Available online at www.derpharmachemica.com

Scholars Research Library

Der Pharma Chemica, 2014, 6(6):211-228
(http://derpharmachemica.com/archive.html)

ISSN 0975-413X
CODEN (USA): PCHHAX

Advances in molecular modeling and docking as a tool for modern drug discovery

Ahmad F. Eweas[*1,2], Ibrahim A. Maghrabi[2] and Ali Ibrahim Namarneh[2]

[1]*Department of Medicinal Chemistry, National Research Center, Dokki, Cairo, Egypt*
[2]*College of Pharmacy, Taif University, Taif, Saudi Arabia*

ABSTRACT

The field of computer aided drug design and discovery (CADDD) is a rapidly growing area that have seen many successes in the last few years. Many giant pharmaceutical companies, in addition to academia, adopt CADDD for drug lead discovery. The explosion of structural informatics, genomics and proteomic plays a major role in leading the efforts towards modern era drug discovery and development. This review discusses the recent advances in two of the major vehicles of CADDD, Molecular modeling and docking and some of the success stories accomplished by both academia and pharmaceutical industry using molecular modeling and docking towards discovery of new drug leads.

Key words: Drug Discovery, Molecular Modeling, Molecular Docking, Success stories

---

INTRODUCTION

The process of discovery of a new drug is a very difficult task. Pharmaceutical and biotechnology companies need to make huge investments in the discovery of a single drug that may cure a disease or simply alleviate the symptoms of another. These are businesses like any other, where profits fuel their growth and provide the investments for future discoveries. Most pharmaceutical or biotechnology companies claim that it costs anywhere between $800 million to $900 million and a time span of twelve to fifteen years. Modern drug discovery is mainly based In-silico -chemico-biological approach where, computer plays very important role in discovery of new drugs, not only it can save money but also time. Use of computational techniques in drug discovery and development process is rapidly gaining in popularity, implementation and appreciation. Both computational and experimental techniques have important roles in drug discovery and development and represent complementary approaches. CADD (Computer Aided Drug Discovery) entails:

1. Use of computing power to streamline drug discovery and development process.

2. Advantage of chemical and biological information about ligands and/or targets to identify and optimize new drugs.

3. Design of *in-silico* filters to eliminate compounds with undesirable properties (poor activity and/or poor Absorption, Distribution, Metabolism, Excretion and Toxicity, (ADMET)) and select the most promising candidates.

Fast expansion in this area has been made possible by advances in software and hardware computational power and sophistication.

4. Identification of molecular targets and an increasing database of publicly available target protein structures like the protein data bank www.pdb.org. CADD (Fig. 1) is being utilized to identify hits (active drug candidates), select leads (most likely candidates for further evaluation), and optimize leads i.e. transform biologically active compounds into suitable drugs by improving their physicochemical, pharmaceutical, ADMET/PK (pharmacokinetic) properties.

5. Virtual screening is used to discover new drug candidates from different chemical scaffolds by searching commercial, public, or private 3-dimensional chemical structure databases. It is intended to reduce the size of chemical space and thereby allow focus on more promising candidates for lead discovery and optimization. The goal is to enrich set of molecules with desirable properties (active, drug-like, lead-like) and eliminate compounds with undesirable properties (inactive, reactive, toxic, poor ADMET/PK). In another words, the use of *in-silico* modelling have significantly minimize time and resource requirements of chemical synthesis and biological testing. The rapid growth of virtual screening is evidenced by increase in the number of citations matching keywords "virtual screening" from 4 in 1997 to 302 in 2004[1]. In his 2003 review article, Green of GlaxoSmithKline concluded that: "The future is bright, the future is virtual" [2] Comparison of traditional and virtual screening in terms of expected cost and time requirements stressed the reality that pharmaceutical industry needs to find means of improving efficiency and effectiveness of drug discovery and development in order to sustain itself. This was recently echoed in 2006 that the current business model would become fundamentally untenable unless there is a significant improvement in efficiency and effectiveness of the process[3].

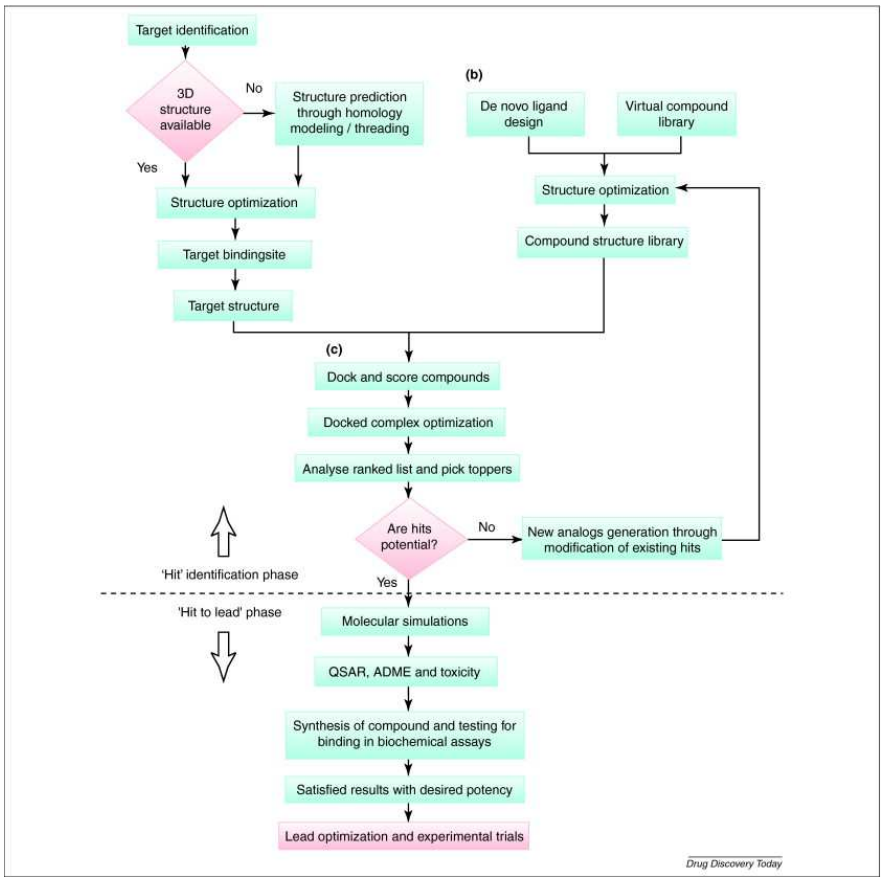

Figure 1: The Computer-Aided Drug Discovery Process

Estimates of time and cost of currently bringing a new drug to market vary, but seven to twelve years and $ 1.2 billion are often cited. Furthermore, five out of forty thousand compounds tested in animals reach human testing and only one of five compounds reaching clinical studies is approved. This represents an enormous investment in terms of time, money and human and other resources. It includes chemical synthesis, purchase, curation, and biological screening of hundreds of thousands of compounds to identify hits followed by their optimization to generate leads which requiring further synthesis. In addition, predictability of animal studies in terms of both efficacy and toxicity is frequently suboptimal. Therefore, new approaches are needed to facilitate, expedite and streamline drug discovery and development, save time, money and resources.

It is estimated that computer modelling and simulations account for ~ 10% of pharmaceutical R&D expenditure and that they will rise to 20% by 2016 [4]. Role of computational models is to increase prediction based on existing Knowledge [5]. Computational methods are playing increasingly larger and more important role in drug discovery and development [6-12] and are believed to offer means of improved efficiency for the industry. They are expected to limit and focus chemical synthesis and biological testing and thereby greatly decrease traditional resource requirements. Modern drug discovery and development process including prominent role of computational modelling represents a brief overview, rather than an exhaustive review, of CADD and the following commonly used computational approaches are discussed: Molecular modeling and structure (target)-based design (docking).

1. MOLECULAR MODELING

Molecular modeling encompasses all theoretical methods and computational techniques used to model or mimic the behavior of molecules. The techniques are used in the fields of computational chemistry, drug design, computational biology and materials science for studying molecular systems ranging from small chemical systems to large biological molecules and material assemblies. The simplest calculations can be performed by hand, but inevitably, computers are required to perform molecular modeling of any reasonably sized system. The common feature of molecular modeling techniques is the atomistic level description of the molecular systems. Most molecular modeling studies involve three stages.

- In the first stage, a model is selected to describe the intra- and inter- molecular interactions in the system. The two most common models that are used in molecular modeling are quantum mechanics and molecular mechanics. These models enable the energy of any arrangement of the atoms and molecules in the system to be calculated, and allow the modeler to determine how the energy of the system varies as the positions of the atoms and molecules change.
- The second stage of a molecular modeling study is the calculation itself, such as an energy minimization, a molecular dynamics or Monte Carlo simulation, or a conformational search.
- Finally, the calculation must be analyzed, not only to calculate properties but also to check that it has been performed properly.

Computational chemistry/molecular modeling is the science (or art) of representing molecular structures numerically and simulating their behavior with the equations of quantum and classical physics. Computational chemistry programs allow scientists to generate and present molecular data including geometries (bond lengths, bond angles and torsion angles), energies (heat of formation, activation energy, etc.), electronic properties (moments, charges, and ionization potential and electron affinity), spectroscopic properties (vibrational modes, chemical shifts) and bulk properties (volumes, surface areas, diffusion, viscosity, etc.). As with all models however, the chemist's intuition and training is necessary to interpret the results appropriately. One of the earliest and still one of the largest uses of computers is to solve complex problems in the natural sciences and engineering disciplines and more specifically to obtain solutions of mathematical models that describe chemical or physical phenomena (or processes). The techniques used to obtain such solutions are part of the general area called Scientific Computing, and the use of these techniques to obtain insights into scientific or engineering problems is called Computational Science. Computational Science is a rapidly emerging trans-disciplinary field at the intersection of the natural sciences, computer science, and mathematics because much scientific investigation now involves computing as well as theory and experiment:

Computational Science = mathematics+ computer science + field of application.

Computational Science typically unifies three distinct elements:

- Modeling, algorithms and simulations
- Software developed to solve natural science, social science, engineering, and medical problems
- Computer and information science that develops and optimizes advanced hardware systems, software, networking and data management components.

1.1. CLASSES OF MOLECULAR MODELING

Molecular modeling applications in drug discovery can classified into two major classes

1.1.1.SMALL MOLECULE (LIGAND) MODELING

This modeling method focuses on the ligand 2D and 3D representations, which is useful in predicting the physiochemical properties of small molecule libraries intended to screening for their drug likeness properties. The approach is old as it started in the late eighties.

2.1.1.1 SOURCES OF MOLECULAR STRUCTURES

Structures of molecules can be determined experimentally or predicted computationally. For small molecules, such as ethanol and cyclohexane, both spectroscopic methods and rigorous quantum chemical computations can provide highly accurate molecular geometries in the vapor phase. Such highly accurate methods are not readily applicable for most drug molecules due to their larger size. However, experiments can give quite accurate structures of drugs in the crystalline state or while bound to target macromolecules. Approximate computations can give sufficiently accurate structures for drugs in the gas or liquid phase. Often, the effects of the medium on the structure are not too significant; in this case, use of experimental crystal structures or computationally predicted gas phase structures is permissible for description of molecules in solution. In other cases, care must be taken to use advanced computational methods that account for effects of the environment.

2.1.1.2. DOWNLOADING MOLECULAR STRUCTURES

Databases such as Klotho provide model structures for many common small molecules. Chemicals with Pharmaceutical Activity from University of Oxford offers access to many drug models and the Protein Data Bank offers access to experimentally determined structures of macromolecules and macromolecular complexes.

2.1.1.3. MOLECULAR STRUCTURES VIA SMILES SERVERS

Several web sites generate 3D molecular structures from the SMILES string-using program like CORINA. CORINA uses built-in tables of standard bond lengths and angles to create a reasonable model for small or rigid molecules. However, the model geometry for larger and flexible molecules is likely to be quite different from the most prevalent geometry in aqueous solution. One such site that generates 3D model structures is Online SMILES Translator by National Institutes of Health.

2.1.1.4. SKETCHING 2D MOLECULAR STRUCTURES

Most chemists are well familiar with drawing 2D molecular structures and several programs allow effortlessly draw 2D representations of three-dimensional molecules. Two of the most popular 2D chemical diagram editors for Windows and Mac OS systems are ChemDraw from CambridgeSoft and MDL Draw from Elsevier MDL. Students can download a fully functional free chemical drawing program MDL Isis/Draw from Elsevier MDL website after registration. Some chemical drawing tools allow generation and export of 3D coordinates of the drawn molecule. The JME Molecular Editor allows sketching simple molecules on-line and exporting these structures into the SMILES string.

2.1.2.MACROMOLECULAR (TARGET) MODELING (HOMOLOGY MODELING)

The prediction of the 3D structure of a protein from its amino acid sequence remains a basic scientific problem. This can often achieved using different types of approaches and the first and most accurate approach is “comparative” or “homology” modeling [13]. Homology modeling methods use the fact that evolutionary related proteins share a similar structure [14, 15]. Determination of protein structure by means of experimental methods such as X-ray crystallography or NMR spectroscopy is time consuming and not successful with all proteins, especially with membrane proteins [16]. Currently, experimental structure determination will continue to increases the number of newly discovered sequences, which grows much faster than the number of structures solved. Currently, 79,356 experimental protein structures are available in the Protein Data Bank (PDB) [17] http://www.rcsb.org/pdb (February 2012).

The process of homology or comparative modeling of proteins can be broken down into four sequential steps (Fig. 2):

1. Identification of known 3D structure(s) of a related protein that can serve as template
2. Sequence alignment of target and template proteins
3. Model building for the target based on the 3D structure of the template and the alignment

4. Refining/validation/evaluation of the models. These steps may be repeated until a satisfactory model is built [18].

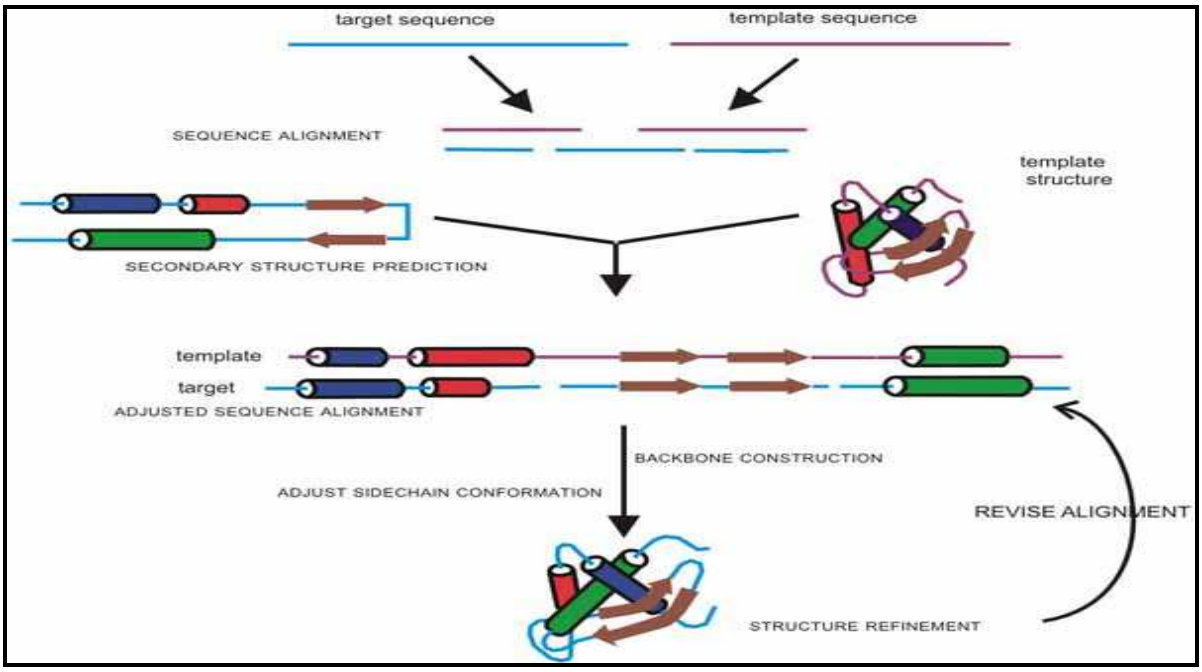

Figure 2: Outline of the homology modeling process and its applications in drug discovery. (http://www.pymol.org)

2.1.3.Homology Moldeing Software

1. ***Abalone*** is designed for macromolecular simulations (proteins, DNA). It supports both explicit and implicit solvent models. In contrast to Ascalaph, tailored to the simulation of small molecules, Abalone is focused on molecular dynamics modeling of biopolymers. It supports such effective methods as the Replica Exchange and hybrid Monte Carlo.
2. ***Ascalaph*** is general purpose molecular modeling software that performs quantum mechanics calculations for initial molecular model development, molecular mechanics and dynamics simulations in the gas or in condensed phase. It can interact with external molecular modeling packages (MDynaMix, NWChem, CP2K, PC GAMESS/Firefly and Delphi).
3. ***Yasara*** is a molecular-graphics, modeling and simulation package for Linux and Windows. Yasara is powered by PVL (Portable Vector Language), a new development framework. PVL allows you to visualize even the largest proteins and enables true interactive real-time simulations with highly accurate force fields on standard PCs
4. ***RasMol*** is a molecular graphics program developed at the University of Edinburgh. The software is intended for the visualization of proteins, nucleic acids and small molecules. The program has the ability to read in PDB as well as several other formats. Coloring schemes including atom type, temperature factor and hydrophobicity.
5. ***MacroModel*** is a computer program for molecular modelling of organic compounds and biopolymers. It features various force fields coupled with energy minimization algorithms for the prediction of geometry and relative conformational energies of molecules. MacroModel also has the ability to perform molecular dynamics simulations to model systems at finite temperatures using stochastic dynamics and mixed Monte Carlo algorithms.
6. ***SYBYL-X*** provides capabilities for crucial small molecular modeling and simulation, includng structure-activity relationship modeling, pharmacophore hypothesis generation, molecular alignment, conformational searching, homology modeling, sequence alignment, and other key tasks required to understand and model the static and dynamic 3D structural properties of proteins and other biological macromolecules.
7. ***Amber*** is a suite of programs for molecular simulation and analysis of proteins, nucleic acids, lipids, carbohydrates. Amber" refers to two things: a set of molecular mechanical force fields for the simulation of biomolecules (which are in the public domain, and are used in a variety of simulation programs); and a package of molecular simulation programs which includes source code and demos.
8. ***MOE*** internal representation of organic chemical structures and flexible architecture provide a solid foundation for molecular modeling and computational chemistry.

9. ***SIMLYS*** is a tool to aid in the analysis of molecular dynamics, Monte Carlo and other stimulations. Its purpose is twofold it is a system performing the actual analysis and it serves as a shell to integrate new analysis functions. *SIMLYS* allows one to analyse the results. from various simulations, as for example from proteins or polymers, by using the trajectories. The program is separated into modules performing the input/output, building the interface to the user, preparing the coordinates and performing the calculations.

In additon to standalone software suiets, several online servers are now avalible for automated homolgy modleing from sequence. These servers are very imoprtant for medicinal chemists who are not familiar with homolgy software as they can bulid a 3D model for any sequence without deep knowledge of the homolgy modeling process.

2.1.4. APPLICATIONS OF HOMOLOGY MODELING

Homology modeling is widely used in structure based drug design process. The importance of homology modeling is increasing as the number of available crystal structures increases. There are several other common applications of homology models:

1. Studying the effect of mutations [19]
2. Identifying active and binding sites on protein (useful for ligand design) [20]
3. Searching for ligands of a given binding site (database mining) [21]
4. Designing novel ligands of a given binding site
5. Modeling substrate specificity [22]
6. Predicting antigenic epitopes [23]
7. Protein–protein docking simulations [24]
8. Molecular replacement in X-ray structure refinement [25]
9. Rationalizing known experimental observations [26]
10. Planning new computational experiments with the provided models.

Typical applications of a homology model in drug discovery require a very high accuracy of the local side chain positions in the binding site. A very large number of homology models have been built over the years. Targets have included antibodies [27] and many proteins involved in human biology and medicine [28, 29]. Clearly, in the absence of crystal structures, homology models are the only alternative to get a 3D representation of the target. Although homology-modeling methods can build reasonably accurate models, refinement methods are needed to get a more accurate characterization of the binding site, and determine the exact side chain conformation, as minor errors may render the model useless for HTD applications.

2.1.5. EXAMPLES OF HOMOLOGY MODELING APPLICATIONS IN DRUG DISCOVERY

2.1.5.1. ESTIMATION OF TARGET DRUGGABILITY FROM IN-SILICO GENERATED STRUCTURES

A druggable target has the ability to bind tightly with small molecules. As most drugs bind to specific binding sites on a protein, it makes sense to identify a priority of such domains as a measure of target druggability. Researchers at Eidogen-Sertanty developed the Target Informatics Platform (TIP), which contains information about protein structure/homology models and binding sites of several protein families [30]. It is important to note that complementing crystal structures with homology models has resulted in 100% structural coverage of some gene families like Nuclear Receptors, Phosphodiesterase and over 98% coverage of protein kinases and trypsin-like proteases [31]. Using TIP and the complex of COX-2 with its inhibitor celecoxib, researchers were able to identify a similar binding site in the $PPAR_x$ receptor, which contained several important binding residues, offering possible clues to design novel PPAR ligands. In another study, Hirayama et al. developed an index termed propensity for ligand binding (PLB) to identify druggable binding sites in homology models, which was later used to successfully predict the druggable cavity in a homology model of tryptophanyl-tRNA synthetase [32, 33].

2.1.5.2. THREE-DIMENSIONAL STRUCTURES OF G PROTEIN COUPLED RECEPTORS AS A PLATFORM FOR COMPUTER-AIDED DRUG DISCOVERY

G protein-coupled receptors (GPCRs) is a large group of evolutionarily related proteins that are expressed on the plasma membrane of animal and other eukaryotic cells and act as sensors for extracellular molecules such as neurotransmitters, hormones and various other signaling compounds. Activated receptors trigger the activation of intracellular proteins, which in turn initiate a biochemical-signaling cascade that dramatically alter the biology of cells, with vast physiological and pathophysiological implications [33]. For these reasons, GPCRs are the object of intense drug discovery efforts aimed at the identification of not only more potent and selective modulators of the receptors that are already validated drug targets, but also novel modulators of the many receptors that are not yet

targeted by drugs [34]. A great deal of interest has always surrounded the structural characterization of GPCRs, since three-dimensional structures of drug targets can serve as the basis for rational computer-aided drug discovery campaigns. However, until the end of the last century, the structure of these membrane receptors proved to be extremely elusive: despite the fact that the human genome comprises about 1000 different GPCRs, until the beginning of this century high-resolution structures were not available for any of the members of the superfamily. Finally, recent progresses in X-ray crystallography brought the number of experimentally solved GPCRs from zero in 1999 to 17 different receptors in complex with different ligands, for about eighty structures, in 2012 [35]. All these structures belong to the largest class of GPCRs, known as 'class A', 'family I', or 'rhodopsin family', while members of the remaining four classes of the superfamily have yet to join the club of experimentally solved receptors [36]. Moreover, as a number of ligand-discovery campaigns illustrate, it can also be implemented based on homology models when experimental structures are not available. For instance, novel modulators of the thyrotropin releasing hormone (TRH) receptor and the free fatty acid receptor 1 (FFA1 or GPR40) were recently identified in such a fashion [37, 38].

3. MOLECULAR DOCKING

In the field of molecular modeling, docking is a method, which predicts the preferred orientation of one molecule to a second when bound to each other to form a stable complex. Knowledge of the preferred orientation in turn may be used to predict the strength of association or binding affinity between two molecules using scoring functions [39].The associations between biologically relevant molecules such as proteins, nucleic acids, carbohydrates, and lipids play a central role in signal transduction. Furthermore, the relative orientation of the two interacting partners may affect the type of signal produced. Therefore docking is useful for predicting both the strength and type of signal produced. Docking is frequently used to predict the binding orientation of small molecule drug candidates to their protein targets in order to in turn predict the affinity and activity of the small molecule. Hence docking plays an important role in the rational design of drugs. The modeling of bimolecular complexes by computational docking using the known structures of their constituents is developing rapidly to become a powerful tool in structural biology. It is especially useful in combination with even limited experimental information describing the interface [40].

Molecular docking involves the prediction of ligand (small molecule) conformation and orientation, referred as 'pose', within the active site of the molecular target (Fig.3). Virtual screening based on molecular docking has become an integral part of many modern structure-based drug discovery efforts. Hence, it becomes a useful endeavor to evaluate existing docking programs, which can assist in the choice of the most suitable docking algorithm for any particular study [41].

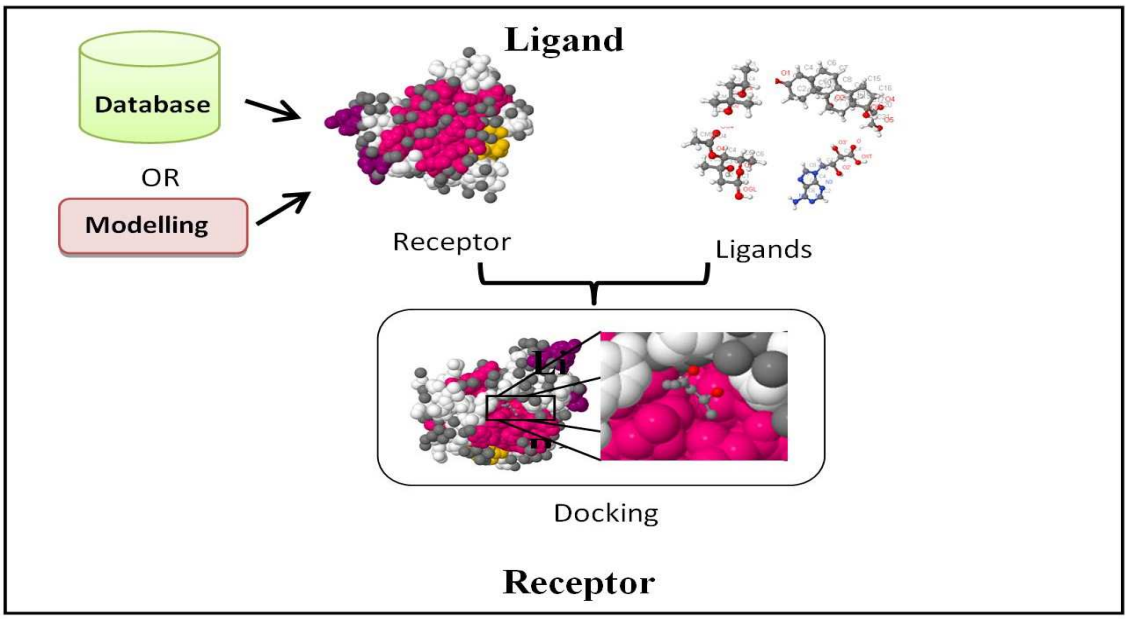

Figure 3: Molecular docking flow chart

Molecular docking represents one of the growing applications in medicinal chemistry where in molecular modeling techniques are used to predict how any macromolecules (typically a protein) interact with other molecules (may be other proteins, nucleic acids or small drug-like molecules). Molecular docking is usually performed between a small molecule and a target macromolecule. This is often referred to as *ligand– protein docking*, but there is growing interest in protein–protein docking. In this review, the focus will be on ligand–protein docking. The ability of a protein to interact with small molecules governs a significant part of the protein's dynamics, which may enhance/ inhibit its biological function. This plays an important role in the rational design of drugs. The ability to bind large molecules such as other proteins and nucleic acids to form supra-molecular complexes is also known to play an important role in controlling biological pathways. Given the biological significance of molecular docking, considerable efforts have been directed in understanding the process of molecular docking [42, 43].

3.1. THEORY

Modeling the interaction of two molecules is a complex problem. Many forces are involved in the intermolecular association, including hydrophobic, van der Waals, or stacking interactions between aromatic amino acids, hydrogen bonding, and electrostatic forces. Modeling the intermolecular interactions in a ligand-protein complex is difficult since there are many degrees of freedom as well as insufficient knowledge of the effect of solvent on the binding association. The process of docking a ligand to a binding site tries to mimic the natural course of interaction of the ligand and its receptor via the lowest energy pathway [44]. There are simple methods for docking rigid ligands with rigid receptors and flexible ligands with rigid receptors, but general methods of docking considering conformationally flexible ligands and receptors are problematic. Docking protocols can be described as a combination of a search algorithm, and the scoring functions (Figure 4).

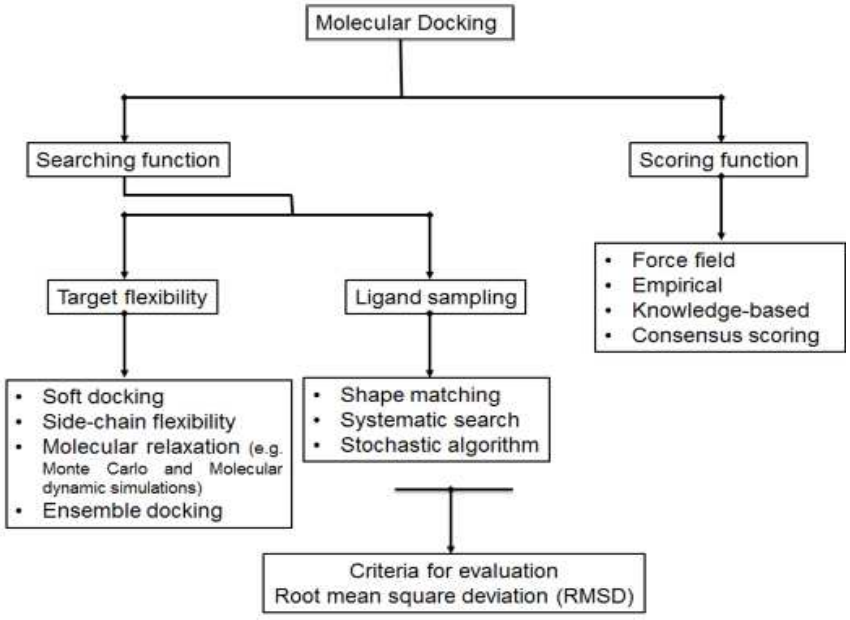

Figure 4: Methods used for protein-ligand docking

The protein-ligand docking procedure can be typically divided into two parts: rigid body docking and flexible docking.

1. ***Rigid Docking***. This approximation treats both the ligand and the receptor as rigid and explores only six degrees of translational and rotational freedom, hence excluding any kind of flexibility. Most of the docking suites employ rigid body docking procedure as a first step.

2. ***Flexible Docking***. A more common approach is to model the ligand flexibility while assuming having a rigid protein receptor, considering thereby only the conformational space of the ligand. Ideally, however, protein flexibility should be taken into account, and some approaches in this regard have been developed. There are three

general categories of algorithms to treat ligand flexibility: systematic methods, random or stochastic methods, and simulation methods. Due to the large size of proteins and their multiple degrees of freedom, their flexibility may be the most challenging issue in molecular docking. The methods to address the flexibility of proteins can be grouped into, soft docking, side-chain flexibility, molecular relaxation and protein ensemble docking. Huang *et al.* [45] described them.

3.2. EXPERIMENTAL DOCKING PROCEDURES

There are a number of excellent reviews of molecular docking methods and a large number of publications comparing the performance of a variety of molecular docking tools [46]. Following, we will describe the four-step procedure adopted in this study to perform the molecular docking.

3.1.1.TARGET SELECTION

Ideally, the target structure should be determined experimentally by either X-ray crystallography or nuclear magnetic resonance, which can be downloaded from PDB; however, docking has been performed successfully in comparison to homology models or threading. The model should have good quality. It can be tested using validation software such as Molprobity [47]. After selecting the model, it must be prepared by removing the water molecules from the cavity, stabilizing charges, filling the missing residues, and generating the side chains, all according to the available parameters. The receptor should be at this point biologically active and in the stable state.

3.1.2.LIGAND SELECTION AND PREPARATION

The type of ligands chosen for docking will depend on the goal: for lead discovery, crude filters such as net charge, molecular weight, polar surface area, solubility, commercial availability, and price-per-compound can be applied to reduce the number of molecules to be docked. For lead optimization, filters such as similarity thresholds, pharmacophores, synthetic accessibility, and absorption, distribution, metabolism, excretion, and toxicology (ADME-Tox) properties are additionally applied. For focused lead optimization, a custom library of analogs that are related to the lead compound(s) is often constructed for docking, to inform and prioritize medicinal chemistry efforts [48]. Most docking tools treat ligands flexibly, with the exception of ring conformations. In general, the more rotatable bonds in a ligand, the more difficult and time consuming the docking will tend to be. This is because the size of the search space increases exponentially with the number of torsions. More highly branched torsion trees lead to more difficult searches than do linear torsion trees. Rotation of conjugated bonds, such as in amides, carbamates, ureas, etc., should be limited.

3.1.3.DOCKING

Molecular docking involves computationally exploring a search space that is defined by the molecular representation used by the method, and ranking candidate solutions to determine the best binding mode. Thus, docking requires both a search method and a scoring function**.**

3.1.4.SCORING FUNCTION

The scoring function provides a way to rank placements of ligands relative to one another. Ideally, the score should correspond directly to the binding affinity of the ligand for the protein, so that the best scoring ligands are the best binders. Scoring functions can be empirical, knowledge based, or molecular mechanics based. In addition, some docking strategies use one scoring function during the docking, and a different one postdocking to rerank the results; such retrospective scoring, however, cannot affect the efficiency and accuracy of the primary scoring function [49]. Scoring functions are fast approximate mathematical methods used to predict the strength of the non-covalent interaction between two molecules after being docked. Most commonly one of the molecules is a small organic compound such as a drug and the second is the drug's biological target such as a protein receptor [50]. Scoring functions have also been developed to predict the strength of other types of intermolecular interactions for example between two proteins [51] or between protein and DNA [52]. Scoring is actually composed of three different aspects relevant to docking and design:

- Ranking of the configurations generated by the docking search for one ligand interacting with a given protein, this aspect is essential to detect the binding mode best approximating the experimentally observed situation.
- Ranking different ligands with respect to the binding to one protein, that is, prioritizing ligands according to their affinity; this aspect is essential in virtual screening.

- Ranking one or different ligands with respect to their binding affinity to different proteins; this aspect is essential for the consideration of selectivity and specificity.

Scoring methods can range from molecular mechanics force fields such as AMBER, OPLS or CHARMM through to empirical free energy scoring functions or knowledge based functions. The currently available docking methods utilize the scoring functions in one of two ways. The first approach uses the full scoring function to rank a protein ligand conformation. The search algorithm then modifies the system, and the same scoring function can be reapplied to rank the new structure.

3.1.5.EVALUATING DOCKING RESULTS

Regardless of the ligand–protein docking tool used, docking results should be evaluated by considering the chemical complementarity between ligand and protein. Are all possible hydrogen bond donors and acceptors in the ligand satisfied?, Are the charged groups in the ligand interacting with oppositely charged side chains in the receptor, or are they buried in hydrophobic pockets?, Are hydrophobic groups in the ligand buried in hydrophobic pockets in the receptor? Furthermore, the parameters chosen for the docking can be judged by the docking tool's ability to reproduce the binding mode of a ligand to protein, when the structure of the ligand–protein complex is known. The criterion usually used is the all-atom RMSD between the docked position and the crystallographically observed binding position of the ligand, and success is typically regarded as being less than 2 °A. If the scoring function were perfect, the docked conformation with the lowest energy would always correspond to the crystallographically observed binding mode, assuming that there are no bad contacts in the crystal structure. This is not always the case, and sometimes a different binding mode is observed significantly more often than the lowest energy-binding mode. Furthermore, current docking methods will tend to find the binding mode with the lowest possible interaction energy for a given ligand: this score does not necessarily indicate whether the ligand even binds. There has been growing interest in developing methods to distinguish binders from nonbinders [53].

3.2. DOCKING SOFTWARE

1. **Auto Dock** uses Monte Carlo simulated annealing and Lamarckian genetic algorithm to create a set of possible conformations. LGA is used as a global optimizer and energy minimization as a local search method.
2. **DOCK** is one of the oldest and best-known ligand-protein docking programs. The initial version used rigid ligands; exibility was later incorporated via incremental construction of the ligand in the binding pocket. As said DOCK is a fragment-based method using shape and chemical complementary methods for creating possible orientations for the ligand. These orientations can be scored using three different scoring functions; however none of them contain explicit hydrogen-bonding terms, solvation/ desolvation terms, or hydrophobicity terms thus limiting serious use. DOCK seems to handle well a polar binding site and is useful for fast docking, but it is not the most accurate software available.

3. **Gold** has won many new users during the last few years because of its good results in impartial tests. It has a good hit rate overall, however it somewhat when dealing with hydrophobic binding pockets. Gold uses genetic algorithm to provide docking of exible ligand and a protein with exible hydroxyl groups. The development of GOLD is currently focused on improving the computational algorithm and adding a support for parallel processing. GOLD has one of the most comprehensive validation test sets and is available for use at CSC [54].

4. ***V life*** has provides following functions:
- Building polypeptides using V Life MDS, Molecular Docking using V Life MDS
- Homology modeling using Biopredicta,
- Protein complex optimization using V Life MDS, Using alignment method in V Life MDS,
- Building molecules using V Life MDS,
- Conformational search using V Life MDS
- Optimizing Molecules using V Life MDS, Using miscellaneous utilities in V Life MDS
- QSAR using V Life MDS.

5. ***ICM (MolSoft LLC)*:** The Internal Coordinate Mechanics (ICM) program is based on a stochastic algorithm that relies on global optimization of the entire flexible ligand in the receptor field (flexible ligand/grid receptor approach [55]. Global optimization is performed in the binding site such that both the intramolecular ligand energy and the ligand receptor interaction energy are optimized. The program combines large-scale random moves of several types

with gradient local minimization and a history mechanism that both expels from the unwanted minima and promotes the discovery of new minima. The random moves include pseudo-Brownian moves, optimally biased moves of groups of torsions, and single torsion changes. The energy calculations are based on the ECEPP/3 force field with Merck molecular force field (MMFF) partial charges. Five potential maps (electrostatic, hydrogen bond, hydrophobic, van der Waals attractive and repulsive) are calculated for the receptor. The location of the receptor-binding pocket can be specified by the user or selected by the cavity detection module implemented in the program.

6. ***Glide (Schrodinger, Inc.)***. The Glide (Grid-Based Ligand Docking with Energetics) algorithm approximates a systematic search of positions, orientations, and conformations of the ligand in the receptor-binding site using a series of hierarchical filters. The shape and properties of the receptor are represented on a grid by several different sets of fields that provide progressively more accurate scoring of the ligand pose. The fields are computed prior to docking. The binding site is defined by a rectangular box confining the translations of the mass center of the ligand. A set of initial ligand conformations is generated through exhaustive search of the torsional minima, and the conformers are clustered in a combinatorial fashion. Each cluster, characterized by a common conformation of the core and an exhaustive set of side-chain conformations, is docked as a single object in the first stage. The search begins with a rough positioning and scoring phase that significantly narrows the search space and reduces the number of poses to be further considered to a few hundred. In the following stage, the selected poses are minimized on pre computed OPLS-AA van der Waals and electrostatic grids for the receptor. In the final stage, the 5–10 lowest-energy poses obtained in this fashion are subjected to a Monte Carlo procedure in which nearby torsional minima are examined, and the orientation of peripheral groups of the ligand is refined. The minimized poses are then rescored using the Glide Score function, which is a more sophisticated version of ChemScore [56] with force field–based components and additional terms accounting for solvation and repulsive interactions. The choice of the best pose is made using a model energy score (Emodel) that combines the energy grid score, GlideScore, and the internal strain of the ligand.

3.3. APPLICATIONS OF MOLECULAR DOCKING

A binding interaction between a small molecule ligand and an enzyme protein may result in activation or inhibition of the enzyme. If the protein is a receptor, ligand binding may result in agonism or antagonism. Docking may be applied to:

1. Hit Identification – docking combined with a scoring function can be used to quickly screen large databases of potential drugs *in-silico*, to identify molecules that are likely to bind to protein target of interest (Virtual Screening).
2. Lead Optimization – docking can be used to predict in where and in which relative orientation a ligand binds to a protein (also referred to as the binding mode or pose).This information may in turn be used to design more potent and selective analogs.
3. Bioremediation– Protein ligand docking can also be used to predict pollutants that can be degraded by enzymes. Estimating the binding affinity.
4. Searching for lead structures for protein targets

3.3.1.APPLICATION EXAMPLES OF MOLECULAR DOCKING FOR DRUG DISCOVERY

Molecular docking has been the most widely employed technique. Though the main application lies in structure-based virtual screening for identification of new active compounds towards a particular target protein, in which it has produced a number of success stories [57], it is actually not a stand-alone technique but is normally embedded in a workflow of different *in-silico* as well as experimental techniques [58]. Several research groups focus on evaluating of the performance of various docking programs or on making improvements to the scoring functions when experimental testing has already been done. Such efforts could give meaningful guidance to choose the methodology for a particular target system.

3.3.1.1 HUMAN G PROTEIN-COUPLED RECEPTORS (GPCRS)

With at least 800 unique full-length members, GPCRs comprise the largest family of cell surface receptors [59]. They are ubiquitous biological control points of the cell. This membrane protein family translates external signals into readable stimuli resulting in precise cell behaviors [60]. Examples of physiological responses controlled by GPCRs are cell growth and differentiation, cardiovascular function, metabolism, immune responses, and neurotransmission. They also represent the largest family of drug targets with about 50% of the existing drugs currently targeting GPCRs for their beneficial action [61], and their therapeutic potential might be even larger [62,

63].The breakthroughs in GPCR crystallography improve dramatically the potential of GPCRs structure-based ligand design approaches. Virtual screening studies against the adenosine receptor A_2A resulted in very high hit rates. Four million "drug-like and lead-like" compounds virtual screening using GOLD resulted indeed in a hit rate of 41% [63]. Out of 56 compounds experimentally tested, 23 showed antagonist activity under 10 mM, among them 11 had submicromolar activity and two compounds had affinity under 60 nM. Nine novel chemotypes were identified supporting high diversity of the hits generated through structure-based virtual screening (Fig. 5).

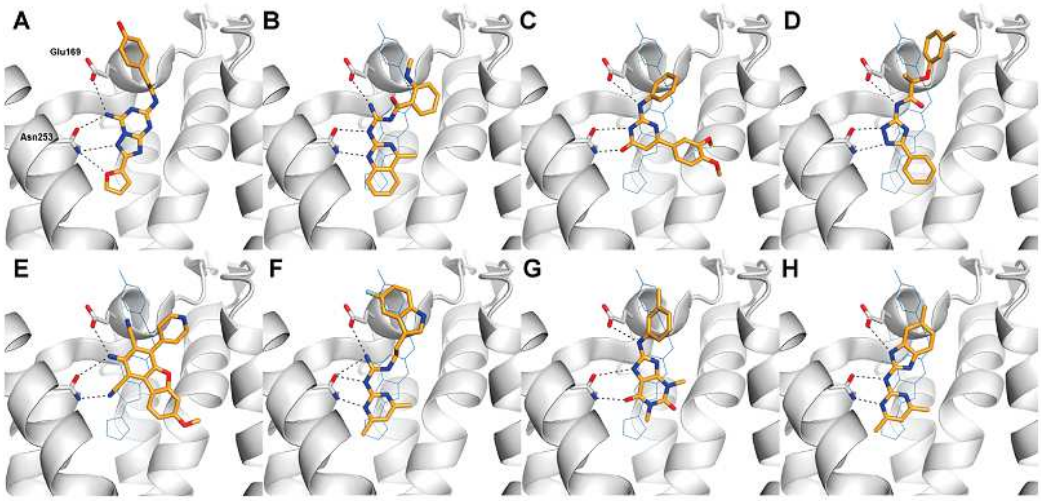

Figure 5. Binding mode of the cocrystallized ligand 6 (A) and the predicted binding modes of the seven ligands discovered in the docking screen

In another study [64], 1.4 million compounds have been screened *in-silico* against the same A_2A X-ray structure, 20 high-ranking novel compounds have been selected and tested experimentally resulting in a hit rate of 35%. The activity range was between 200 nM and 9 mM. These studies suggest practical applicability of receptor-based virtual screening in GPCR drug discovery (Fig. 6). Furthermore extraordinary high hit rates and high activity have been identified suggesting the high potential of the X-ray diffraction crystals compared to the earlier homology models.

3.3.1.2. ANTIVIRAL DRUG DISCOVERY

The reverse transcriptase (RT) of HIV-1 is one of the major targets of the antiretroviral drug therapies used for the treatment of AIDS. RT is responsible for the retro transcription of RNA to DNA in the first phase of the intracellular viral replication. In an attempt to target this protein, Bustanji et al. selected 2800 fragment-like compounds from the NCI database (Fig. 7), performed a high throughput docking and selected the four best hits for testing based on a consensus docking score of which four were found to inhibit RT in biological testing [65].

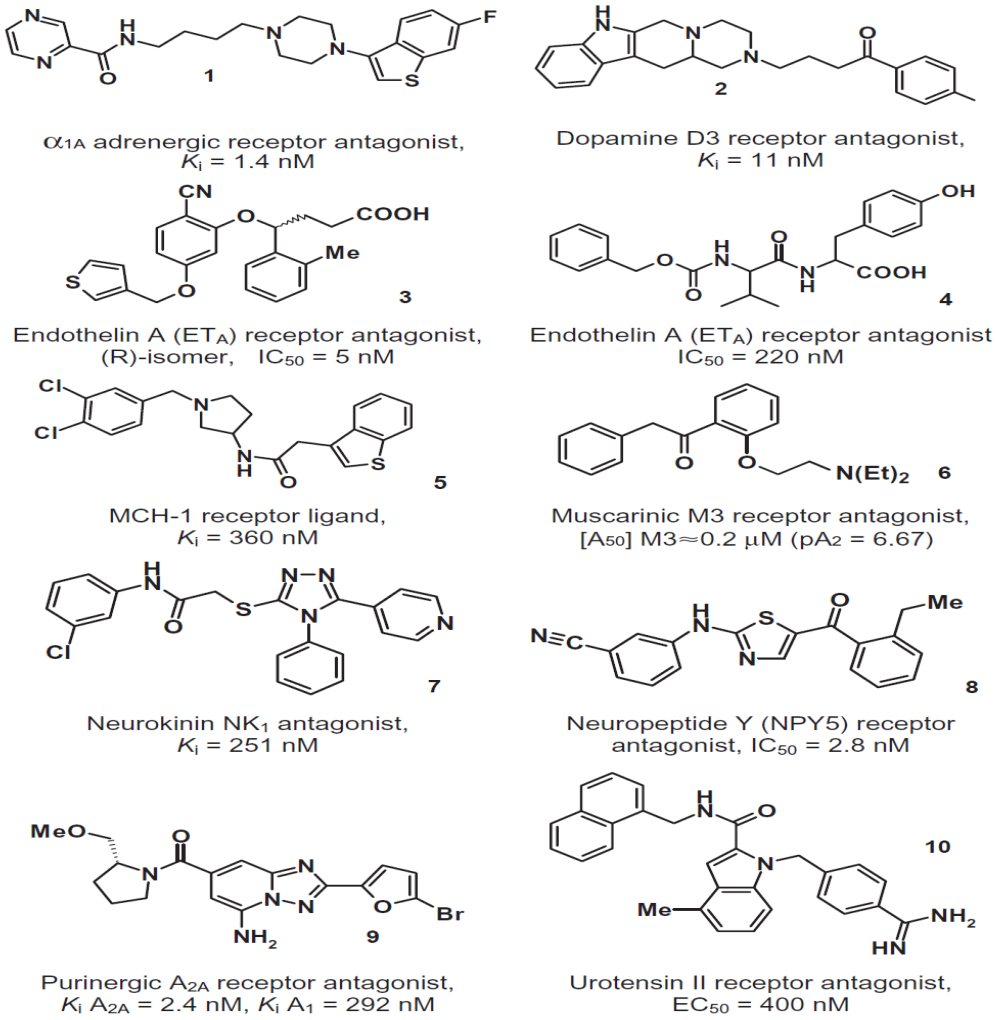

Figure 6: GPCR ligands from docking virtual screening

Another VS screening study was performed to identify novel compounds targeting RT as well as other viral functions associated with RNA transcription. A shape-based screening was applied on the NCI database in a first VS run using dihydroxy benzoyl naphthyl hydrazone, a Known RT inhibitor, as query compound. The most active Hits identified through this process was employed for a Second VS now using a combined ligand-based strategy comprising 3D-, 2D-similarity searches and ligand-based pharmacophore screening (Fig. 8). When tested on the RT functions, several of the selected compounds characterized by new scaffolds were shown to inhibit both RT-associated ribonuclease H and RT activities in a low micro molar range [66].

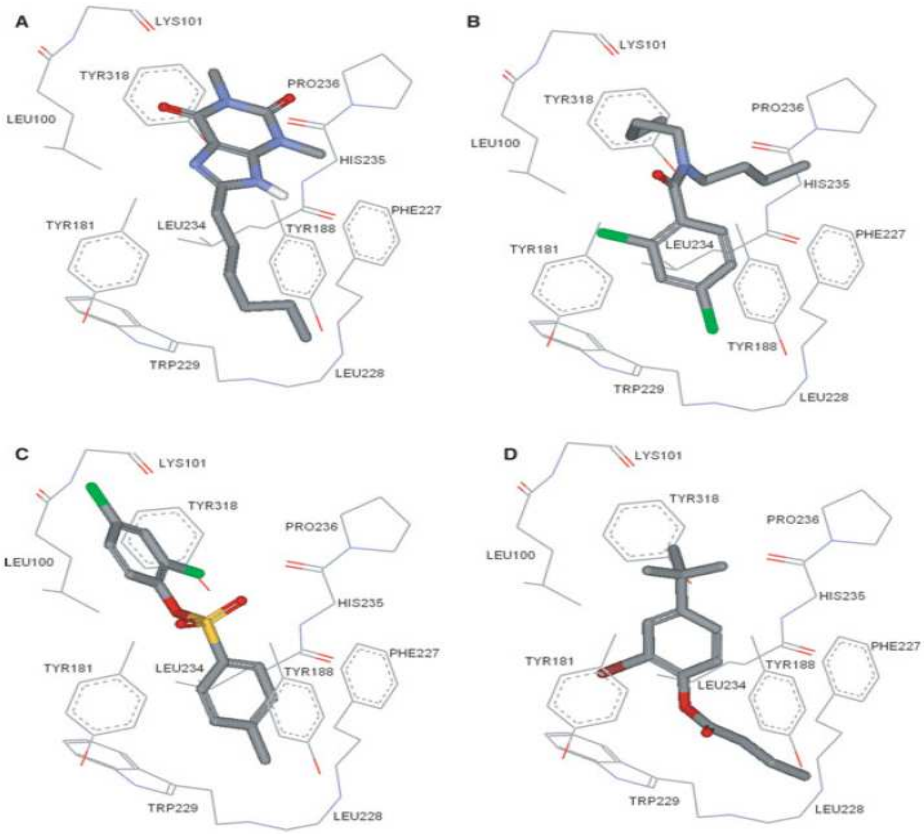

Figure 7: The top-ranking consensus pose of the active captured lead compounds docked inside the allosteric site of HIV-1 R

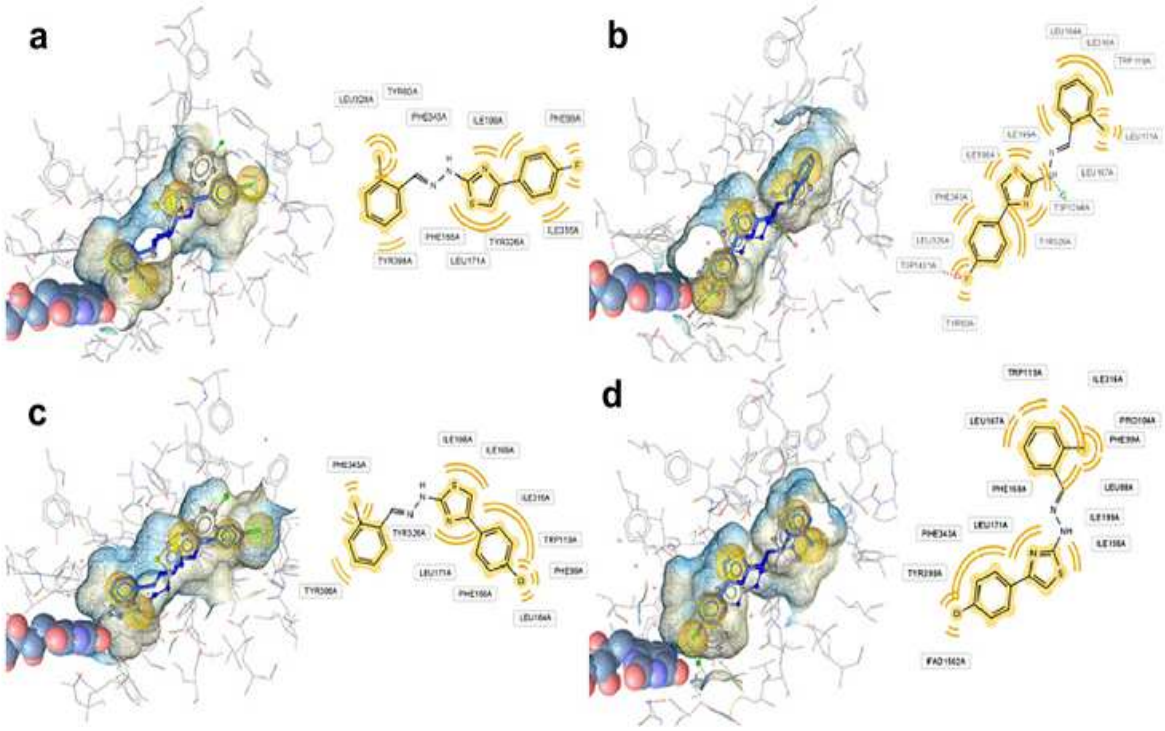

Figure 8: Average conformations of Molecular docking of best-hit compounds

3.3.1.3. ANTICANCER DRUG DISCOVERY (KINASES)

Protein tyrosine kinases (tks) are enzymes that catalyze the transfer of phosphate from ATP to tyrosine residues in polypeptides. The human genome contains about 90 TK and 43 TK-like genes, the products of which regulate cellular proliferation, survival, differentiation, function, and motility. Protein kinases represent attractive targets in oncology drug discovery[67]. An interesting class of targets is the erythropoietin-producing human hepatocellular carcinoma receptors (Eph), the largest family of receptor tyrosine kinases. The Eph receptors have been implicated in sprouting angiogenesis and blood vessel remodeling during vascular development. In a recent, study Caflisch *et al*. have identified three potential tyrosine kinase inhibitors after sequence of virtual screening and docking steps starting with a library of 9 million compounds in the ZINC library. The docked library consisted of about 175 000 compounds derived from nearly 9 million molecules using two-dimensional chemical descriptors and three-dimensional geometric constraints (i.e., relative distance and orientation of pairs of functional groups). Using this procedure, they have identified a series of 5-(piperazine-1-yl)isoquinoline derivatives that exhibited low micromolar affinities for unphosphorylated Abl1 in a competition binding assay [68].

3.3.1.4. NUCLEAR RECEPTORS (RETINOIC ACID RECEPTOR).

Nuclear receptors are a class of proteins found within cells that are responsible for sensing steroid and thyroid hormones and certain other molecules. In response, these receptors work with other proteins to regulate the expression of specific genes, thereby controlling the development, homeostasis, and metabolism of the organism. The retinoic acid receptor (RAR) is a type of nuclear receptor which can also act as a transcription factor [69] that is activated by both all-trans retinoic acid and 9-cis retinoic acid. There are three retinoic acid receptors (RAR), RAR-alpha, RAR-beta, and RAR-gamma, encoded by the *RARA*, *RARB*, *RARG* genes, respectively. Each receptor isoform has several splice variants: two- for alpha, four- for beta, and two- for gamma. A 3D structural model of the inactive conformation of the retinoic acid receptor (RAR) α subtype (RARα) was developed from the RARγ 3D structure, bound to the agonist all-*trans*-retinoic acid, and the estrogen receptor α-subtype (ERα), bound to an antagonist. After validation of the method with known agonists and antagonists, 153,000 ACD compounds were docked into the RAR binding site with full flexibility of the ligand and the amino acid side chains of the protein, using the Molsoft Internal Coordinates Mechanics (ICM 2.7) program. Two novel RAR antagonists were discovered, for example, compound **18** (55% inhibition at 20 µM) [70]. Comparable results were obtained with all three human isoforms: RARα, RARß, and RARγ. In a similar investigation [71], a model of the active RARα conformation was developed from the agonist-bound RARγ conformation. Docking of the ACD compounds as above but with a refined procedure, considering all atoms of the binding site, resulted in 5364 high-scoring hits. The 300 compounds with the lowest binding energy (i.e., highest predicted affinity) were visually inspected for shape complementarity, hydrogen bonding network, ligand conformations, and possible van der Waals clashes. Finally, 30 compounds were selected for biological testing. Despite the fact that an RARα 3D model was used for the docking, the two most active hits have a higher affinity to RARβ than to RARα, for example, compound **19** (EC_{50} RARβ = 200 nM, EC_{50} RARα = 4 µM) [72].

3.3.1.5. TNF INHIBITORS

Tumor necrosis factor a (TNF-α) is a multifunctional cytokine that acts as a central biological mediator for critical immune functions, including inflammation, infection, and antitumor responses [73]. Dysregulation of TNF-α has been implicated in cases of tumorigenesis, diabetes, and especially in autoinflammatory diseases such as rheumatoid arthritis, psoriatic arthritis, and Crohn's disease[74]. In a recent study Over 20 000 compounds from a chemical library of natural-product and natural-product-like structures were screened in-silico. The continuously flexible ligands were docked to a grid representation of the receptor and assigned a score reflecting the quality of the complex according to the internal coordinate mechanics (ICM) method [ICM-Pro 3.6-1d molecular docking software (Molsoft)]. The highest scoring 16 compounds from the virtual screening results were tested in a preliminary ELISA to assess their ability to inhibit the binding of TNF-α to TNFR-1. Two chemically distinct structures, the pyrazole-linked quinuclidine and the indolo[2,3-a]quinolizidine, emerged as the top candidates [75]. The binding poses of these two compounds overlap well with the crystallographic pose of SPD304 to TNF-a (Figure 13). Like SPD304, compounds 1 and 2 are large enough to interact with the residues from both subunits of the TNF-a dimer, thereby occupying and blocking the binding site for the third TNF-a subunit.

CONCLUSION

Virtual chemical library screening by docking has become a method routinely used in chemoinformatics to identify ligands for targets of therapeutic interest. With the development of significantly more sophisticated molecular modeling, tools and a growth in the use of high throughput X-ray crystallography of the target alone or in complex with small molecules, structure-based drug design techniques have become an indispensable tool for the development of target-based therapies. Importantly, these newly appreciated approaches are being supported and/or driven by rapidly improving computational platforms that are more reproducibly docking, scoring and ranking drug-like compounds, which has allowed many drug discovery scientists to carry out more focused, hypothesis-driven discovery initiatives limiting the number of compounds that are synthesized. It is important to note that the adoption of early stage PK and PD studies has also contributed greatly to the significantly reduced late-stage attrition rate of clinical candidates. In the particular case of drugs acting as inhibitors of specific target proteins, these advances together with a more careful attention to the conformation, mechanism of action, and drug-like property of the inhibitor are expected to result in novel therapeutic agents that are more potent, selective and bioavailable.

REFERENCES

[1] A. Pozzan, *Current pharmaceutical design*, **2006**, 12 2099-2110
[2] D. Green, *Progress in medicinal chemistry*, **2003**, 41 61
[3] R. Shankar, X. Frapaise, B. Brown, *Drug Discov Develop*, **2006**, 9 57-60
[4] H. Van de Waterbeemd, E. Gifford, *Nature reviews Drug discovery*, **2003**, 2 192-204
[5] N. Kumar, B.S. Hendriks, K.A. Janes, D. de Graaf, D.A. Lauffenburger, *Drug discovery today*, **2006**, 11 806-811
[6] M.M. Hann, T.I. Oprea, *Current opinion in chemical biology*, **2004**, 8 255-263
[7] T.I. Oprea, H. Matter, *Current opinion in chemical biology*, **2004**, 8 349-358
[8] O. Roche, W. Guba, *Mini reviews in medicinal chemistry*, **2005**, 5 677-683
[9] D. Chin, C. Chuaqui, J. Singh, *Mini Reviews in Medicinal Chemistry*, **2004**, 4 1053-1065
[10] A. Jain, *Current opinion in drug discovery & development*, **2004**, 7 396
[11] M.J. Stoermer, *Medicinal Chemistry*, **2006**, 2 89-112
[12] M. Stahl, W. Guba, M. Kansy, *Drug discovery today*, **2006**, 11 326-333
[13] D. Oranit, S.-P. Alexandra, N. Ruth, H.J. Wolfson, *Current medicinal chemistry*, **2004**, 11 71-90
[14] D. Schneidman-Duhovny, R. Nussinov, H.J. Wolfson, *Current medicinal chemistry*, **2004**, 11 91-107
[15] C.N. Cavasotto, S.S. Phatak, *Drug discovery today*, **2009**, 14 676-683
[16] J.M. Chandonia, S.E. Brenner, *PROTEINS: Structure, Function, and Bioinformatics*, **2005**, 58 166-179
[17] D. Vitkup, E. Melamud, J. Moult, C. Sander, *Nature Structural & Molecular Biology*, **2001**, 8 559-566
[18] C. Floudas, H. Fung, S. McAllister, M. Mönnigmann, R. Rajgaria, *Chemical Engineering Science*, **2006**, 61 966-988
[19] J. Westbrook, Z. Feng, L. Chen, H. Yang, H.M. Berman, *Nucleic acids research*, **2003**, 31 489-491
[20] M.A. Martí-Renom, A.C. Stuart, A. Fiser, R. Sánchez, F. Melo, A. Šali, *Annual review of biophysics and biomolecular structure*, **2000**, 29 291-325
[21] J.-P. Boissel, W.-R. Lee, S. Presnell, F. Cohen, H.F. Bunn, *Journal of Biological Chemistry*, **1993**, 268 15983-15993
[22] Y. Sheng, A. Sali, H. Herzog, J. Lahnstein, S.A. Krilis, *The Journal of Immunology*, **1996**, 157 3744-3751
[23] C.S. Ring, E. Sun, J.H. McKerrow, G.K. Lee, P.J. Rosenthal, I.D. Kuntz, F.E. Cohen, *Proceedings of the national academy of Sciences*, **1993**, 90 3583-3587
[24] L.Z. Xu, R. Sánchez, A. Sali, N. Heintz, *Journal of Biological Chemistry*, **1996**, 271 24711-24719
[25] A. Sali, R. Matsumoto, H. McNeil, M. Karplus, R. Stevens, *Journal of Biological Chemistry*, **1993**, 268 9023-9034
[26] I.A. Vakser, *Proteins: Structure, Function, and Bioinformatics*, **1997**, 29 226-230
[27] P.L. Howell, S. Almo, M. Parsons, J. Hajdu, G. Petsko, *Acta Crystallographica Section B: Structural Science*, **1992**, 48 200-207
[28] B. Guenther, R. Onrust, A. Sali, M. O'Donnell, J. Kuriyan, *Cell*, **1997**, 91 335-345
[29] M.T. Mas, K.C. Smith, D.L. Yarmush, K. Aisaka, R.M. Fine, *Proteins: Structure, Function, and Bioinformatics*, **1992**, 14 483-498
[30] P.V. Sahasrabudhe, R. Tejero, S. Kitao, Y. Furuichi, G.T. Montelione, *Proteins: Structure, Function, and Bioinformatics*, **1998**, 33 558-566

[31] I.T. Weber, *Proteins: Structure, Function, and Bioinformatics*, **1990**, 7 172-184
[32] P.J. Kundrotas, M.F. Lensink, E. Alexov, *International journal of biological macromolecules*, **2008**, 43 198-208
[33] K. Hambly, J. Danzer, S. Muskal, D.A. Debe, *Molecular diversity*, **2006**, 10 273-281
[34] S. Soga, H. Shirai, M. Kobori, N. Hirayama, *Journal of chemical information and modeling*, **2007**, 47 2287-2292
[35] K.L. Pierce, R.T. Premont, R.J. Lefkowitz, *Nature Reviews Molecular Cell Biology*, **2002**, 3 639-650
[36] R. Lappano, M. Maggiolini, *Nature reviews Drug discovery*, **2011**, 10 47-60
[37] K.A. Jacobson, S. Costanzi, *Molecular pharmacology*, **2012**, 82 361-371
[38] R. Fredriksson, H.B. Schiöth, *Molecular pharmacology*, **2005**, 67 1414-1425
[39] S. Engel, A.P. Skoumbourdis, J. Childress, S. Neumann, J.R. Deschamps, C.J. Thomas, A.-O. Colson, S. Costanzi, M.C. Gershengorn, *Journal of the American Chemical Society*, **2008**, 130 5115-5123
[40] I.G. Tikhonova, C.S. Sum, S. Neumann, S. Engel, B.M. Raaka, S. Costanzi, M.C. Gershengorn, *Journal of medicinal chemistry*, **2008**, 51 625-633
[41] T. Lengauer, M. Rarey, *Current opinion in structural biology*, **1996**, 6 402-406
[42] A.D. van Dijk, R. Kaptein, R. Boelens, A.M. Bonvin, *Journal of biomolecular NMR*, **2006**, 34 237-244
[43] S. Joy, P.S. Nair, R. Hariharan, M.R. Pillai, *In silico biology*, **2006**, 6 601-605
[44] A.N. Jain, *Current Protein and Peptide Science*, **2006**, 7 407-420
[45] S.-Y. Huang, X. Zou, *International journal of molecular sciences*, **2010**, 11 3016-3034
[46] R.D. Taylor, P.J. Jewsbury, J.W. Essex, *Journal of computer-aided molecular design*, **2002**, 16 151-166
[47] V.B. Chen, W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, *Acta Crystallographica Section D: Biological Crystallography*, **2009**, 66 12-21
[48] J. Gasteiger, M. Marsili, *Tetrahedron Letters*, **1978**, 19 3181-3184
[49] V. Mohan, A.C. Gibbs, M.D. Cummings, E.P. Jaeger, R.L. DesJarlais, *Current pharmaceutical design*, **2005**, 11 323-333
[50] M.F. Lensink, R. Méndez, S.J. Wodak, *Proteins: Structure, Function, and Bioinformatics*, **2007**, 69 704-718
[51] T.A. Robertson, G. Varani, *PROTEINS: Structure, Function, and Bioinformatics*, **2007**, 66 359-374
[52] A. Ruvinsky, A. Kozintsev, *Proteins: Structure, Function, and Bioinformatics*, **2006**, 62 202-208
[53] A. Kaapro, J. Ojanen, *URL http://www. lce. hut. fi/teaching/S-114.500/k2002/Protdock. pdf*, **2002**,
[54] B. Kramer, M. Rarey, T. Lengauer, *Proteins: Structure, Function, and Bioinformatics*, **1999**, 37 228-241
[55] G. Nemethy, K.D. Gibson, K.A. Palmer, C.N. Yoon, G. Paterlini, A. Zagari, S. Rumsey, H.A. Scheraga, *The Journal of Physical Chemistry*, **1992**, 96 6472-6484
[56] M.D. Eldridge, C.W. Murray, T.R. Auton, G.V. Paolini, R.P. Mee, *Journal of computer-aided molecular design*, **1997**, 11 425-445
[57] B. Wang, S. Ekins, Computer applications in pharmaceutical research and development, John Wiley & Sons2006.
[58] R.T. Kroemer, *Current Protein and Peptide Science*, **2007**, 8 312-328
[59] D.E. Gloriam, R. Fredriksson, H.B. Schiöth, *BMC genomics*, **2007**, 8 338
[60] M.J. Marinissen, J.S. Gutkind, *Trends in pharmacological sciences*, **2001**, 22 368-376
[61] J.P. Overington, B. Al-Lazikani, A.L. Hopkins, *Nature reviews Drug discovery*, **2006**, 5 993-996
[62] P.G. Strange, *Current opinion in drug discovery & development*, **2008**, 11 196-202
[63] J. Carlsson, L. Yoo, Z.-G. Gao, J.J. Irwin, B.K. Shoichet, K.A. Jacobson, *Journal of medicinal chemistry*, **2010**, 53 3748-3755
[64] V. Katritch, V.-P. Jaakola, J.R. Lane, J. Lin, A.P. IJzerman, M. Yeager, I. Kufareva, R.C. Stevens, R. Abagyan, *Journal of medicinal chemistry*, **2010**, 53 1799-1809
[65] Y. Bustanji, I.M. Al-Masri, A. Qasem, A.G. Al-Bakri, M.O. Taha, *Chemical biology & drug design*, **2009**, 74 258-265
[66] S. Distinto, M. Yáñez, S. Alcaro, M.C. Cardia, M. Gaspari, M.L. Sanna, R. Meleddu, F. Ortuso, J. Kirchmair, P. Markt, *European journal of medicinal chemistry*, **2012**, 48 284-295
[67] O. Fedorov, S. Müller, S. Knapp, *Nature chemical biology*, **2010**, 6 166-169
[68] H. Zhao, D. Huang, A. Caflisch, *ChemMedChem*, **2012**, 7 1983-1990
[69] P. Germain, P. Chambon, G. Eichele, R.M. Evans, M.A. Lazar, M. Leid, A.R. De Lera, R. Lotan, D.J. Mangelsdorf, H. Gronemeyer, *Pharmacological reviews*, **2006**, 58 712-725
[70] G. Allenby, M.-T. Bocquel, M. Saunders, S. Kazmer, J. Speck, M. Rosenberger, A. Lovey, P. Kastner, J.F. Grippo, P. Chambon, *Proceedings of the National Academy of Sciences*, **1993**, 90 30-34

[71] M. Schapira, B.M. Raaka, H.H. Samuels, R. Abagyan, *Proceedings of the National Academy of Sciences*, **2000**, 97 1008-1013
[72] M. Schapira, B.M. Raaka, H.H. Samuels, R. Abagyan, *BMC structural biology*, **2001**, 1 1
[73] H. Wajant, K. Pfizenmaier, P. Scheurich, *Cell Death & Differentiation*, **2003**, 10 45-65
[74] B.B. Aggarwal, *Nature Reviews Immunology*, **2003**, 3 745-756
[75] D.S.H. Chan, H.M. Lee, F. Yang, C.M. Che, C.C. Wong, R. Abagyan, C.H. Leung, D.L. Ma, *Angewandte Chemie*, **2010**, 122 2922-2926