necessarily exist physically in an investigators collection but which can be readily obtained through purchase or synthesis. While such methods which include pharmacophore searching (PS) and high-throughput docking (HTD) are less than perfect at prediction of biological activity, 'enrichment of actives' is the desired aim and has been achieved for many targets that have been studied in silico the development of the first algorithms such as UCSF Dock virtual screening continues to increase in credibility as a source of chemical starting points in the drug discovery process [10]. Computational chemistry has partially fulfilled its promise but the reality of drug–receptor interactions, at the molecular level, continues to be too complex to provide a failsafe in silico technology for drug discovery [13].

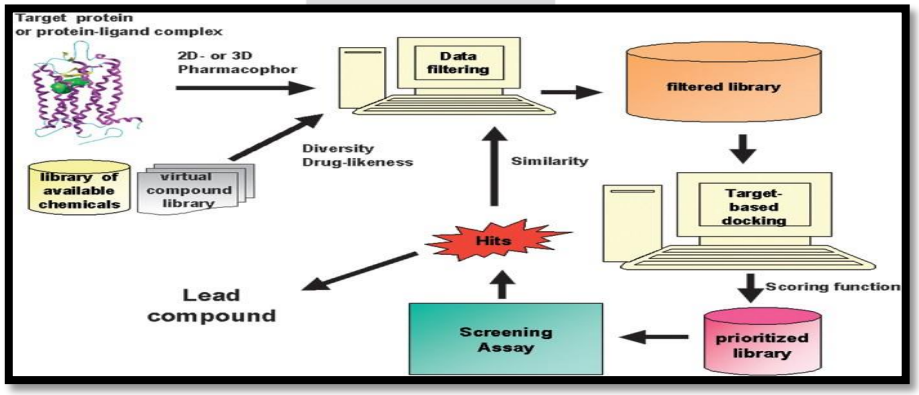


"Figure 2. Virtual screening methods"

Analysis of a specific biological activity for an entire screening library can be done by high throughput screening in a matter of days. In order to control costs, time and waste, the computational chemist is encouraged to develop some kind of computer programmer capable of automatically evaluating very large libraries of compounds and integrate it into the drug discovery process. This is called "virtual screening" (VS). VS is a set of computational methods or *in silico* analogues of biological screening [11].

1. VIRTUAL SCREENING METHODS INTO FOUR MAIN CLASSES BASED ON THE AMOUNT OF STRUCTURAL AND BIOACTIVITY DATA AVAILABLE, AS FOLLOWS:

i. If just a single active molecule is available, then similarity searching can be used.
ii. If several active molecules are known, it is possible to define what is known as a common 3D pharmacophore leading to a 3D database search.
iii. Machine learning methods should be used for VS only if it is not possible to identify a common pharmacophore and there is a sufficient number of active and inactive molecules available. They can be used to derive a structure activity relationship from the known active molecules for use in predicting biological activity. An example of machine learning for virtual screening is shown in figure 3.
iv. 4) A docking study can be employed if the 3D structure of the biological target is known. It involves the prediction of the binding mode of individual molecules. It aims to identify the orientation that is closest in geometry to the observed structure [11].



"Figure 3. An example of machine learning for virtual screening"

Examples

Examples of drugs that came to the market with the assistance of VS include captopril (antihypertensive drug), saquinavir, ritonavir, and indinavir (three drugs for the treatment of human immunodeficiency virus), tirofiban (fibrinogen antagonist), dorzolamide (used to treat glaucoma), zanamivir (a selective antiviral for influenza virus), aliskiren (antihypertensive drug), boceprevir (protease inhibitor used for the treatment of hepatitis C), nolatrexed (in phase III clinical trial for the treatment of liver cancer)[9]

2. Advantages of virtual screening

- The use of complementary experimental and VS techniques increases the chance of success in many stages of the discovery process.
- VS has emerged as a reliable, cost effective and time-saving technique for the discovery of lead compounds in recent times.
- The main advantages of this method compared to laboratory experiments are noted below [14].

3. Methods of virtual screening

The type of method/methods used in VS depends upon information available as input and the type of results required for output. For example, if a 3-D structure of target protein is available, molecular docking or combinatorial drug design can be used to perform receptor based, fine grained compounds sieving. If a 3-D receptor structure is unavailable, then pharmacophore model derived from bioactive ligands or molecular property profiles, such as molecular weight, lipophilicity, ADME properties or drug-like properties, can be used as filters in VS. In the following sections, recent developed techniques used in VS and the advances made in VS are reviewed [13]. Following are the methods used in virtual screening,

i. Structural based virtual screening
ii. Ligand based virtual screening
iii. Pharmacophore based virtual screening
iv. Machine learning virtual screening
v. Virtual screenings based on molecular docking
vi. Virtual screening based on ADME/TOX

i. Structure based virtual screening: -

Structure based virtual screening (SBVS) is a robust, useful and promising in silico technique for drug design. It is also called as target based virtual screening. This technique emerged in the 1980s, when designed and tested a set of algorithms that could explore the geometrically feasible alignments of a ligand and target [9]. It attempts to predict the best interaction between ligands against a molecular target to form a complex. As a result, the ligands are ranked according to their affinity to the target, and the most promising compounds are shown at the top of the list. SBVS methods require that the 3D structure of the target protein be known so that the interactions between the target and each chemical compound can be predicted in silico. In this strategy, the compounds are selected from a database and classified according to their affinity for the receptor site. During the execution of SBVS, the evaluated molecules are sorted according to their affinity to the receptor site. Hence, it is possible to identify ligands that are more likely to present some pharmacological activity with the molecular target. Score functions are used to verify the likelihood of a binding site describing the affinity between the ligand and target. In this process, a reliable scoring function is the critical component of the docking process [9]. Structure-based VS can be implemented if the 3D shape binding of the biological target is known. An example of structure-based VS is docking – the fourth class of VS. One focus in their model is the consideration of 3D distribution of atomic lipophilicity derived from a quantum mechanics-based continuum solvation model [15]. Ven a 3D structure of the target, TBVS relies on Docking and scoring to provide potential candidates for further analysis. AutoDock, DOCK, FlexX, FRED, GOLD and ICM are some of the most used docking programs in the field. Most of these programs can dock libraries of single structures or even multi conformer libraries, and are thus suitable for high-throughput searches. Regardless of the choice of the docking software, one is limited in TBVS by the ability to score the ligands in an appropriate manner. There are four major categories of scoring functions. First, knowledge-based methods use Boltzmann- weighted potentials of mean force derived from statistical analyses of ligand–receptor inter-atomic contacts, based on available complexes in the Protein Data Bank (PDB). SMoG, Mugger's, and Drug- Score are implementations of this approach. Second, 'master equation' approaches estimate the energetic contributions of various interaction types in a semi quantitative manner. Third, regression-based methods take advantage of the available biological activity for training sets of ligand–receptor complexes extracted from the PDB. Finally, Poisson–Boltzmann equation solvers that address electrostatics incorporating solvent effects. Consensus scoring can be used when a particular scheme fails. We recently showed that pharmacokinetic awareness can be integrated into a regression-based scoring scheme, using VolSurf [10].

The use of SBVS has advantages and disadvantages

Advantages

- There is a decrease in the time and cost involved in the screening of millions of small molecules.
- There is no need for the physical existence of the molecule, so it can be tested computationally even before being synthesized.
- There are several tools available to assist SBVS.

Disadvantages

- Some tools work best in specific cases, but not in more general cases.

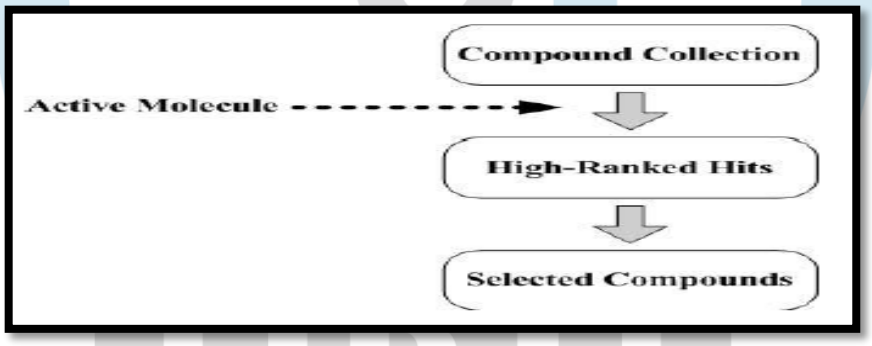
- It is difficult to accurately predict the correct binding position and classification of compounds due to the difficulty of parameterizing the complexity of ligand-receptor binding interactions.
- It can generate false positives and false negatives [9].
- Docking protocols are essential for achieving accurate SBVS. These protocols are composed of two main components: the search algorithm and the score function.

Successful applications of structural-based virtual screening

Novel, potent and selective CK2 (casein kinase II) inhibitors were obtained by screening a subset (ca. 400 000 molecules) of the Novartis database using DOCK. The ATP binding site of a human CK2a was inferred via homology modeling (X-ray structure of Zea maysCK2a, PDB entry 1DAW). Filters related to molecular weight, to the number of rotatable bonds, and to undesired substructures were applied before TBVS. Protonation states were adjusted using a rule-based method. The compounds with best DOCK scores were further refined using known pharmacophore information (e.g. hydrogen bonds to the kinase hinge region). Novel BCR-ABL tyrosine kinase inhibitors were identified with a related TBVS workflow, using DOCK and a database of 200 000 commercially available compounds. Human carbonic anhydrase II inhibitors were identified using a set of hierarchical filters and the FlexX docking engine, applied to a database of _100 000 compounds. The top 100 compounds after this step were subject to flexible docking using FlexX [10].

ii. Ligand based virtual screening: -

A ligand is a molecule that is able to bind to and form a complex with a biomolecule to serve a biological purpose. Ligand-based VS involves using information available from a single or set of compounds which have been identified as potential leads. A lead compound is a compound that exhibits pharmacological properties which suggest its value as a starting point for drug development. Ligand-based VS is conducted by identifying molecules that share some similarity or properties with the single/multiple active molecules. It aims to score database molecules based on their overall shape similarity to query molecules. Example of ligand-based VS are substructure/similarity searching, pharmacophore-based designs, and machine learning techniques, which correspond to the first, second, and third class of VS methods grouped by respectively [11]. An example workflow for ligand-based VS is shown in figure 4.



"Figure 4. Example work flow of ligand based virtual screening"

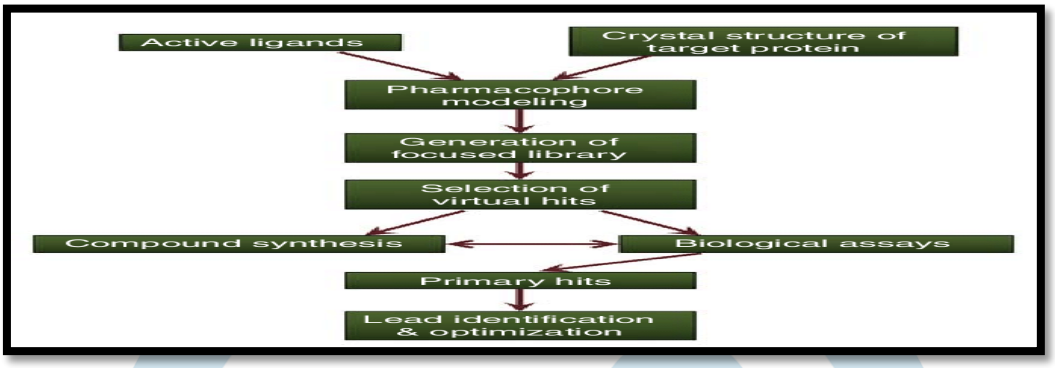
Successful applications of ligand–based virtual screening

CATS (chemically advanced template search), a topologically-based pharmacophore screening method, has the ability to perform 'scaffold-hopping' (i.e. to identify iso-functional structures that differ in chemo type). When using mibefradil, a known T-type calcium channel blocker (IC50 ¼ 1.7 mM) as a query, CATS identified one significant hit (clopimozid; 5, with an IC50 < 1 mM, among the top 12 ranked molecules. The same technique, CATS, was applied to identify structurally novel glycogen synthase kinase-3 inhibitors, first by identifying the oxadiazol–pyridyl moiety, a new chemotype, then by synthesizing additional analogs. Compounds with inhibitory activity below 1 mM were identified. Pharmacophore-searching techniques have also result in several successful LBVS applications. Potent and novel integrin a4b1 antagonists were identified using Catalyst, employing a pharmacophore query obtained from a tetrapeptide inhibitor. A similar LBVS workflow was followed to identify nonpeptidic urotensin II (U-II) receptor antagonists, using a G-protein coupled receptor-biased subset of the Aventis database [10].

iii. Pharmaco-phore based virtual screening: -

Gund was probably the first who described that functional groups (pharmacophores) could be used for searching databases to identify molecules that may share the same structural features. A pharmacophore is the spatial arrangement of key structural features of a set of known ligands or of the target receptor. This has led to the successful development and application of 3D-database pharmacophore searching for discovering novel lead compounds in drug discovery. Pharmacophore-based virtual screening is involved in two steps: identification of pharmacophore model and 3-D search based on the specific constraints. Pharmacophore models are typically used when some active compounds have been identified but the three dimensional (3D) structure of the target protein or receptor is still unknown. The main difficulty in pharmacophore generation is in handling the conformational flexibility since the active conformations of the molecules are usually postulated. A pharmacophore model not only can be obtained from a set of active ligands, but also can be derived from the 3-D structure of receptor. The information provided by the 3-D structure can be analyzed to identify interaction points in the binding site as pharmacophore model, which can be used as a query in a 3-D database screening. For example, when a crystal complex is available, the atoms of ligand contributing to receptor-ligand interaction

can be defined as features in a pharmacophore model. Both of UNITY and CATALYST can define pharmacophore model based on the structure of protein. But sometimes, due to the complexity of the receptor structure, it is very necessary to find important features for the definition of pharmacophore. The LUDI interaction site identification procedure was then used to generate the interaction sites. Finally, a set of 3-D queries is then derived from the interaction map and this database is searched with the set of queries. These hits are scored using LUDI. After the pharmacophore model is identified, 3-D database search is performed to find compounds bearing these pharmacophoric features. Now, several programs are available for 3-D database search based on pharmacophore model. The most widely used programs include catalyst, unity, MDL, Chem-X [13].



"Figure 5. Pharmaco-phore based virtual screening"

iv. Machine learning virtual screening: -

Machine learning involves the design and development of algorithms and techniques in order to allow computers to learn and understand data. Machine learning research focuses on extracting the relationship from available data by computational and statistical methods. There are many reasons why engineers need to identify or model complex relationships e.g. understanding the world, predicting the future, classification, decision support and control. A machine learns when it changes its structure or data based on its input, hence, broadly say, performance of the machine is expected to be improved, assuming that data are plentiful.

Sub structural Analysis

Sub structural analysis is the first machine learning method used in cheminformatics. It is a class of QSAR techniques which assume that a defined molecular fragment gives a constant contribution to an activity. It aims to assign a weight for each sub structural fragment which reflects its possibility of being active or inactive.

Linear Discriminant Analysis

Here, linear discriminant analysis (LDA) aims to separate molecules into defined classes. The simplest type of LDA is the binary classification problem. It finds the linear combination of features and draws the hyper plane which best separates the data. It was first applied to a set of biologically active molecules by. Their study involves a relatively small dataset containing only 20 compounds with general structure of the aminotetralins and aminoindans.

Neural Networks

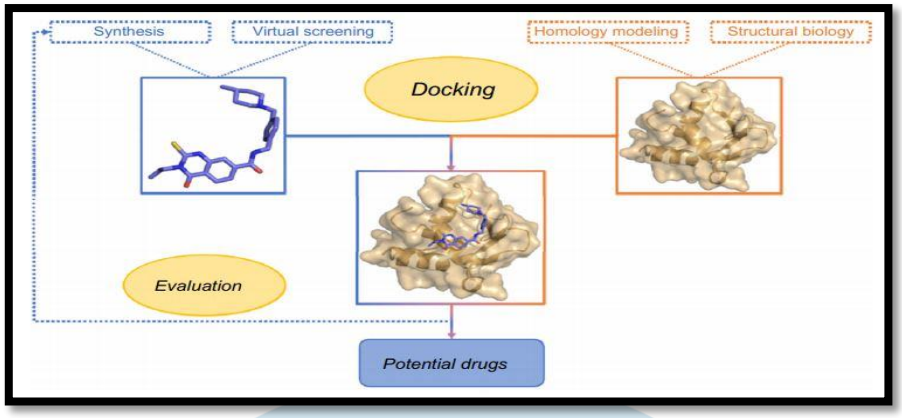
Neural networks (NNs) are a type of artificial intelligence that attempt to mimic the way a human brain works. They work by creating connections between artificial neurons which are arranged in layers and associated with weights. The weights are initially set to random values. The NN must first be trained in order to adjust the weights by using a known set of inputs and corresponding set of outputs.

Decision Trees

Decision trees consist of a set of rules which can associate specific molecular features with an activity or property of interest. A decision tree represents a Boolean function and is described as a tree-like structure. Each node corresponds to a specific rule. Once decision trees have been trained from a training set, an unknown sample can be classified by starting at the root node then following the edge appropriate to the rule. This is repeated until a terminal node is reached. This approach allows easy determination of the most relevant chemical features to the target biological property [11].

v. Virtual screening based on molecular docking: -

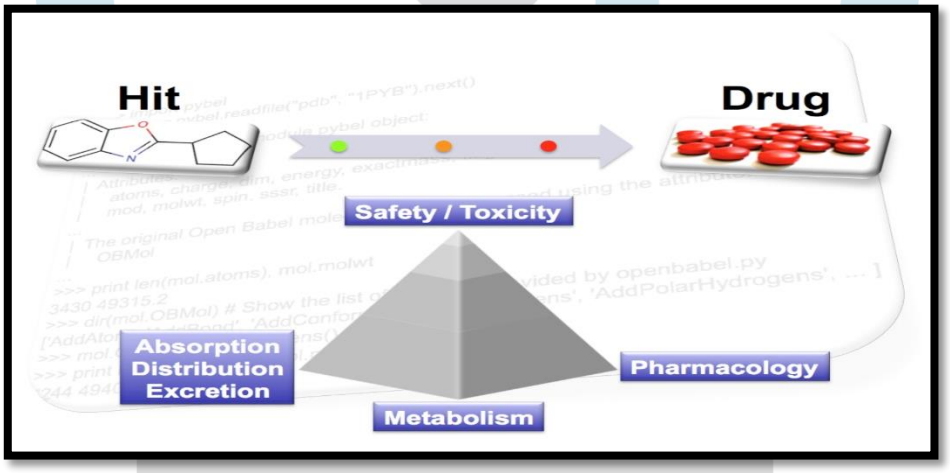
Molecular docking can fit molecules together in favorable configuration to form a complex system. It has been practically applied as a very efficient way in the studies of protein-ligand interactions. Structural information from the theoretically modeled complex may help us to clarify the mechanism of molecular recognition, and even can instruct us to discover novel ligand according to the predicted binding affinities between receptor and ligands. The attraction of receptor-ligand docking is that it represents the most detailed and relevant computational model for identifying a receptor-focused subset of database. There are a large number of docking programs available for use in virtual screening, differing in the sampling algorithm, scoring functions, the treatment of flexibility of ligand and receptor, and the CPU time required to dock a molecule to a given target [13].



"Figure 6. Virtual screening based on molecular docking"

vi. Virtual screening based on ADME/TOX: -

The significant failure rate of drug candidates in later developmental stages is driving the need for predictive tools that can eliminate inappropriate compounds before substantial time and money are invested in testing. It has been estimated that about 50% of such failures are caused by ADME/Tox deficiencies. Virtual screening should not be restricted to the scenarios with respect to optimizing binding affinity, and the pharmacokinetic properties should also be treated as important filtering protocol. Many computational approaches have been developed for ADME parameters including bioavailability, aqueous solubility, intestinal permeability, blood-brain barrier penetration, metabolism, drug-drug interactions, drug transport and toxicity. The predictions of these properties are involved in two aspects of modeling methods: data modeling and molecular modeling. For molecular modeling, molecular mechanics, pharmacophore modeling, molecular docking, or even quantum mechanics are used to explore the potential interactions between the small molecules under consideration and proteins known to be involved in ADME processes, such as cytochrome P450s. For data modeling, quantitative structure-activity relationship (QSAR) approaches are typically applied [13].



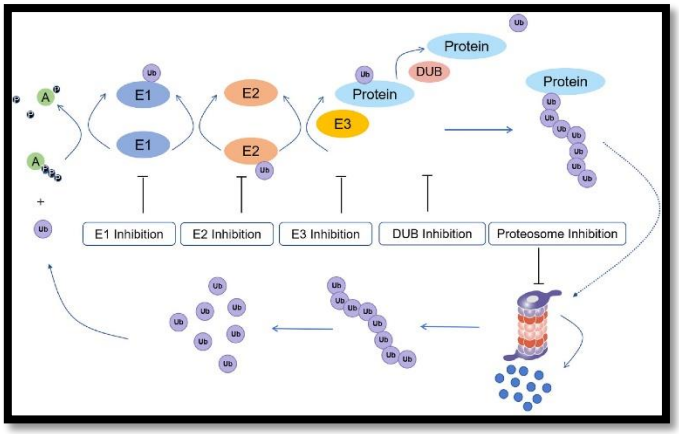
"Figure 7. Virtual screening based on ADME/TOX"

VIII. Ubiquitination Regulator Discovery by Virtual Screening in Cancer [16]

Definition

Ubiquitination Regulators identified by Virtual Screening for the Treatment of Cancer specifically for Breast cancer. The ubiquitin-proteasome system transpacific cellular protein degradation is having ability to regulate various critical processes, such as cell cycle control and DNA repair.

Virtual screening usually carried out in hierarchical workflow with sequential steps that filter and remove unwanted molecules. Molecules successfully pass through all stages of the virtual screening are called the "hit compounds." One another method of virtual screening is to hit compounds is clustering, which is an unsupervised learning technique were input data is fed into the algorithm in order to identify patterns and classify the data into several categories. By using above mentioned method clustering to group hits based on their structure, the biased choice of molecules can be circumvented and representative samples of compounds can be obtained. Another method like Hierarchical clustering, HDBSCAN and k-means clustering are different clustering methods for hit selection Finally, hit compounds need to also be verified through experiments to validate their bioactivity.



"Figure 8. Ubiquitination Regulator Discovery by Virtual Screening"

Types of Ubiquitination Regulator

a. Identification of E1 Regulators by Virtual Screening

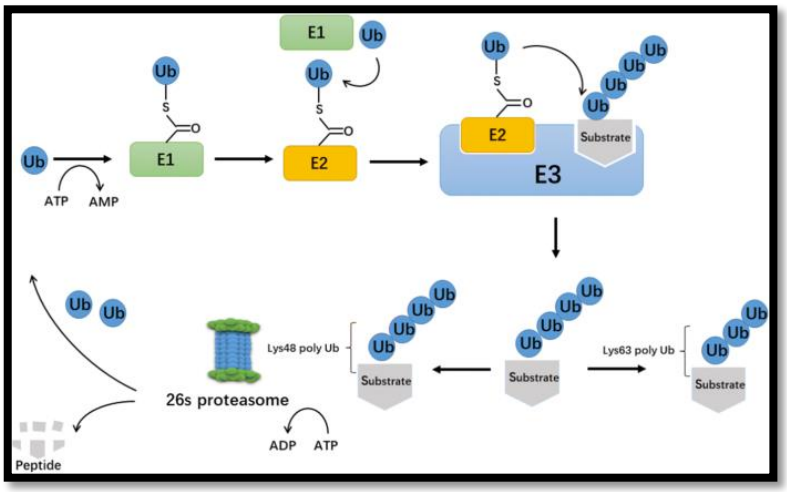
Inhibitors of E1s are the potentially new class of cancer drug. Two types of ubiquitin-activating E1 enzymes are mainly known as the main is UBA1 and another one the newly discovered UBA6 which is currently has unclear functions. but, due to the lack of specificity of E1s, there are few reports on E1 inhibitors discovered by Virtual Screening.

b. Identification of E2 Regulators by Virtual Screening

Another one method is E2 Regulators by Virtual Screening which is intermediate between the E1 and E3 proteins, the E2s play an important function in determining the kind of polyubiquitin chain connected. Presently, around 40–50 genes encoding ubiquitin-conjugating enzymes have been found in the human genome, which is much greater than the number of E1 enzymes. After E2s bind to the activated ubiquitin, ubiquitin molecules are transfered to ubiquitin ligase E3. Each E2 can only associate and cooperate with a specific set of E3 enzymes, the E2–E3 interaction gives potentially selective anticancer approach. However, E2 enzyme inhibitors are currently in developing stage and pre-clinical testing stage as they still lack sufficient specificity for E2s.

c. Identification of E3 Regulators by Virtual Screening

The E3s are responsible for the direct binding of ubiquitin to the protein, thus conferring substrate specificity and selectivity. Mutations or down-regulation of E3 enzymes can often be detected in different tumors. As the substrate recognition components of UPS pathway, 500–1000 ubiquitin ligases in the human body are present in human body Hence, targeting specific E3 ligases will ultimately affect only a particular subset of ubiquitin substrates, without affecting for the entire ubiquitination pathway because of this E3 enzymes have great potential for cancer treatment.



"Figure 9. Types of Ubiquitination Regulator"

IX. CONCLUSION

This topic gives detail idea about In-Silico methods the steps involve. Virtual screening, especially the structure-based virtual screening, has emerged as a reliable, cost-effective and time-saving technique for the discovery of lead compounds. Here, the basic ideas and computational tools for virtual screening have been briefly introduced, and emphasis is placed on aspects of recent development of docking-based virtual screening, scoring functions in molecular docking and ADME/Tox-based virtual screening. This paper gives an overview of drug discovery process and emphasizes in the area of virtual screening and detail classification of virtual screening method. This topic gives detail idea about Structural based virtual screening, Ligand based virtual screening, Pharmacophore based virtual screening, Machine learning virtual screening, Virtual screenings based on molecular docking, Virtual screening based on ADME/TOX

X. ACKNOWLEDGMENT

Authors are over helmed in all humbleness and gratefulness to acknowledge our depth to all those who have helped to put these ideas, well above the level of simplicity and into something concrete. We like to express special thanks of gratitude our principal who gave us the golden opportunity to do this wonderful article.

XI. REFERENCES

[1] Deepali Gangrade, Gauri Sawant and Ashish Mehta Re-Thinking Drug Discovery: In Silico Method Journal of Chemical and Pharmaceutical Research, 2016, 8(8):1092-1099.2
[2] Wadood A, Ahmed N, Shah L, Ahmad A, Hassan H, Shams S. In-Silico Drug Design: An Approach Which Revolutionarised the Drug Discovery Process. Oa Drug Design & Delivery 2013 Sep 01;1(1):3.
[3] Arun Bahadur Gurung, Mohammad Ajmal Ali, Joongku Lee, Mohammad Abul Farah and Khalid Mashay Al-Anazi, An Updated Review of Computer-Aided Drug Design and Its Application to Covid-19, Hindawi Biomed Research International Volume 2021, Article Id 8853056, 18 Pages.
[4] Bag Mh, Ahmad K, Adil M, Khan Za, Khan Mi Et Al., Drug Discovery and in Silico Techniques: A Mini-Review. Enz Eng 2014 3: 123.
[5] Prakash, Nutan, And Patel Devangi. "Drug Discovery." Journal of Antivirals and Antiretrovirals 2, No. 4 2010: 063-068.
[6] V. Srinivasa Rao and K. Srinivas, Modern Drug Discovery Process: An in Silico Approach, Journal of Bioinformatics and Sequence Analysis, June 2011.Vol. 2(5), Pp. 89-94,
[7] S Ekins, J Mestres and B Testa, In Silico Pharmacology for Drug Discovery: Methods for Virtual Ligand Screening and Profiling, British Journal of Pharmacology, 2007 152, 9–20.
[8] Georg C. Terstappen and Angelo Reggiani, In Silico Research in Drug Discovery, Trends in Pharmacological Sciences January 2001vol. 22 No.1.
[9] Eduardo Habib Bechelane Maia, Letícia Cristina Assis, Tiago Alves De Oliveira2, Alisson Marques Da Silva2 And Alex Gutterres Taranto1, Structure-Based Virtual Screening: From Classical To Artificial Intelligence, Frontiers In Chemistry, 1 April 2020, Volume 8, Article 343.
[10] Tudor I Opreaa, And Hans Matterb, Integrating Virtual Screening in Lead Discovery: Current Opinion in Chemical Biology 2004, 8:349–358.
[11] Kitsuchart Pasupa, The Review of Virtual Screening Techniques: Kmitl Information Technology Journal (Jan. – Jun. 2012).
[12] Campbell Mines, Virtual Screening Strategies in Drug Discovery: Current Opinion in Chemical Biology 2007, 11:494–502.
[13] Ting Jun Hou and Xiaojie Xu, Recent Development and Application of Virtual Screening in Drug Discovery: An Overview, Current Pharmaceutical Design, 2004, 10, 1011-1033.
[14] Supratik Kar & Kunal Roy, How Far Can Virtual Screening Take Us in Drug Discovery: Expert Opinion On Drug Discovery, 8:3, 245-261
[15] J. Chem., New Trends in Virtual Screening: Inf. Model. 2020, 60, 4109−4111
[16] Song Y. Q., Wu, C., Wu, K. J., Han, Q. B., Miao, X. M., Ma, D. L., & Leung, C. H. Ubiquitination Regulators Discovered by Virtual Screening for The Treatment of Cancer. Frontiers in Cell and Developmental Biology, 2021. 9.