"COMPUTER AIDED DRUG DESIGN, ITS SUCCESS AND LIMITATION"

Submitted in partial fulfillment of the requirements for the degree of Bachelor of Pharmacy
by

BENGAL GAJANAN BHUJANGRAO (Roll No.17PH05)
DABIR SAFA RAFIQUE (Roll No.17PH07)
KARBARI TUBA SALIM (Roll No.17PH16)
SAYYED GUFRAN ALI MEER AJAGAR ALI (Roll No. 17PH42)

Supervisor
PROF. NIKHAT FARHANA ANSARI
Department of Pharmaceutical chemistry ,
School of Pharmacy,

Anjuman-I-Islam's Kalsekar Technical Campus Plot No. 2 3,
Sector -16, Near Thana Naka, Khanda Gaon,
New Panvel, Navi Mumbai. 410206
Academic Year: 2020-2021

CERTIFICATE

Department of Pharmaceutics,
School of Pharmacy,
Anjuman-I-Islam's Kalsekar Technical Campus
Khanda Gaon, New Panvel, Navi Mumbai, 410206

This is to certify that the project entitled " **COMPUTER AIDED DRUG DESIGN , ITS SUCCESS AND LIMITATIOM** " is a bonafide work of Dabir Safa Rafique (Roll No.17PH07) submitted for the appreciation of the degree of Bachelor of Pharmacy in Department of Pharmaceutics.

Name of supervisor: **Prof. NIKHAT ANSARI**

Dean **Director**

APPROVAL FOR BACHELOR OF PHARMACY

This project entitled "**COMPUTER AIDED DRUG DESIGN** " by **Dabir Safa Rafique** is approved for the degree of Bachelor of Pharmacy in Department of Pharmaceutical chemistry.

Examiners

Supervisors

DECLARATION

I hereby declare that this written submission represents my ideas in my own words and I have adequately and referenced the original sources. I also declare that I have adhered to all the principles of academic honesty and integrity, and have not misrepresented or fabricated or falsified any idea/fact/data/source in my submission. I understand that any violation of the above will cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not properly cited or from whom proper permission has not been taken when needed.

Student Name: **DABIR SAFA RAFIQUE**

Roll No: **17PH07**

ACKNOWLEDGEMENT

I would like to take this opportunity to express my sincere gratitude to our guide **Prof. NIKHAT ANSARI**, Department of Pharmaceutical chemistry , AIKTC, School of Pharmacy, New Panvel; for her valuable support and guidance throughout my project work.

I am grateful for her encouragement and innovative ideas that helped me to complete the project more efficiently.

I would also like to thank Dr. Abdul Razzak Honnutagi, Director of AIKTC, New Panvel, for providing us an outstanding academic environment.

I am thankful to Dr. Shariq Syed, Dean, School of Pharmacy, New panvel and all my B.Pharm. faculties for providing advice and valuable guidance.

I am also thankful to all my family members whose constant support and encouragement in every aspect helped me to complete my project.

Student Name: **Dabir Safa Rafique**
Roll No: **17PH07**
Department of Pharmaceutics,
University of Mumbai.

COMPUTER ADDED DRUG DESIGN SUCCESS AND LIMITATION

1. Abstract

Modern drug discovery is characterized by the production of vast quantities of compounds and the need to examine these huge libraries in short periods of time
The need to store, manage and analyze these rapidly increasing resources has given rise to the field known as computer-aided drug design (CADD).

CADD represents computational methods and resources that are used to facilitate the design and discovery of new therapeutic solutions.

Digital repositories, containing detailed information on drugs and other useful compounds, are goldmines for the Study of chemical reactions capabilities. Design libraries, with the potential to generate molecular variants in their entirety, allow the selection and sampling of chemical compounds with diverse characteristics.

Structure-based drug design and ligand-based drug design
are two methods commonly used in computer-aided drug design. In this article, we discuss the theory behind both methods, as well as
their successful applications and limitations.

2. Background

Developing new drugs is a very expensive and time-consuming process that dates back millions of years to when only herbal remedies were in use [1]. Drugs with synthetic/semi-synthetic origins only came into existence in the last century [2]. Compounds developed prior to this time were not very effective in terms of potency or safety, and must therefore be optimized. In the era of trial-and-error processes, rational strategies were developed to improve the potency of compounds [3-6]. In the 1980s, the use of computers was extended from data handling to a more prominent role in drug discovery [7]. The use of computers in the field of pharmaceutical research is typically designated as computer-aided drug design (CADD) [8, 9]; although it is also referred to as computer-assisted molecular design (CAMD). CADD methods have emerged as an effective tool for drug discoveries. CADD is a specialized discipline that uses computational methods to simulate drug receptor interactions to determine if a given molecule will bind to a target, and if so, what its affinity would be [10]. This method has become the most widely used technique to significantly decrease the number of potential medicinal compounds from a large library by predicting which will be inactive and active. This method requires significantly less cost and time for high throughput screening without compromising the quality of lead discovery. Binding of ligands to the receptor may occur via hydrophobic, electrostatic, and hydrogen-bonding interactions [11]. In addition, solvation energies of the ligand and receptor site also play major roles in this process because partial to complete desolvation must occur prior to binding [12]. There are two major types of drug design techniques: ligand-based drug design (LBDD) and structure-based drug design (SBDD).

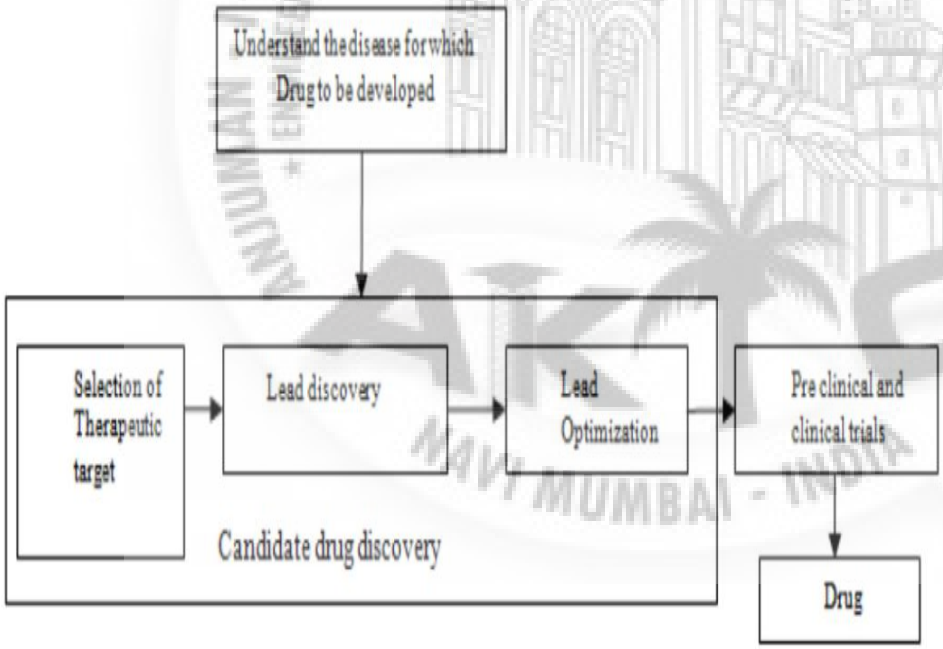
SBDD is used when the three-dimensional structures of target proteins are available, while LBDD design is employed in cases in which the structures are unknown. CADD methods are dependent on bioinformatics tools, applications and databases [13]. MD simulation has become one of the most influential tools to predict the conformation of small molecules, as well as for modelling conformational changes within a biological target upon binding by small molecules [14, 15]. Semi-empirical methods such as ab initio methods or density functional theory are most often used to provide the expected opt(electrostatic potential, polarizability, etc.) of the drug candidate that influence its binding affinity [16]. The advantages of using CADD in drug discovery include: a) cost savings; b) time-to-market, the predictive power of CADD facilitates selection of promising lead candidates, thereby preventing time from being wasted on dead end compounds; c) better insight, one of the intangible benefits of CADD is the deep insight that researchers acquire into drug receptor interactions. Computer-aided drug design may be used to identify hits using structure or ligand-based virtual screening, optimization of hit-to-lead for affinity and selectivity (SBDD, LBDD, etc.) and optimization of the other pharmaceutical properties of leads while maintaining its affinity. Fig. 1 shows different methods used in CADD. In this review, we provide detailed information regarding different methods used in CADD, as well as some of their major successes and limitations imized parameters for molecular mechanics calculations and to estimate important electronic properties

3. Introduction

Computer aided drug design (CADD) provides several tools and techniques that helps in various stages of drug design thus reducing the cost of research and development time of the drug. Drug discovery and developing a new medicine is a long, complex, costly and highly risky process that has few peers in the commercial world. This is why computer-aided drug design (CADD) approaches are being widely used in the pharmaceutical industry to accelerate the process. The cost benefit of using computational tools in the lead optimization phase of drug development is substantial. The cost and time invested by the pharmacological research laboratories are heavy during the various phases of drug discovery, starting from therapeutic target identification[1,2] , candidate drug discovery, drug optimization through pre clinical and extensive clinical experiments to assess the effectiveness and safety of newly developed drugs. The major pharmaceutical companies have invested heavily in the routine ultra-High Throughput Screening (uHTS) of vast numbers of drug-like' molecules. [3,4] In parallel with this, drug design and optimization increasingly uses computers for virtual screening. [5-7] Recent advancements in DNA microarray experiments explore thousands of genes involved in a disease can be used for gaining in depth knowledge about the disease targets, metabolic pathways and toxicity of the drugs. [8] The theoretical tools include empirical mo-lecular mechanics, quantum mechanics and, more recently, statistical mechanics. This latest advance has permitted explicit solvent effects to be incorporated. All this work is the availability of high quality computer graphics, largely supported on workstations . [9]

Two distinct categories of research are clearly distinguishable 1) Crystallography, NMR or homology modelling. A detailed molecular structure of the target macromolecule, the drug receptor, is known from x-ray. 2) Variable activity of otherwise similar molecules. The target receptor binding site has properties which can only be inferred from a knowledge of the both these types of approach. Drug Discovery Process Drug discovery is a series of processes which when followed identify the drug compounds for the effective treatment or control of disease targets. It starts with the screening of large number of chemical compounds to optimize the disease targets. It requires insight information about the structure of the drug receptor so that the drug molecules can be adjusted to the binding site.

Figure 1: Drug Discovery Process



Drug discovery process starts with understanding the disease for which the drug to be designed. It consists of the following steps.

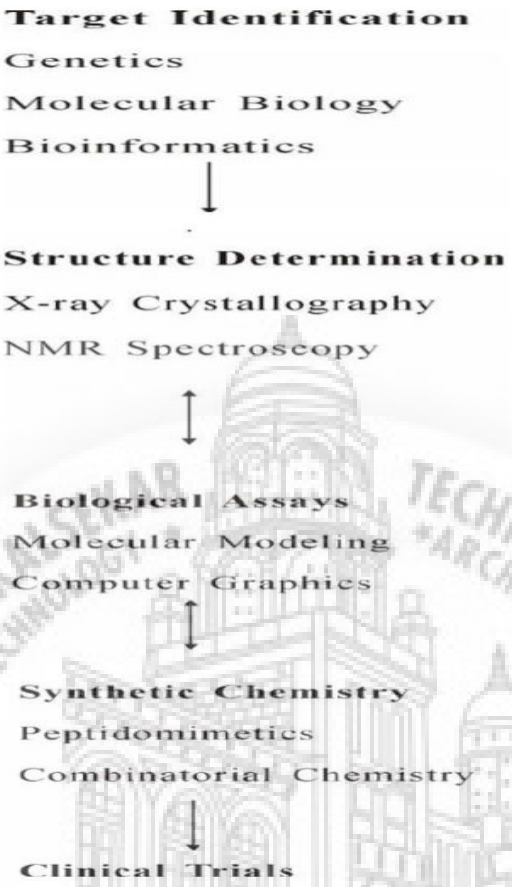
1. **Candidate Drug Discovery**
   - *Selection of Therapeutic Target*
   - *Lead Discovery*
   - *Lead Optimization*
2. **Pre clinical and clinical trials to evaluate the safety, efficacy and adverse effects of the drug**
   - *Animal Studies*
   - *Clinical Trials*
3. **FDA approval process for the newly discovered drug and bringing the drug to market for public use.**
   - *Additional post marketing testing*
   - *Furtherimprovement of the drug.*

In general, it takes 3-6 years for new drug discovery and pre-clinical development. The clinical trials can last up to 10 years or more before the product reaches the market.[10] Approximately it takes 12-15 years and costsmore than $1.3 billion to bring a successful drug to market.[11] On an average, among the 5000-10000 screened compounds about 250 compounds are selected for preclinical trials. From them only 5 survive to enter into clinical trials while only one approved by the FDA after strenuous review of the newly discovered drug.

CADD Strategies in the Drug Discovery Process

Strategies for CADD vary depending on the extent of structural and other information available regarding the target (enzyme/receptor) and the ligands. —Direct‖ and —indirect‖ design are the two major modeling strategies currently used in the drug design process. In the indirect approach the design is based on comparative analysis of the structural features of known active and inactive com- pounds. In the direct design the three- dimensional fea-tures of the target (enzyme/receptor) are directly con- sidered.

Working of CADD[13]



Preparation of a Target Structure

Success of virtual screening depends upon the amount and quality of structural information known about both the target and the small molecules being docked. The first step is to evaluate the target for the presence of an appropriate binding pocket.[12-13] This is usually done through the analysis of known target-ligand co-crystal structures or using in-silico methods to identify novel binding sites.[14]

A target structure experimentally determined through X-ray crystallography or NMR techniques and deposited in the PDB is the ideal starting point for docking. Structural genomics has accelerated the rate at which target structures are being determined. In the absence of experimentally determined structures, several successful virtual screening campaigns have been reported based on comparative models of target proteins[15-17]

Homology Modeling

In the absence of experimental structures, computational methods are used to predict the 3D structure of target proteins. Comparative modeling is used to predict target structure based on a template with a similar sequence, leveraging that protein structure is better conserved than sequence, i.e., proteins with similar sequences have similar structures. Homology modeling is a specific type of comparative modeling in which the template and target proteins share the same evolutionary origin. Comparative modeling involves the following steps: (1) identification of related proteins to serve as template structures, (2) sequence alignment of the target and template proteins, (3) copying coordinates for confidently aligned regions, (4) constructing missing atom coordinates of target structure, and (5) model refinement and evaluation. Fig. 1.4 illustrates the steps involved in homology modeling. Several computer programs and web servers exist that automate the homology modeling process e.g., PSIPRED[18] and MODELER.[19]

Molecular dynamics-based detection

The dynamic nature of biomolecules sometimes makes it insufficient to use a single static structure to predict putative binding sites. Multiple conformations of target are often used to account for structural dynamics of target. Classic molecular dynamic (MD) simulations can be used for obtaining an ensemble of target conformations beginning with a single structure. The MD method uses principles of Newtonian mechanics to calculate a trajectory of conformations of a protein as a function of time. Classic MD methods tend to get trapped in local energy minima. To overcome this, several advanced MD algorithms such as targeted- MD[20], conformational folding simulations[21], temperature accelerated MD simulations[22], and replica exchange MD[23] have been implemented for traversing multiple minima energy surface of proteins.

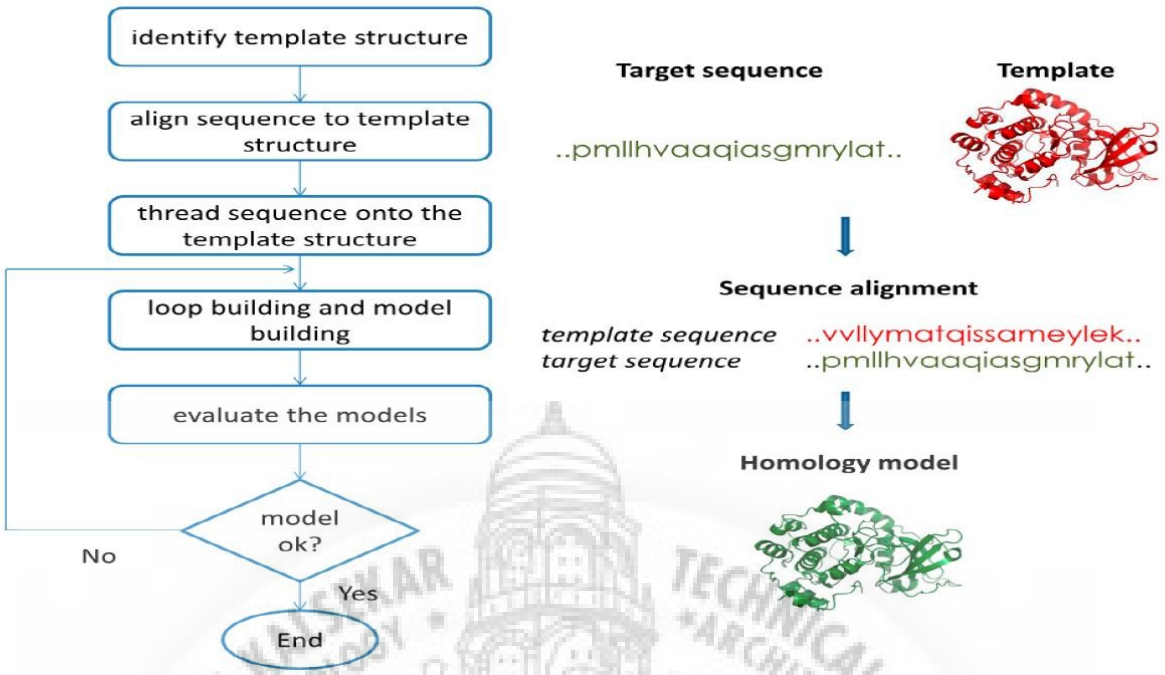


Figure: 1.4 Steps involved in homology model building process.38-39

Monte Carlo Search with Metropolis Criterion (MCM) Simulations

MCM samples conformational space faster than molecular dynamics in that it requires only energy function evaluation and not the derivative of the energy functions. Although traditional MD drives a system toward a local energy minimum, the randomness introduced with Monte Carlo allows hopping over the energy barriers, preventing the system from getting stuck in local energy minima. MCM simulations have been adopted for flexible docking applications such as in MCDOCK.[24]

Genetic Algorithms

Genetic algorithms introduce molecular flexibility through recombination of parent conformations to child conformations. In this simulated evolutionary process, the ―fittest‖ or best scoring conformations are kept for another round of recombination. In this way, the best possible set of solutions evolves by retaining favorable features from one generation to the next. In docking, a set of values that describe the ligand pose in the protein are state variable. State variables may include set of values describing translation, orientation, conformation, number of hydrogen bonds, etc.

The state corresponds to the genotype; the resulting structural model of the ligand in the protein corresponds to the phenotype, and binding energy corresponds to the fitness of the individual. Genetic operators may swap large regions of parent's genes or randomly change (mutate) the value of certain ligand states to give rise to new individuals. Genetic Optimization for Ligand Docking (GOLD)[25] explores full ligand flexibility with partial target flexibility using a genetic algorithm.

Scoring Functions for Evaluation of Protein Ligand Complexes

Docking applications need to rapidly and accurately assess protein-ligand complexes, i.e., approximate the energy of the interaction. A ligand docking experiment may generate hundreds of thousands of target-ligand

complex conformations, and an efficient scoring function is necessary to rank these complexes and differentiate valid binding mode predictions from invalid predictions.

Force-Field or Molecular Mechanics-Based Scoring Functions

Force-field scoring functions use classic molecular mechanics for energy calculations.[26] These functions use parameters derived from experimental data and *ab initio* quantum mechanical calculations. The binding free energy of protein-ligand complexes are estimated by the sum of van der Waals and electrostatic interactions. DOCK uses the AMBER force fields in which van der Waals energy terms are represented by the Lennard- Jones potential

Knowledge-Based Scoring Function

Knowledge based scoring functions use the information contained in experimentally determined complex structures. They are formulated under the assumption that interatomic distances occurring more often than average distances represent favorable contacts. On the other hand, interactions that are found to occur with lower frequencies are likely to decrease affinity. Several knowledge based potentials have been developed to predict binding affinity like potential of.

Consensus-Scoring Functions

Consensus approaches rescore predicted poses several times using different scoring functions. These results can then be combined in different ways to rank solutions.[31] Some strategies for combining scores include (1) weighted combinations of scoring functions, (2) a voting strategy in which cut-offs established for each scoring method is followed by decision based on number of poses a molecule has, (3) a rank by number strategy ranks each compound by its average normalized score values and (4) a rank by rank method sorts compounds based on average rank determined by individual scoring functions.[32]

Structure-Based Virtual High-Throughput Screening

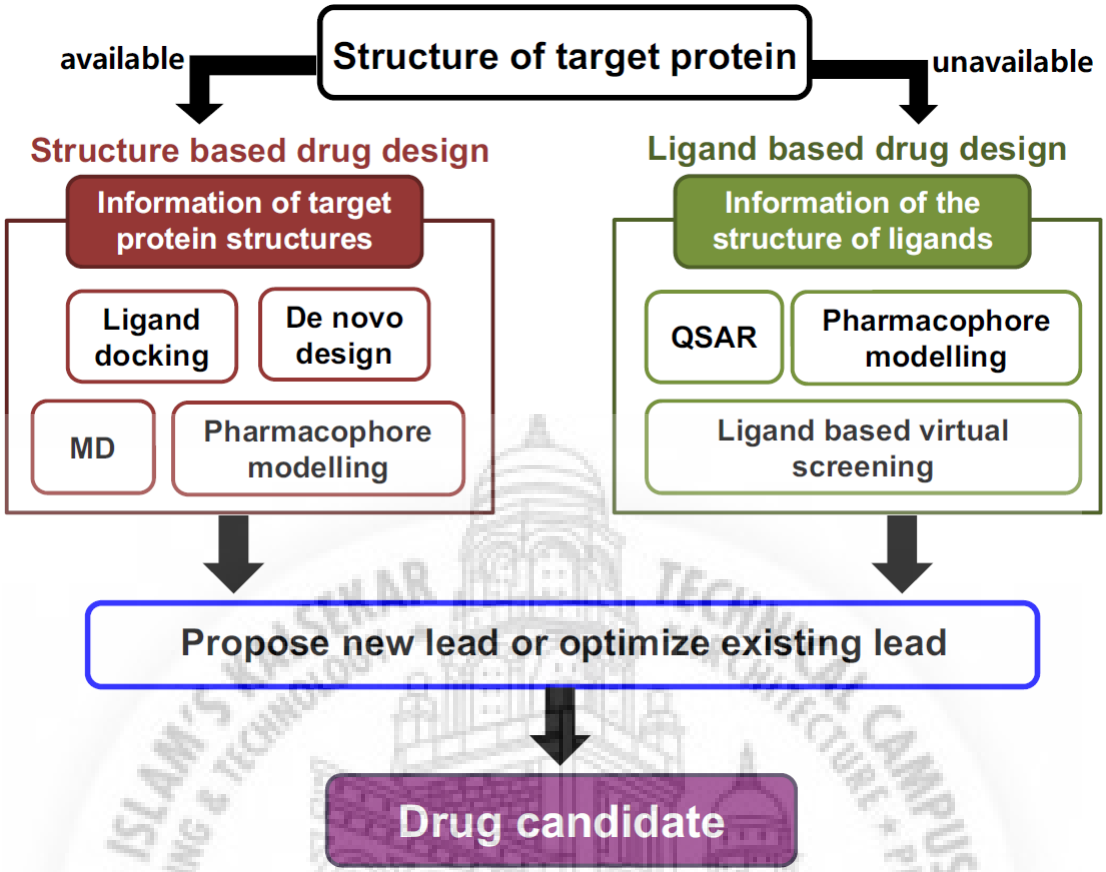
Structure-based virtual high-throughput screening (SB-vHTS), the *in silico* method for identifying putative hits out of hundreds of thousands of compounds to the targets of known structure, relies on a comparison of the 3D structure of the small molecule with the putative binding pocket. SB-vHTS selects for ligands predicted to bind a particular binding site as opposed to traditional HTS that experimentally asserts general ability of a ligand to bind, inhibit, or allosterically alter the protein's function. To make screening of large compound libraries within finite time feasible. SB-vHTS often uses limited conformational sampling of protein and ligand and a simplified approximation of binding energy that can be rapidly computed. The key steps in SB-vHTS are: (1) preparation of the target protein and compound library for docking, (2) determining a favorable binding pose for each compound and (3) ranking the docked structures.[33]

Ligand-Based Computer-Aided Drug Design

The ligand-based computer-aided drug discovery (LBDD) approach involves the analysis of ligands known to interact with a target of interest. These methods use a set of reference structures collected from compounds known to interact with the target of interest and analyse their 2D or 3D structures. The overall goal is to

for novel compounds possessing the biological activity of interest, hit-to-lead and lead-to drug optimization, and also for the optimization of DMPK/ADMET properties. LBDD is based on the similar property principle which states that molecules that are structurally similar are likely to have similar properties.[34] LBDD approaches in contrast to SBDD approaches can also be applied when the structure of the biological target is unknown. Additionally, active compounds identified by ligand -based virtual high-throughput screening (LB-vHTS) methods are often more potent than those identified in SB-Vhts.[35] represent these compounds in such a way that the physicochemical properties most important for their desired

interactions are retained, whereas extraneous information not relevant to the interactions is discarded. It is considered as an indirect approach to the drug discovery in that it does not necessitate knowledge of the structure of the target of interest. The two fundamental approaches of LBDD are (1) selection of compounds based on chemical similarity to known actives using some similarity measure or (2) the construction of a quantitative structure activity relationship (QSAR) model that predicts biological activity from chemical structure.



Molecular Descriptors

Molecular descriptors can include properties such as molecular weight, geometry, volume, surface areas, ring content, rotatable bonds, interatomic distances, bond distances, atom types, planar and nonplanar systems,indices, functional group composition, aromaticity indices, solvation properties, and many others.[36] These descriptors are generated through knowledge-based, graph-theoretical methods, molecular mechanical, or quantum-mechanical tools[37-38] and are classified according to the dimensionality‖ of the chemical representation from which they are computed[39]: 1-dimensional (1D), scalar physicochemical properties such as molecular weight; 2D, molecular constitution-derived descriptors; 2.5D, molecular configuration-derived descriptors; 3D, molecular conformation-derived descriptors. These different levels of complexity, however, are overlapping with the more complex descriptors, often incorporating information from the simpler ones.

Docking

Docking is the computational determination of binding affinity between a protein structure and a ligand. This method involves proficient sampling of all possible poses of the ligand in the binding pocket of the target protein to ease optimal binding geometry, as measured by the defined scoring functions [39, 41]. Docking of small molecules is generally performed in one of three ways: (a) rigid docking, in which the target and ligand are treated as rigid; (b) flexible ligand docking, in which the target is held rigid; or (c) flexible docking, in which both the target and ligand are considered flexible [42]. Molecular docking protocols can also be defined as a blend of a search algorithm and a scoring function [43-46]. Many scoring functions and algorithms are currently available. The search algorithm is supposed to provide support and freedom to the protein- ligand coordination to enable accurately and sufficient sampling, including the binding modes. Logically, the search algorithm is supposed to have good speed and effectiveness, while the scoring function must be able to analyze physicochemical properties of molecules and thermodynamics of interaction. The complexity of docking increases in the order of rigid docking, flexible ligand docking, and flexible docking [47]. A reliable docking algorithm should exhaustively search all possible binding modes between the ligand and target; however, this is impractical because of the large size of the search space. Therefore, constraints, restraints, and approximations are applied to reduce the dimensionality of the problem in an attempt to locate the global minima as efficiently as possible. Since large conformational space is available to protein structures, partial flexibility (side chain) has recently been incorporated into some docking algorithms, *e.g.*, GLIDE [37], GOLD [38], AUTODOCK [48], FlexX [49], etc. Genetic algorithms (AUTODOCK, GOLD) and Monte Carlo simulated annealing algorithms (GLIDE) are widely used. The genetic algorithm is an iterative process that sustains a population of individuals that are candidates of the solutions to the problem being elucidated. However, simulated annealing is an iterative procedure that constantly apprises one candidate solution until it reaches a termination condition [50].

Pharmacophore

A pharmacophore is the ensemble of steric and electronic features including 1D (physical or biological properties), 2D (substructures) and 3D (charged/ionizable groups, hydrophobic groups, and hydrogen

bond acceptors/donors) aspects that are necessary to ensure the optimal supramolecular interactions with a specific biological target structure and considered to be responsible for a desired biological activity [51- 56]. The concept of a pharmacophore has become an important tool in CADD. In a pharmacophore, each atom that exhibits certain properties related to molecular recognition is bridged to a pharmacophore feature. These molecular features are labeled hydrogen bond donors (HBD), hydrogen bond acceptors (HBA), hydrophobic aromatics, *etc*. [57]. Pharmacophore fingerprinting compares different molecules at the pharmacophore level. When only a few pharmacophore features are considered in a 3D model, the pharmacophore is sometimes described as a query. A pharmacophore model can be established in both ligand-based and structure-based manners. In the ligand-based approach, this method can be used by superposing a set of active molecules and extracting the common essential chemical features required for their bioactivity. In SBDD, this approach can be used by probing possible interaction points between the target protein and ligands. Pharmacophore techniques have been used extensively in virtual screening, de novo design, lead optimization, multitarget drug design, etc. With advances in computational chemistry, a large number of automated tools have been developed and made available for pharmacophore modeling.

Pharmacophore/Ligand Based Virtual Screening (PBVS/LBVS)

Pharmacophore based virtual screening (PBVS) uses a pharmacophore modeling approach to screen large databases to identify molecules of desired biological effects. To accomplish this, a query (pharmacophore model) that encodes the correct 3D organization of the required interaction pattern in the most likely manner is created. Different options are available for constructing a pharmacophore
model (query) depending on the information available for the particular protein target. Examples of some programs that perform pharmacophore based searches include UNITY [58], MACCS-3d [59], Catalyst [60], PHASE [62], and ROCS [63]. Table 1 shows some of the major tools used in LBDD. In general, PBVS is conducted in two consecutive steps, checking the atom type and/or functional group required by the pharmacophore and checking whether the spatial arrangement of these compounds matches these queries [64]. PBVS is superior to SBVS with respect to its ability to screen multi-conformational databases consisting of millions of compounds in comparatively less time and yielding high quality and structurally diverse hits/leads [65]. Moreover, the production of several false positive hits/leads has been a major obstruction to both PBVS and SBVS based drug discovery processes. Poor identification of important physicochemical parameters is one possible reason for the failures of both PBVS and SBVS [66].

Scoring Functions: Concept and Application

Scoring function is the most important component in structurebased drug design for evaluating the efficacy of ligands binding to their target proteins [67]. In molecular docking experiments, protein-ligand complexes need to be rapidly and accurately assessed [68]. As molecular docking experiments generate thousands of ligand binding orientations/conformations, scoring functions are used to rank these complexes and differentiate the accurate binding mode predictions from inaccurate predictions [69, 70, 71].

The goal of an ideal scoring function is to rank the complex as determined empirically [72]. Additionally, the scoring function should be able to predict the absolute binding affinity of the complex to facilitate identification of the potential hits/lead candidates against any therapeutic target from a large library of compounds as used in virtual screening. Scoring functions are very helpful to screening libraries of compounds or individual compounds based on their binding mode and affinity. Over the years, various scoring functions that exhibit different accuracies and computational efficiencies have been developed [73]. In this section, we briefly review the scoring functions in literature developed for protein-ligand interactions in molecular docking. Fig. **2** shows different scoring functions currently in use. Scoring functions have been categorized into four different types:

1. Force-field or molecular mechanics-based scoring functions.
2. Empirical scoring functions.
3. Knowledge-based scoring functions.
4. Consensus scoring functions.

Force-Field or Molecular Mechanics-Based Scoring Functions

Classic molecular mechanics are used by force-field scoring functions for energy calculations [74]. These scoring functions use various physical features such as van der Waals (VDW) interactions, electrostatic interactions, and bond stretching/bending/torsional forces. Force-field or molecular mechanics-based scoring functions utilize parameters derived from both experimental and *ab initio* quantum mechanical calculations [75]. These scoring functions estimate the binding free energy of protein-ligand complexes by the sum of the van der Waals (VDW) interactions and electrostatic interactions [76]. Despite its various successful applications, a major challenge associated with force field scoring functions is their inability to treat solvent molecules in ligand binding [77]. To overcome this shortcoming, variables from the empirical scoring functions are often taken into consideration along with force-field functions.

Empirical Scoring Functions

These scoring functions are based on counting the number of different types of interactions between two binding partners [78, 79]. These functions count the number of atoms within a ligand and receptor that are in contact with each other or calculate changes in the solvent accessible surface area (_SASA) in the complex and the uncomplexed structure of the protein and ligand. These interaction terms of the function may include favorable contacts (hydrophobichydrophobic), unfavorable contacts (hydrophobic- hydrophilic), favorable contributions to affinity (especially if shielded from solvent), no contribution if solvent exposed (number of hydrogen bonds), and unfavorable conformational entropy contribution (number of rotatable bonds immobilized in complex formation).

Knowledge-Based Scoring

This scoring function attempts to capture knowledge about the receptor (target) - ligand binding available in the protein data bank (PDB) by statistical analysis of structural data alone [80-81]. Frequency of

occurrence of individual contacts is assumed to measure their energetic contribution to the binding. A specific contact that occurs more frequently than an average or random distribution indicates

attractive interaction, whereas less frequent occurrence indicates repulsive interaction, e.g., PMF score (potentials of mean force) [82].

Consensus Scoring Function

Despite the availability of some good scoring functions, consensus scoring functions have been developed. Every scoring function currently in use has some limitations and advantages. The consensus scoring function was developed while considering the advantages of different scoring functions to achieve high accuracy [83]. Consensus scoring functions, which are the most advanced scoring technique, improve the probability of finding the correct solution via a combination of different scoring functions [84]. The best aspect of consensus scoring functions is their ability to score predicted binding poses using different scoring functions [85]. Commonly used consensus scoring strategies include: (1) Weighted combinations of scoring functions, vote by number strategy. (2) Vote by number strategy in which a cutoff value is established for each scoring method used and the final decision is made based on the number of passes a molecule has (3) Rank by number strategy in which each compound is ranked by its average normalized score. (4) Rank by rank strategy in which the compounds are sorted on the basis of their average rank and predicted by individual scoring functions.

Molecular Dynamics

Molecular dynamics simulation, also referred as MD, is one of the principal tools for the theoretical study of biological molecules [86]. In MD, Newtonian mechanics are applied to calculate the trajectory of a system [87]. However, standard MD methods depend on the initial conformation and are not inherently suitable for simulation of ligand-target interactions [88]. This results in MD being

unable to cross the high-energy barricades within the simulation's lifespan and prevents it from efficiently traversing the rough surface of protein in complex with ligand. Simulated annealing strategies are applicable for more efficient use of MD in docking [89]. This process computationally calculates the behavior of a molecular system with respect to time. A great deal of detailed information regarding the variations and conformational changes within proteins and nucleic acids has been provided by molecular dynamics. These computational methods are now commonly used to investigate the dynamics behavior of biological molecules and their complexes [90]. These methods are also widely applied to determine structures from x-ray crystallography and NMR experiments.

Tool	Year	Technique used	Important feature	Availability	References
CoMFA (Comparative Molecular Field Analysis)	1988	Molecular field based	3D QSAR technique based on data from known active molecules	Commercial	[128]
APOLLO (Automated PharmacOphore Location through Ligand Overlap)	1989	Feature-based method	Identification of interaction points belonging to the receptor site and creating a pseudo receptor from a set of ligands	Available upon request (Not Commercialized)	[129]
ALADDIN	1989	NA	3D database searching method	NA	[130]
DISCO	1993	Bron-Kerbosh clique-detection algorithm	Each molecule is characterized by ligand points and site points.	Commercial	[131]
XED (extended electron distribution)	1994	Molecular field based method	Field points are used as simple and effective descriptions of the electrostatic and van der Waals maxima and minima surrounding a molecule	Commercial	[132]
COMSIA (Comparative Molecular Similarity Indices Analysis)	1994	Molecular field based method	Worked on the concept of COMFA with an extra feature of the use of Gaussian-type physicochemical properties	Commercial	[133]
Apex-3D	1995	Feature-based method	Takes into consideration both conformationaldependent structural parameters and physicochemical properties	Not available (replaced by Catalyst)	[134]
GASP (Genetic Algorithm Similarity Program)	1995	Atom-based method	Uses genetic algorithms for pharmacophore identification. Automatically allows conformational flexibility and maps features among molecules	Commercial	[135]
HipHop	1996	Feature-based method	Uses pruned exhaustive search to identify common features	Commercial	[136]
MOE (Molecular Operating Environment)	2004	Property-based algorithm	Pharmacophoric structural features are represented by labeled points in space	Commercial	[137]
PHASE	2006	Feature-basedmethod	Uses fine-grained conformational sampling and a range of scoring techniques to identify common pharmacophore hypotheses	Commercial	[138]
HypoGen	2000	Feature-based method	Allows identification of hypotheses that are common to the active molecules in the training set, but not present in the inactive molecule	Commercial	[139]
LigandScout	2004	Pattern-matching based	Model 3D pharmacophore models from structural data of macromolecule/ligand complexes or from training and test sets of organic molecules	Commercial	[140]
ROCS (Rapid Overlay of Chemical Structures)	2005	Molecular field based method	Perceive similarity between molecules based on their three-dimensional shape.	Commercial	[141]
PharmaGist	2008	Feature-based method	First Webserver for elucidating 3D pharmacophores from a set of drug-like molecules that are known to bind to a target receptor	Free access	[142]

SOFTWARE FOR GENERAL PURPOSE MOLECULAR MODELING [40]

For workstations, minicomputers, and supercomputers (SGI, Sun, Cray, etc.)

- AMBER—Peter Kollman and coworkers, UCSF.

- Computer assisted model building, energy minimiza-tion, molecular dynamics, and free energy perturbation calculations.
- Midas Plus—UCSF Computer Graphics Laboratory.
- CHARMM—Martin Karplus and cowrokers, Har- vard.
- QUANTA/CHARMm—Molecular Simulations Inc. (MSI) molecular/drug design, QSAR, quantum chemis- try.
- X-ray & NMR data analysis Insight/DISCOVER— Biosym, Inc. Now MSI and Biosym became Accelrys Inc.
- SYBYL—Tripos, Inc.
- ECEPP—Harold Scheraga and coworkers, Cornell
- MM3—Norman Allinger and coworkers, Georgia For personal computers (Apple, Compaq, IBM, etc.)
- Alchemy III—Tripos, Inc.
- Desktop Molecular Modeller—Oxford Elec. Publish- ing Molecular Modeling Pro— WindowChem Software Energy minimization, QSAR (surface area, volume, logP), etc.
- PC MODEL—Serena Software.

Success Story of CADD

There is very large list describing the successful applications of CADD in the development of novel and potent drug candidates in drug discovery. The development of drugs for HIV and flu (influenza) during the 1990s is amongst the greatest acknowledged successful applications of CADD. Relenza (which treats influenza and was a predecessor to Tamiflu) and HIV protease inhibitors are the two most successful outcomes of CADD [91-92]. Relenza is a neuraminidase inhibitor that was licensed to GlaxoSmithKline Inc. (GSK) in 1990 and approved by the FDA in 1999 [93]. HIV protease inhibitors were developed several years before the neuraminidase inhibitors, but Relenza was approved first owing to the pressing medical need. The first HIV protease inhibitor, ritonavir, was synthesized with adequate oral bioavailability in 1991 [94]. The FDA approved this compound in 1996, in record time (72 days).

This drug required eight years for development, which is about half that of a typical drug. This achievement was due to application of a structure-based approach and the FDA's rapid review. A number of other HIV proteases were identified around the same time, including saquinavir (Roche) and nelfinavir (developed by Agouron, now a subsidiary of Pfizer) [95, 96]. These drugs helped transform the treatment of HIV. A large number of drugs identified using CADD already existed in the form of patent medicines. Captopril, the angiotensin- converting enzyme (ACE) inhibitor, is an antihypertensive drug that was approved in 1981 [97]. Dorzolamide, a carbonic anhydrase inhibitor, was approved in 1995 [98]. Additionally, Saquinavir was approved in 1995 [99], and a combination of three therapeutics for treatment of HIV, Saquinavir, Indinavir and Ritonavir was approved in 1996 [100]. Tirofiban, a fibrinogen antagonist that was approved in 1998 [101], and zanamivir, oseltamivir, aliskiren, boceprevir, nolatrexed,

and rupintrivirare are also the results of CADD [102]. A recent study by Kokkonen *et al*. reported the successful use of CADD for identification of inhibitors of Sirtuins, a NAD dependent deacetylase and well-known drug target in neurodegenerative diseases and cancer [102-110]. Another recent successful application of CADD was reported against tuberculosis when a combination of LBDD, SBDD and MD simulation studies were used. The outcome of this study was the identification of a novel and very potent inhibitor (NRB04248) of mycobacterium tuberculosis. This compound was found to have the potential to inhibit PknG (an attractive drug target in mycobacterium tuberculosis) without any cytotoxic effects against host macrophages [111]. CADD has been extremely successful in design and identification of inhibitors against several important diseases, including cancer [112-120], diabetes [121-129], MDR [130- 135], and neurodegenerative disorders [136-142]. A list of some successful inhibitors developed using CADD is given in Table **2**.

List of clinically approved drugs discovered by CADD.

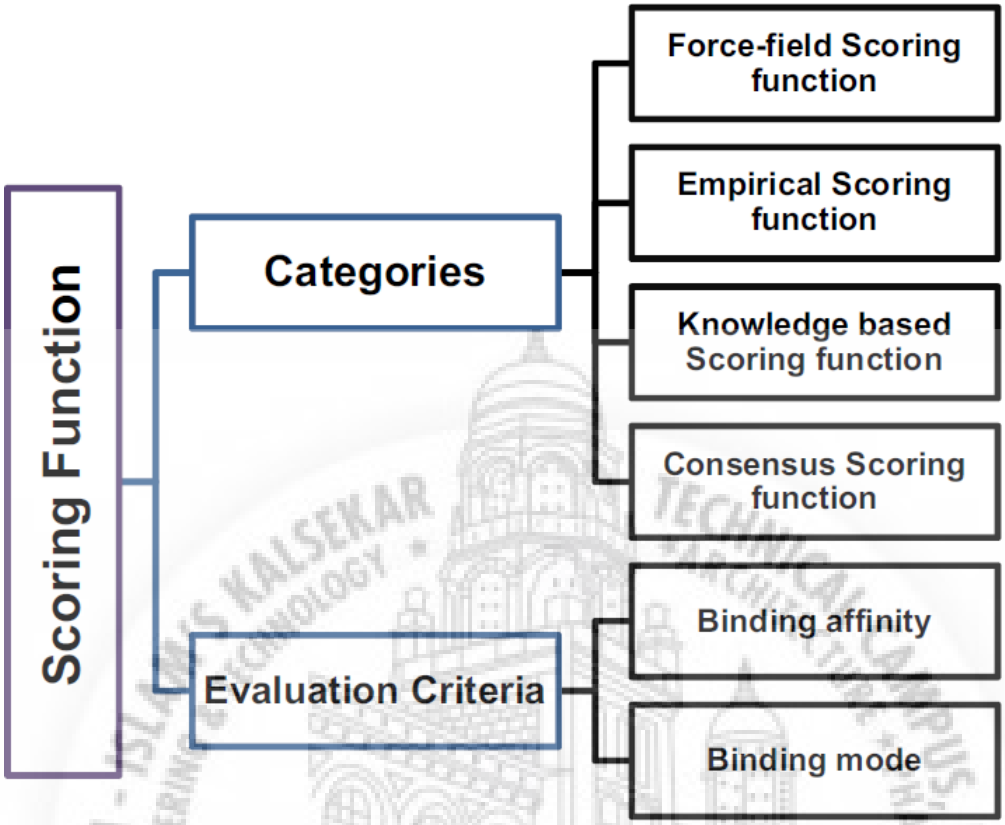
Drug	Approved in Year	Biological Action	References
Captopril	1981	Antihypertensive	[143]
Dorzolamide	1995	Carbonic anhydrase inhibitor	[144]
indinavir	1996	Human immunodeficiency virus (HIV)	[145]
ritonavir	1996	Human immunodeficiency virus (HIV)	[145]
Saquinavir	1995	Human immunodeficiency virus (HIV)	[145]
Triofiban	1998	fibrinogen antagonist	[146]
Raltegravir	2007	Human immunodeficiency virus (HIV)	[147]
Zanamivir	1999	Neuraminidase inhibitor	[148]
Aliskiren	2007	Human renin inhibitor	[149]
Boceprevir	phase III clinical trials	Hepatitis C virus (HCV) inhibitor	[[143]
Nolatrexed	phase III clinical trials	Liver cancer.	[143]
TMI-005	phase II clinical trials	Rheumatoid arthritis	[150]
Oseltamivir	1999	Active against influenza A and B viruses.	[151]
LY-517717	phase II clinical trials	Serine protease Inhibitor	[150]
NVP-AUY922	phase I clinical trials	Inhibitor for HSP90	[152]

CHALLENGES IN CADD

Biological systems are complex and governed by several significant parameters. Accordingly, there are certain limitations and it is not possible to copy and simulate the complete biological system on a computer system using state of the art techniques. One of the biggest remaining challenges in drug discovery is target flexibility. Most molecular docking tools provide high flexibility to the ligand, while the protein is kept fixed or provided with limited flexibility to the residues present within or near the active site. It is very difficult to provide complete molecular flexibility to the protein as this increase the space and time complexity of the computation [153]. However, efforts are being made to add as many parameters as possible. Receptor and target molecules are highly flexible in solution because of conformation changes [154, 159]. Therefore, designing an inhibitor blindly to identify a single, rigid structure may lead to the wrong result. Docking tools supply enough flexibility to the ligand, with limited flexibility to the residues near binding sites of protein. Proteins and ligand molecules possess high flexibility in solution because of their conformational changes [160].

Therefore, keeping in mind a single, rigid structure while designing inhibitors or drug molecules may also lead to an incorrect result. Water molecules are considered to play a crucial role in cellular systems. Accordingly, it is necessary to incorporate the effects of water molecules and other solvents into docking algorithms [161, 168]. Occasionally, ligand-based investigations produce a model with a good R square value, indicating it is reliable for prediction. We can then use this ligand-based model to predict the activities of potent candidates. Unfortunately, most structure-based results do not seem to be consistent with ligand-based results [169]. It has been suggested that such differences might be because of the dependence of all virtual screening methods on databases, even though they can vary greatly for particular targets [170]. In computer-aided drug design, the system is treated by force field models in which the molecules
are treated as point charges bound by spring-like Lennard-Jones and potential interactions. Despite providing speed to computation, there are several pitfalls to this method. In this system, the electronic degrees of freedom (polarization) are neglected and unable to feature and analyze the breaking of bonds within the systems [171, 174]. Sensitivity to parameters, neglecting electronic degrees of
freedom (polarization), and inability to model bond breaking are some of the major pitfalls of the system [175]. Nevertheless, given enough samples, force fields can be used to model processes including protein- ligand binding and protein folding. One of the major limitations of pharmacophore based LBDD is its dependence on pre-computed databases that contain a limited number of low-energy conformations per molecule. This limits the probability of identifying an active molecule because many conformations are missing; especially those for rotatable bonds of small functionalities such as are found in hydroxyl groups.

This restricts the ability of this approach to distinguish between different rotations during the conformer generation and thus affects sampling [176]. Missing different conformations may be possible because an active molecule cannot be identified. This is especially true for the many different conformations of rotatable bonds of small molecular functionalities such as hydroxyl groups. It is difficult to distinguish between different rotations in terms of root mean square deviation differences, which affect their proper sampling. There is no clear process for constructing a pharmacophore query. However, several studies have reported that different molecules were created for similar targets, i.e., screening a similar dataset produces different molecules, which were found to be inactive. A previous study reported one example of the failure of pharmacophores when different pharmacophores were created for similar targets [177, 179]. Identification of very different molecules from a similar dataset has also been reported. Another possible shortcoming is identification of kinase inhibitors possessing similar structures, but different activity profiles against a kinome [180]. Taken together, these findings indicate that pharmacophore approaches for identifying kinase inhibitors do not provide a clear picture of their activity against the targeted kinase [181]. It is important to note that molecular dynamics have several limitations. For example, the method is computationally very demanding and dependent on the size of the system simulation, with times limited to hundreds of nanoseconds or a few microseconds at most [182]. This time period is too short for analysis since the complete folding of a protein requires a time period ranging from milliseconds to seconds [183]. Accordingly, this limitation can lead to inadequate sampling of conformations.



Different categories and the criteria of evaluations of scoring functions in protein–ligand interaction. The quality of the force field is an important feature for observation of certain properties of a system, and it is very important to parameterize force fields for the system. The force field being used needs to be well parameterized and very accurate to distinguish between various conformations at different time steps. However, it is not clear if the force-field being used will attain the accuracy required by the system, especially when some very crucial effects such as polarization of the atoms by their environment are not considered based on the electrostatic potential. Classical descriptions of the particles used comprise another important limitation of MD simulation that restricts investigation of some important quantum mechanical based phenomena, such as electron transfer or bond breaking/formation.

CONCLUSION

Overall, there are several reasons to use the modern techniquesof CADD for drug design and development. Structure-based and ligand-based drug design methods along with molecular dynamics,simulation studies are the backbone of modern CADD processes. We discussed several success stories of these techniques and theirlimitations. The clear concept and advanced knowledge of CADD methods will improve research quality and facilitate identification of new chemical entities, leading to development of useful drugs.

CADD = Computer-Aided Drug Design

GOLD = Genetic Optimisation for Ligand Docking HIV = Human Immunodeficiency Virus
HTS = High Throughput Screening LBDD = Ligand-based drug design MD = Molecular dynamics
PBVS = Pharmacophore based virtual screening QSAR = Quantitative structure-activity relationship SBDD = Structure-based Drug Design (SBDD) SBVS = Structure-based virtual screening
VS = Virtual Screening

REFERENCES

1. Whittaker. The role of bioinformatics in target validation. Drug Discovery To-Clinical trial registration: a statement from the International Committee of Medical Journal Editors. Medical Journal of Australia, 2004; 181: 293-4.
2. Lengauer. Bioinformatics. From Genomes to Drugs. Wiley- VCH, Weinheim, Germany, 2002.
3. Lipinski. Lead and drug-like compounds: the rule-of-five revolution. Drug Discovery Today: Technologies, 2004; 1(4): 337-341.
4. Leeson P D, Davis A M, Steele J. Drug-like properties: Guiding principles for design – or chemical prejudice? Drug Discovery Today: Technologies, 2004; 1(3): 189-195.
5. Hou T. Xu X. Recent development and Application of Virtual Screening in Drug Discovery: An Overview. Current Pharmaceutical Design, 2004; 10: 1011-1033.
6. Klebbe G. Lead Identification in Post-Genomics: Computers as a Complementary Alternative. Drug Discovery Today: Technologies, 2004; 1(3): 225-215.
7. Gisbert Schneider. Uli Fechner. Computer-based de novo design of drug-like molecules. Nature. Reviews. Drug Discovery, 2005; 4(8): 649-663.
8. Butte A. The use and analysis of microarray data. Nature Reviews Drug Discovery, 1(12): 951-960.
9. Richards W. G. Computer-Aided Drug Design.*Pure and Applied Chemistry, 1994; 6(68): 1589-1596.*
10. Kitchen D B. Decornez H. Furr J R. Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. Nature reviews in drug discovery, 2004; 3: 935-949.

11 DiMasi J A. Grabowski H G. The cost of biopharmaceutical R&D: is biotech different? Managerial and Decision Economics, 2007; 28: 469-479.

12. Hajduk PJ, Huth JR, and Tse C. Predicting protein druggability. Drug. Discov. Today, 2005; 10: 1675- 1682.

13. Fauman EB, Rai BK, and Huang ES. Structure-based druggability assessment-identifying suitable targets for small molecule therapeutics. Curr. Opin. Chem. Biol., 2011; 15: 463-468.

14. Laurie AT, Jackson RM. Methods for the prediction of protein-ligand binding sites for structure -based drug design and virtual ligand screening. Curr. Protein. Pept. Sci., 2006; 7: 395-406.

15. Becker OM, Dhanoa DS, Marantz Y, Chen D, Shacham S, Cheruku S, Heifetz A, Mohanty P, Fichman M, Sharadendu A. An integrated in silico 3D model-driven discovery of a novel, potent, and selective amidosulfonamide 5-HT1A agonist (PRX-00023) for the treatment of anxiety and depression. J. Med. Chem., 2006; 49: 3116-3135.

16. Warner SL, Bashyam S, Vankayalapati H, Bearss DJ, Han H, Mahadevan D, Von Hoff DD, Hurley LH. Identification of a lead small-molecule inhibitor of the Aurora kinases using a structure-assisted, fragment-based approach. Mol. Cancer. Ther., 2006; 5: 1764-1773.

17. Budzik B, Garzya V, Walker G, Woolley-Roberts M, Pardoe J, Lucas A, Tehan B, Rivero RA, and Langmead CJ. Novel N-substituted benzimidazolones as potent, selective, CNS-penetrant, and orally active M(1) mAChR agonists. Med. Chem. Lett., 2010; 1: 244-248.

18. Buchan DW, Ward SM, Lobley AE, Nugent TC, Bryson K, and Jones DT. Protein annotation and modelling servers at University College London. Nucleic. Acids. Res., 2010; 38: 563-568.

19. Martí-Renom MA, Stuart AC, Fiser A, Sánchez R, Melo F, and Sali A. Comparative protein structure modeling of genes and genomes. Annu. Rev. Biophys. Biomol. Struct., 2000; 29: 291-325.

20. Schlitter J, Engels M, Krüger P. Targeted molecular dynamics: a new approach for searching pathways of conformational transitions. J. Mol. Graph, 1994; 12: 84-89.

21. Grubmüller H. Predicting slow structural transitions in macromolecular systems: Conformational
flooding. Phys. Rev. E. Stat. Phys. Plasmas. Fluids. Relat. Interdiscip. Topics, 1995; 52: 2893-2906.
22. Abrams CF, Vanden-Eijnden E. Large-scale conformational sampling of proteins using temperatureaccelerated molecular dynamics. Proc. Natl. Acad. Sci. U.S.A., 2010; 107: 4961-4966.
23. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. Chem. Phys.
Lett., 1999; 314: 141-151.
24. Liu M, Wang SM. MCDOCK: a Monte Carlo simulation approach to the molecular docking problem. J.
Comput. Aided. Mol. Des., 1999; 13: 435-451.
25. Jones G, Willett P, Glen RC. A genetic algorithm for flexible molecular overlay and pharmacophore
elucidation. J. Comput. Aided. Mol. Des., 1995; 9: 532-549.
26. Halgren TA. Merck molecular force field. 1. Basis, form, scope, parameterization, and performance of
MMFF94. J. Comput. Chem., 1996; 17: 490-519.
27. Böhm HJ. The computer program LUDI: a new method for the de novo design of enzyme inhibitors. J.
Comput. Aided. Mol. Des., 1992; 6: 61-78.
28. Rarey M, Kramer B, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. J. Mol. Biol., 1996; 261: 470-489.
29. Jain AN. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search
engine. J. Med. Chem., 2003; 46: 499-511.

30.Shimada J, Ishchenko AV, Shakhnovich EI. Analysis of knowledge-based protein-ligand potentials using a self-consistent method. Protein. Sci., 2000; 9: 765-775.

31.Velec HFG, Gohlke H, Klebe G. DrugScore (CSD)-knowledge-based scoring function derived from small molecule crystal data with superior recognition rate of near-native ligand poses and better affinity prediction. J. Med. Chem., 2005; 48: 6296-6303.
32.DeWitte RS, Shakhnovich E. SMoG: De novo design method based on simple, fast and accurate free energy estimates. J. Am. Chem. Soc., 1997; 119: 4608-4617.
33.Mitchell JBO, Laskowski RA, Alex A, Forster MJ, Thornton JM. BLEEP-Potential of mean force describing protein-ligand interactions: II. Calculation of binding energies and comparison with experimental data. J. Comput. Chem., 1999; 20: 1177-1185.
34.Feher M. Consensus scoring for protein-ligand interactions. Drug. Discov. Today, 2006; 11: 421-428.
35.O'Boyle NM, Liebeschuetz JW, Cole JC. Testing assumptions and hypotheses for rescoring success in protein-ligand docking. J. Chem. Inf. Model, 2009; 49: 1871-1878.
36.Becker OM, Dhanoa DS, Marantz Y, Chen D, Shacham S, Cheruku S, Heifetz A, Mohanty P, Fichman M, Sharadendu A. An integrated in silico 3D model-driven discovery of a novel, potent, and selective amidosulfonamide 5-HT1A agonist (PRX-00023) for the treatment of anxiety and depression. J. Med. Chem., 2006; 49: 3116-3135.
37.Johnson MA, Maggiora GM. Concepts and Applications of Molecular Similarity, Wiley, New York, 1990.
38.Stumpfe D, Bill A, Novak N, Loch G, Blockus H, Geppert H, Becker T, Schmitz A, Hoch M, Kolanus W. Targeting multifunctional proteins by virtual screening: structurally diverse cytohesin inhibitors with differentiated biological functions. Chem. Biol., 2010; 5: 839-849.
39.Newman DJ, Cragg GM. Natural products as sources of new drugs over the last 25 years. J Nat Prod 2007; 70(3): 461-77.

40.Lourenço AM, Ferreira LM, Branco PS. Molecules of natural origin, semi-synthesis and synthesis with anti-inflammatory and anticancer utilities. Curr Pharm Des 2012; 18(26): 3979-4046.

41. Wikberg JES, Spjuth O, Eklund M, Lapins M. Chemoinformatics Taking Biology into Account: Proteochemometrics. In: Guha R, Bender A, Eds. Computational Approaches in Cheminformatics and Bioinformatics. Hoboken: John Wiley & Sons 2011; pp. 57-92.

42. Reardon S. Project ranks billions of drug interactions. Nature 2013; 503(7477): 449-450.

43. Hughes JP, Rees S, Kalindjian SB, Philpott KL. Principles of early drug discovery. Br J Pharmacol 2011; 162(6): 1239-49.

44. Krasavin M, Karapetian R, Konstantinov I, et al. Discovery and potency optimization of 2 -amino-5- arylmethyl-1,3-thiazole derivatives as potential therapeutic agents for prostate cancer. Arch Pharm (Weinheim) 2009; 342(7): 420-7.

45. Kaul P. Drug discovery: Past, present and future. In: Jucker E, editor. Progress in Drug Research, Volume 50. Berlin: Springer Science and Business Media 1998; pp. 9-105.

46. Veselovsky AV, Zharkova MS, Poroikov VV, Nicklaus MC. Computer-aided design and discovery of protein-protein interaction inhibitors as agents for anti-HIV therapy. SAR QSAR Environ Res 2014; 25(6): 457-71.

47. Song CM, Lim SJ, Tong JC. Recent advances in computer-aided drug design. Brief Bioinform 2009; 10(5): 579-91.

48. Pârvu L. QSAR - a piece of drug design. J Cell Mol Med 2003; 7(3): 333-5.

49. Nolte RT, Wisely GB, Westin S, et al. Ligand binding and coactivator assembly of the peroxisome proliferator-activated receptor-gamma. Nature 1998; 395(6698): 137-43.

50. Rognan D. Chemogenomic approaches to rational drug design. Br J Pharmacol 2007; 152(1): 38-52.

51. Rao VS, Srinivas K. Modern drug discovery process: an in silico approach. J Bioinform Seq Anal 2011; 2(5): 89-94.

52. Freddolino PL, Harrison CB, Liu Y, Schulten K. Challenges in protein folding simulations: Timescale, representation, and analysis. Nat Phys 2010; 6(10): 751-8.

53. Baig MH, Sudhakar DR, Kalaiarasan P, et al. Insight into the effect of inhibitor resistant S130G mutant on physico-chemical properties of SHV type beta-lactamase: a molecular dynamics study. PloS One 2014; 9(12): e112456.

54. Comba P, Hambley TW, Martin B. Molecular modeling of inorganic compounds. USA: John Wiley & Sons 2009.

55. Huang HJ, Lee KJ, Yu HW, et al. Structure-based and ligand-based drug design for HER 2 receptor. J Biomol Struct Dyn 2010; 28(1): 23-37.

56. Wold S, Dunn WJ III. Multivariate quantitative structure-activity relationships (QSAR): conditions for their applicability. J Chem Inform Comput Sci 1983; 23(1): 6-13.

57. Klebe G, Abraham U, Mietzner T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. J Med Chem 1994; 37(24): 4130- 46.

58. Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP. Random forest: a classification and regression tool for compound classification and QSAR modeling. J Chem Inform Comput Sci 2003; 43(6): 1947-58.

59. Tropsha A. Best practices for QSAR model development, validation , and exploitation. Mol Inform 2010; 29(67): 476-88.

60. Tropsha, A. Variable selection qsar modeling, model validation, and virtual screening. Ann Rep Comput Chem 2006; 2: 113-26.

61. Roy, K., & Mitra, I. On various metrics used for validation of predictive QSAR models with applications in virtual screening and focused library design. Combinat Chem High Throughput Screening 2011; 14(6): 450-74.

62. Gao Q, Yang L, Zhu Y. Pharmacophore based drug design approach as a practical process in drug discovery. Curr ComputerAided Drug Des 2010; 6(1): 37-49.

63. Dror O, Shulman-Peleg A, Nussinov R, Wolfson HJ. Predicting molecular interactions in silico: I. A guide to pharmacophore identification and its applications to drug design. Curr Med Chem 2004; 11(1): 71-90.

64. Holliday JD, Willett P. Using a genetic algorithm to identify common structural features in sets of ligands. J Mol Graph Model 1997; 15(4): 221-32.

65. Guner OF. History and evolution of the pharmacophore concept in computer-aided drug design. Curr Top Med Chem 2002; 2(12): 1321-32.

66. Scapin G. Structural biology and drug discovery. Curr Pharm Des 2006; 12(17): 2087-97.

67. Jones S, Thornton JM. Analysis of protein-protein interaction sites using surface patches. J Mol Biol 1997; 272(1): 121-32.

68. Gilson MK, Zhou HX. Calculation of protein-ligand binding affinities. Ann Rev Biophys Biomol Struct 2007; 36: 21-42.

69. Looger LL, Dwyer MA, Smith JJ, Hellinga HW. Computational design of receptor and sensor proteins with novel functions. Nature 2003; 423(6936): 185-90.

70. Mohan V, Gibbs AC, Cummings MD, Jaeger EP, DesJarlais RL. Docking: successes and challenges. Curr Pharm Des 2005; 11(3): 323-33.

71.Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov 2004; 3(11): 935-49.

72.Bissantz C, Folkers G, Rognan D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. J Med Chem 2000; 43(25): 4759-67.

73.Baig MH, Ahmad K, Hasan Q, et al. Interaction of glucagon Gprotein coupled receptor with known natural antidiabetic compounds: multiscoring in silico approach. Evid-Based Complement Alternat Med 2015; 2015: 497253.

74.Islam B, Sharma C, Adem A, Aburawi E, Ojha S. Insight into the mechanism of polyphenols on the activity of HMGR by molecular docking. Drug Des Develop Ther 2015; 9: 4943-51.

75.Friesner RA, Banks JL, Murphy RB, et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 2004; 47(7): 1739-49.

76.Jones G, Willett P, Glen RC. Molecular recognition of receptor sites using agenetic algorithm with a description of desolvation. J Mol Biol 1995; 245(1): 43-53.

77.Morris GM, Huey R, Olson AJ. Using AutoDock for ligandreceptor docking. Curr Protoc Bioinformatics 2008; Chapter 8: Unit 8.14.

78.Kramer B, Rarey M, Lengauer T. Evaluation of the FLEXX incremental construction algorithm for protein-ligand docking. Proteins 1999; 37(2): 228-41.

79.Wolff ME. Burger's medicinal chemistry and drug discovery. Am J Ther 1996; 3(8): 608.

80.Sun H. Pharmacophore-based virtual screening. Curr Med Chem 2008; 15(10): 1018-24.

81 .Waszkowycz B, Perkins TDJ, Sykes RA, Li J. Large-scale virtual screening for discovering leads in the postgenomic era. IBM Syst J 2001; 40(2): 360-76.

82.Dror O, Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. Novel approach for efficient pharmacophore-based virtual screening: method and applications. J Chem Inf Model 2009; 49(10): 2333-43.

83.Sandor M, Kiss R, KeserGM. Virtual fragment docking by Glide: A validation study on 190 proteinfragment complexes. J Chem Inform Model 2010; 50(6): 1165-72.

84.Shoichet BK. Virtual screening of chemical libraries. Nature 2004; 432(7019): 862-5.

85.Lyne PD. Structure-based virtual screening: an overview. Drug Discov Today 2002; 7(20): 1047-55.

86.Lyne PD, Kenny PW, Cosgrove DA, et al. Identification of compounds with nanomolar binding affinity for checkpoint kinase-1 using knowledge-based virtual screening. J Med Chem 2004; 47(8): 1962-8.

87.Prathipati P, Dixit A, Saxena AK. Computer-aided drug design: Integration of structure-based and ligand-based approaches in drug design. Curr Comput-Aid Drug Des 2007; 3(2): 133-48.

88.UNITY Chemical Information Software 4.3. St. Louis, MO: Tripos Inc. 2001.

89.Güner OF, Hughes DW, Dumont LM. An integrated approach to three-dimensional information management with MACCS-3D. J Chem Inf Comput Sci 1991; 31(3): 408-14.

90.Sprague PW, Hoffmann R. CATALYST pharmacophore models and their utility as queries for searching 3D databases. In: Computer-Assisted Lead Finding and Optimization: Current Tools Medicinal Chemistry USA: witey online library 1997; pp. 223-40.

91.Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA. PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1.Methodology and preliminary results. J Comput Aided Mol Des 2006; 20(10-11): 647-71.

92.Rush TS 3rd, Grant JA, Mosyak L, Nicholls A. A shape-based 3-D scaffold hopping method and its application to a bacterial proteinprotein interaction. J Med Chem 2005; 48(5): 1489-95.

93.Fang X, Wang S. A web-based 3D-database pharmacophore searching tool for drug discovery. J Chem Inf Comput Sci 2002; 42(2): 192-8.

94.Roy KK, Singh S, Saxena AK. Integration-mediated prediction enrichment of quantitative model for Hsp90 inhibitors as anticancer agents: 3D-QSAR study. Mol Divers 2011; 15(2): 477-89.

95.Wang R, Lu Y, Wang S. Comparative evaluation of 11 scoring functions for molecular docking. J Med Chem 2003; 46(12): 2287- 303.

96.Erickson JA, Jalaie M, Robertson DH, Lewis RA, Vieth M. Lessons in molecular recognition: the effects of ligand and protein flexibility on molecular docking accuracy. J Med Chem 2004; 47(1): 45-55.

97.Gohlke H, Hendlich M, Klebe G. Knowledge-based scoring function to predict protein-ligand interactions. J Mol Biol 2000; 295(2): 337-56.

98.Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD. Improved protein-ligand docking using GOLD. Prot Struct Funct Bioinform 2003; 52(4): 609-23.

99.Krovat EM, Steindl T, Langer T. Recent advances in docking and scoring. Curr Comput-Aided Drug Des 2005; 1(1): 93-102.

100.Huang SY, Grinter SZ, Zou X. Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions. Phys Chem Chem Phys 2010; 12(40): 12899-908.

101.Charifson PS, Corkery JJ, Murcko MA, Walters WP. Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. J Med Chem 1999; 42(25): 5100-5109.

102.Huang SY, Grinter SZ, Zou X. Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions. Phys Chem Chem Phys 2010; 12(40): 12899-908.

103.Halperin I, Ma B, Wolfson H, Nussinov R. Principles of docking: An overview of search algorithms and a guide to scoring functions. Proteins 2002; 47(4): 409-43.

104.Böhm HJ. Prediction of binding constants of protein ligands: a fast method for the prioritization of hits obtained from de novo design or 3D database search programs. J Comput Aided Mol Des 1998; 12(4): 309-23.

105.Eldridge MD, Murray CW, Auton TR, Paolini GV, Mee RP. Empirical scoring functions: I. The development of a fast empirical scoring function to estimatethe binding affinity of ligands in receptor complexes. J Comput Aided Mol Des 1997; 11(5): 425-45.

106.Gohlke H, Hendlich M, Klebe G. Knowledge-based scoring function to predict protein-ligand interactions. J Mol Biol 2000; 295(2): 337-56.

107.Muegge I. A knowledge-based scoring function for protein-ligand interactions: Probing the reference state. In Virtual Screening: An Alternative or Complement to High Throughput Screening? Netherlands: Springer 2002; pp. 99-114.

108.Huang SY, Zou X. An iterative knowledge-based scoring function to predict protein-ligand interactions: I. Derivation of interaction potentials. J Comput Chem 2006; 27(15): 1866-75.

109.Muegge I. PMF scoring revisited. J Med Chem 2006; 49(20): 5895- 902.

110.McGaughey GB, Sheridan RP, Bayly CI, et al. Comparison of topological, shape, and docking

111.Lindorff-Larsen K, Piana S, Dror RO, Shaw DE. How fast-folding proteins fold. Science 2011; 334(6055): 517-20.

112.Buch I, Giorgino T, De Fabritiis G. Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. Proc Natl Acad Sci USA 2011; 108(25): 10184-9.

113.Faver JC, Ucisik MN, Yang W, Merz KM Jr. Computer-aided Drug Design: Using Numbers to your Advantage. ACS Med Chem Lett 2013; 4(9).

114.Scior T, Bender A, Tresadern G, et al. Recognizing pitfalls in virtual screening: a critical review. J Chem Inf Model 2012; 52(4): 867-81.

115.De Luca L, Barreca ML, Ferro S, et al. Pharmacophore-based discovery of small-molecule inhibitors of protein-protein interactions between HIV-1 integrase and cellular cofactor LEDGF/p75. ChemMedChem 2009; 4(8): 1311-6.

116.Christ F, Voet A, Marchand A, et al. Rational design of smallmolecule inhibitors of the

LEDGF/p75-integrase interaction and HIV replication. Nat Chem Biol 2010; 6(6): 442-8.

117.Vancraenenbroeck R, De Raeymaecker J, Lobbestael E, et al. In silico, in vitro and cellular analysis with a kinome-wide inhibitor panel correlates cellular LRRK2 dephosphorylation to inhibitor activity on LRRK2. Front Mol Neurosci 2014; 7: 51.

118.Tieleman DP, Biggin PC, Smith GR, Sansom MS. Simulation approaches to ion channel structure- function relationships. Q Rev Biophys 2001; 34(4): 473-561.

119.Michalet X, Weiss S, Jäger M. Single-molecule fluorescence studies of protein folding and conformational dynamics. Chem Rev 2006; 106(5): 1785-813.

120.Cramer RD, Patterson DE, Bunce JD. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc 1988; 110(18): 5959-67.

121.Snyder JP, Rao SN, Koehler KF, Vedani A, Pelliciari R. APOLLO Pharmacophores and the pseudoreceptor concept. Trends QSAR Mol Model 1993; 92: 44-51.

122.Van Drie JH, Weininger D, Martin YC. ALADDIN: an integrated tool for computer-assisted molecular design and pharmacophore recognition from geometric, steric, and substructure searching of three-dimensional molecular structures. J Comput Aided Mol Des 1989; 3(3): 225-51.

123.Martin YC, Bures MG, Danaher EA, DeLazzer J, Lico I, Pavlik PA. A fast new approach to pharmacophore mapping and its application to dopaminergic and benzodiazepine agonists. J Comput Aided Mol Des 1993; 7(1): 83-102.

124.Vinter JG. Extended electron distributions applied to the molecular mechanics of some intermolecular interactions. J Comput Aided Mol Des 1994; 8(6): 653-68.

125.Klebe G, Abraham U, Mietzner T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. J Med Chem 1994; 37(24): 4130- 46.

126.Golender V, Vesterman B, Eliyahu O, et al. Knowledgeengineering approach to drug design and its implementation in the Apex-3D expert system. In: Sanz F, Giraldo J, Manaut F, Eds. QSAR and Molecular Modeling: Concepts, Computational Tools and Biological Applications: Proceedings of the 10th European Symposium on Structure-Activity Relationships, QSAR and Molecular Modeling: 1994 September 4-9; Barcelona, Spain: J. R. Prous Science Publishers 1995; pp. 246-51.

127.Barnum D, Greene J, Smellie A, Sprague P. Identification of common functional configurations among molecules. J Chem Inf Comput Sci 1996; 36(3): 563-71.

128.Jones G, Willet P. GASP: Genetic algorithm superimposition program. In: Güner OF, Ed. Pharmacophore perception, development, and use in drug design. La Jolla, CA: International University Line (IUL) 2000; pp. 85-106.

129.Molecular Operating Environment (MOE). Chemical Computing Group. www.chemcomp.com. retrieved August 2009.

130.Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA. PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. J Comput Aided Mol Des 2006; 20(10-11): 647-71.

131.Li H, Sutter J, Hoffmann R. HypoGen: an automated system for generating 3D predictive pharmacophore models. Pharmacop Percep Develop Use Drug Des 2000; 2: 171.

132.Wolber G, Langer T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. J Chem Inf Model 2005; 45(1): 160-9.

133.Rush TS 3rd, Grant JA, Mosyak L, Nicholls A. A shape-based 3-D scaffold hopping method and its application to a bacterial proteinprotein interaction. J Med Chem 2005; 48(5): 1489-95.

134.Schneidman-Duhovny D, Dror O, Inbar Y, Nussinov R, Wolfson HJ. PharmaGist: a webserver for ligand-based pharmacophore detection. Nucleic Acids Res 2008; b36(Web Server issue): W223-8. doi: 10.1093/nar/gkn187.

135.Talele TT, Khedkar SA, Rigby AC. Successful applications of computer aided drug discovery: moving drugs from concept to the clinic. Curr Top Med Chem 2010; 10(1): 127-41.

136.Vijayakrishnan R. Structure-based drug design and modern medicine. J Postgrad Med 2009; 55(4): 301-4. [145] Van Drie JH. Computer-aided drug design: the next 20 years. J Comput Aided Mol Des 2007; 21(10-11): 591-601

137.Hartman GD, Egbertson MS, Halczenko W, et al. Non-peptide fibrinogen receptor antagonists. 1. Discovery and design of exosite inhibitors. J Med Chem 1992; 35(24): 4640-2.

138.Schames JR, Henchman RH, Siegel JS, Sotriffer CA, Ni H, McCammon JA. Discovery of a novel binding trench in HIV integrase. J Med Chem 2004; 47(8): 1879-81.

139.Kim CU, Lew W, Williams MA, et al. Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. J Am Chem Soc 1997; 119(4): 681-90.

140.Cohen NC. Structure-based drug design and the discovery of aliskiren (Tekturna): perseverance and creativity to overcome a R&D pipeline challenge. Chem Biol Drug Des 2007; 70(6): 557-65.

141.Levin JI, Chen JM, Laakso LM, et al. Acetylenic TACE inhibitors. Part 3: Thiomorpholine sulfonamide hydroxamates. Bioorg Med Chem Lett 2006; 16(6): 1605-9.

142.An J, Lee DC, Law AH, et al. A novel small-molecule inhibitor of the avian influenza H5N1 virus determined through computational screening against the neuraminidase. J Med Chem 2009; 52(9): 2667-72.

143.Eccles SA, Massey A, Raynaud FI, et al. NVP-AUY922: a novel heat shock protein 90 inhibitor active against xenograft tumor growth, angiogenesis, and metastasis. Cancer Res 2008; 68(8): 2850- 60.

144.Jansen JM, Cornell W, Tseng YJ, Amaro RE (2012) Teach–Discover–Treat (TDT): collaborative computational drug discovery for neglected diseases. J Mol Graph Modell 38:360–2 11.

145.Riniker S, Landrum GA, Montanari F, Villalba SD, Maier J, Jansen JM, Walters WP, Shelat AA (2017) Virtual-screening workfow tutorials and prospective results from the Teach–Discover–Treat competition 2014 against malaria. F1000Research 6:1136 12.

146.Riniker S, Landrum GA, Montanari F, Villalba SD, Maier J, Jansen JM, Walters WP, Shelat AA (2017) Tutorial for the Teach–Discover–Treat (TDT) Competition 2014—Challenge 1: anti-malaria hit fnding using classifier fusion boosted predictive models. https://github.com/sriniker/TDT-tutor ial- 2014. Accessed 18 Dec 2018 13.

147.Kluyver T, Ragan-Kelley B, Pérez F, Granger B, Bussonnier M, Frederic J, Kelley K, Hamrick J, Grout J, Corlay S, Ivanov P, Avila D, Abdalla S, Willing C, Team Jupyter Development (2016) Jupyter Notebooks—a publishing format for reproducible computational workfows. Agents and

agendas. In: Loizides F, Schmidt B (eds) Positioning and power in academic publishing: players. IOS Press, Amsterdam, pp 87–90 14

148.Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, Light Y, McGlinchey S, Michalovich D, Al-Lazikani B, Overington JP (2012) ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res 40:1100–7 15.

149.Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28:235–42 16.

150.RDKit (2018) RDKit: Open-Source Cheminformatics, Version 2018.09.1. http://www.rdkit.org

17. Davies M, Nowotka M, Papadatos G, Dedman N, Gaulton A, Atkinson F, Bellis L, Overington JP (2015) ChEMBL web services: streamlining access to drug discovery data and utilities. Nucleic Acids Res 43:W612–W620 18.

151.Gilpin W (2015) PyPDB: a Python API for the protein data bank. Bioinformatics 32:159–60 19. Raschka S (2017) BioPandas: working with molecular structures in pandas DataFrames. J Open Source Softw 2:279 20.

152.Schrödinger L (2015) The PyMOL molecular graphics system. Version 1.8 21. Oliphant T (2006) A guide to NumPy. Trelgol Publishing 22. van der Walt S, Colbert SC, Varoquaux G (2011) The NumPy array: a structure for efcient numerical computation. Comput Sci Eng 13(2):22–30 23.

153.McKinney W (2010) Data structures for statistical computing in Python. In: van der Walt S, Millman J (eds) Proceedings of the 9th Python in science conference, pp 51–56 24.

154.McKinney W (2011) pandas: a foundational Python library for data analysis and statistics 25. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: machine learning in Python. J Mach Learn Res 12:2825–2830

155.Hunter JD (2007) Matplotlib: a 2D graphics environment. Comput Sci Eng 9:90–95 27.

156.Waskom M (2018) seaborn v0.9.0 28. Continuum Analytics Inc (dba Anaconda Inc) (2017) conda. https://www. anaconda.com. Accessed 18 Dec 2018 29.

157.Chen J, Zeng F, Forrester SJ, Eguchi S, Zhang MZ, Harris RC (2016) Expression and function of the epidermal growth factor receptor in physiology and disease. Physiol Rev 96:1025–1069 30.

158.Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH (2016) PubChem substance and compound databases. Nucleic Acids Res 44:D1202–D1213 31.

159.Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, Chang Z, Woolsey J (2006) DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res 34:D668–D672

160.Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug

Deliv Rev 23:3–25 33.

161.Brenk R, Schipani A, James D, Krasowski A, Gilbert IH, Frearson J, Wyatt PG (2008) Lessons learnt from assembling screening libraries for drug discovery for neglected diseases.

ChemMedChem 3:435–444 34.

162.Baell JB, Holloway GA (2010) New substructure flters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. J Med Chem 53:2719–2740

163Johnson MA, Maggiora GM (1990) Concepts and applications of molecular similarity, 1st edn. Wiley, New York 36. Bender A, Glen RC (2004) Molecular similarity: a key technique in molecular informatics. Org Biomol Chem 2:3204

164.Bajorath J (2017) Representation and identifcation of activity clifs. Expert Opin Drug Discov 12:879–883 38. Accelrys Inc, San Diego, CA, USA (2011) MACCS structural keys Morgan HL (1965) The generation of a unique machine description for chemical structures—a technique developed at Chemical Abstracts Service. J Chem Doc 5:107–113

165.Rogers D, Hahn M (2010) Extended-connectivity fngerprints. J Chem Inf Model 50:742–754 41.

166.Maggiora G, Vogt M, Stumpfe D, Bajorath J (2014) Molecular similarity in medicinal chemistry. J Med Chem 57:3186–3204 42.

167.Butina D (1999) Unsupervised data base clustering based on Daylight's fngerprint and Tanimoto similarity: a fast and automated way to cluster small and large data sets. J Chem Inf and Model 39:747– 750

168.Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20:273–297

169.van Gerven M, Bohte S (2017) Editorial: artifcial neural networks as models of neural information processing. Front Comput Neurosci 11:114 49

170.Wermuth CG, Ganellin CR, Lindberg P, Mitscher LA (1998) Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). Pure Appl Chem 70:1129–1143

171.Seidel T, Wolber G, Murgueitio MS (2018) Pharmacophore perception and applications. Applied chemoinformatics. Wiley, Weinheim, pp 259–282

172.Macqueen J (1967) Some methods for classifcation and analysis of multivariate observations. In: 5th Berkeley symposium on mathematical statistics and probability, pp 281–297

173.Daina A, Blatter MC, Baillie Gerritsen V, Palagi PM, Marek D, Xenarios I, et al. Drug Design Workshop: A Web-Based Educational Tool To Introduce Computer-Aided Drug Design to the General Public. Journal of Chemical Education. 2017; 94(3):335-44.

174.Xiang M, Cao Y, Fan W, Chen L, Mo Y. Computer-aided drug design: lead discovery and optimization. Combinatorial chemistry & high throughput screening. 2012; 15(4):328-37.

175.Hopfinger AJ. Computer-assisted drug design. Journal of medicinal chemistry. 1985; 28(9):1133- 9

176.Imam SS, Gilani SJ. Computer Aided Drug Design: A Novel Loom To Drug Discovery. Org. Med. Chem. 2017; 1(4):1-6.

177.Ferreira LG, dos Santos RN, Oliva G, Andricopulo AD. Molecular docking and structure-based drug design strategies. Molecules. 2015; 20(7):13384-421.

178.Anderson AC. The process of structure-based drug design. Chemistry & biology. 2003; 10(9):787- 97.

179.Hoque I, Chatterjee A, Bhattacharya S, Biswas R. An Approach of Computer-Aided Drug

Design (CADD) Tools for In Silico Pharmaceutical Drug Design and Development. Int. J. Adv. Res. Biol. Sci. 2017; 4(2) 60-71.

180.Macalino SJ, Gosu V, Hong S, Choi S. Role of computeraided drug design in modern drug discovery. Archives of pharmacal research. 2015; 38(9):1686-701.

181.Lill M. Virtual screening in drug design. In Silico Models for Drug Discovery. 2013 Vol. 993, pp. 1-12. Humana Press, Totowa, NJ.

182.Maithri G, Manasa B, Vani SS, Narendra A, Harshita T. Computational drug design and molecular dynamic studies—a review. Biomedical Data Mining. 2016; 5:1-7. 11. Kar S, Roy K. How far can virtual screening take us in drug discovery? Expert Opin. Drug Discov. 8(3), pp. 245-261.