Review Article

Hacettepe University Journal of the Faculty of Pharmacy
Volume 40 / Number 1 / January 2020 / pp. 34-47

Molecular Modelling and Computer-Aided Drug Design: The Skill Set Every Scientist in Drug Research Needs and Can Easily Get

Moleküler Modelleme ve Bilgisayar Destekli İlaç Tasarımı: İlaç Araştırmalarında Yer Alan Her Bilim İnsanının İhtiyaç Duyduğu ve Kolayca Edinebileceği Beceri Seti

Suat SARI*

Hacettepe University, Faculty of Pharmacy, Department of Pharmaceutical Chemistry, Ankara, Turkey

Corresponding author:
Suat Sari
Hacettepe University, Faculty of Pharmacy, Department of Pharmaceutical Chemistry, 06100 Sihhiye, Ankara,Turkey
Tel.: + 90 312 305 18 72
Fax: + 90 312 305 32 72
E-mail: suat.sari@hacettepe.edu.tr , suat1039@gmail.com

Received date : 18.04.2020
Accepted date : 26.05.2020

ABSTRACT

The overwhelming advances in molecular modelling witnessed for a couple of decades go hand in hand with the booming computer and information technologies. Computer-aided drug design (CADD) is probably the most important field of molecular modelling given the time scale and cost for turning a chemical entity into an approved drug. In this review we provide a brief definition of molecular modelling and CADD with historical corner stones. In this review methods, tools, and applications of molecular modelling in different stages of CADD were focused on by referring to a number of success stories. Useful databases and non-commercial software for different purposes are also introduced. The review aims to provide a glimpse of these methods for scientists taking part in any field of drug research and to show that everyone can and should make the best of these methods with a vast amount of available free tools and documentation.

Keywords: molecular modelling, computer-aided drug design, virtual screening, molecular docking, pharmacophore modelling, shape similarity, molecular dynamics simulations, ADMET prediction

ÖZET

Moleküler modellemede son birkaç on yıldır tanıklık ettiğimiz baş döndürücü gelişmeler ilerleyen bilgisayar ve enformasyon teknolojileri ile birlikte gerçekleşmektedir. Bilgisayar destekli ilaç tasarımı (BDİT), bir kimyasalın ilaca dönüştürülmesi için gerekli zaman ve masraflar göz önüne alındığında, muhtemelen moleküler modellemenin en önemli alanıdır. Mevcut derlemede BDİT'nın farklı aşamalarındaki moleküler modelleme yöntemleri, araçları ve uygulamalarına, bazı başarı hikayelerine de atıfta bulunularak odaklanılmıştır. Ayrıca, farklı amaçlar için kullanılan faydalı veri tabanları ve ticari olmayan yazılımlar tanıtılmıştır. Derleme, bu yöntemlerle ilgili, ilaç araştırmalarının herhangi bir alanında görev alan bilim insanlarına fikir vermeyi ve herkesin bu yöntemlerden, mevcut çok sayıda ücretsiz yazılım ve dokümantasyon ile en iyi şekilde faydalanabileceğini göstermeyi amaçlamaktadır.

Anahtar Kelimeler: moleküler modelleme; bilgisayar destekli ilaç tasarımı; sanal aktivite tarama; moleküler kenetleme; farmakofor modelleme; şekil benzerliği; moleküler dinamik simülasyonları; ADMET tahmini

1. Introduction

The overwhelming advances in information technology observed in the past couple of decades, among many other things, made it possible to understand and predict molecules at every level of complexity by simply using our PCs or even our mobile devices. Thanks to multicore processors and high-end graphics available to almost everyone, molecular modelling is no longer the expertise of only those with highly technical computing skillset. Diversity of the methods in molecular modelling enabled this approach to reach to various fields in life sciences. Any researcher who crosses molecules' path can get to know of many things *in silico* before or after an experiment (1, 2). Is my molecule water soluble enough to run an assay or how can I make it water soluble? Can this compound pass blood-brain barrier? Does it contain any functional groups that can react nonspecifically with the assay medium? Can it inhibit CYP3A4? Could this peptide assume an α-helix conformation? Which mutations could possibly affect the function of my protein? Could this antagonist have triggered a desensitized state for my receptor? Questions like these can find more accurate answers in shorter time with the hands of the inquirers themselves thanks to molecular modelling.

Of course, molecular modelling has a special place in drug research. Given the time span and costs for a chemical entity to be labelled as "drug", the questions such as "which compound", "which targets", and "which off-targets" definitely need to be addressed at the very beginning. At this point, molecular modelling is put in use to work out what we call "computer-aided drug design" (CADD) to save time and money, to dodge pit falls and dead ends, and to detect blind spots (3).

In this review, molecular modelling and virtual screening is defined and a historical background is provided with drug discovery perspective. Steps of CADD are presented with state-of-the-art methodologies, applications, and success stories. The review also compiles resources for the most common tools of CADD.

2. Molecular Modelling and Its Historical Background

Molecular modelling is all the computational methods used to predict molecular structure and behavior. From material science to structural biology, molecular modelling methodologies are used in many fields to understand systems made up of a wide range of complexity from small molecules to biological macromolecules such as proteins, receptors, and nucleotide chains. These systems can be modelled by treating atoms as particles with charges and potential energy using forcefields (molecular mechanics) or by applying wave function at atomic and subatomic scales (quantum mechanics) (4). From drawn representations of chemical structures to millisecond-long simulations of biological systems, molecular modelling has undergone a massive progress through history (5). Following are some of the milestones of molecular modelling:

- The term chemical structure was introduced between 1858 and 1861 by identification of valence rules in organic chemistry and representing bonds as lines in molecules with carbon chains.
- In 1865, August Wilhelm Hofmann for the first time used physical models in his organic chemistry lecture, in which organic compounds such as methane and chloroform were represented with croquet balls joined by sticks.
- Hofmann also established today's commonly accepted color scheme for atoms: black for carbon, white for hydrogen, blue for nitrogen and green for chlorine.
- Crum-Brown in 1865 and Sir Edward Frankland and B. F. Duppa in 1867 were the scientists to use 2D drawn structures in ball and stick models for the first time.
- That carbon compounds have tetrahedral geometry was first suggested by E. Paterno (1869), Jacobus Henricus van't Hoff, and Joseph Achille LeBel (1874), which is considered as the emergence of three-dimensional molecular structure elaboration.
- In 1898, van't Hoff suggested that each carbon-carbon bond had a favored conformation about a torsional angle, setting the foundations of what we refer to as the global minimum energy conformations today.
- Urey and Bradley introduced a formulation which included quadratic Hooke's Law potential equations to describe harmonic vibrations in simple molecules and found the Morse potential to give the best fit to empirical data for bond stretching

ISSN: 2458 - 8806

in 1931. This was a breakthrough in forcefield concept.

- The first examples of molecular mechanics came in 1946 with Hill's force field to treat the repulsive and attractive nonbonding terms, and quadratic terms for bond stretch and angle bend; Dostrovsky, Hughes and Ingold's study on the repulsive and attractive nonbonding terms; and Westheimer and Mayer's analyses of the conformations of hindered biphenyls. These methods let molecular modelling spread out of physical chemistry society.
- The importance of 3D aspects of molecules in understanding structure, stability, conformation, and reactivity was appreciated with the emergence of conformation analysis led by the study of Barton on the conformations of steroids in 1950.
- The most famous physical molecular model of all time, the structure of DNA double strand, was elucidated by Watson and Crick in 1953.
- The first molecular dynamics (MD) calculations study was reported in the same year under the name Equation of State Calculations by Fast Computing Machines, which featured simulated annealing and aided the groundwork for Monte Carlo simulations.
- Potentials of nonbonding interactions in organic structure modelling were first applied by Kitaigorodsky in 1960.
- The first use of computer for empirical force field calculations was reported in 1961 by Hendrickson.
- In 1965, Wiberg developed a steepest descent algorithm for geometry optimization to address conformational analysis.
- Scientists from Oak Ridge National Laboratory in the US designed a molecular drawing program called ORTEP (Oak Ridge National Laboratory) in 1965. The US government also funded the first computer network in 1969, which is accepted as the ancestor of internet.
- In 1971, Lee and Richards described the molecular surface in protein structure context and elucidated an algorithm to derive it.
- In 1974, computer modelling of oligosaccharides starting from crystal structure was reported for the first time. In the same year, force field calculations of synthetic macromolecules started to appear.
- At the beginning of 1980s, with the booming personal computer (PC) industry, molecular modelling became accessible from PCs and the use of the graphical user interface (GUI) started. This marks the advent of personal molecular modelling for the average chemist.
- The World Wide Web kickstarted in 1993, which probably marks the beginning of web servers for molecular modelling.

3. Computer Aided Drug Design (CADD)

Drug design and discovery is highly complex and expensive process that requires contribution of a wide range of disciplines. The general estimation is that it takes more than 10 years and a billion US dollars for a chemical entity to be used in the clinic. Although *in silico* methods are usually adopted in the early-to-mid-stage drug discovery studies, the selection of candidates passed from these stages to preclinical and clinical phases greatly affects the attrition rates. Therefore, the use of *in silico* methods in drug discovery has increased reasonably for the past couple

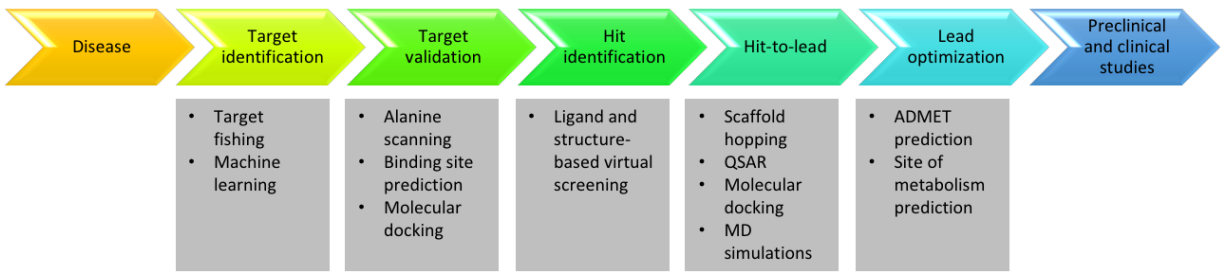

Figure 1. Steps of drug discovery and applications of CADD.

of decades (Figure 1). CADD applications mainly assist experimental studies in decision making by following the major drug discovery stages, such as target validation, hit identification, lead generation, and optimization (3, 6).

3.1. Target Identification and Validation

Drug development process usually starts with identification of a druggable target relevant to the disease of interest. Introduction of such targets prompts efforts to find potential compounds to modulate the target pathway and finally to elicit a phenotypic response (7). *In silico* approaches are usually helpful to understand the structure, topology, and ligand binding sites of the target, as well as their key residues and ligand binding interactions. *In silico* methods such as homology modelling, molecular docking, and MD simulations have become paramount to assist *in vitro* studies for target validation such as protein structure elucidation (e.g. X ray diffraction), alanine scanning, site-directed mutagenesis, and radioligand binding (8-10).

Chemical and biological databases play crucial role in bringing the wet lab and computers together. Currently, there is a vast number of databases which archive and process countless sorts of data belonging to millions of molecules ranging from small molecules to proteins, genes, and nucleotide chains (Table 1) (11).

3.2. Hit Identification

Once a target is validated for a certain disorder, campaigns are run for identification of hit matter, i.e. chemical entities that display an intended activity against the target, by both pharma companies and academia. Just as the robotics technology made it possible to randomly screen biological activity of thousands of compounds *in vitro* at the same time (high-throughput screening), the computer technology did the same for screening millions of compounds *in silico*. With the so-called high-throughput virtual screening (or briefly virtual screening), biological screening of small compound sets rationally selected from huge libraries became more common. Virtual screening features a number of filters to narrow down the libraries step by step and to eventually suggest potential hits for *in vitro* tests (12). As the filters go from rough to exhaustive, the computational burden and time required increases dramatically, which, however is balanced with shrinking library throughout the steps. Evaluation of druglike chemical space and PAINS (Pan-assay Interference Compounds) is an example for rough filters, which is commonly applied at the beginning of virtual screening campaigns (13, 14). This step is crucial for eliminating entities with potentially poor pharmacokinetics and those with non-specific reactive functionalities, since such compounds account for a good amount of failed clinical candidates (15). This step is usually followed by a ligand- and/or structure-based virtual screening (16).

3.3. Lead Generation and Optimization

Finding hit compounds with ability to modulate an isolated target is the tip of the iceberg in drug discovery issues. When evaluated thoroughly (plotting dose-response curves, profiling for related targets, testing against cell cultures, *in vivo* and toxicity profiling, cytochrome P450 and efflux pump interactions, pharmacokinetics etc.) these hits will most likely fail at a certain point. This is where medicinal chemists take the stage to tailor these hits to fit specific requirements (17). *In silico* drug discovery methods enable medicinal chemists to create libraries of countless virtual compounds envisaged accordingly. The so-called scaffold-hopping methodology utilizes both ligand- and structure-based approaches to model structurally relevant molecules as synthetic candidates (18). Establishing structure-activity relationships (SARs) at this stage is crucial for creating virtual libraries. Quantitative SARs (QSAR) is an inevitable computation tool for decision making regarding SARs (19). The synthesized compounds are then subjected to a set of *in vitro* and *in vivo* tests to obtain leads.

Lead optimization stage mainly focuses on pharmacokinetic and toxicity issues, i.e. ADME+T (Absorption, Distribution, Metabolism, Excretion + Toxicity), therefore *in silico* methods are less likely to be included at this level. However molecular modelling offers a wide range of tools to assist decision making for various scenarios. There are many web servers and free tools to evaluate ADME+T profile of compounds through fast *in silico* predictions, as well as more sophisticated, specific, and exhaustive models. Site of metabolism, hERG channel affinity, blood-brain barrier permeability, human serum albumin binding, CYP and P-glycoprotein affinity are among the properties that scientists can predict without a high level of expertise in computational chemistry

ISSN: 2458 - 8806

Table 1. Examples of useful chemical and biological databases

Database [1]	Source	Description
ZINC	https://zinc.docking.org	Indexes more than 230M commercial compounds in ready-to-dock format with diverse data including predicted and biological properties and vendor information.
ChEMBL	https://www.ebi.ac.uk/chembl/	Archives 2M compounds with detailed experimental data, target, assay, and literature information.
DrugBank	www.drugbank.ca	Includes 13,551 drug entries (approved and experimental) with detailed information including drug targets.
PubChem	https://pubchem.ncbi.nlm.nih.gov	Archives 103M compounds with relevant chemical, bioactivity, and literature information.
ChemSpider	www.chemspider.com	Provides 81M compounds with 278 data sources including supplier information.
DUD-E	http://dude.docking.org	Provides 22K active compounds with target information and 50 decoys for each active compound for virtual screening enrichment studies.
BindingDB	www.bindingdb.org/bind	Stores measured affinities of 820K small molecules to 7K protein targets.
Protein data banks	www.wwpdb.org www.pdbe.org www.rcsb.org www.bmrb.wisc.edu www.pdbj.org	Store three-dimensional structural data of biological macromolecules. Currently, more than 160K structurers are available.
UniProt	www.uniprot.org	Provides sequence and functional information for some 177M proteins.
NCBI	www.ncbi.nlm.nih.gov	Along with literature (PubMed) and many other databases, NCBI is one of the largest protein, DNA, RNA, genome, and gene databases.
SCOP	http://scop.mrc-lmb.cam.ac.uk	Provides a detailed and comprehensive description of structural and evolutionary relationships between 532K proteins with known structure.
BioGRID	https://thebiogrid.org	Biological database of protein-protein interactions, genetic interactions, chemical interactions, and post-translational modifications.
PROSITE	https://prosite.expasy.org	Consists of documentation entries describing protein domains, families and functional sites as well as associated patterns and profiles to identify them.

[1]in alphabetical order

and bioinformatics (20-23).

4. Approaches in CADD

Rapid development in information technologies brought about a vast amount of data, which increased our ability to create highly predictive models. In CADD, data is related to either small molecules or their potential targets, i.e. biological macromolecules. Virtual screening approaches in CADD can be classified as ligand-based and structure/receptor-based methods according to the type of employed data. Ligand-based methods use models built exclusively from ligand data, while structure-based or receptor-based approaches benefit from structural data of target macromolecules. Virtual screening methods, in most cases, incorporate both approaches, which are referred to as hybrid methods. This, of course, has much to do with the extend of the data available (16).

4.1. Ligand-Based Methods

The general consideration in CADD is that similar ligands are supposed to exert similar biological effects, which lies beneath the rationale of ligand-based

CADD (24). Similarity can be measured in diverse ways, i.e. molecular descriptors, which incorporate molecular properties. According to the complexity and organization of the data, ligand-based methods are classified as 1D-, 2D- or 3D-methods (25). Descriptors for 1D-methods include physicochemical properties of molecules without structural data, such as molecular weight, H bond acceptor count, and polar surface area. The idea of 1D-methods suggests that compounds belong to a specific group, such as orally available small-molecule drugs, possess similar set of physicochemical properties. This kind of methods are helpful to filter libraries in terms of drug-likeness or lead-likeness or to predict certain pharmacokinetic properties to help decision making (26). 2D-methods employ physicochemical properties assigned to 2D-structure of compounds. Molecular fingerprints are widely used to determine similarity between compounds regarding fragment connectivity independent from their spatial orientations, which is handled by 3D-methods, such as pharmacophore modelling (Figure 2) and shape similarity (27-29). The increase in the order of dimensions causes computational burden. For example, in 3D-methods, similarity between two compounds is calculated by finding the best molecular volume alignment, which is selected from possible conformations of the compounds. This requires coordination data for each atom, which is re-calculated for each geometry. However, increasing computational burden is usually the result of demands for higher precision in molecular modelling. In addition, 3D-geometry and spatial volume are valuable data in CADD considering drug-target affinity. In an effort to reduce computation burden by maintaining precision, new methods were introduced in which 3D-molecular fields are reduced into 1D-descriptors or molecular volume information is compressed and described as vectors (30, 31).

4.2. Structure-Based Methods

Two things gave rise to the emergence of structure-based methods as powerful tools for CADD: advances in biomolecular spectroscopic methods, such as X ray diffraction and NMR, and rapid development of computer technologies, especially processors and graphics. X ray diffractometers have become widely accessible, which triggered an avalanche of structural data of biological macromolecules deposited in the web sites called protein data bank. Handling macromolecule data requires high-performing processors and graphics, which is available to individuals, except clusters of parallel processors required in the case complex dynamics simulations. The impor-

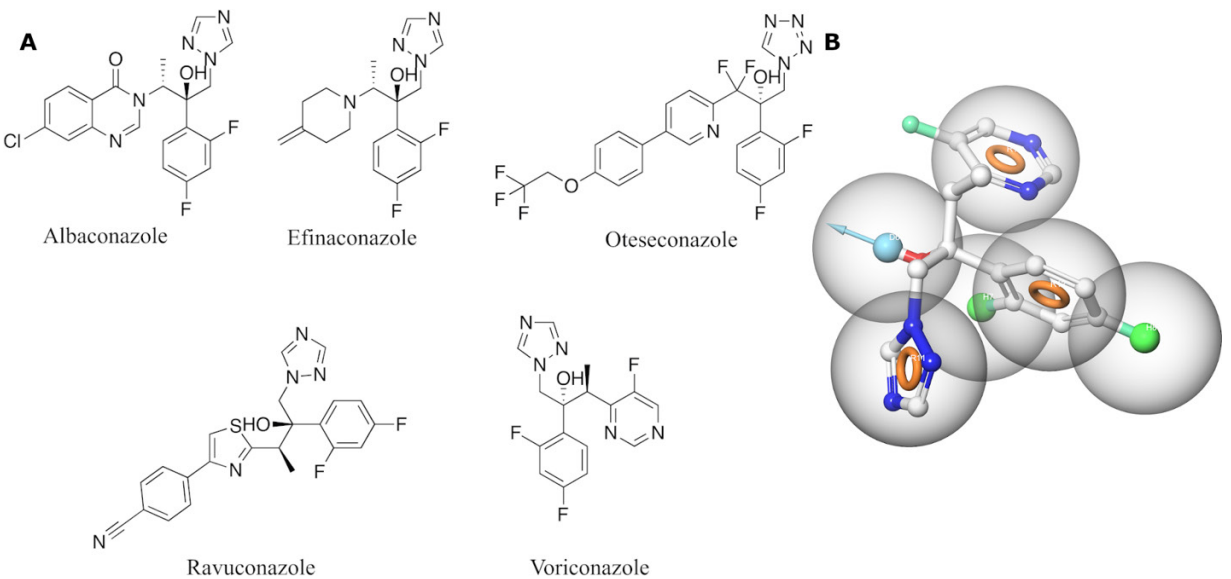

Figure 2. Last-generation azole antifungals (A) and a pharmacophore model of the last-generation azole antifungals aligned with voriconazole (B). The model consists of three rings, two hydrophobic groups, and one H bond donor represented as brown rings, green spheres and a blue sphere with an arrow, respectively. The cut-off space for each pharmacophore is highlighted with a transparent sphere. Voriconazole is represented as gray sticks and balls.

ISSN: 2458 - 8806

tance of target macromolecule structure data became apparent with the ever expanding the knowledge of ligand-receptor affinity and interactions at molecular level. CADD benefits from these developments in several situations some of which are listed below (32):

- Creating a structural model for a protein when there is no spectroscopic or experimental structural data
- Identification of possible druggable sites for a receptor
- Ligand-receptor interactions and how these interactions manipulate signal transduction
- Possible effects of certain residues, co-factors, ions, and solvents on conformational changes in macromolecule structure and its gating
- The effects of charge polarization in ligand-receptor binding

For these issues and more, there are powerful and popular tools, such as homology modelling, molecular docking, and MD simulations, which are available either free of charge or through paid licenses (32).

4.2.1. Homology Modelling

Proteins with similar amino acid sequences tend to have similar tertiary structures. Homology modelling is a method to predict three-dimensional structure of a protein when there is no experimental structural data available. This is performed by the help of homologous proteins with experimental structural data. The process starts with sequence alignment of the query structure with one or more template protein(s). This alignment and the atomic coordinates of the template protein(s) are then used to thread the structural model of the query protein. The model is further optimized by loop refinement and side chain minimization, and subjected to various structural assessment methods for validation (33, 34).

Apart from creating a whole structure model, homology modelling can be used to fill in loops and missing side chains, something usually encountered with the structures obtained from protein data bank. Also, *in silico* modifications in protein structure such as alanine scanning is possible via homology modelling. Thus, homology modelling is a powerful tool for structure-based modelling employed in target validation, hit generation, and optimization studies (35-37).

4.2.2. Molecular Docking

Molecular docking is an *in silico* technique to predict the preferred binding orientation and affinity of two molecules, a ligand and a receptor, to form a stable complex. Receptor in molecular docking is usually a macromolecule such as protein, DNA, RNA, or peptide, which is kept rigid; while ligand is a flexible small molecule. However, techniques for docking two macromolecules to each other or making both ligand and receptor flexible are available. Molecular docking predicts countless ligand-receptor complexes (search space) and ranks them according to a score (docking score) which is calculated by a scoring function (Figure 3). Docking score is a metric used to predict how much affinity a ligand is bound to a receptor with, which is usually governed by non-bonded interactions, although it is possible to model possible covalent bonds between ligand and receptor by covalent docking method (38).

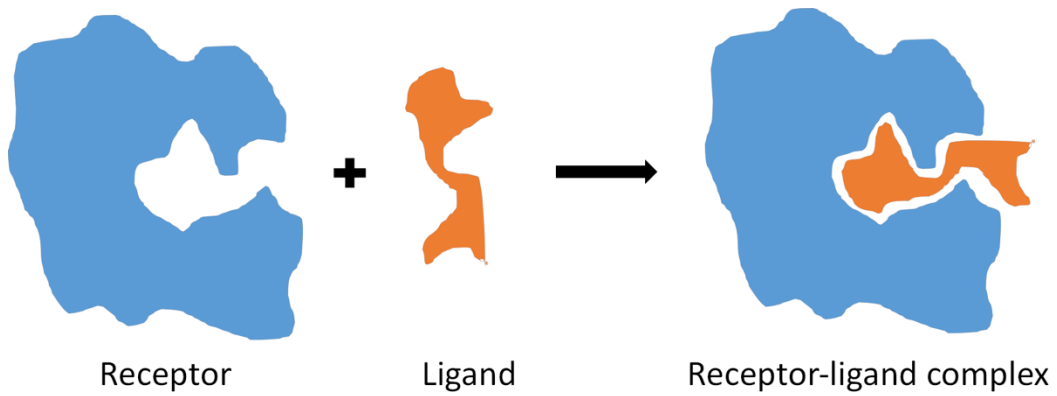

Figure 3. Molecular docking.

Biological systems work by means of signal transduction, which is pretty much affected by the attributes of ligand-receptor complexes. By predicting ligand-receptor complexes it is possible to predict the outcomes of signal transduction, which makes molecular docking a precious tool for drug design. In recent years, molecular docking has become an indispensable component of virtual screening. On the other hand, it is routinely applied to study ligand-receptor interactions at atomic level to understand the importance of certain amino acids, cofactors, chelating with metals, hydrogen and halogen bonds, solvation effects, and more (Figure 4) (39).

4.2.3. MD Simulations

While molecular docking provides a picture of a biological process, it is the dynamics of this process that actually matters for its biological consequences. MD of a system is the physical movements of all its atoms, which is simulated by adding Newtonian mechanics to the initial conditions of each atom, i.e. energies and coordinates. Then, forces between atoms or particles and potential energies are calculated at each given time period to determine trajectories for atoms or molecules, which reflect the dynamic evolution of the system (42).

Depending on the number of the atoms of the system and the presumed time scale, MD simulations require a great amount of data usage and time, especially compared to molecular docking or other target and ligand-based methods. To overcome this computational burden in a reasonable amount of time, parallel processors were introduced, some of which are available to researchers via remote access. Lately, an evolutionary method called Graphics Processing Unit (GPU) acceleration has made use of high-end graphics processors to greatly reduce computation time, making researchers less dependent to costly CPU clusters (43).

In addition to target validation, MD simulations are gradually becoming a part of virtual screening campaigns thanks to the advances mentioned above. MD simulations are usually considered in connection with molecular docking in virtual screening, for example to determine the stability of a ligand-receptor complex, however its connected use with ligand-based methods is becoming increasingly popular (44, 45).

5. CADD: Success Stories

In contrast with what one would expect, most of the chemical entities labelled as “drug” today come from serendipitous studies such as combinatorial chemistry (46). This is partly due to a sharp decrease in the speed of new entities hitting the market observed past couple of decades, when CADD is most

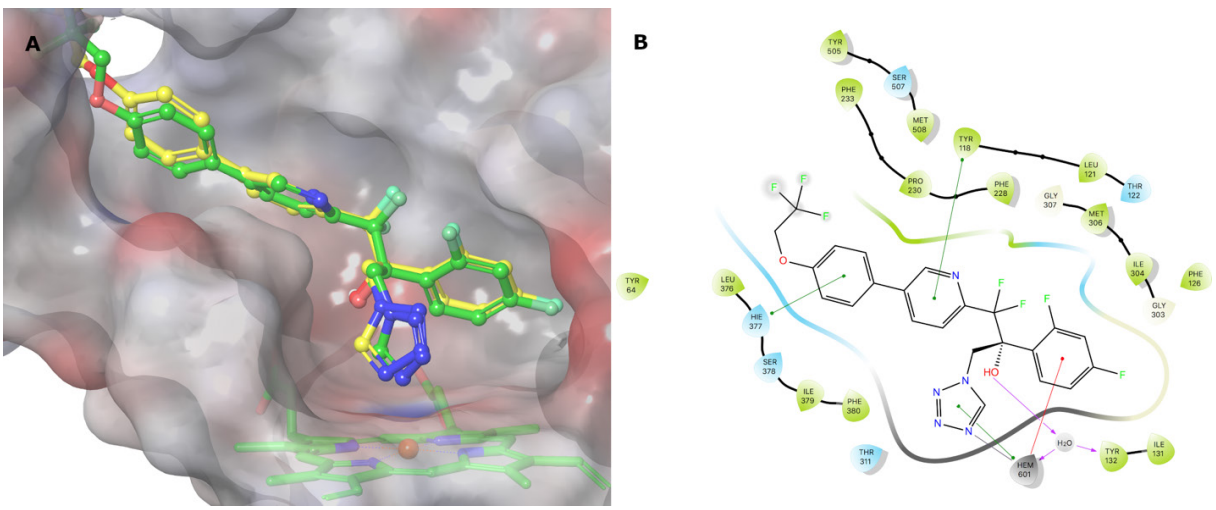

Figure 4. Superposition of the co-crystallized ligand (oteseconazole) and its docking pose obtained by Glide (2019-4, Schrödinger, LLC, NY, USA) (40) in *Candida albicans* lanosterol 14α demethylase (CACYP51) active site (A) and binding interactions of oteseconazole with CACYP51 according to the crystallographic study (B) (41). The images were rendered using Maestro (2019-4, Schrödinger, LLC, NY, USA).

ISSN: 2458 - 8806

expected to bear fruit. Still, examples of drugs discovered through CADD inspire researchers from different backgrounds, three of which are provided as example in this review.

Dorzolamide (Merck): Marketed as an eye-drop for glaucoma, dorzolamide was developed through structure-based drug design (47). This compound was developed as a human carbonic anhydrase II (hCAII) inhibitor and tailored to fit in the specific active site of hCAII starting from a parent racemate compound (MK-927), whose enantiomers show 100-fold difference in affinity to hCAII (48, 49). X ray crystallography studies and *ab initio* calculations identified the conformational variations that caused the observed potency difference between the enantiomers. Therefore, further modifications were taken on for the *S* enantiomer of MK-927 using structure-based modelling and introduction of a methyl group to the 6th position yielded a more potent derivative. The 4-isobuthylamino group was replaced with ethylamino group to counter the reduced water solubility. The resulting four enantiomers were evaluated through *in vitro* and crystallographic studies and the *trans S,S* configuration (dorzolamide) was found to have the greatest potency (50).

Zanamivir (GSK) and Oseltamivir (Gilead Sciences): The antiviral drug zanamivir acts through viral neuraminidase inhibition and is used against influenza infections (51). Following the structural elucidation of neuraminidase via X ray crystallography, structure-based virtual screening campaigns were conducted to find potential anti-viral inhibitors and zanamivir is the result of one of these campaigns using the software GRID (52). Zanamivir was designed from 2-deoxy-2,3-dehydro-*N*-acetylneuraminic acid by substitution of the 4-hydroxy group with 4-guanidino. Further structure-based design to utilize an extra binding gorge on neuraminidase led to the discovery of more potent and orally available oseltamivir (47).

Aliskiren (Novartis): Aliskiren was designed as a renin inhibitor for the treatment of hypertension. It is also the first member of the class called direct renin inhibitors. Aliskiren's design starts with the efforts to find renin inhibitors by mimicking the natural peptide substrate of the renin system (53). The first non-peptide derivative was developed by Goschke et al. (1997), which was a success compared to the previous peptide derivatives in terms of pharmacokinetic profile (54). With a structure-based modelling using a crystal structure of renin, the lead designed by Goschke et al. was further optimized to improve potency. Identification of an additional pocket in the renin catalytic site through the following X ray crystallography studies led to rational design of new derivatives with better affinity and selectivity to renin over other aspartic peptidases and further SAR studies to optimize *in silico* interactions with renin resulted in a derivative, which was aliskiren, with potency at sub-nanomolar concentration (55).

6. Resources for Molecular Modelling Tools

Thanks to the age of information and the growing "open source, open data" trend, it has never been easier to understand and utilize molecular modelling tools of different sorts. The internet is full of free-to-use applications or web servers as well as tutorial materials (Table 2).

Increasing popularity of molecular modelling created demands for hands-on training and workshops, which are now a common part of scientific meetings of related fields. These workshops, some of which are supported by the leading molecular modelling software companies, offer important opportunities, especially for postgraduates and postdoctoral fellows. In this respect, we organized the 1st Molecular Modelling Workshop in Hacettepe University Faculty of Pharmacy with Hacettepe University Medicinal Chemistry Research, Development, and Application Center (MAGUM) on December 5-6, 2019. The workshop focused on state-of-the-art techniques in molecular modelling for CADD (56). The participants had the opportunity to know academic free modelling tools and experience hands-on training. The workshop also provided a general idea, opportunities, catches, and pitfalls of virtual screening.

7. Conclusions

The overwhelming developments in computer and information technologies for the past couple of decades have boosted molecular modelling and its key application, CADD. As computers became an indispensable material of research, so did molecular modelling techniques in CADD at all levels, from sketching a molecule to running millisecond MD simula-

Table 2. List of example non-commercial tools for molecular modelling and CADD[1] in alphabetical order

Software [1]	Type	Application	Source	License	Developer(s)
AutoDock	Stand-alone software	Molecular docking	http://autodock.scripps.edu	Free	The Scripps Research Institute
AutoDock Vina	Stand-alone software	Molecular docking	http://vina.scripps.edu	Free	The Scripps Research Institute
Avogadro	Stand-alone software	3D molecular edition and molecular mechanics calculations	https://avogadro.cc	Free	University of Pittsburgh
DOCK	Stand-alone software	Molecular docking	http://dock.compbio.ucsf.edu	Academic free	University of California-San Francisco
Chimera	Stand-alone software	Interactive structure visualization	https://www.cgl.ucsf.edu/chimera	Free	University of California-San Francisco
DS Visualizer	Stand-alone software	Interactive structure visualization	https://www.3dsbiovia.com/products/collaborative-science/biovia-discovery-studio/visualization-download.php	Academic free	BIOVIA
GROMACS	Stand-alone software	MD simulations	http://www.gromacs.org	Free	Science for Life Laboratory, Stockholm University
HADDOCK	Stand-alone software	Protein-protein docking	http://milou.science.uu.nl/services/HADDOCK2.2	Academic free	Centre Bijvoet Center for Biomolecular Research
I-TASSER	Web server	Comparative protein modelling	https://zhanglab.ccmb.med.umich.edu/I-TASSER	Free	Zhang Lab, University of Michigan
KNIME	Stand-alone software	Descriptor calculation, QSAR modelling	www.knime.com	Free	KNIME AG
Maestro	Stand-alone software	Interactive structure visualization	https://www.schrodinger.com/freemaestro	Academic free	Schrödinger
MarvinSketch	Stand-alone software	Small molecule sketch and edition	https://chemaxon.com/products/marvin	Academic free	ChemAxon
Memoir	Web server	Membrane protein modelling	http://opig.stats.ox.ac.uk/webapps/memoir	Free	Oxford Protein Informatics Group, Oxford University
Modeller	Stand-alone software	Comparative protein modelling	https://salilab.org/modeller	Free	University of California San Francisco
Molinspiration	Web server	Descriptor calculation	www.molinspiration.com	Free	Molinspiration Cheminformatics
MolProbity	Stand-alone software	Fixing and hydrogen modelling for PDB structures	http://molprobity.biochem.duke.edu	Free	School of Medicine, Duke University

ISSN: 2458 - 8806

Table 2. continued...

NAMD and VMD	Stand-alone software	MD simulations	https://www.ks.uiuc.edu/Development	Academic free	Theoretical and Computational Biophysics Group
OpenEye	Stand-alone software	Molecular modelling and CADD platform	www.eyesopen.com		OpenEye Scientific
PharmaGist	Web server	Pharmacophore modelling	http://bioinfo3d.cs.tau.ac.il/PharmaGist	Free	School of Computer Science, Tel Aviv University
Pharmer	Stand-alone software	Pharmacophore modelling	http://smoothdock.ccbb.pitt.edu/pharmer	Free	Camacho Lab, University of Pittsburgh
PharmMapper	Web server	Pharmacophore modelling	www.lilab-ecust.cn/pharmmapper	Free	Honglin Li's Lab, East China University of Science & Technology
PPM Server	Web server	Prediction of membrane protein orientation	https://opm.phar.umich.edu/ppm_server	Free	Lomize Group, University of Michigan
PROCHECK	Web server	Protein structure validation	https://servicesn.mbi.ucla.edu/PROCHECK	Free	University of California-San Francisco
QSAR Toolbox	Stand-alone software	QSAR modelling	https://qsartoolbox.org	Free	The Laboratory of Mathematical Chemistry, Prof. Dr. As. Zlatarov University
QSARINS-Chem	Stand-alone software	QSAR modelling	http://www.qsar.it	Free	The QSAR Research Unit in Environmental Chemistry and Ecotoxicology, University of Insubria
Rosetta	Stand-alone software	De novo protein modelling	https://www.rosettacommons.org/software	Academic free	RosettaCommons
Silicos-it	Collection of stand-alone software	Ligand-based virtual screening	http://silicos-it.be.s3-website-eu-west-1.amazonaws.com	Free	Silicos-it
SwissADME	Web server	Descriptor calculation, ADMET prediction	www.swissadme.ch	Free	Swiss Institute of Bioinformatics
SwissDock	Web server	Molecular Docking	www.swissdock.ch	Free	Swiss Institute of Bioinformatics
SWISS-MODEL	Web server	Comparative protein modelling	https://swissmodel.expasy.org	Free	Swiss Institute of Bioinformatics
SwissSimilarity	Web server	Ligand-based virtual screening	www.swisssimilarity.ch	Free	Swiss Institute of Bioinformatics

tions, for scientists of drug research. There are many examples, in which the cost and time needed for drug discovery is reduced, complex decision-making processes are eased, disease pathways and interplay of diverse molecules are understood. In addition to availability of fast-processing computers, exponentially increasing amount of open data in this field is encouraging to learn and utilize molecular modelling in CADD. Actually, today there are more molecular modelling tools that are free or academic free than commercial software with paid license. The ocean called "internet" is full of documentations, tutorials, and forums to help through getting familiar with these tools and troubleshooting. Also, molecular modelling workshops, separate or as part of scientific meetings, became very common and wide-spread. Thus, every scientist in drug research can and should make the best of molecular modelling.

References

1. Dalkas GA, Vlachakis D, Tsagkrasoulis D, Kastania A, Kossida S: State-of-the-art technology in modern computer-aided drug design. Briefings in Bioinformatics 2013, 14(6):745-752.
2. Ooms F: Molecular modeling and computer aided drug design. Examples of their applications in medicinal chemistry. Current Medicinal Chemistry 2000, 7(2):141-58.
3. Kapetanovic IM: Computer-aided drug discovery and development (CADDD): in silico-chemico-biological approach. Chemico-Biological Interactions 2008, 171(2):165-176.
4. Leach AR: Molecular Modelling : Principles and Applications. Pearson Education Ltd.; Essex, UK, 2001.
5. Schlecht MF: Molecular Modeling on the PC. Wiley-VCH; New York, USA, 1997.
6. Ramirez D: Computational Methods Applied to Rational Drug Design. The Open Medicinal Chemistry Journal 2016, 10:7-20.
7. Blake RA: Target validation in drug discovery. Methods in Molecular Biology 2007, 356:367-77.
8. Agamah FE, Mazandu GK, Hassan R, Bope CD, Thomford NE, Ghansah A, et al.: Computational/in silico methods in drug target and lead prediction. Briefings in Bioinformatics 2019, bbz103.
9. Sahu TK, Pradhan D, Rao AR, Jena L: In silico site-directed mutagenesis of neutralizing mAb 4C4 and analysis of its interaction with G-H loop of VP1 to explore its therapeutic applications against FMD. Journal of Biomolecular Structure and Dynamics 2019, 37(10):2641-2651.
10. Anand P, Nagarajan D, Mukherjee S, Chandra N: ABS-Scan: In silico alanine scanning mutagenesis for binding site residues in protein-ligand complex. F1000Research 2014, 3:214.
11. Audouze K, Taboureau O: Chemical biology databases: from aggregation, curation to representation. Drug Discovery Today: Technologies 2015, 14:25-29.
12. Hughes JP, Rees S, Kalindjian SB, Philpott KL: Principles of early drug discovery. British Journal of Pharmacology 2011, 162(6):1239-1249.
13. Schneider N, Jackels C, Andres C, Hutter MC: Gradual in silico filtering for druglike substances. Journal of Chemical Information and Modeling 2008, 48(3):613-628.
14. Baell J, Walters MA: Chemistry: Chemical con artists foil drug discovery. Nature 2014, 513(7519):481-483.
15. Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J: Clinical development success rates for investigational drugs. Nature Biotechnology 2014, 32(1):40-51.
16. Katsila T, Spyroulias GA, Patrinos GP, Matsoukas MT: Computational approaches in target identification and drug discovery. Computational and Structural Biotechnology Journal 2016, 14:177-184.
17. Keseru GM, Makara GM: Hit discovery and hit-to-lead approaches. Drug Discovery Today 2006, 11(15-16):741-748.
18. Zhao H. Scaffold selection and scaffold hopping in lead generation: a medicinal chemistry perspective. Drug Discovery Today 2007, 12(3-4):149-155.
19. Subramaniyan B: QSAR and Lead Optimization. In: Puratchikody A, Lakshmana Prabu S, Umamaheswari A (eds.), Computer Applications in Drug Discovery and Development, IGI Global; Pennsylvania, USA. 2018: pp 21.
20. Hosea NA, Jones HM: Predicting pharmacokinetic profiles using in silico derived parameters. Molecular Pharmaceutics 2013, 10(4):1207-1215.
21. Daina A, Michielin O, Zoete V: SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific Reports 2017, 7:42717.
22. Chemi G, Gemma S, Campiani G, Brogi S, Butini S, Brindisi M: Computational Tool for Fast in silico Evaluation of hERG K(+) Channel Affinity. Frontiers in Chemistry 2017, 5:7.
23. Bikadi Z, Hazai I, Malik D, Jemnitz K, Veres Z, Hari P, et al.: Predicting P-glycoprotein-mediated drug transport based on support vector machine and three-dimensional crystal structure of P-glycoprotein. PLoS One 2011, 6(10):e25815.
24. Martin YC, Kofron JL, Traphagen LM: Do structurally similar molecules have similar biological activity? Journal of Medicinal Chemistry 2002, 45(19):4350-4358.

ISSN: 2458 - 8806

25. Zoete V, Daina A, Bovigny C, Michielin O. SwissSimilarity: A Web Tool for Low to Ultra High Throughput Ligand-Based Virtual Screening. Journal of Chemical Information and Modeling 2016, 56(8):1399-1404.
26. Bajorath J: Selected concepts and investigations in compound classification, molecular descriptor analysis, and virtual screening. Journal of Chemical Information and Modeling 2001, 41(2):233-245.
27. Cereto-Massague A, Ojeda MJ, Valls C, Mulero M, Garcia-Vallve S, Pujadas G: Molecular fingerprint similarity search in virtual screening. Methods 2015, 71:58-63.
28. Caporuscio F, Tafi A: Pharmacophore modelling: a forty year old approach and its modern synergies. Current Medicinal Chemistry 2011, 18(17):2543-2553.
29. Kumar A, Zhang KYJ: Advances in the Development of Shape Similarity Methods and Their Application in Drug Discovery. Frontiers in Chemistry 2018, 6:315.
30. Thijs G, Langenaeker W, De Winter H: Application of spectrophores™ to map vendor chemical space using self-organising maps. Journal of Cheminformatics 2011, 3(1):P7.
31. Ballester PJ, Richards WG: Ultrafast shape recognition to search compound databases for similar molecular shapes. Journal of Computational Chemistry 2007, 28(10):1711-1723.
32. Batool M, Ahmad B, Choi S: A Structure-Based Drug Discovery Paradigm. International Journal of Molecular Sciences 2019, 20(11): E2783.
33. Muhammed MT, Aki-Yalcin E: Homology modeling in drug discovery: Overview, current applications, and future perspectives. Chemical Biology & Drug Design 2019, 93(1):12-20.
34. Schwede T: Homology Modeling of Protein Structures. In: Roberts GCK (eds), Encyclopedia of Biophysics. Springer; Berlin, Germany. 2013: pp 992-998.
35. Fiser A, Sali A: ModLoop: automated modeling of loops in protein structures. Bioinformatics 2003, 19(18):2500-2501.
36. Liang S, Grishin NV: Side-chain modeling with an optimized scoring function. Protein Science 2002, 11(2):322-331.
37. Saranyah K, Kalva S, Mukund N, Singh SK, Saleena LM: Homology modeling and in silico site directed mutagenesis of pyruvate ferredoxin oxidoreductase from Clostridium thermocellum. Combinatorial Chemistry & High Throughput Screening 2015, 18(10):975-989.
38. Meng XY, Zhang HX, Mezei M, Cui M: Molecular docking: a powerful approach for structure-based drug discovery. Current Computer-Aided Drug Design 2011, 7(2):146-157.
39. de Ruyck J, Brysbaert G, Blossey R, Lensink MF: Molecular docking as a popular tool in drug design, an in silico travel. Advances and Applications in Bioinformatics and Chemistry 2016, 9:1-11.
40. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, et al.: Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. Journal of Medicinal Chemistry 2004, 47(7):1739-1749.
41. Hargrove TY, Friggeri L, Wawrzak Z, Qi A, Hoekstra WJ, Schotzinger RJ, et al.: Structural analyses of Candida albicans sterol 14alpha-demethylase complexed with azole drugs address the molecular basis of azole-mediated inhibition of fungal sterol biosynthesis. Journal of Biological Chemistry 2017, 292(16):6728-6743.
42. Haile JM: Molecular Dynamics Simulation: Elementary Methods. Wiley; Chichester, UK, 1992.
43. Friedrichs MS, Eastman P, Vaidyanathan V, Houston M, Legrand S, Beberg AL, et al.: Accelerating molecular dynamic simulation on graphics processing units. ournal of Computational Chemistry 2009, 30(6):864-872.
44. Salmaso V, Moro S: Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. Frontiers in Pharmacology 2018, 9:923.
45. Wieder M, Perricone U, Seidel T, Langer T: Pharmacophore Models Derived From Molecular Dynamics Simulations of Protein-Ligand Complexes: A Case Study. Natural Product Communications 2016, 11(10):1499-1504.
46. Ban TA: The role of serendipity in drug discovery. Dialogues in Clinical Neuroscience 2006, 8(3):335-344.
47. Talele TT, Khedkar SA, Rigby AC: Successful applications of computer aided drug discovery: moving drugs from concept to the clinic. Current Topics in Medicinal Chemistry 2010, 10(1):127-141.
48. Eriksson AE, Jones TA, Liljas A: Refined structure of human carbonic anhydrase II at 2.0 A resolution. Proteins 1988, 4(4):274-282.
49. Baldwin JJ, Ponticello GS, Anderson PS, Christy ME, Murcko MA, Randall WC, et al.: Thienothiopyran-2-sulfonamides: novel topically active carbonic anhydrase inhibitors for the treatment of glaucoma. Journal of Medicinal Chemistry 1989, 32(12):2510-2513.
50. Greer J, Erickson JW, Baldwin JJ, Varney MD: Application of the three-dimensional structures of protein target molecules in structure-based drug design. Journal of Medicinal Chemistry 1994, 37(8):1035-1054.
51. McKimm-Breschkin JL: Influenza neuraminidase inhibitors: antiviral action and mechanisms of resistance. Influenza and Other Respiratory Viruses 2013, 7 Suppl 1:25-36.
52. von Itzstein M, Wu WY, Kok GB, Pegg MS, Dyason JC, Jin B, et al.: Rational design of potent sialidase-based inhibitors of influenza virus replication. Nature 1993, 363(6428):418-423.

53. Wolfenden R: Analog approaches to the structure of the transition state in enzyme reactions. Accounts of Chemical Research 1972, 5(1):10-18.

54. Göschke R, Cohen NC, Wood JM, Maibaum J: Design and synthesis of novel 2,7-dialkyl substituted 5(S)-amino-4(S)-hydroxy-8-phenyl-octanecarboxamides as in vitro potent peptidomimetic inhibitors of human renin. Bioorganic & Medicinal Chemistry Letters 1997, 7(21):2735-2740.

55. Rahuel J, Rasetti V, Maibaum J, Rueger H, Goschke R, Cohen NC, et al.: Structure-based drug design: the discovery of novel nonpeptide orally active inhibitors of human renin. Chemical Biology 2000, 7(7):493-504.

56. Sarı S. Birinci Moleküler Modelleme Çalıştayı. Hacettepe ÜniversitesiEczacılıkFakültesiMedi□si□nalKi□myaAraştırma, Geli□şti□rmeveUygulamaMerkezi□.5-6Aralık2019,Ankara, Türkiye.

ISSN: 2458 - 8806