MOLECULAR MODELİNG METHODS USED İN DRUG DESİGN

Hulya Celik, Kübra Akman

Ağrı Ibrahim Cecen University, Faculty of Pharmacy Basıc Pharmaceutıcal Scıences Department / Fundamental Scıences of Pharmacy 03200 Agri/TURKEY
Corresponding author: Hülya ÇELİK
E-mail: hycelik@agri.edu.tr

ABSTRACT

Computer-aided design is used in biomedical fields to accelerate and assist the determination of targeted drug active ingredient, assist in precursor drug selection, optimize absorption, distribution, metabolism, excretion and toxicity profile, and prevent safety problems. Commonly used computational approaches include ligand-based drug design (pharmacophore, a 3-dimensional spatial arrangement of chemical properties required for biological activity), structure-based drug design (drug-target placement), and quantitative structure-activity and quantitative structure-property. In addition, computer-aided drug design is applied at an early stage of the drug development process for target verification, hit determination, target achievement optimization and potential drug agent optimization. A detailed description of how computer-aided applications are used in drug design, such as molecular docking, structure-based design, drug-receptor interactions, effector-based drug design, target-based drug design, molecular modeling methods, docking method, De Novo design, and fragment-based design, which are frequently used in these optimization stages. It has been investigated. In addition, we tried to explain with examples how molecular docking is used to rationalize biological activity applications and guide their optimization.

Keywords: Computer aided design, docking, drug design, molecular modeling, molecule discovery

INTRODUCTION

The molecular modeling method is a method that provides interactions with real molecules, is guided by the knowledge and intuition of the modeler, and allows the creation and evaluation of a large number of trial forms by utilizing countless possibilities to explore possible conformations and configurations (Leach 2001). An explicit application of this method allows the user to quantitatively, visually and concretely sample the binding states of a number of

molecules. The modeling method is the interpretation of the energy barriers of molecules based on molecular mechanical calculations (Vriend 1990). For this purpose, promising suitability for offline molecular modeling and drug design can be obtained by using offline modeling programs such as Charmm, Gromos, Amber, Gold (Figure 1.1.) (Blinov et al. 2004).

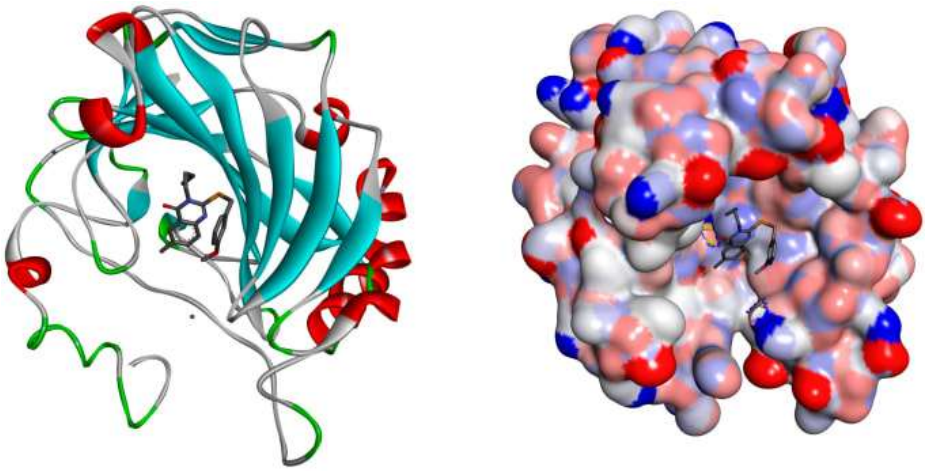


Figure 1.1. Molecular modeling simulation

Molecular dynamics simulations, on the other hand, are a powerful tool for incorporating the nature of biomolecules and are used to explore the minimum energy state at optimum conditions. Atomic parameters are general and approximate simplified quantum mechanical values based on a classical atom (Muhammed and Aki-Yalcin 2019). Since these parameters will change in a fluidic structure, quantum molecular dynamics is needed. This field has become one of the popular areas of scientific study with the great development of today's supercomputers and the spread of the advanced desktop computer. The fact that industrial companies' comprehensive databases and computational computer programs are available to academic scientists has led to an increase in the widespread work area in this field (Bharath et al. 2011).

The development of potential therapeutic new drugs is one of the difficult and complex processes in the pharmaceutical industry world. The discovery of new therapeutic agents entails significant economic costs. Rational drug design has been an important scientific topic for centuries, as a drug's activity is the result of a combination of factors such as bioavailability, toxicity, and metabolism. (Mandal and Mandal 2009; Baldi 2010). In this century, technological advances in the structural characterization of biological macromolecules, computer science, and molecular biology have been impressive and have enabled rational drug design. Molecular modeling has been observed to assist in the discovery of new drugs and guide most research in medicine and pharmacy. These tools or technologies

that enhance the design and effectiveness of drug discovery are invaluable for the potential benefits to human health and reduce the huge cost of drug discovery in terms of time and money (Güner 2008).

Computer-aided drug design (CAIT) is one of these tools that can be used to increase the efficiency of drug discovery. CIT can provide predictions when experiments are difficult, expensive, or impossible, and coordinate the available experimental data, making a valuable prediction with the experiment and forming a prediction with them. This helps CCI pharmacists or drug design scientists better understand the details of their problems and improve their approach. It also provides valuable information for drug design experiments and helps guide further experimental planning (Veselovsky and Ivanov 2003; Van Drie 2007; Huang et al., 2010). This makes the process potentially more efficient. However, CIT is not a direct route to new drugs, but provides a somewhat more detailed map and approach to target (Figure 1.2.).

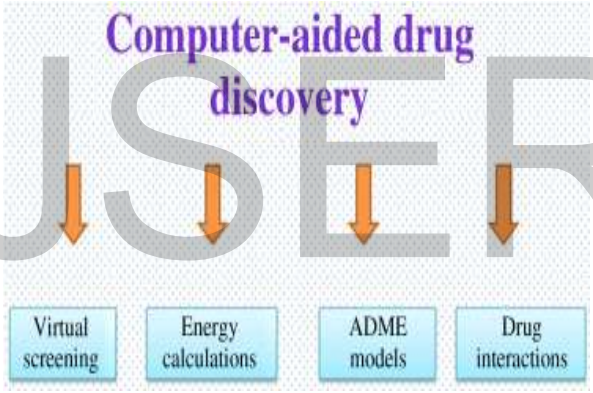


Figure 1.2. Computer-assisted drug discovery map

Ultimately, CIT will help us save days and money for drug discovery projects. Strategies for BDIT vary depending on the extent of available structural and other information regarding the target (enzyme/receptor) and ligands. With the strategies currently used, there are two main models in the drug design process as direct and indirect design. In the indirect approach, drug design is based on the comparative analysis of the structural properties of known active and inactive compounds. In direct design, it is modeling based on the direct acceptance of the target's three-dimensional properties (enzyme / receptor). At the early stage of a drug discovery process, researchers may be faced with little or no structure-activity relationship information. At this point, drug structure development and screening should be done by performing high-throughput screening with BDIT, and as a result of this detailed screening,

drug developers should follow any screening clues or other first sources of information. The compounds screened are commercially available, emerging from natural products, previously synthesized groups of compounds, or from a computational library (Ooms 2000; Amaro et al., 2008). However, those skilled in molecular modeling can assist in selecting compounds to be selected for high-throughput screening. Rather than scanning randomly, specific physicochemical properties can be selected to find a diverse set of compound analogs. The purpose of these computational drug analyzes is to select and test less selective compounds while obtaining as much information as possible about the dataset (Figure 1.3.).

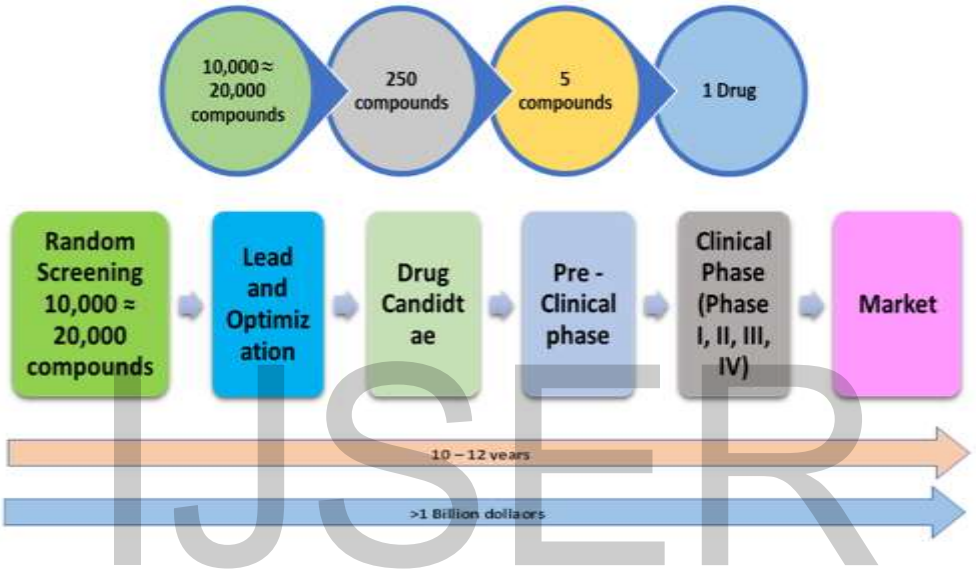


Figure 1.3. Stages of a new drug discovery

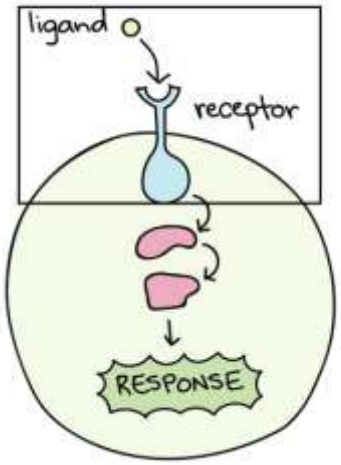
Any reduction in the number of these compounds to be investigated by reducing the amount of redundancy in the database has a significant impact on search efficiency and associated costs. Recently, the use of sensible design to maximize the structural diversity of databases has been explored to advance CIT's findings and improvements. Hierarchical clustering and maximum variability methods are approaches to drug design that compare three-dimensional databases with randomization approaches to test their effectiveness in increasing diversity (Kapetanovic 2008). Studies using two-dimensional fingerprint as a molecular identifier were validated and the performance of rational selection methods was compared with the performance of the random approach.

Accordingly, CIT approaches aim to increase the speed and efficiency of the drug discovery process. As shown in Figure 1.3, CIT is not a direct route to new drugs, but provides researchers with a slightly more detailed map of the target. It informs target drug design and helps facilitate the CCI drug design process by helping to coordinate information. Many

success stories of the use of CDI in new drug discovery and the usefulness of such testing in close association with traditional medicine and pharmaceutical techniques have been widely documented in the literature. The performance of rational selection methods is compared with the performance of the random approach.

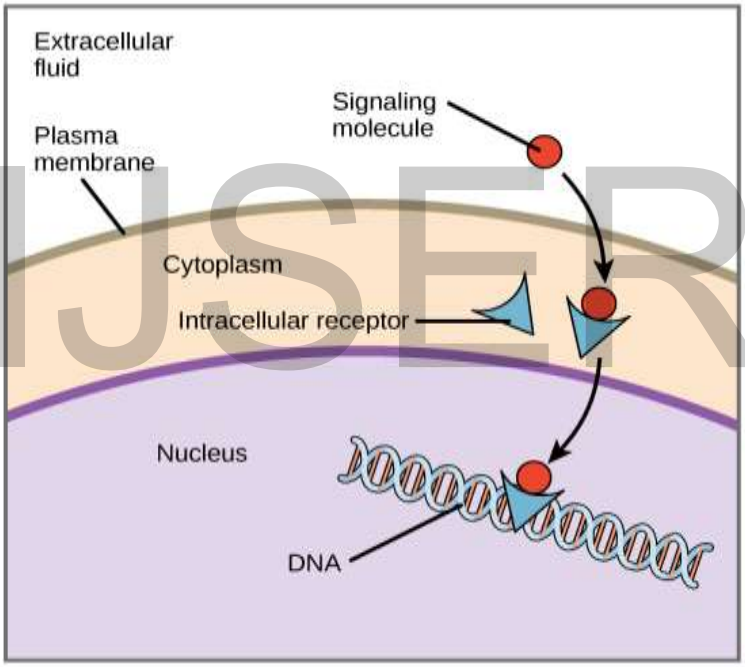
Drug-Receptor Interactıons

Just as a drug-receptor interaction begins a journey of thousands of kilometers in a single step, a complex signaling pathway in a cell begins with a single key event. The binding of a signal molecule or ligand to its receptor or receptor molecule can be called a drug-receptor interaction (Figure 2.1.). Receptors and ligands come in many forms, but they all have one thing in common: They come in tightly bound pairs, with a receptor that only recognizes a particular ligand (or more) and a ligand that binds to one (or more) target receptors. Binding of the ligand to the receptor changes its shape or activity, allowing it to signal or induce changes directly in the cell. (Bongrand 1999). Identifying receptors of key ligands can often lead to new research directions and provide valuable mechanistic information about signal transduction, drug action or off-target effects. However, despite the emergence of mass spectrometry-based techniques to identify intracellular protein-protein and small molecule-protein interactions, fair and accurate identification of ligand-receptor interactions remains a very difficult task. This is mainly due to the transient nature of this interaction and the problems found in the analysis of plasma membrane proteins that are generally hydrophobic and relatively low in abundance.



Şekil 2.1. Ligand-receptor interaction

However, in order to exhibit the characteristic binding properties of cell surface proteins, they often need to be integrated into their natural environment, i.e. living cells or tissues (Kong et al., 2006). ; Fang 2012). There are many types of receptors, but they can be divided into two categories: intracellular receptors located inside the cell (in the cytoplasm or nucleus) and cell surface receptors located in the plasma membrane. Intracellular receptors are receptor proteins found inside cells, usually in the cytoplasm or nucleus. In most cases, intracellular receptor ligands are hydrophobic small molecules because they must be able to cross the plasma membrane to reach their receptors. For example, they are corpus intracellular receptors for hydrophobic steroid hormones such as the steroid hormones estradiol (an estrogen) and testosterone (Frei et al. 2012). As shown in Figure 2.2, when a hormone enters a cell and binds to its receptor, it changes the shape of the receptor, allowing the receptor-hormone complex to enter the nucleus and regulate word activity.



Şekil 2.2. Signaling molecules and cell receptors

Hormonal binding exposes regions of the receptor that have DNA binding activity, which means they can bind to specific DNA sequences. These sequences are located next to certain genes in a cell's DNA, and when receptors bind next to these genes, they change their transcription levels. Multiple signaling pathways involving both intracellular and cellular receptors induce changes in gene transcription. However, intracellular receptors are unique in that they induce these changes very directly, bind to DNA, and alter transcription themselves. (Licitra and Liu 1996).

Cell surface receptors are lipid-binding proteins that bind to ligands on the outer surface of the cell. In this type of signaling, the ligand does not need to cross the plasma membrane. Thus, many different types of molecules can act as ligands.

A typical cell surface receptor has three distinct domains or protein domains. The ligand-binding extracellular domain is a hydrophobic domain that extends across the membrane, while the intracellular domain normally transmits signals. The size and structure of these regions can vary greatly depending on the type of receptor, and the hydrophobic region consists of several amino acid sequences that cross the membrane (Kasahara et al. 1994; Ferrante and Gorski 2012). There are many types of cell surface receptors, such as ligand-dependent ion channels, G protein-coupled receptors, and tyrosine kinase receptors.

MEDICINE DESIGN METHODS

It is a multi-purpose and challenging issue. The problem is characterized by a large and complex solution space, which is further complicated by the presence of conflicting goals. Multi-objective optimization methods designed to solve such problems have been presented to drug discovery and accepted by the scientific community. (Nicolaou and Brown 2013). Research on multi-objective optimization technology has experienced a major resurgence in the last two decades, as there is an urgent need among the many scientific communities working on problems to find solutions that exist many times over. The technology was incorporated into computerized drug discovery over a decade ago and has slowly gained acceptance since then. Studies on this topic first emerged, explaining its use in quantitative structure-activity relationship (QSAR) models, molecular linkage multi-target optimization technology, design Molecular library design, and novo design. (Talevi 2018). As a result of these initiatives, many methods and application examples have emerged, which have been developed to meet the drug discovery needs of researchers, especially in the field of drug design.

Effector-Based Drug design

Quantitative structure-activity relationships

Quantitative structure–activity relationship (QSAR) models, first proposed in the 1960s, have become widely accepted by the drug discovery community. QSAR models associate molecular descriptors with biological features using statistical techniques and/or computational intelligence algorithms. QSAR models were used to describe structure-activity

relationships in the available data for interpretation purposes. They have been presented as predictive models for the prediction of the biological property of untested chemical structures (Gramatica 2007). Multi-objective optimization methods in drug design have also been used in QSAR modeling over the past decade, taking into account many conflicting goals, including model accuracy and complexity using a multi-objective genetic programming method and Pareto sequencing. Advances in drug chemistry have provided an important basis for the search for new drug candidates with a combination of optimized pharmacodynamic and pharmacokinetic properties. Drug discovery is currently driven by innovation and knowledge using a combination of experimental and computational methods. Understanding the structure and function of the target as well as the mechanism of interaction with potential drugs is crucial to this approach. Quantitative structure-activity relationships (QSARs) are a vital part of modern drug design, as they represent a much cheaper and faster alternative to in vitro media yield and a much cheaper and faster alternative to in vivo experiments, which are often limited later in the exploratory stage. plays a role (Tropsha 2010; Andrade et al., 2010). Today, it can be said that no drug has been developed without previous QSAR analyses. As seen in Figure 3.1., it highlights the important role of QSAR in drug design by showing the flow chart of the process up to optimization of prodrug synthesis.

Figure 3.1. Possible process roadmap of QSAR in new drug discovery

The emergence of modern QSAR formalism is attributed to the 1964 work of Hansch and Fujita and Free and Wilson. The QSAR methodology is based on the concept that observed differences in the biological activity of a number of compounds can be quantitatively correlated with differences in their molecular structures. Therefore, the biological activity of similar molecular structures correlates with specific molecular features using regression

techniques to estimate the relative importance of features that contribute to biological effect. Classical QSAR methods were descriptively experimentally derived from molecular parameters (Physicochemical data) and those calculated from the molecular linkage table (2D structure). It is clear that the experimental features are a result of the entire three-dimensional structure (3D). However, they cannot be measured for unsynthesized compounds. On the other hand, 2D descriptors that can be calculated for idealized compounds, all information in the 3D structure have been calculated. Thus, in the 1980s, when the study of 3D molecular structure became a practical routine with the parallel development of several computational molecular modeling techniques, Computerized/Assisted Drug Design or Computer-Assisted/Assisted Molecular Design of the drug design process emerged, and the QSAR methodology is a broad subset of CBIT. created the field. Since then, several QSAR methodologies have been proposed. The proposed methodologies are characterized by having specific approaches for calculating and selecting molecular descriptors and specific statistical algorithms for generating the resulting models (Green and Marshall 1995; Salum and Andricopulo 2009).Similar to the direct, namely, receptor-based or structure-based and indirect, that is, ligand-based approaches currently used in the BDIT process, QSAR studies can be grouped into two main groups: receptor-independent and receptor-dependent QSAR analyses. In the first group, either the geometry of the receptor is absent or it is neglected in QSAR analysis due to uncertainty in the geometry of the receptor and/or the ligand binding mode. This group includes classical (zero-dimensional), one-dimensional (1D), two-dimensional (2D), three-dimensional (3D) and four-dimensional QSAR approaches. Calculated descriptors such as atomic and molecular counts, molecular weight, sum of atomic properties (0D-QSAR); number of parts (1D-QSAR); opological descriptors (2D-QSAR); geometric, atomic coordinates or energy grid descriptors (3D-QSAR); and recognizable molecular features such as combination of atomic coordinates and sampling of conformations (RI-4D-QSAR). In RD-QSAR analysis, models are derived from the 3D structure of multiple ligand-receptor complex conformations. This approach provides a clear simulation of the induced fitness process using the structure of the ligand-receptor complex, which is allowed to be fully flexible using molecular dynamics simulation of both the ligand and the receptor (Santos-Filho et al., 2009). RD-QSAR is used to collect descriptively binding interaction energies from the interaction between analog molecules and the receptor. Computational methods play a pivotal role in modern medicinal chemistry and offer unique potential to transform the early stages of drug research, particularly in terms of time and cost. Many of the techniques used in structure-based drug design have experienced significant advances over the

past few years. This provided a remarkable increase in the speed and effectiveness of the approach.

Pharmacophore Based Drug Screening

Research into the development of new drugs usually begins with target selection and then continues with the determination of their biological activity, optimization of their pharmacological properties, and selection of clinical candidates. Although precursor detection is the first step in an overall drug discovery program, the availability of a structurally diverse set of precursor compounds is critical for successful candidate optimization. High-throughput screening assays have served as key tools for rapid identification of new parent compounds in many laboratories. However, difficulties in establishing or accessing existing collections of compounds have hindered the effective identification of new lead compounds. Combined with high-throughput analysis technologies, virtual analysis has become an effective tool for early detection of potential hit structures. Unlike the physical library of chemical compounds required for high-throughput screening, virtual screening searches databases in silico to identify drug candidates and requires chemical databases. When the three-dimensional structure of the target is unknown, a virtual ligand-based screening can be applied as a compound selection filter to identify biologically active compounds. Ligand-based virtual scanning involves two different methods.

1 Flexible alignment of molecules taking into account only the contribution of atoms

2 Using other chemical properties unrelated to 3D pharmacopoeia representations, such as hydrogen bonding and lipophilicity, as input data for flexible alignment

In contrast, if the 3D structure of the target protein is available, both high-throughput docking and receptor-based pharmacophore virtual screening can be applied to identify new drug candidates. A pharmacophore is defined as an arrangement of molecular properties or structural elements related to biological activity.

Recently, the term has become one of the most popular symbols in drug discovery. As a useful tool for drug design, pharmacophore-based virtual screening has proven useful for hit detection and lead optimization in the initial phase of new drug development programs. The main advantage of this approach is that nearly millions of compounds can be scanned for hit detection. The disadvantage is the failure of systematic approaches. Also, important interactions may not be well represented in a given chemical property model, increasing the possibility of significant information loss in the resulting three-dimensional pharmacophore. As a result, it is possible to predict the binding free energy contributions of certain chemical properties (Wolber et al. 2008).

Pharmacophore properties are usually represented by dots in 3D space. A pharmacophore property may consist of functional groups such as hydrogen bond donor (HBD), hydrogen bond acceptor (HBA), cations, anions, aromatics, and hydrophobic domain (Hyp). Extraction of a pharmacophore property from a group of bioactive compounds does not require the 3D structure of the target protein (Guner et al., 2004; Langer and Wolber 2004). In fact, extracting the so-called common chemical properties and interaction with target proteins is a critical step for pharmacophore production. Care should be taken when considering conformational flexibility for the production of pharmacophores, in which active conformation of the molecules is assumed. Catalyst (Gunner 2005), DISCO (Walters et al., 1998), GASP (Lipinskive et al., 1997), Phase (Huuskonen et al., 2000), MOE (Zuegge et al., 2001), and LigandScout (Abagyan and Totrov 2001) There are several commercially available programs for the automatic generation of pharmacophore models, including These programs have their own algorithms for handling alignment and pharmacophore generation as well as conformational flexibility.

For pharmacophore-based virtual scanning, it is necessary to create a 3D structural ligand database, define pharmacophore properties, and search for biologically active conformations. The creation of a 3D structure is usually achieved using automated software programs to convert 2D structures into 3D formats such as SMILE arrays, SLN strings, and MDL SD link tables. The most widely used 3D compatible rendering programs are CONCORD (Diller et al., 2001) and CORINA (Schneider and Böhm 2002). These programs allow rapid identification of reliable compatibilities with output in a variety of file formats. In pharmacophore-based virtual screening, the flexibility of small molecules is governed and driven by multiple conformations for each molecule in the database. A critical confirmation of the quality of such multi-conformers is their ability to reproduce previously known bioactive conformations. Various commercial programs are available to generate multiple conformer ligands (Schneider and Böhm 2002; Krier et al., 2005). Figure 3.2 shows the ligand or receptor-based pharmacophore process virtual screening. The virtual scanning process includes several sequential computational steps such as target selection, database preparation, pharmacophore modeling, 3D scanning, and prioritization of compounds for final confirmation of biological activity.

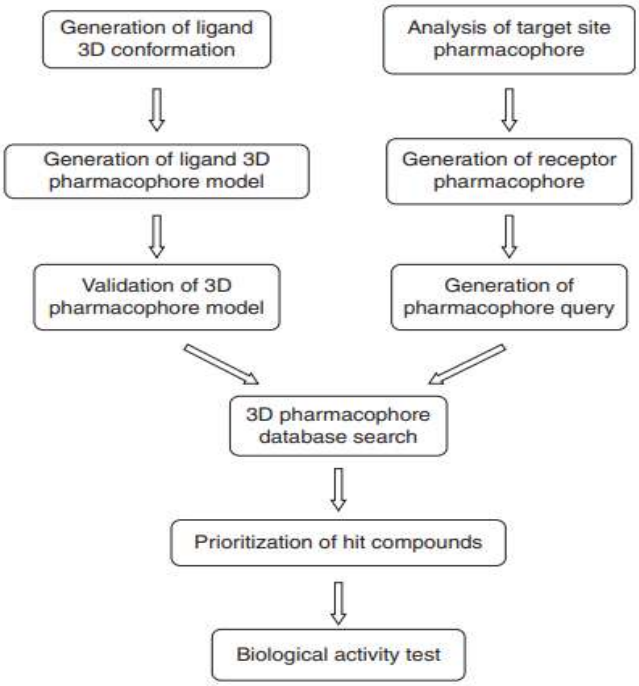


Figure 3.2. Ligand and receptor pharmacophore-based virtual screening process

An increasing influence on rational drug design over the past few years has been exerted by various software programs from major software companies (Accelrys, Tripos, Chemical Computing Group, and Schrödinger) based on the concept of chemical property-based pharmacophore models. In this context, rather than comparing substructures with each other, the pattern of binding of a ligand to its receptor is characterized by position and tolerance constraints in three-dimensional space and coding for different types of interactions. These include vectors for H bonds, aromatic Pi stacking planes, or spheres for hydrophobic or electrostatic interactions. Each interaction feature represents the region in three dimensions. This kind of generalization is highly effective for database mining. As with organic molecules, different structural motifs may express a similar chemical behavior and thus the same biological effect. Despite having no experimental information in most cases, the biological structure of the ligand or the target protein present, the ligand-based pharmacophore approach can provide essential information for drug chemists.

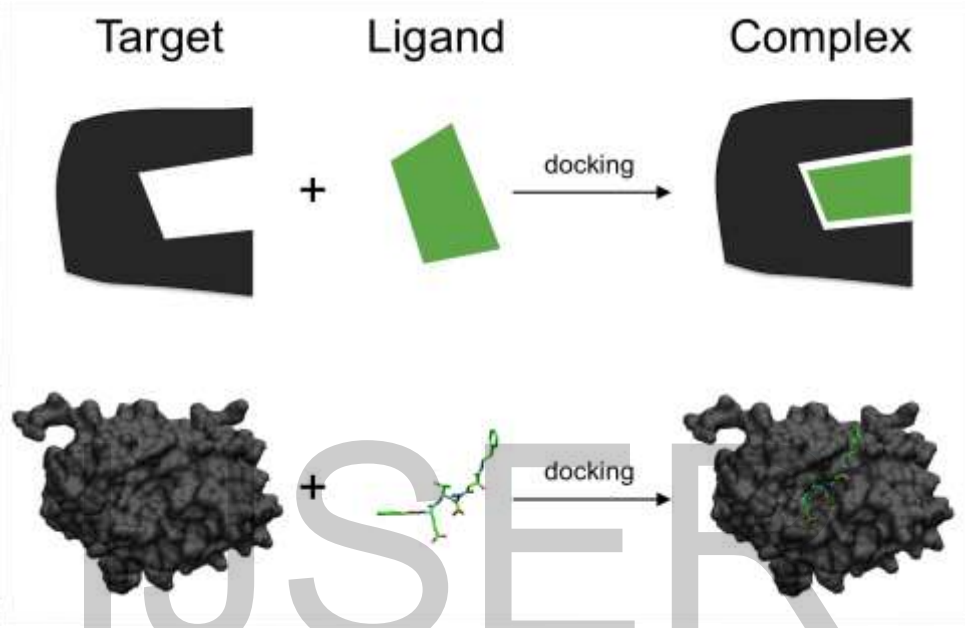
The number of articles published in the last few years is a result of researchers' growing interest in the re-emerging field of pharmacophore modeling in drug discovery, as well as the fact that approaches such as structure-based docking do not fully meet people's expectations. Recent literature studies have shown progress in clamping in terms of speed and accuracy of clamping pose estimation (Ferrara et al., 2004; Kitchen et al., 2004; Perola et al., 2004). However, the biggest problem is still the accurate estimation of the free binding energy. While the docking scoring functions used for this task may work well in the specific application situations for which they are set, they are more likely to fail in other target groups. The simple three-dimensional pharmacophore concept has gained renewed interest, as placement and scoring are computationally expensive and sequencing of hits is still not possible with satisfactory accuracy. The concept of pharmacophore is always used keeping in mind the need to understand, explain and predict molecular interactions and structure-activity relationships with targets. Its practical applicability for pharmaceutical chemists has made it an excellent communication tool between concept modelers and synthetic chemists. Pharmacophores have absolute simplicity and utility for searching structural databases. Pharmacophore-based Drug screening will actually allow rapid bioactivity profiling of compounds before they are synthesized and will also significantly improve the library design process. However, it has been noted in the literature that there is still much need for improvement in this area, in research on improving methods and even designing new algorithms. Demand for experts in the field who interface between pharmaceutical chemistry and computer science will increase over the next decade, both in the pharmaceutical industry and in software companies specializing in computer-aided molecular design. There is no doubt that we will experience significant progress in pharmacophore-based screening technologies in the near future.

Targeted Drug Design Molecular Modeling

a) Docking Method

Docking aims to accurately predict the structure of a ligand within the constraints of a receptor binding site and accurately predict the strength of binding. In modern drug discovery, protein-ligand or protein-protein coupling plays an important role in predicting using electrostatic interactions to measure the orientation of the ligand when it binds to a protein receptor or enzyme (Figure 3.3.). Van der Waals interactions also play an important role in addition to Coulomb interactions and the formation of hydrogen bonds. The sum of all these interactions is approximated by a docking score representing the binding potential (Shoichet

and Kuntz 1991). In the simplest solid body systems, the ligand is sought in a six-dimensional rotational or translational domain to fit into the binding site, which can serve as a precursor compound for drug design. In the solid-body approximation, the coupling accuracy is much greater for bonded complexes than for uncomplex molecules. While the structural changes observed between bounded and freeforms are small, the difference in accuracy indicates that the rigidity assumption is not fully warranted.



Şekil 3.3. Molecular modeling simulation

Also, the difference between near native structures and others far from native cannot be distinguished even by simple scoring functions e.g. surface complementarity, solvent accessible surface area (SASA) embedding, unsolvable energy, electrostatic interaction energy or total molecular mechanical energy etc. (Mezei 2003). Therefore, docking procedures were developed by several groups, allowing for receptor and ligand flexibility. The earliest reported docking methods were based on the lock-and-key assumption proposed by Fischer, who stated that both the ligand and the receptor can be treated as rigid bodies, and their affinity is directly proportional to a geometric fit between their shapes. Later, the Binduced-fit theory proposed by Koshland suggested that ligand and receptor should be handled flexibly during insertion (Koshland 1963). Each backbone action affects multiple side chains as opposed to relatively independent side chains. Therefore, the sampling procedure in a fully flexible receptor/ligand insertion has a higher order of magnitude in terms of the number of degrees of freedom than in flexible insertion with a rigid receptor. As a result,

these flexible docking algorithms not only predict the binding mode of a molecule more accurately than rigid body algorithms, but also predict its binding affinity relative to other compounds. Over the last two decades, more than 60 different docking tools and programs have been developed for both academic and commercial purposes. For example, some docking programs are; It can be said AutoDock, FlexX, Surflex, GOLD, ICM, Glide, Cdocker, LigandFit, FRED, MOE-Dock, LeDock, AutoDock Vina, rDock, UCSF Dock. While the strategies for ligand insertion differ, these programs are broadly classified into a wide range of incremental cluster structure approaches such as shape-based algorithms, genetic algorithms, systematic search techniques, and Monte Carlo simulations. With the exception of the GOLD program, almost all current flexible ligand insertion programs treat the receptor as rigid. These programs were evaluated to test their ability to generate the correct mode of binding of a ligand to its biological target and to identify known compounds with the highest scores in virtual screening experiments. Among these programs, AutoDock Vina, GOLD, and MOE-Dock predicted the top-ranked poses with the best scores. GOLD and LeDock were able to identify the correct ligand binding poses. In addition, both Glide (XP) and GOLD consistently predict poses with 90.0% accuracy. GOLD has been shown to produce higher enrichment factors than Glide in a virtual screening trial against Factor Xa, whereas Glide outperformed GOLD against the same target in a similar virtual screening trial (Wang et al., 2016) . Overall, it was recently reported that these docking programs are able to predict experimental poses with root mean square deviations (RMSDs) ranging from 1.5 to 2 Å on average. However, flexible receptor insertion, particularly spine flexibility in receptors, still poses a major challenge for current insertion methods. In conclusion, structure-based drug design is a powerful technique for rapid identification of small molecules against the 3D structure of macromolecular targets, which can be obtained by X-ray, NMR or homology models. Because of the abundant information regarding the sequences and structures of proteins, the structural information of individual proteins and their interactions has become crucial for further drug therapy. Although many insertion programs exist for conformational search and binding pose estimation, their scoring functions are not accurate and need further development. However, despite the disadvantages of each docking strategy such as scoring, open protein flexibility, open water, active research is widely conducted to address all relevant issues.

b) De Novo Design

De novo design is an attractive approach to create engineered proteins with predetermined structures and functions. As seen in Figure 3.4., De novo design methods propose new

scaffolds that have been tested for their ability to form structures similar to known inhibitors and are consequently synthesized and tested for activity (Schneider and Fechner 2005). De novo design requires a deep understanding of the structure and shape of a protein and its functions. It is performed from scratch or with a modular approach.

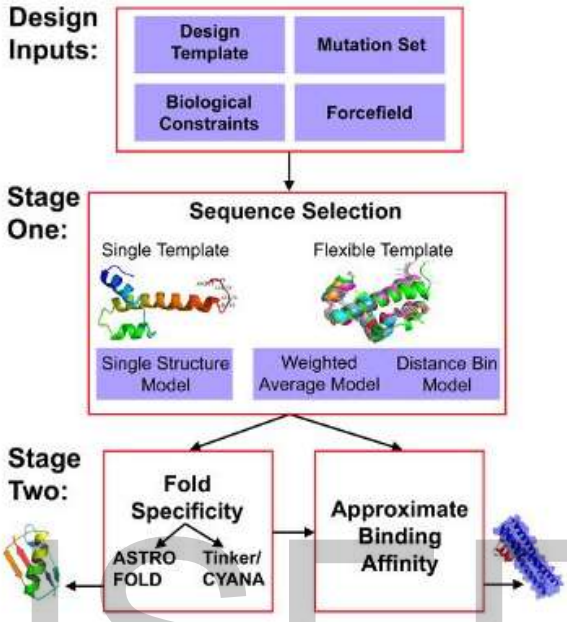


Figure 3.4. The overall workflow of the De Novo Design strategy

Linear amino acid chains have a unique 3D structure in the natural environment to realize proteins. In de novo design, the ultimate goal is to identify amino acid sequences that fold into proteins with desired functions. There are several approaches to predict protein structure, including comparative modeling and fold recognition (Hay 2010). Sequence similarity in proteins refers to structural similarity. Accordingly, the comparative modeling method was created. The structure of a protein can be predicted by comparing its amino acid sequence with that of the known natural 3D structure. Predictions are of high quality when the target and template share more than 50% of the array. Sequence similarity refers to structural similarity, but similar structures can be found for proteins with different sequences. As a complementary method, fold recognition aims to predict the three-dimensional folded structure of a protein with known sequence. The structure is more evolutionarily conserved than the array.

As a result, the repertoire of different folds is more limited. Fold recognition methods mainly include advanced sequence comparison and secondary structure prediction and comparison. Also included is the estimation of the loop structure, which includes helices and helixes.

Consistent with native 3D targets with minimal energy interaction, de novo-designed proteins can often fold very quickly. De novo protein design can increase the stability of the target protein and has also been used to lock proteins in certain useful ways (Pegg et al., 2001). Numerous successes have been achieved in the development of computational algorithms for protein design.

Common approaches used by de novo design include:

✓ Structural motif: such as "four α-helical bundles" motif, "helical loop-helix" motif

✓ Protein structures known as natural scaffolds

✓ Molecular templates

The effectiveness of the design approach is demonstrated in several cases where computer-defined host molecules are subsequently synthesized and function as efficient anion hosts. Therefore, de novo design can be considered as a complement to other virtual techniques such as database searching and non-virtual techniques such as high-throughput scanning (Mauser and Guba 2008). In addition, successful de novo design examples in the hit and precursor finding stages of the drug discovery process are used to demonstrate that de novo design provides a method for the identification of lead compounds.

c) Trailer Based Design

The search for new drugs is grappling with high attrition rates at all stages of research and development. Chemists have the opportunity to address this problem, because weathering can be traced in part to the quality of chemical cues. Fragment-based drug discovery (FBDD) is a new approach increasingly used in the pharmaceutical industry to reduce attrition and provide leads for previously recalcitrant biological targets (Erlanson et al., 2004). FBDD has evolved significantly over the last 10 years and is now considered a tangible alternative to more traditional hit identification methods such as high-throughput scanning. FBDD identifies low molecular weight ligands (~150 Da) that bind to biologically important macromolecules. The three-dimensional experimental bonding mode of these parts is determined using X-ray crystallography or NMR spectroscopy. It is used to facilitate their optimization into potent molecules with drug-like properties. Compared to high-throughput screening, the fragment approach requires fewer compounds to be screened and offers more efficient and destructive optimization situations, despite the lower initial power of screening hits (Figure 3.5.). The number of commercial and academic groups actively involved in FBDD-based research has

increased, and as a result, the development and refinement of techniques and method continues(Congreveet al., 2008).

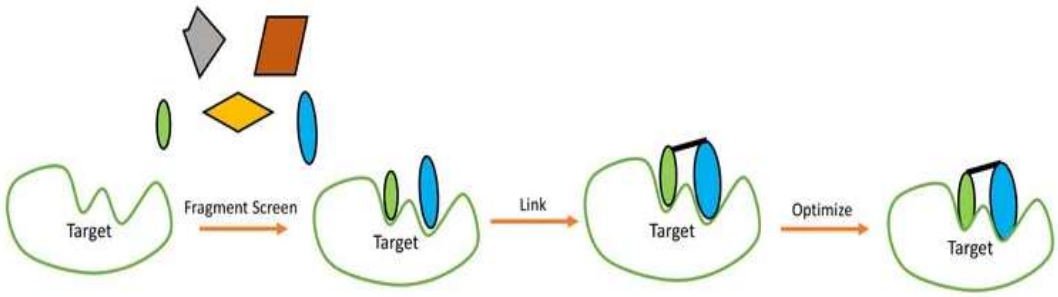


Figure 3.5. Fragment-based drug design

From its inception, the FBDD approach had two key principles that were critical to its success, distinguishing it from high-throughput screening and other hit identification techniques. The first is the concept that the chemical space can be explored more efficiently by scanning collections of small fragments rather than larger libraries of molecules. The potential number of fragments containing up to 12 heavy atoms (not including three- and four-membered ring structures) is estimated at 107, while the number of potential drug-like molecules containing up to 30 heavy atoms is estimated to exceed 1060 (Bembenek et al., 2009) . Therefore, a much larger proportion of the "fragment-like" chemical area can be viably screened in FBDD compared to the "drug-like" chemical area covered in a high virtual scan where the molecular size is much larger.

The second key idea is that, since fragment molecules by definition are small in size (typically less than 250 Da), they typically need to bind with lower affinity (nanomolar to micromolar range) to their target proteins (micromolar to millimolar range) than drug-like molecules that are more capable of interacting. . However, the per-atom binding efficiency is at least as high as for larger hit molecules (Zartler and Shapiro 2005). Indirectly, screening techniques used in FBDD must be much more sensitive than bioassay in a high-throughput screening.

In general, sensitive biophysical techniques are used to detect these weak binding events and to characterize fragment interactions with the target active site. Nuclear magnetic resonance (NMR) and protein X-ray crystallography have been widely used in fragment-based research because these techniques are highly sensitive in detecting low-affinity fragment binding and also provide information about the resulting fragment-protein interactions. Track-based scanning has an intuitive appeal. The success of pharmaceutical companies such as Abbott and biotechnology companies such as Astex Therapeutics, SGX Pharmaceuticals, and

Plexxikon in developing fragments for clinical candidates has influenced the chemical community and led to fragment screening efforts at many other industrial and academic institutions (Coyle and Walser 2020). A great deal of effort has been made in the industry to establish fragment-based screening over the last 3-4 years and is now being implemented as a complementary strategy to high-throughput screening. This is in part due to the fact that investments in high-throughput screening and combinatorial chemistry in the 1990s did not yield success for the more demanding classes of drug targets. However, despite obvious efforts to implement part scanning, there are significant cultural and practical challenges that must be overcome in large companies to effectively implement this new methodology. In particular, after the identification of fragment hits, optimization to a more conventional range of effects will often be difficult without structural knowledge. Significant upfront investment in structural biology is required both to establish the binding modes of fragments within the active site of target proteins and to eliminate false positives. This commitment to timely structural biology can be difficult to achieve in practice in large organizations, particularly where only a fraction of the targets are readily amenable to 3D structure determination. Another problem is that fragment hits with low or undetectable potency in a biological assay may initially seem less attractive to medicinal chemists, compared to traditional high-throughput screening hits with higher potency. On the contrary, in an academic setting that assembles a small library of fragments and screens using a biophysical technique such as surface plasmon resonance (SPR), protein-ligand NMR or even X-ray crystallography, compared to combining and screening a large library in a bioassay. available (Li 2020).

As a result, a number of compounds evolved from fragments have entered the clinic, and this approach is increasingly accepted as an additional way to identify new hit compounds in pharmaceutical discovery and inhibitor design. FBDD is a powerful method to develop strong small molecule compounds, starting with fragments that weakly bind to targets. Therefore, FBDD will play more important roles in drug discovery as it can be easily carried out with different biophysical and computer-based methods. FBDD becomes an attractive strategy in targeted drug discovery as it exhibits several advantages over high-throughput screening campaigns. Many potent compound inhibitors of various targets have been developed using this approach. The methods used in fragment scanning and understanding fragment binding modes are critical in FBDD. Before application, it is necessary to know the fragment libraries in fragment identification, validation and the methods used, the strategies applied to amplify the identified fragments into drug-like precursor compounds, and the explanations of the applications of FBDD to different targets. FBDD will play more important roles in drug

discovery as it can be easily carried out with different biophysical and computer-based methods. In addition, in the development of parts, attractive properties can be integrated that can translate into compounds with favorable physical, pharmacokinetic and toxicity (absorption, distribution, metabolism, excretion and toxicity) properties. Various in silico approaches are used to support FBDD strategies and the optimization of both from the follower to the lead. These fragment expansion strategies include hotspot analysis, drug availability estimation, SAR (structure-activity relationships) with catalog methods, application of machine learning deep learning models for virtual screening, and various de novo design methods to propose new synthesizable compounds.

CONCLUSION AND RECOMMENDATIONS

Molecular modeling has become a valuable and fundamental tool for scientific studies in pharmacy and medicine in the drug design process. The molecular modeling samples shown in the presented thesis show that computer aided drug design computation methods are especially useful for the development of biological activity applications.

The drug discovery and development process is long and expensive. Drug discovery starts with target identification, then confirms targets and identifies drug candidates before any newly discovered drug is released. Reducing the production cost level is the main objective. In this context, computer-aided drug design is indeed a very useful tool for pharmaceutical companies and academic research groups to search for potential drug candidates with low cost and time. However, there is still room for further improvement in computer-aided drug design, including more accurate docking score values and docking functions, target flexibility in the docking procedure, and faster evaluation of solvent effects, and improving computational efficiency. Continuous improvements in chemical and structural biology, bioinformatics, and computational technology are required to improve current computer-aided drug design. It has been observed that it is a scientific field that needs to be studied and researched in order to develop the computer aided design methods presented in this thesis and to continue to develop biological applications from different perspectives.

Suggestions to be made as a result of this thesis:

- A course called drug design and molecular discovery at the Faculty of Pharmacy should be added to the curriculum and taught to students.
- Computer-aided drug design programs, namely software and advanced computers, should be owned by the Faculty of Pharmacy.
- Pharmaceutical design laboratories should be taught.

- Scientific organizations should be made on drug design methods and the participation of faculty students should be ensured.
- Academicians trained in this field in Turkey should be invited to the Faculties of Pharmacy to benefit from their knowledge.
- People working on this issue should be supported economically.

Acknowledgement

This study was prepared from Kübra Akman's Research Project Thesis.

REFERENCES

Abagyan, R., Totrov, M. 2001. High-throughput docking for lead generation. Current opinion in chemical biology, 5(4), 375-382.

Amaro, R. E., Baron, R., McCammon, J. A. 2008. An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. Journal of computer-aided molecular design, 22(9), 693-705.

Andrade, C. H., Pasqualoto, K. F., Ferreira, E. I., Hopfinger, A. J. 2010. 4D-QSAR: perspectives in drug design. Molecules, 15(5), 3281-3294.

Baldi, A. 2010. Computational approaches for drug design and discovery: An overview. Systematic reviews in Pharmacy, 1(1), 99.

Bembenek, S. D., Tounge, B. A., Reynolds, C. H. 2009. Ligand efficiency and fragment-based drug discovery. Drug discovery today, 14(5-6), 278-283.

Bharath, E. N., Manjula, S. N., Vijaychand, A. 2011. In silico drug design tool for overcoming the innovation deficit in the drug discovery process. Chemistry, 18(23.3), 1-0.

Coyle, J., Walser, R. 2020. Applied Biophysical Methods in Fragment-Based Drug Discovery. Slas Discovery: Advancing the Science of Drug Discovery, 25(5), 471-490.

Diller, D. J., Merz Jr, K. M. 2001. High throughput docking for library design and library prioritization. Proteins: Structure, Function, and Bioinformatics, 43(2), 113-124.

Güner, O. F. 2000. Pharmacophore perception, development, and use in drug design (Vol. 2). Internat'l University Line.

Blinov, M. L., Faeder, J. R., Goldstein, B., Hlavacek, W. S. 2004. BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. Bioinformatics, 20(17), 3289-3291.

Bongrand, P. 1999. Ligand-receptor interactions. Reports on progress in Physics, 62(6), 921.

Congreve, M., Chessari, G., Tisi, D., Woodhead, A. J. 2008. Recent developments in fragment-based drug discovery. Journal of medicinal chemistry, 51(13), 3661-3680.

Erlanson, D. A., McDowell, R. S., O'Brien, T. 2004. Fragment-based drug discovery. Journal of medicinal chemistry, 47(14), 3463-3482.

Fang, Y. 2012. Ligand–receptor interaction platforms and their applications for drug discovery. Expert opinion on drug discovery, 7(10), 969-988.

Ferrante, A., Gorski, J. 2012. Enthalpy-entropy compensation and cooperativity as thermodynamic epiphenomena of structural flexibility in ligand–receptor interactions. Journal of molecular biology, 417(5), 454-467.

Frei, A. P., Jeon, O. Y., Kilcher, S., Moest, H., Henning, L. M., Jost, C., Wollscheid, B. 2012. Direct identification of ligand-receptor interactions on living cells and tissues. Nature biotechnology, 30(10), 997.

Gramatica, P. 2007. Principles of QSAR models validation: internal and external. QSAR combinatorial science, 26(5), 694-701.

Green, S. M., Marshall, G. R. 1995. 3D-QSAR: a current perspective. Trends in pharmacological sciences, 16(9), 285-291.

Guner, O., Clement, O., Kurogi, Y. 2004. Pharmacophore modeling and three dimensional d1atabase searching for drug design using catalyst: recent advances. Current medicinal chemistry, 11(22), 2991-3005.

Guner, O. F. 2005. The impact of pharmacophore modeling in drug design. IDrugs, 8(7), 567.

Ferrara, P., Gohlke, H., Price, D. J., Klebe, G., Brooks, C. L. 2004. Assessing scoring functions for protein-ligand interactions. Journal of medicinal chemistry, 47(12), 3032-3047.

Hay, B. P. 2010. De novo structure-based design of anion receptors. Chemical Society Reviews, 39(10), 3700-3708.

Huang, H. J., Yu, H. W., Chen, C. Y., Hsu, C. H., Chen, H. Y., Lee, K. J.,Chen, C. Y. C. 2010. Current developments of computer-aided drug design. Journal of the Taiwan Institute of Chemical Engineers, 41(6), 623-635.

Huuskonen, J., Rantanen, J., Livingstone, D. 2000. Prediction of aqueous solubility for a diverse set of organic compounds based on atom-type electrotopological state indices. European journal of medicinal chemistry, 35(12), 1081-1088.

Kapetanovic, I. M. 2008. Computer-aided drug discovery and development (CADDD): in silico-chemico-biological approach. Chemico-biological interactions, 171(2), 165-176.

Karasar, N. 2005. Bilimsel araştırma yöntemi,Ankara: Nobel Yayın Dağıtım.

Kasahara, N., Dozy, A. M., Kan, Y. W. 1994. Tissue-specific targeting of retroviral vectors through ligand-receptor interactions. Science, 266(5189), 1373-1376.

Kitchen, D. B., Decornez, H., Furr, J. R., Bajorath, J. 2004. Docking and scoring in virtual screening for drug discovery: methods and applications. Nature reviews Drug discovery, 3(11), 935-949.

Krier, M., de Araújo-Júnior, J. X., Schmitt, M., Duranton, J., Justiano-Basaran, H., Lugnier, C., Rognan, D. 2005. Design of small-sized libraries by combinatorial assembly of linkers and functional groups to a given scaffold: application to the structure-based optimization of a phosphodiesterase 4 inhibitor. Journal of medicinal chemistry, 48(11), 3816-3822.

Kong, H. J., Boontheekul, T., Mooney, D. J. 2006. Quantifying the relation between adhesion ligand–receptor bond formation and cell phenotype. Proceedings of the National Academy of Sciences, 103(49), 18534-18539.

Koshland, D. E. 1963. Correlation of structure and function in enzyme action. Science, 142(3599), 1533-1541.

Langer, T., Wolber, G. 2004. Pharmacophore definition and 3D searches. Drug Discovery Today: Technologies, 1(3), 203-207.

Langer, T., Hoffmann, R. D. 2006. Pharmacophore modelling: applications in drug discovery. Expert opinion on drug discovery, 1(3), 261-267.

Leach, A. R., Leach, A. R. 2001. Molecular modelling: principles and applications. Pearson education.

Li, Q. 2020. Application of fragment-based drug discovery to versatile targets. Frontiers in Molecular Biosciences, 7.

Licitra, E. J., Liu, J. O. 1996. A three-hybrid system for detecting small ligand–protein receptor interactions. Proceedings of the National Academy of Sciences, 93(23), 12817-12821.

Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced drug delivery reviews, 23(1-3), 3-25.

Mandal, S., Mandal, S. K. 2009. Rational drug design. European journal of pharmacology, 625(1-3), 90-100.

Mauser, H., Guba, W. 2008. Recent developments in de novo design and scaffold hopping. Current opinion in drug discovery development, 11(3), 365-374.

Mezei, M. 2003. A new method for mapping macromolecular topography. Journal of Molecular Graphics and Modelling, 21(5), 463-472.

Muhammed, M. T., Aki- Yalcin, E. 2019. Homology modeling in drug discovery: Overview, current applications, and future perspectives. Chemical biology drug design, 93(1), 12-20.

Nicolaou, C. A., Brown, N. 2013. Multi-objective optimization methods in drug design. Drug Discovery Today: Technologies, 10(3), e427-e435.

Ooms, F. 2000. Molecular modeling and computer aided drug design. Examples of their applications in medicinal chemistry. Current medicinal chemistry, 7(2), 141-158.

Pegg, S. C. H., Haresco, J. J., Kuntz, I. D. 2001. A genetic algorithm for structure-based de novo design. Journal of computer-aided molecular design, 15(10), 911-933.

Perola, E., Walters, W. P., Charifson, P. S. 2004. A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. Proteins: Structure, Function, and Bioinformatics, 56(2), 235-249.

Santos-Filho, O. A., Hopfinger, A. J., Cherkasov, A., de Alencastro, R. B. 2009. The receptor-dependent QSAR paradigm: an overview of the current state of the art. Medicinal Chemistry, 5(4), 359-366.

Salum, L. B., Andricopulo, A. D. 2009. Fragment-based QSAR: perspectives in drug design. Molecular diversity, 13(3), 277-285.

Schneider, G., Böhm, H. J. 2002. Virtual screening and fast automated docking methods. Drug discovery today, 7, 64-70.

Schneider, G., Fechner, U. 2005. Computer-based de novo design of drug-like molecules. Nature Reviews Drug Discovery, 4(8), 649-663.

Shoichet, B. K., Kuntz, I. D. 1991. Protein docking and complementarity. Journal of molecular biology, 221(1), 327-346.

Talevi, A. 2018. Computer-aided drug design: An overview. Computational drug discovery and design, 1-19.

Tropsha, A. 2010. Best practices for QSAR model development, validation, and exploitation. Molecular informatics, 29(6- 7), 476-488.

Van Drie, J. H. 2007. Computer-aided drug design: the next 20 years. Journal of computer-aided molecular design, 21(10), 591-601.

Veselovsky, A. V., Ivanov, A. S. 2003. Strategy of computer-aided drug design. Current Drug Targets-Infectious Disorders, 3(1), 33-40.

Vriend, G. 1990. WHAT IF: a molecular modeling and drug design program. Journal of molecular graphics, 8(1), 52-56.

Wang, Z., Sun, H., Yao, X., Li, D., Xu, L., Li, Y., Hou, T. 2016. Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: the prediction accuracy of sampling power and scoring power. Physical Chemistry Chemical Physics, 18(18), 12964-12975.

Walters, W. P., Stahl, M. T., Murcko, M. A. 1998. Virtual screening-an overview. Drug discovery today, 3(4), 160-178.

Wolber, G., Seidel, T., Bendix, F., Langer, T. 2008. Molecule-pharmacophore superpositioning and pattern matching in computational drug design. Drug discovery today, 13(1-2), 23-29.

Zartler, E. R., Shapiro, M. J. 2005. Fragonomics: fragment-based drug discovery. Current opinion in chemical biology, 9(4), 366-370.

Zuegge, J., Schneider, G., Coassolo, P., Lavé, T. 2001. Prediction of hepatic metabolic clearance. Clinical pharmacokinetics, 40(7), 553-563.