COMPUTER AIDED DRUG DESIGNING – A NOVEL APPROACH METHOD

Abstract

Today, there possesses different types of drugs against different diseases and various sophisticated technologies but still it is not enough to develop drugs instantly against a new emerging disease as drug designing and development is a complex and lengthy process and requires a huge amount of money. In 21st generation, other reliable methodsare being used to speed up the drug discovery process following the combinatorial chemistry emergence, known as computer aided drug designing where drugs leads are being predicted among the huge number of molecules by checking the interaction between the target and various ligands and also their drug likeliness property through insilico approaches. It involves methods such as homology modeling de novo approach, virtual screening approach, molecular docking, QSAR, simulation and various others. ChemDraw and other molecular structure drawing programs, descriptor generators like DRAGON, MolConnZ, and OEChem, text editors like UltraEdit and EMACS, databases like ZINC, PubChem, and ChemDiv, all QSAR programs, Perl and Python modules, the AutoDock3 program, the AutoDock Tools package, the OpenBabel program, and many more programs are used for all of these tasks. With the help of this softwares drug designing using insilico approach has minimized the cost and the time required as well as reduced the chance of failure or rejection of drugs at the clinical trial phases.Till date many drugs have been discovered using insilico methods for different diseases such as cancer, hypertension, diabetes which falls among the top category in the list of global causes of deaths along with various other diseases and even the diseases that are caused by pathogens.

Keywords: In-silico, Drug design, CADD, ADMET, QSAR.

Authors

Jyotirmoy Das
Department of Life Science and Bioinformatics'
Assam University,
Silchar , Assam, India.

Basudev Tassa
Department of Life Science and Bioinformatics'
Assam University,
Silchar , Assam, India.

Biplab Kumar Das
Department of Zoology,
Jengraimukh College,
Jengraimukh, Majuli, Assam, India.
biplabkumar1987@gmail.com

I. INTRODUCTION

The development and discovery of drugs is a complex, lengthy, time consuming and interdisciplinary course. It is also highly expense and bestows many challenges and changes since decades (Danishuddin & Khan, 2015). But new diseases and disorders are discovered eventually and even some leads to an emergency like that of the covid19. Thus, new potential drugs are of great demand to overcome and mitigate the risk of new diseases which requires multidisciplinary aspects to accomplish this challenging process. In recent times, the use of in-silico methods and molecular modeling for computer aided drug design has gained much popularity in this field due to its advantage of being cost effective and time saver over the traditional method of drug development. The traditional drug designing method is the one in which the foremost task is to identify suitable drug target molecules which include nucleic acids and protein receptors, ion channels transporters, enzymes. Drugs were traditionally found by manufacturing compounds through laborious, multi-step processes against a battery of in vivo biological screens, then pursuing the most promising candidates to learn more about their metabolism, pharmacokinetic characteristics, and possible toxicity (Kapetanovic, 2008). High attrition rates have been the outcome of this type of development method, with failures ascribed to inadequate pharmacokinetics, ineffectiveness, toxicity to animals, negative effects in people, and other commercial and incidental considerations. The development of genomics, proteomics, bioinformatics, and effective technologies such as combinatorial chemistry, high throughput screening (HTS), virtual screening, de novo design, in vitro, in silico ADMET screening, and structure-based drug design has transformed the drug discovery process today (Kapetanovic, 2008).

In silico drug design comprises computational methods and resources that serve to enhance the prospects for future drug lead discovery. Bioinformatic techniques offer significant potential in various areas, including target identification (typically proteins/enzymes), target validation, comprehension of protein evolution and phylogeny, and protein modeling. These approaches not only expedite the identification of drug targets and the screening and refinement of drug candidates but also aid in the characterization of side effects and the prediction of drug resistance. A key focus of contemporary bioinformatics methods involves the prediction and recognition of biologically active candidates, along with the mining and storage of pertinent information. Additionally, it furnishes strategies and algorithms for forecasting new drug targets and for the storage and management of existing drug target data. The rapid expansion in this field has been made feasible through advancements in both software and hardware, resulting in enhanced computational power and sophistication. Additionally, the identification of molecular targets and the growing database of publicly accessible target protein structures have contributed to this progress. Computer-Aided Drug Design (CADD) is being employed to identify active drug candidates (hits), select the most promising candidates for further evaluation (leads), and optimize these leads. This optimization involves the transformation of biologically active compounds into suitable drugs by enhancing their physicochemical, pharmaceutical, and ADMET/PK (pharmacokinetic) properties, as outlined by Kapetanovic in 2008.

Virtual screening is a technique used to uncover new drug candidates from diverse chemical scaffolds by exploring three-dimensional chemical structure databases, whether commercial, public, or private. The primary objective is to reduce the chemical space's size, enabling a more concentrated focus on candidates with higher potential for lead discovery

and optimization. The ultimate goal is to enrich the pool of molecules with desirable properties (active, drug-like, lead-like) while eliminating compounds with undesirable characteristics (inactive, reactive, toxic, poor ADMET/PK), as also noted by Kapetanovic in 2008.

In essence, in silico modeling plays a significant role in minimizing the time and resources required for chemical synthesis and biological testing. The remarkable growth of virtual screening is evident in the increasing number of citations related to the keywords "virtual screening," which rose from four in 1997 to 302 in 2004, as highlighted by Pozzan in 2006. In his 2003 review article, Green of GlaxoSmithKline expressed an optimistic view of the future, stating that "The future is bright. The future is virtual" [Green, 2003].

A wide array of software tools are employed in in-silico drug design, encompassing grid computing, window-based general PBPK/PD modeling software, PKUDDS tailored for structure-based drug design, along with programming languages like APIS, JAVA, Perl, and Python. In the realm of in-silico drug design, these software resources are complemented by software libraries and various techniques, such as drug design visualization, homology modeling, molecular dynamics simulations, energy minimization, molecular docking, and quantitative structure-activity relationship (QSAR) analysis, as noted by Oa et al. in 2013.
In-silico drug design plays a significant role across all stages of drug development, from the early preclinical discovery phase to the advanced stages of clinical development. Its application in drug development aids in the identification of potent lead molecules, thus mitigating the risk of late-stage clinical failures. Consequently, this approach can lead to substantial cost reductions, as emphasized by Oa et al. in 2013.

II. DEVELOPMENT OF DRUG AND DRUG CANDIDATE: AN INSILCO APPROACH

Regardless of the fact that a copious number of drugs are available in the market and have been routinely used against various diseases, the fight between humans and the emerging new diseases are ongoing and will be so for the foreseeable future. Thus, there is an immense need for the advancement in the drug discovery process and that is being met by the latest approach of drug designing i.e. Insilco approach which is also known as Computer aided drug design, which saves both time and money. In the contemporary landscape, the adoption of computer-aided drug discovery (CADD) techniques has become an imperative for leading pharmaceutical companies and research groups. It is particularly vital in the initial phases of drug discovery, with the aim of expediting the drug development process in a more cost-effective manner and reducing the likelihood of failures in the later stages. The utilization of rational drug design, a core aspect of CADD, follows a knowledge-driven approach that can furnish critical insights into the interaction patterns between proteins and ligands (complexes) and the determination of binding affinities. This approach is significantly empowered by the accessibility of supercomputing, parallel processing capabilities, and advanced software tools, which have substantially accelerated lead identification in the field of pharmaceutical research. The ultimate objective of the drug discovery process is to identify novel drug molecules that can bind to specific disease-associated targets and modify the target's function, as elucidated by Baig et al. in 2014. Computer-Aided Drug Design, in silico approaches, have found extensive application in both Lead Identification and Lead Optimization stages of drug development against a variety of targets over the years. The

classification of Drug Design can be broadly categorized into two types: Analog-based studies and Structure-based studies, depending on the availability of three-dimensional structural information of the target.

1. **Analog/Ligand Based Studies:** Analog-based studies involve the collection of data from established drugs or ligands that exhibit activity against a target biological molecule, be it a protein or DNA/RNA. Using this gathered information, a set of guidelines is formulated to either create a new ligand or make adjustments to an existing one, with the aim of augmenting its biological activity, as outlined by Sandala in 2013.

2. **Structure Based Studies:** Structure-based methods, relying on the three-dimensional structure of the target, surmount several of the constraints associated with analog-based investigations. These techniques facilitate the creation of a comprehensive theoretical model for protein-ligand interactions, which in turn enables the proactive design of new lead compounds tailored for a specific biological target, as demonstrated by Marrone et al. in 1997. A notable example of a success story in structure-based design is the development of the antihypertensive drug captopril, which acts as an inhibitor of angiotensin-converting enzyme (ACE).

   These two kinds of CADD techniques have developed independently throughout the years and kept becoming better. However, it has been shown that combining both ligand- and structure-based design methodologies is a more effective approach than using either strategy alone in the drug development process because both may be used to complement each other's advantages and disadvantages.

III. METHODS USED IN COMPUTER AIDED DRUG DESIGN AND DEVELOPMENT

Several fields, including chemical and structural biology, computational chemistry, organic synthesis, and pharmacology, are involved in the present drug discovery process scenario. As such, it consists of multiple stages:

- In order to study the functions and relationships with a particular disease, target identification entails the identification and isolation of individual targets (Anderson, 2003).
- The phase known as "target validation" occurs when a drug's target is connected to a disease of interest and its ability to control biological processes in the body after binding to a partner molecule
- Lead identification is the process of identifying a synthetic compound that exhibits some potency and specificity against a biological target and is thought to have the potential to develop into a medication that can treat the desired condition.
- Lead optimization involves evaluating lead compounds and their analogs iteratively in order to increase potency and other important features (Macalino et al., 2015). Therefore, to identify and rank candidates with the best chance of developing into a safe and effective medication, both in vitro and in vivo studies are carried out.Additionally, pharmacokinetic and pharmacodynamic features that can be applied to analogs that will be synthesized for evaluation are determined through the development of structure-activity relationships, or SARs (Andricopulo et al., 2009).
- Drug synthesis and formulation research, in vivo animal investigations for potency and toxicity, and characterization of mechanistic toxicity are all part of the preclinical

stage (Macalino et al., 2015).

- Clinical trials are conducted on human volunteers and consist of three phases that evaluate the proposed drug's safety, potential side effects, dose, effectiveness, and pharmacokinetic and pharmacological qualities (Silverman and Holladay, 2014).

For the production or screening of putative ligands (modulators), insilico techniques to drug discovery are divided into structure-based and ligand-based (enzyme/receptor), followed by methodologies. The target's 3D structure is used in the structure-based method along with biological testing and optimization. On the other hand, the ligand-based approach involves developing theoretical predictive models by subjecting a collection of molecules with various structures to known potential cytocomputational modeling methodologies.These models are then utilized for virtual screening of a sizable chemical database to identify new chemical entities and for structural tweaking to improve potency (Macalino et al., 2015).

1. **Structure-Based Drug Design (SBDD):** In Structure-Based Drug Design (SBDD), insights gleaned from the binding site of a three-dimensional macromolecular structure are leveraged to craft and assess ligands based on their anticipated interactions with the protein binding site. Consequently, the initial crucial steps in SBDD involve the identification of a valid drug target and the acquisition of its structural data. In this context, research in structural and computational biology has played an instrumental role in generating protein structures through techniques such as X-ray crystallography, nuclear magnetic resonance (NMR), cryo-electron microscopy (EM), homology modeling, and molecular dynamics (MD) simulations, as expounded upon by Macalino et al. in 2015.

SBDD can be categorized into two primary approaches: the de novo method and the virtual screening approach. De novo drug design harnesses information from the 3D receptor structure to identify small fragments that closely align with the binding site. These fragments are then linked together according to connection rules, ensuring synthetic feasibility, ultimately yielding a structurally novel ligand that can be synthesized for subsequent screening, as detailed by Macalino et al. in 2015. Conversely, virtual screening (VS) leverages available libraries of small molecules to identify compounds with specific bioactivity. These compounds may serve as replacements for existing ligands targeting specific biomolecules or be used to discover compounds for previously unexplored targets with available structural information, as described by Andricopulo et al. in 2009.

2. **Homology Modelling:** Homology modeling, also known as comparative protein modeling, is a technique that enables the construction of an atomic-resolution model for a "target" protein. This model is generated using the amino acid sequence of the target protein and the three-dimensional (3D) structure of a related homologous protein, referred to as the "template." Homology modeling involves the identification of one or more known protein structures that exhibit similarity to the structure of the query sequence. It further entails creating an alignment that maps the residues in the query sequence to those in the template sequence. Notably, research indicates that protein structures tend to be more conserved than protein sequences among homologous proteins. Even when sequences share less than 20% sequence identity, they can still have significantly different structures, as outlined by Ahmad in 2013.

Proteins related by evolution often have similar sequences, and naturally occurring homologous proteins tend to share common protein structures. Research findings have shown that the evolutionarily conserved protein three-dimensional structure exceeds expectations due to the conservation of sequences, making it possible to generate a structural model for the target using sequence alignment and template structure. Since protein structures exhibit a higher degree of conservation compared to DNA sequences, detectable levels of sequence similarity generally indicate substantial structural similarity, as highlighted by Ahmad in 2013. To perform this modeling, bioinformatics software tools are employed, utilizing the known 3D structures of templates as a basis for generating the 3D structure of the target, as elucidated by Park et al. in 2008. Notably, Modeller is a widely used tool in homology modeling, and the SWISS-model repository serves as a database of protein structures created through homology modeling, as described by Ahmad in 2013.

3. **Virtual High-Throughput Screening:** Virtual screening is a computational method utilized to assess extensive compound libraries for their potential to bind to specific sites on target molecules, such as proteins. Compounds that match well are then subjected to experimental testing. In the realm of drug discovery, virtual screening (VS) is an integral part of the process, serving as a rapid computational technique for systematically exploring large collections of chemical structures. The primary goal is to identify structures with a high likelihood of binding to a drug target, typically a protein receptor or enzyme, as documented by Ahmad in 2013.

   Virtual screening plays a pivotal role in drug discovery, even though the term "virtual screening" is relatively recent when compared to the broader and more established concept of database searching. Walters et al. define virtual screening as the "automated evaluation of extensive compound libraries" through the use of computer programs, as previously described by Bohacek et al. in 1996. As this definition implies, virtual screening deals with the challenge of sifting through the vast chemical space, which comprises over 10^60 possible compounds, to arrive at a manageable number that can be synthesized, procured, and experimentally evaluated.

   While addressing the entire chemical universe is an intriguing question, practical virtual screening scenarios are more focused on designing and enhancing targeted combinatorial libraries and enriching the sets of compounds available in in-house repositories or those offered by vendors. Virtual screening offers notable advantages, as it is a cost-effective alternative to High-Throughput Screening (HTS), operates at a faster pace than traditional screening methods, and can swiftly assess a large number of potential drug-like molecules. HTS, by itself, often involves a trial-and-error approach but can be effectively complemented by the application of virtual screening techniques, as explained by Ahmad in 2013.

4. **De novo Design:** The term "de novo" originates from Latin and signifies "from the beginning." The structural characterization of drug target active sites provides valuable insights into their binding characteristics. This insight into the composition of the active site and the spatial arrangement of various amino acids within the binding site serves as a foundation for designing ligands tailored to that specific target. De novo design methods make extensive use of computational tools capable of analyzing protein active sites and proposing potential compounds. Numerous promising approaches focused on ligand design have been documented, as highlighted by Aparoy

et al. in 2012. Additionally, Murcko conducted a comprehensive examination of computer-aided ligand design techniques, categorizing them into six major classes.

- **Fragment Location Methods:** help identify the ideal sites for atoms or tiny pieces inside the active site.

- **Site Point Connection Methods:** to identify sites (also known as "site points") and then insert fragments into the active site such that the appropriate atoms occupy those positions.

- **Fragment Connection Methods:** After fragments are placed, they are connected and held in a desired orientation with the help of "linkers" or "scaffolds.".

- **Sequential Buildup Methods:** Construct a ligand atom by atom, or fragment byfragment.

- **Whole Molecule Methods:** Different conformations of compounds are inserted into the active site to evaluate electrostatic complementarity and/or shape.

- **Random Connection Methods:** a unique class of approaches that incorporates bond disconnection techniques, randomness introduction techniques, and some of the characteristics of fragment connection and sequential accumulation methods.

- Over the years, a range of de novo methods, particularly whole molecule techniques such as docking, have been seamlessly incorporated into diverse fields, including chemistry, pharmacology, molecular biology, and computer modeling. The calculation of critical electrostatic and solvation terms necessary for accurately assessing binding energies is often challenging and time-consuming. However, advancements in algorithm sophistication are steadily enhancing the accuracy of these parameter approximations, as noted by Gane and Dean in 2000. It is evident from recent literature that the drug design process has evolved into an indispensable component of drug discovery initiatives.

Binding site prediction or identification

- According to Anderson (2003) and Kalyaanamoorthy & Chen (2011), the perfect binding site is a concave area with a variety of chemical functions that interact with a ligand to provide the desired effect (activation, modulation, or inhibition). Important information for SBDD is provided by proteins co-crystallized with their substrates or recognized inhibitors, as well as through mutation studies that pinpoint important residues for interaction. However, further investigations are required to carry out structure-based rational drug discovery when no information about the binding site is available (Macalino et al., 2015). Presently, numerous in silico methods have been documented in various scientific papers and are accessible for the identification of binding regions within proteins. Tools like PASS, Q-SiteFinder, LIGSITEcsc, SiteMap, FPocket, ConCavity, MED-SuMo, MDPocket, FTMAP, POOL, and many others can predict the binding site of a small molecule compound. Alternatively, for the identification of peptide binding sites, specific approaches such as PepSite, PeptiMap, and PEP-SiteFinder are available (Saladin et

al., 2014).

- Additionally, the challenge of detecting allosteric sites has been addressed through the development of open-access web servers, including SPACER. A prior report systematically evaluated a range of web servers and standalone programs for predicting protein-ligand binding sites. The report revealed that while these methods can be valuable for identifying potential binding sites, their predictive accuracy may depend on various factors, such as the similarity to templates and the size of the ligand (K. Chen et al., 2011).

5. **Molecular Docking:** In the field of molecular modeling, docking is a technique used to anticipate the preferred arrangement of one molecule when it binds to another, creating a stable complex. Molecular docking is specifically concerned with the binding of a ligand to its receptor or target protein. Its primary objective is the identification and enhancement of potential drug candidates by analyzing and simulating the molecular interactions between the ligand and the target macromolecules. The process of molecular docking involves generating various ligand conformations and orientations, after which the most suitable ones are selected. Numerous molecular docking tools are available for this purpose, such as ArgusDock, DOCK, FRED, eHITS, AutoDock, and FTDock.

   Within the domain of molecular modeling, scoring methods are employed to rank the affinity of ligands for binding to the active site of a receptor. In the context of virtual high-throughput screening, compounds are subjected to docking into the active site and subsequently evaluated based on their potential to bind tightly to the target macromolecule (A. Wadood et al., 2013).

6. **Molecular Dynamic (MD) Simulation:** Molecular dynamics is a highly effective procedure that hinges on simulating the motion of molecules. This simulation is achieved by solving Newton's equations of motion for each individual atom, incrementally updating the speed and position of each atom with small time steps. MD simulations offer an alternative approach to explore the configuration space, following the principles mentioned above. This method is often conducted at elevated temperatures, typically in the range of several hundred to a few thousand degrees. This enables the exploration of the local area surrounding the sampled point and overcoming relatively small energy barriers, usually on the order of a few tens of kJ/mol. The generation of configurations in MD simulations can be achieved through the selection of appropriate configurations at specific intervals during the simulation, thus focusing on minimizing these particular structures. MD methods harness the inherent dynamics of the system to explore low-energy deformation modes and are well-suited for sampling the conformational space of large, confined systems (A. Wadood et al., 2013).

7. **Ligand- Based Drug Design:** In situations where the three-dimensional structure of the target protein is unavailable, data derived from a collection of active ligands that interact with a relevant target, such as a receptor or enzyme, can be utilized to identify critical structural and physicochemical characteristics (referred to as molecular descriptors) responsible for the observed biological activity. In this context, the assumption is that compounds with similar structural features exhibit analogous biological responses and interactions with the target. It is essential for the compound dataset to span a broad

concentration range, ideally covering at least four orders of magnitude, to establish a robust ligand-based screening model. Common approaches in ligand-based design encompass techniques such as quantitative structure-activity relationships (QSARs) and pharmacophore-based methods (Macalino et al., 2015).

IV. QUANTITATIVE-STRUCTURE ACTIVITY RELATIONSHIP (QSAR)

The foundation of QSAR research is the idea that alterations in a set of chemicals' bioactivity are linked to structural and molecular changes (Macalino et al., 2015). This correlation is used to create a statistical model that can be used to theoretically anticipate the biological properties of new chemicals (Melo-Filho et al., 2014). A few limitations must be met in order to produce a trustworthy QSAR model: The bioactivity data should meet certain requirements. Firstly, there should be a minimum of 20 compounds with activity in the data set. Secondly, the compounds should be selected appropriately for the training and test sets. Thirdly, the ligands' molecular descriptors should not exhibit autocorrelation to prevent over-fitting. Finally, the model should be validated through internal and/or external validation to ascertain its predictability and applicability (Macalino et al., 2015). According to Cramer et al. (1988), comparative molecular field analysis (CoMFA), Even after more than three decades since its inception, the 3D-QSAR method remains one of the most widely utilized techniques. Recent developments in the realm of 3D-QSAR have introduced additional strategies, including Topomer CoMFA (Cramer, 2003), Spectral Structure-Activity Relationship (S-SAR), the adaptation of fields for molecular comparison (AFMoC), and Comparative Residue Interaction Analysis (CoRIA). Notwithstanding its significant achievements in the field of drug discovery, 3D-QSAR is still beset by several limitations that can potentially be resolved through the adoption of more advanced multidimensional QSAR approaches, such as 4D, 5D, and 6D-QSAR. 4D-QSAR, for instance, was developed to address issues related to ligand conformation and orientation within the target binding site, while 5D-QSAR accounts for factors like receptor flexibility and induced fit effects (Macalino et al., 2015). Lastly, 6D-QSAR takes into consideration solvation effects due to their pivotal role in receptor-ligand interactions (Damale et al., 2013). With advancements in computational power and software performance, efforts have been directed toward enhancing QSAR model development and validation through techniques like Discovery Bus (Cartmell et al., 2005) and Auto QSAR. These methods enable the objective discovery, updating, and validation of numerous highly predictive statistical models by continually integrating new machine learning agents and descriptors into the system (Macalino et al., 2015).

1. **Pharmacophore Modeling:** The goal of pharmacophore screening is to find molecules with distinct scaffolds but comparable three-dimensional configurations of important functional groups that interact (Vuorinen & Schuster, 2015). By taking use of the bioactive conformation of potential drugs, binding site information can be added to the pharmacophore model. In the molecular alignment phase of QSAR modeling investigations, pharmacophore modeling is also frequently carried out (Melo-Filho et al., 2014). A number of frequently utilized software applications for the automatic production of pharmacophores are MOE, PHASE b), LigandScout, and Discovery Studio. Numerous reviews have already been done on these programs and other algorithms (Vuorinen & Schuster, 2015). In order to decrease false negative and false positive outcomes, respectively, it is crucial to create a model with carefully calibrated sensitivity and specificity. Areas where inactive compounds are present can utilize spatial limitations, which can be adjusted to prevent the model from becoming overly restricted. Moreover,

characteristics that aren't regularly seen in active compounds ought to be eliminated from the model or made optional.To ascertain a model's sensitivity and specificity against an external test set, validation tests must be carried out following model refinement (Vuorinen & Schuster, 2015). Sequence-derived 3D pharmacophore models can be produced even in the lack of both receptor 3D data and a selection of active ligands. based on the idea that comparable ligands and receptors can bind together. Pharma3D can identify shared sequence patterns for ligand biomolecular recognition in protein families and build a single-feature pharmacophore database by using homology models and 3D crystal structures. This strategy has been effectively used in virtual GPCR (family A) screening. Theoretically, sequences with identical motifs can identify the same ligand functionality at a comparable geographical region. This allows for the identification of the sequence patterns associated with particular single-feature pharmacophores, which are then used to create the 3D pharmacophore model. Despite its drawbacks, this is a compelling method for therapeutic targets for which there is little to no knowledge on receptors and ligands (Macalino et al., 2015).

2. **Compound Selection:** Before moving on to the in vitro and in vivo evaluation of lead potency, compound selection, sometimes known as "cherrypicking," is typically carried out. One of the major drawbacks of the existing scoring systems is that they rank compounds erroneously, which makes it harder to determine whether hits are indeed hits, as was described in the Docking and Scoring section. An other method is to evaluate the interaction of a ligand in the binding pocket and find compounds exhibiting comparable interaction patterns, in addition to rescoring poses and generating a consensus list of hits (Macalino et al., 2015). For the purpose of pose grouping and rating of virtual screening hits, the Automatic analysis of Poses using Self-Organizing Map (AuPosSOM) (Bouvier et al., 2010) can be utilized. Their approach is predicated on the idea that a target and an active molecule must make precise contact in order for the intended activity to manifest. Furthermore, clustering is frequently carried out in order to choose a representative chemical for each cluster and look for shared scaffolds among the hit compounds. A more economical and time-effective way to explore a broad chemical space is through the evaluation of structurally varied molecules (Vuorinen & Schuster, 2015).

3. **Specificity Target:** One essential factor in the quest for effective medications is specificity. Drug discovery and development are hampered by the frequent occurrence of false positives caused by aggregation, ligand promiscuity, and chemical reactivity that are still seen throughout the experimental evaluation of lead candidates. While reactive compounds are found by utilizing reactive group filters to increase hit list quality, surfactants are used in screening experiments to prevent compound aggregation (Roche et al., 2002). Furthermore, compounds that frequently perform well in different HTS studies can be eliminated using the pan-assay interference compounds (PAINS) filter, eliminating them from usage as particular inhibitors.

   However, since PAINS was created using a small sample size of HTS data and might not be appropriate for application in other screening experiments, its use should be carefully reviewed (Macalino et al., 2015).

V. MOLECULAR PROPERTIES PREDICTION

Molecular properties constitute a delicate equilibrium of various structural characteristics that determine a molecule's similarity to known drugs. In the development of orally administered drugs, achieving good drug absorption and appropriate drug delivery holds paramount importance. Molecular structure underpins physicochemical properties, drug metabolism, pharmacokinetics (DMPK), and toxicity characteristics (Mannhold & Kubinyi, 2009). A staggering 30% of oral drugs experience failure during development due to unfavorable pharmacokinetics, primarily resulting from low and highly variable bioavailability. High oral bioavailability becomes a crucial consideration in the development of bioactive therapeutic agents (Sandala, 2013).

Consequently, predicting bioavailability-related properties such as solubility, lipophilicity, optimal drug absorption, low polar surface area, hydrogen bond donors and acceptors counts, molecular weight, and LogP (partition coefficient) is essential before embarking on actual synthesis. Creating an in-silico model for forecasting oral bioavailability becomes imperative, striking the right balance between solubility and partitioning properties. Various software tools like Mol inspiration, Osiris, MolSoft, and ALOGPS enable the calculation of molecular properties, lipophilicity, and solubility parameters (Sandala, 2013).

Oral drug-likeness encompasses several key attributes:

- **Oral Bioavailability:** Ensuring that a compound can be effectively absorbed via the oral route.
- **Appropriate Toxicity:** Meeting the criteria for proceeding to phase-I clinical trials. Minimal interaction with therapeutic targets: Avoiding excessive potency.
- **Aqueous Solubility:** Facilitating dissolution in bodily fluids.
- **Permeability:** The ability to penetrate biological membranes.
- **Pharmacokinetic viability:** Suitability for the body's processes.
- **Blood-Brain Barrier Permeability:** The capacity to cross the blood-brain barrier.

The "Rule-of-5," derived from the World Drug Index, has widely been accepted as a practical guideline for oral drug likeness. It specifies criteria such as molecular weight (MW) $\leq$ 500, ClogP (logarithm of the partition coefficient) $\leq$ 5, hydrogen bond donors $\leq$ 5, and hydrogen bond acceptors $\leq$ 10. In general, lead-like properties tend to be less stringent, with suggested thresholds of MW < 350 and ClogP < 3 as good starting points for leads. For the screening of small fragments, a "Rule-of-3" has been proposed, recommending that promising lead fragments should meet criteria like MW $\leq$ 300, ClogP $\leq$ 3, $\leq$ 3 hydrogen bond donors and acceptors, $\leq$ 3 rotatable bonds, and a polar surface area (PSA) $\leq$ 60.

Molecular size and hydrogen bonding significantly influence logP. Octanol-water partition (logP) and distribution (logD) coefficients are frequently employed to

estimate membrane penetration, permeability, and gastrointestinal absorption (Sandala, 2013).

VI. ABSORPTION, DISTRIBUTION, METABOLISM, EXCRETION, AND TOXICITY (ADMET) TEST

Determining the ADMET qualities of leads during the initial phases of drug screening becomes necessary due to high attrition rates caused by subpar pharmacokinetic profiles. On the other hand, it is not cost-effective or time-efficient to evaluate the pharmacokinetic properties of millions of chemicals experimentally. Thus, virtual screening can be used to filter hits and exclude compounds with unwanted properties in order to rapidly evaluate a lead compound's drug-likeness before doing comprehensive experimental testing (Bajorath, 2002). In silico ADMET filters are used to predict drug-like features of substances, much as QSAR. They are created from chemical or molecular descriptors. The most basic and well-known models include Veber rules (Veber et al., 2002), the Rule of Three for Fragments (Congreve et al., 2003), and the Lipinski Rule of Five (Lipinski et al., 1997). A sizable chemical database or a list of possible leads can be filtered using publicly accessible web servers such as ChemBioServer and Free ADMET Filtering-Drugs2 (FAF-Drugs2) (Lagorce et al., 2008) and (Lagorce et al., 2011). ChemBioServer can do substructure search, cluster compounds, show and graph molecular properties, filter compounds according to various chemical properties, steric clashes, and toxicity, and suggest a representative for each group. As an alternative, the user can select from a number of pre-defined filters in FAF-Drugs2, including the ones listed above as well as others like the central nervous system (CNS) filter and reactive group filter. Beyond these, molecules with unfavorable moieties can also be found using pharmacophore models made from toxicity-causing inhibitors. Reactivity models, like the ones used in SMARTCyp, are useful in addressing the problem of drug metabolism. According to Macalino et al. (2015), SMARTCyp is a free web service and downloadable tool that identifies locations in 2D compound structures that are most likely to experience Phase I CYP450-mediated metabolism.. In order to identify potential sites of metabolism, it computes the reactivity of ligand fragments using quantum chemical calculations and the accessibility of atoms in the molecule. As an alternative, MetaSite also uses a similar technique to find putative metabolic reactivity sites; however, it uses the compound's 3D configuration as the query input. It is important to have in mind that while using these in silico ADMET models, the tools are more useful for qualitative hit or compound set analysis than for precise quantitative value prediction (Macalino et al., 2015). These techniques are useful for ranking a class of compounds that have been identified for assessment or evaluation of a certain descriptor and SAR in vitro or in vivo (Gleeson & Montanari, 2012).

VII. DISCUSSION

Medications are substances that can either prevent diseases or aid in the restoration of health to individuals afflicted with ailments, making them an indispensable component of modern medicine. In the distant past, the process of designing a new drug through the modification of the molecular structure of an existing one was a slow, trial-and-error endeavor. However, today's technology allows a computer to visualize the molecular structure of any drug from a vast database. Even minor molecular adjustments can lead to significant alterations in the original drug, affecting aspects such as absorption, metabolism, half-life, therapeutic efficacy, and potential side effects. Before substantial time and resources are invested in extensive testing, computers can also predict which chemicals are unlikely to be

effective in treating a specific disease.

Leveraging computers to manipulate chemicals at the molecular level and create novel drugs is rooted in the field of molecular pharmacology. This field is concerned with studying the chemical compositions of drugs and their interactions at the molecular level within cells, including the DNA within the nucleus. In the conventional drug discovery process, new compounds are typically synthesized through a laborious, multi-step procedure and subjected to a battery of in vivo biological assays. Promising candidates are further examined for their pharmacokinetic properties, metabolism, and potential toxicity (Anh Vu et al., 2015).

Top of Form

The process of drug discovery and development is highly intricate and time-consuming, with numerous factors contributing to the failure of various drugs, such as ineffectiveness, side effects, poor pharmacokinetics, and market viability. Over the past three and a half decades, the expenses associated with this process have significantly increased. Consequently, computer-aided drug design (CADD) techniques are now widely adopted in the pharmaceutical industry to expedite the process. The utilization of computational tools during the lead optimization phase of drug development offers substantial cost benefits.

On average, it takes 10-15 years and an investment of US $500-800 million to bring a drug to market, with the synthesis and testing of lead analogs representing a significant portion of these costs (Basak, 2012). Therefore, leveraging computational tools in the hit-to-lead optimization stage is advantageous, as it allows for the exploration of a broader chemical space while reducing the number of compounds that need to be synthesized and experimentally tested. The computational optimization of a hit compound encompasses a structure-based analysis of docking poses and energy profiles for analogs, ligand-based screening for compounds with similar chemical structures or improved predicted biological activity, and the prediction of favorable binding affinity or optimization of drug metabolism and pharmacokinetics (DMPK) or absorption, distribution, metabolism, excretion, and potential toxicity (ADMET) properties.

The relatively lower cost of CADD in comparison to the chemical synthesis and biological characterization of compounds makes these methods attractive for focusing, streamlining, and diversifying the chemical space exploration (Sujit. G, n.d.).

Till now many drug and drug candidates have been discovered through insilico method against various diseases. Even some drugs are repurposed via drug repurporsing method and is being used for more than one disease which saves time and also cost as it had already gone through various procedures insilico as well as clinical and thus few methods can be skipped. In the recent days with the advancement of technologies insilico approach can be more effectively and efficiently used in drug designing resulting in discovery of drugs and fighting against the newly emerging disease. In this study we could found various insilico designed drugs starting from captropil against hypertension, and others like imatinib for leukemia, boceprevir for hepatitis C leading to the present-day repurposed drugs against the corona virus such as Sofasuvir,Remdisivir.

VIII. CONCLUSION

The process of finding and developing new drugs is multidisciplinary, costly, and time-consuming. The process of creating new bioactive compounds has evolved according to scientific developments. In silico methods have been made possible by developments in parallel hardware support and computational methodologies. Computational methods have relentlessly helped in designing new, safe and effective therapeutics. Successful implementation of software-based techniques has provided an opportunity to identify in vitro biologically active agents without much effort. Since the development of insilico methods, many drugs have been designed. The in-silico methods has led to the development of new drugs in mitigating the new emerging diseases within a short span of time and cost. As such captopril was the first ligand-based ACE inhibitor designed through in silico. Later many other novel drugs have been developed. Thus, insilico methods have reduced the problems of drug designing by saving time and money and led to discovery of many drugs and drugleads.

REFERENCES

[1] Ahmad, A. (2013). In-silico drug design: An approach which revolutionarised the drug discovery process. *Oa Drug Design and Delivery*,*1*.

[2] Anderson, A. C. (2003). The process of structure-based drug design. *Chemistry & Biology*, *10*(9), 787–797.https://doi.org/10.1016/j.chembiol.2003.09.002

[3] Andricopulo, A. D., Salum, L. B., & Abraham, D. J. (2009). Structure-based drug design strategies in medicinal chemistry. *Current Topics in Medicinal Chemistry*, *9*(9), 771–790. https://doi.org/10.2174/156802609789207127

[4] Anh Vu, L., Thi Cam Quyen, P., & Thuy Huong, N. (2015). In silico Drug Design: Prospective for Drug Lead Discovery. *International Journal of Engineering Science Invention*, *4*(10), 60–70.www.ijesi.org

[5] Aparoy, P., Reddy, K. K., & Reddanna, P. (2012). Structure and ligand based drug design strategies in the development of novel 5- LOX inhibitors. *Current Medicinal Chemistry*, *19*(22), 3763–3778.https://doi.org/10.2174/092986712801661112

[6] Baig, M. H., Ahmad, K., Adil, M., Khan, Z., Khan, M., Lohani, M., Khan, M., & Kamal, M. (2014). Drug Discovery and In Silico Techniques: A Mini-Review. *Enzyme Engineering*. https://doi.org/10.4172/2329-6674.1000123

[7] Bajorath, J. (2002). Integration of virtual and high-throughput screening. *Nature Reviews. Drug Discovery*, *1*(11), 882–894.https://doi.org/10.1038/nrd941

[8] Basak, S. C. (2012). Chemobioinformatics: the advancing frontier of computer-aided drug design in the post-genomic era. In *Current computer-aided drug design* (Vol. 8, Issue 1, pp. 1–2).https://doi.org/10.2174/157340912799218507

[9] Bohacek, R. S., McMartin, C., & Guida, W. C. (1996). The art and practice of structure- based drug design: A molecular modeling perspective. *Medicinal Research Reviews*, *16*(1), 3–50.https://doi.org/10.1002/(SICI)1098-1128(199601)16:1<3::AID-MED1>3.0.CO;2-6

[10] Bouvier, G., Evrard-Todeschi, N., Girault, J.-P., & Bertho, G. (2010). Automatic clustering of docking poses in virtual screening process using self-organizing map. *Bioinformatics (Oxford, England)*, *26*(1), 53–60.https://doi.org/10.1093/bioinformatics/btp623

[11] Cartmell, J., Enoch, S., Krstajic, D., & Leahy, D. (2005). Automated QSPR through Competitive Workflow. *Journal of Computer-Aided Molecular Design*, *19*, 821–833. https://doi.org/10.1007/s10822-005-9029-8

[12] Chen, K., Mizianty, M. J., Gao, J., & Kurgan, L. (2011). A critical comparative assessment of predictions of protein-binding sites for biologically relevant organic compounds. *Structure(1993)*, *19*(5), 613–621. https://doi.org/10.1016/j.str.2011.02.015

[13] Congreve, M., Carr, R., Murray, C., & Jhoti, H. (2003). A "rule of three" for fragment-based lead discovery? *Drug Discovery Today*, *8*(19), 876–877. https://doi.org/10.1016/s1359- 6446(03)02831-9

[14] Cramer, R. D. (2003). Topomer CoMFA: A Design Methodology for Rapid Lead Optimization. *Journal of Medicinal Chemistry*, *46*(3),374–388. https://doi.org/10.1021/jm020194o

[15] Cramer, R. D., Patterson, D. E., & Bunce, J. D. (1988). Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *Journal of the American Chemical Society*, *110*(18), 5959–5967.https://doi.org/10.1021/ja00226a005

[16] Damale, M., Harke, S., Khan, F., Shinde, D., & Sangshetti, J. (2013). Recent advances in multidimensional QSAR (4D-6D): A critical review. *Mini Reviews in Medicinal Chemistry*, *14*.https://doi.org/10.2174/13895575113136660104

[17] Danishuddin, M., & Khan, A. U. (2015). Structure based virtual screening to discover putative drug candidates: Necessary considerations and successful case studies. *Methods*, *71*(C), 135–145.https://doi.org/10.1016/j.ymeth.2014.10.019

[18] Foote, B. S., Spooner, L. M., & Belliveau, P. P. (2011). Boceprevir: Un inhibidor de proteasa para el tratamiento de hepatitis C crónica. *Annals of Pharmacotherapy*, *45*(9), 1085–1093.https://doi.org/10.1345/aph.1P744

[19] Gane, P. J., & Dean, P. M. (2000). Recent advances in structure-based rational drug design. *Current Opinionin Structural Biology*, *10*(4), 401–404. https://doi.org/https://doi.org/10.1016/S0959-440X(00)00105-6

[20] Gleeson, M., & Montanari, D. (2012). Strategies for the generation, validation and application of in silico ADMET models in lead generation and optimization. *Expert Opinion onDrugMetabolism&Toxicology*,*8*, 1435–1446. https://doi.org/10.1517/17425255.2012.711317

[21] Green, D. (2003). Virtual Screening of Virtual Libraries. *Progress in Medicinal Chemistry*, *41*, 61–97.https://doi.org/10.1016/S0079-6468(02)41002-8

[22] Kalyaanamoorthy, S., & Chen, Y.-P. P. (2011). Structure-based drug design to augment hit discovery. *Drug Discovery Today*, *16*(17),831–839. https://doi.org/https://doi.org/10.1016/j.drudis.2011.07.006

[23] Kapetanovic, I. M. (2008). Computer-aided drug discovery and development (CADDD): In silico-chemico-biological approach. *Chemico-Biological Interactions*, *171*(2), 165–176. https://doi.org/10.1016/j.cbi.2006.12.006

[24] Lagorce, D., Maupetit, J., Baell, J., Sperandio, O., Tuffery, P., Miteva, M., Galons, H., & Villoutreix, B. (2011). The FAF-Drugs2 server: A multistep engine to prepare electronic chemical compound collections. *Bioinformatics (Oxford, England)*, *27*, 2018–2020. https://doi.org/10.1093/bioinformatics/btr333

[25] Lagorce, D., Sperandio, O., Galons, H., Miteva, M. A., & Villoutreix, B. O. (2008). FAF- Drugs2: free ADME/tox filtering tool to assist drug discovery and chemical biology projects. *BMC Bioinformatics*, *9*, 396.https://doi.org/10.1186/1471-2105-9-396

[26] Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.*, *23*,3.

[27] Macalino, S. J. Y., Gosu, V., Hong, S., & Choi, S. (2015). Role of computer-aided drug design in modern drug discovery. *Archives of Pharmacal Research*, *38*(9), 1686–1701. https://doi.org/10.1007/s12272-015-0640-5

[28] Marrone, T. J., Briggs James M., and, & McCammon, J. A. (1997). STRUCTURE-BASED DRUG DESIGN:Computational Advances. *Annual Review of Pharmacology and Toxicology*, *37*(1), 71–90.https://doi.org/10.1146/annurev.pharmtox.37.1.71

[29] Melo-Filho, C. C., Braga, R. C., & Andrade, C. H. (2014). 3D-QSAR approaches in drug design: perspectives to generate reliable CoMFA models. *Current Computer-Aided Drug Design*, *10*(2), 148–159. https://doi.org/10.2174/1573409910666140410111043

[30] Oa, L., London, P., Commons, C., & License, A. (2013). *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.1*(1).

[31] Park, D., Kim, J., Lee, Y., & Ahn, C. (2008). QSID Tool: a new three-dimensional QSAR environmental tool. *Journal of Computer-Aided Molecular Design*, *22*, 873–883. https://doi.org/10.1007/s10822-008-9219-2

[32] Pozzan, A. (2006). Molecular Descriptors and Methods for Ligand Based Virtual High Throughput Screening in Drug Discovery. *Current Pharmaceutical Design*, *12*, 2099–2110. https://doi.org/10.2174/138161206777585247

[33] Rasheed, O. F. A.-. (2019). *Molecular Docking Studies for Design, Synthesis and Characterization of New Imatinib Analogues*. *December*.https://doi.org/10.35248/2153 2435.19.10.616

[34] Roche, O., Schneider, P., Zuegge, J., Guba, W., Kansy, M., Alanine, A., Bleicher, K., Danel, F., Gutknecht, E.-M., Rogers-Evans, M., Neidhart, W., Stalder, H., Dillon, M., Sjögren, E., Fotouhi, N., Gillespie, P., Goodnow, R., Harris, W., Jones, P., … Schneider, G. (2002). Development of a Virtual Screening Method for Identification of "Frequent Hitters" in Compound Libraries. *Journal of Medicinal Chemistry*, *45*(1), 137–142. https://doi.org/10.1021/jm010934d

[35] Saladin, A., Rey, J., Thévenet, P., Zacharias, M., Moroy, G., & Tufféry, P. (2014). PEP-SiteFinder: a tool for the blind identification of peptide binding sites on protein surfaces. *Nucleic Acids Research*, *42*(Web Server issue), W221-6. https://doi.org/10.1093/nar/gku404

[36] Sandala, A. (2013). *Rational design and synthesis of benzimidazole derivatives for the evaluation of*

biological activity.http://hdl.handle.net/10603/21288

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

[38] Vetrichelvan, T., Jegadeesan, M., & Devi, B. (2002). Antidiabetic Activity of Alcoholic Extract of Celosia argentea LINN. Seeds in Rats. *Biological & Pharmaceutical Bulletin*, *25*, 526–528.https://doi.org/10.1248/bpb.25.526

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

