International Journal of Advanced Research in Biological Sciences
ISSN: 2348-8069
www.ijarbs.com
DOI: 10.22192/ijarbs Coden: IJARQG(USA) Volume 4, Issue 2 - 2017

**Review Article**


**DOI:** http://dx.doi.org/10.22192/ijarbs.2017.04.02.009

# An Approach of Computer-Aided Drug Design (CADD) Tools for *In Silico* Pharmaceutical Drug Design and Development

**Injamul Hoque*, Arkendu Chatterjee, Somenath Bhattacharya, Raj Biswas**
Department of Pharmacy, Bengal College of Pharmaceutical Sciences and Research, Bidhannagar, Durgapur-713212, West Bengal, India
*Corresponding author

## Abstract

Computer-aided drug design (CADD) depends on the extent of structure and other information available regarding the target (enzyme/receptor/protein) and the ligands. The theoretical basis of CADD involves molecular mechanics, quantum mechanics, molecular dynamics, structure-based drug design (SBDD), ligand-based drug design (LBDD), homology modeling, ligplot analysis, molecular docking, de novo drug design, pharmacophore modeling and mapping, virtual screening (VS), quantitative structure-activity relationships (QSARs), *In silico* ADMET (absorption, distribution, metabolism, excretion and toxicity) prediction etc. CADD centre was created to foster collaborative research between biologist, biophysicists, structural biologists and computational scientists. The major goal of the CADD centre is to initiate these collaborations leading to the establishment of research projects to discover novel chemical entities with the potential to be developed into novel therapeutic agents.

**Keywords:** Bioinformatics, Softwares, Homology modeling, Ligplot analysis, Molecular docking, De novo drug design, Pharmacophore modeling, Virtual screening (VS), Quantitative structure-activity relationships (QSARs), Lipinski's rule.

## Introduction

Advances in the field of biochemistry, molecular biology and cell biology, facilitated by developments in genomics and proteomics, are producing a large number of novel biological targets that may be exploited for therapeutic intervention. To facilitate the discovery of novel therapeutic agents, rational drug design methods in combination with structural biology offer great potential. The latest technological advances are (QSAR/QSPR, structure-based design and bioinformatics). Drug discovery and developing a new medicine is a long, complex, costly and highly risky process that has few peers in the commercial world. This is why computer-aided drug design (CADD) approaches are being widely used in the pharmaceutical industry to accelerate the process. The cost benefit of using computational tools in the lead optimization phase of drug development is substantial. On an average, it takes 10-15 years and US $500-800 million to introduce a drug into the market, with synthesis and testing of lead analogues being a large contributor to that sum. Therefore, it is beneficial to apply computational tools in hit-to-lead optimization to cover a wider chemical space while reducing the number of compounds that must be synthesized and tested *in vitro*.

Computational methods of drug design are based on a postulate that pharmacologically active compounds act by interaction with their macromolecular targets, mainly proteins or nucleic acids. Major factors of such interactions are surfaces of molecules, electrostatic force, hydrophobic interaction and hydrogen bonds formation. These factors are mainly considered during analysis and prediction of interaction of two molecules [(1)].

## COMPUTER-AIDED DRUG DESIGN (CADD):

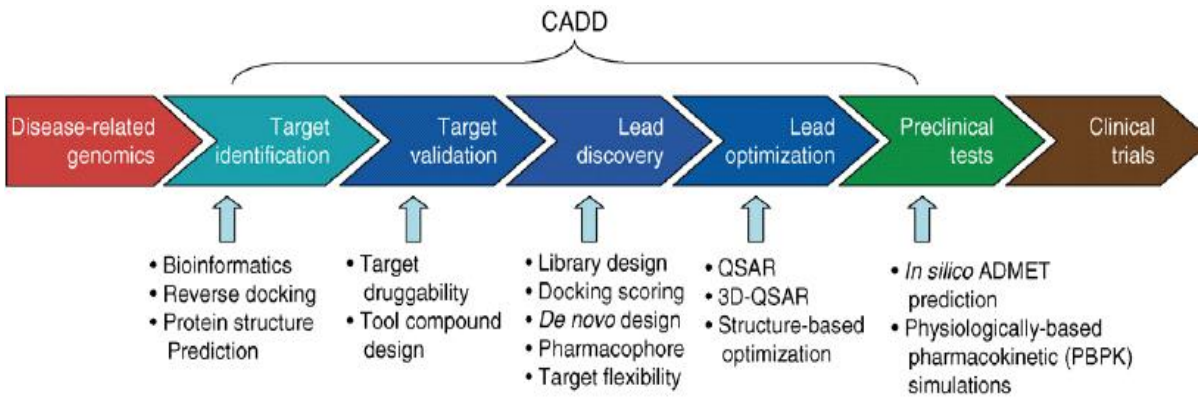


**Figure 1:** *In silico* Computer-aided drug design [2,3]

Computer-aided drug design is a computer technology that designs a product and documents the design's process. CADD may facilitate the manufacturing process by transferring detailed diagrams of a products materials, processes, tolerances and dimensions with specific conventions for the product in question [2]. It can be used to produce either two-dimensional or three-dimensional diagrams, which can then when rotated to be viewed from any angle, even from the inside looking out. The channel of drug discovery from idea to market consists of seven basic steps: disease selection, target selection, lead compound identification, lead optimization, pre-clinical trial testing, clinical trial testing and pharmacogenomic optimization. In practice, the last five steps required to pass repeatedly. The compounds for testing can be obtained from natural source (Plants, animals, microorganisms) and by chemical synthesis. These compounds can be rejected as perspectives owing to absence or low activity, existence of toxicity or carcinogenicity, complexity of synthesis, insufficient efficiency etc. As a result only one of 100000 investigated compounds may be introduced to the market and one average cost of development of new drug rose up to 800 million dollars. The reduction of time-consuming and cost of the last stages of drug testing is unlikely due to strict state standard on their realization. Therefore main efforts to increasing efficiency of development of drugs are directed to stages of discovery and optimization of ligands [3].

*In silico* methods can help in identifying drug targets via bioinformatics tools. They can also be used to analyze the target structure for possible binding/active sites, generate candidate molecules, check for their likeness, dock these molecules with the target, rank them according to their binding affinities, further optimize the molecule to improve binding characteristic. The uses of computers and computational methods permeate all aspects of drug discover today an forms the core of (a) structure-based drug design and (b) ligand-based drug design [4].

**(a) STRUCTURE-BASED DRUG DESIGN (SBDD):** Structure-based drug design is the technique to be used in drug design. Structure-based drug design helps in the discovery process of new drugs [5].

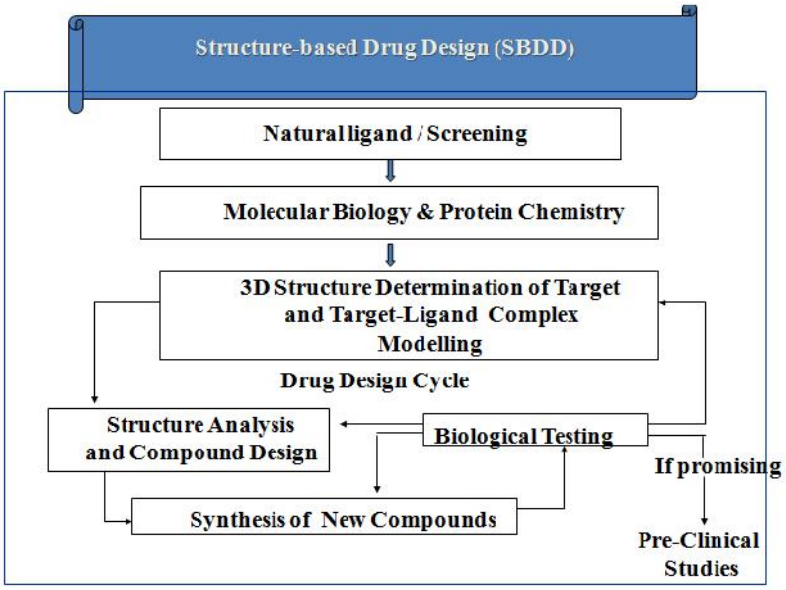


**Figure 2:** Structure-based drug design [5]

## (b) LIGAND- BASED DRUG DESIGN (LBDD):

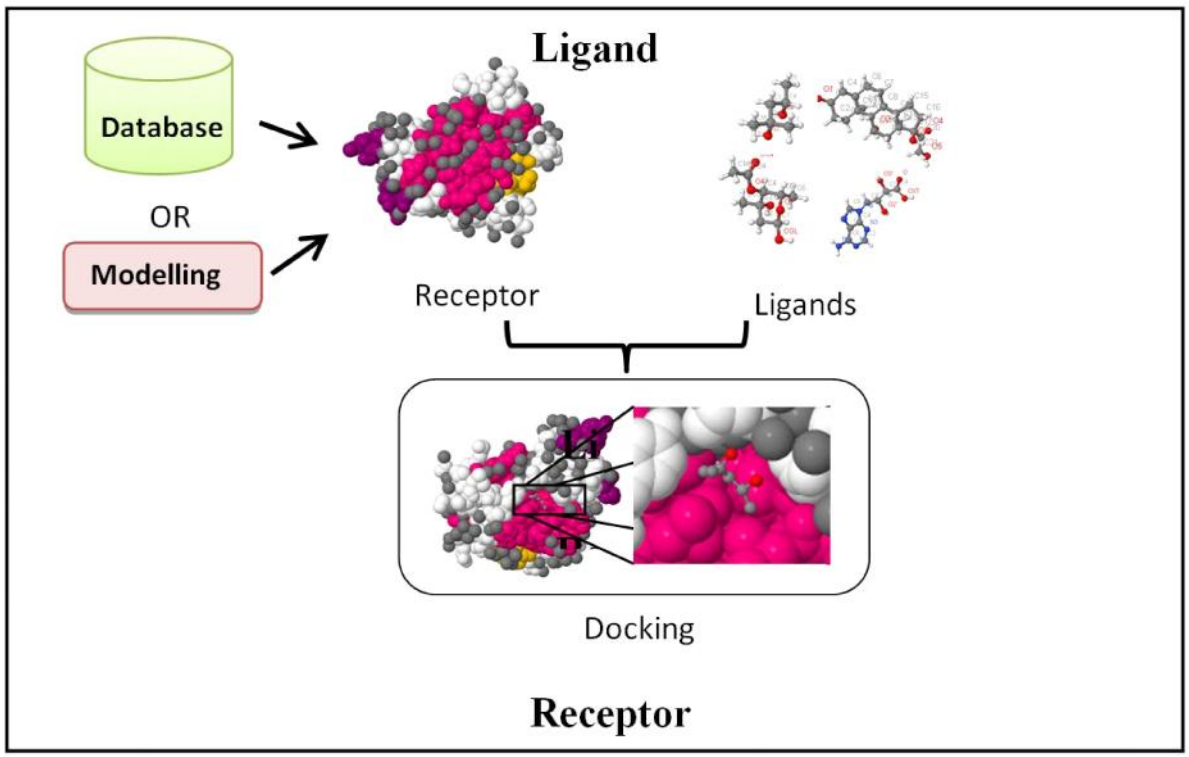


**Figure 3:** Ligand-based drug design [6]

The ligand-based drug design approach involves the analysis of ligands known to interact with a target. These methods use a set of reference structure collected from compounds known to interact with the target of interest and analysis their 2D or 3D structure [7]. In some cases, usually in which data pertaining to the 3D structure of a target protein are not available, drug design can instead be based on process using the known ligands of a target protein as the starting point. This approach is known as "ligand-based drug design" [8].

## BIOINFORMATICS IN COMPUTER-AIDED DRUG DESIGN:

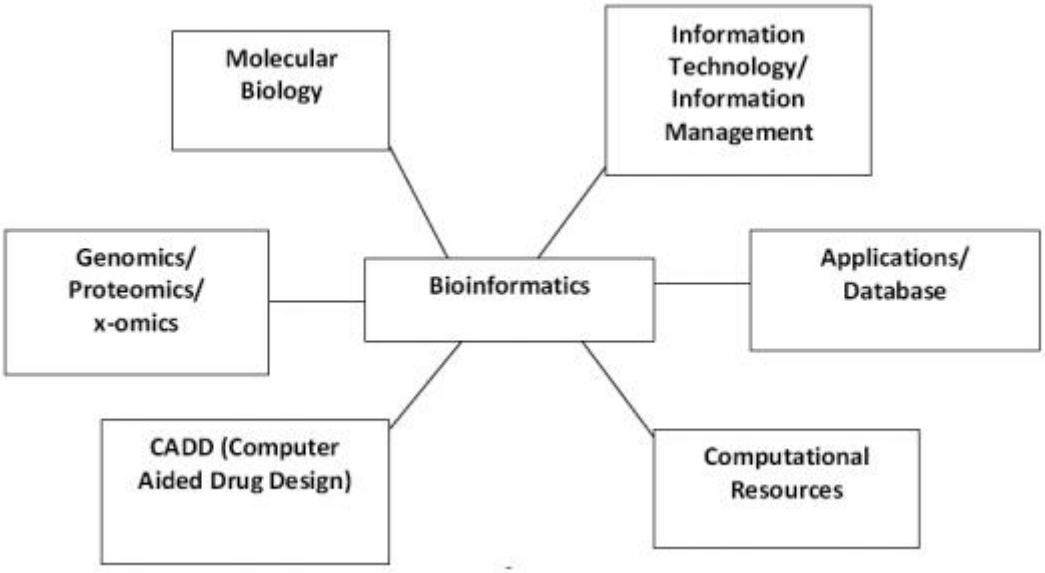


**Figure 4:** Bioinformatics in computer-aided drug design [7]

A few years ago, National Institutes of Health (NIH) created Biomedical Information Sciences and Technology Initiative (BISTI) to examine the current state of bioinformatics in the United States. Computer-aided drug design is a specialized discipline that uses computational methods to simulate drug-receptor interactions as there is considerable overlap in CADD research and bioinformatics [7].

**VARIOUS TYPES OF SOFTWARES USED FOR *IN SILICO* COMPUTER-AIDED DRUG DESIGN [8]:** Auto dock tools, UCSF chimera 1.10, Ligand scout 3.12, Rasmol, Chem draw utra 12, Chem sketch, Marvin sketch, Padel-descriptor, NCSS 10, Analyse-it.

**PARAMETERS:** Some important parameters of computer-aided drug design are described as below.

### (a) HOMOLOGY MODELING:

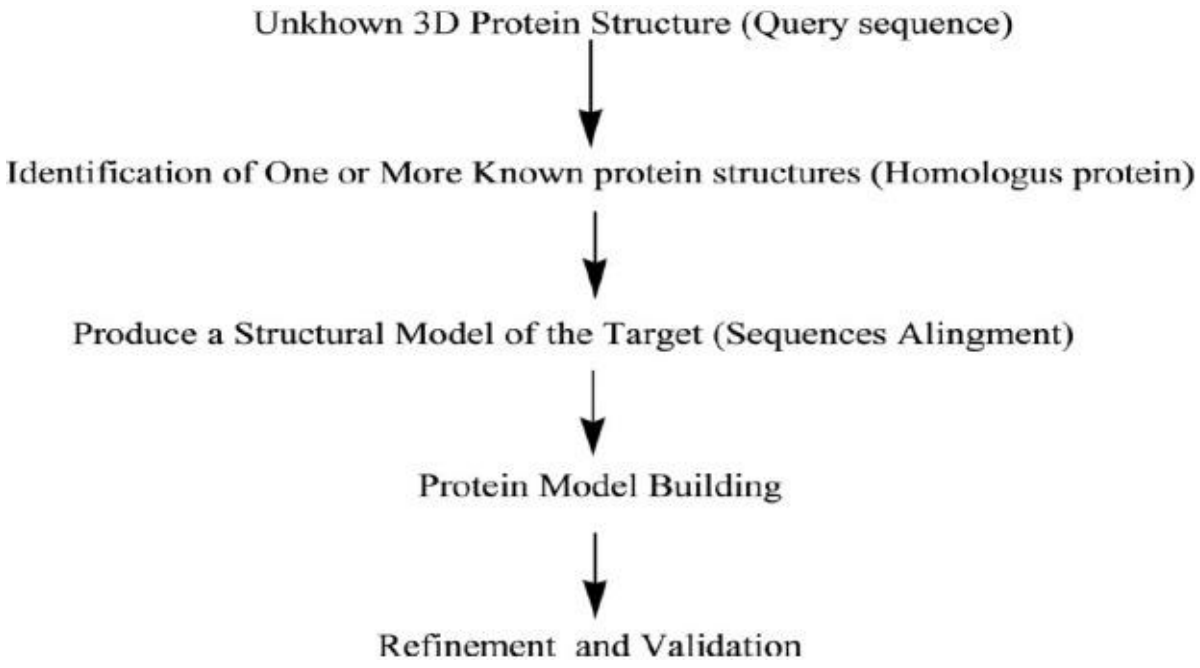


**Figure 5:** Structure prediction by homology modeling [9]

In the absence of experimental structures, computational methods are used to predict the 3D structure of target proteins. Comparative modeling is used to predict target structure-based on a template with a similar sequence, leveraging that protein structure is better conserved than sequence, i.e., proteins with similar sequences have similar structures. Homology modeling is a specific type of comparative modeling in which the template and target proteins share the same evolutionary origin. Comparative modeling involves the following steps:

(1) Identification of related proteins to serve as template structures,
(2) Sequence alignment of the target and template proteins,
(3) Copying coordinates for confidently aligned regions,
(4) Constructing missing atom coordinates of target structure,
(5) Model refinement and evaluation.

Several computer programs and web servers exist that automate the homology modeling process e.g., PSIPRED and MODELLER**.** Major goal of structural biology involve formation of protein-ligand complexes; in which the protein molecules act energetically in the course of binding. Therefore, perceptive of protein-ligand interaction will be very important for structure-based drug design. Lack of knowledge of 3D structures has hindered efforts to understand the binding specificities of ligands with protein. With increasing in modeling software and the growing number of known protein structures, homology modeling is rapidly becoming the method of choice for obtaining 3D coordinates of proteins. Homology modeling is a representation of the similarity of environmental residues at topologically corresponding positions in the reference proteins. In the absence of experimental data, model building on the basis of a known 3D structure of a homologous protein is at present the only reliable method to obtain the structural information. The knowledge of the 3D structures of proteins provides invaluable insights into the molecular basis of their functions [9].

**(b) LIGPLOT ANALYSIS:** Ligplot analysis a computer program that generates schematic 3D representations of protein-ligand complexes from standard ‘protein data bank (PDB)’ file input.

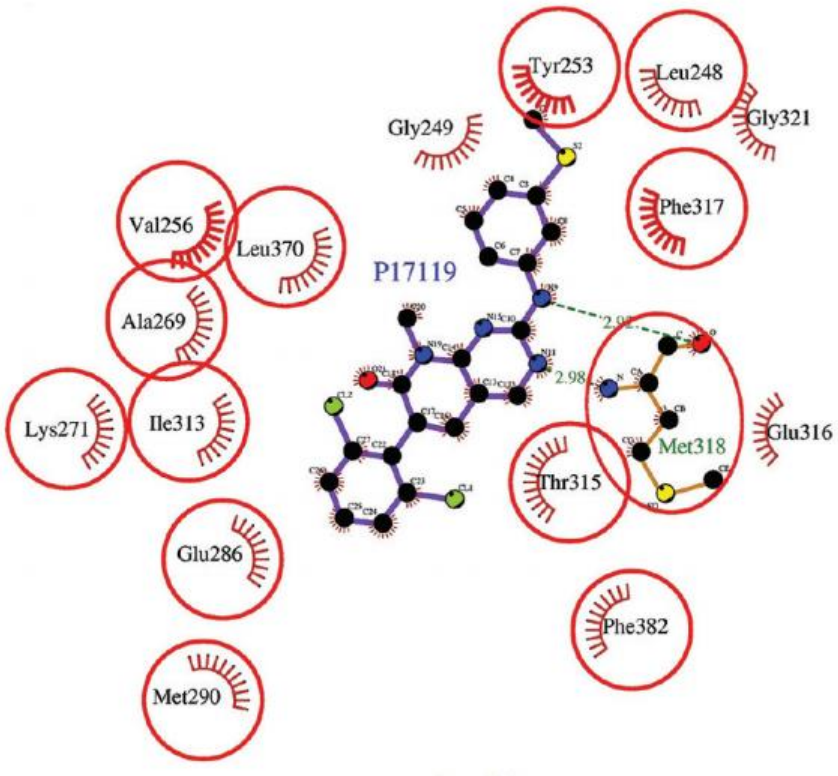


**Figure 6:** Ligplot analysis [The red circles and ellipses in each plot indicate protein residues that are in equivalent 3D positions to the residues in the first plot. Hydrogen bonds are shown as green dotted lines, while the arcs represent residues making non-bonded contacts with the ligand] [10].

## (c) MOLECULER DOCKING:

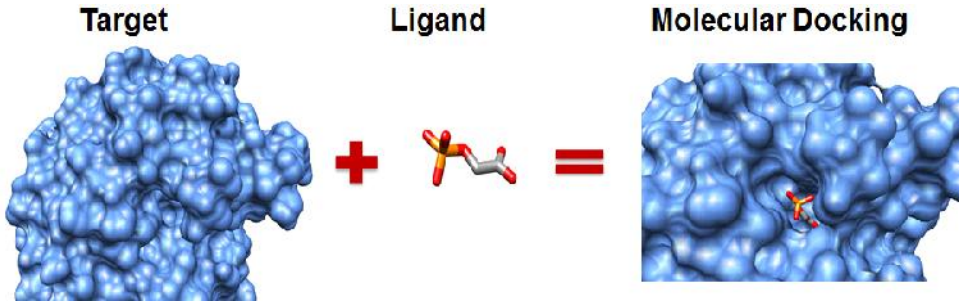


**Figure 7:** Molecular docking [11]

Molecular docking is the computational modeling of the structure of complexes formed by two or more interacting molecules. The goal of molecular docking is the prediction of the three dimensional structure. Docking plays an important role in the rational design of drugs. The aim of molecular docking is to achieve an optimized conformation for both the protein and ligand so and relative orientation between protein and ligand so that the free energy of the overall system is minimized. Molecular recognition plays a key role in promoting fundamental biomolecular events such as enzyme-substrate, drug-protein and drug-nucleic acid interaction [11].

- **Docking theory:** The following docking theory topics are available [11]:

1. CDOCKER: Uses a random preliminary ligand placement and full CHARMm forcefield-based docking.
2. LibDock: Fast docking-based on binding site features ('hotspots').
3. LigandFit: Docking-based on an initial shape match to the binding site.
4. MCSS: Uses CHARMm to dock fragments by using a unique computationally efficient Multiple Copy Simultaneous Search algorithm.

Drug-receptor interactions occur on atomic scales. To form a deep understanding of how and why drug compounds bind to protein targets, we must consider the biochemical and biophysical properties of both the drug itself and its target at an atomic level. Swiss PDB (protein data bank) is an excellent tool for doing this. It can predict key physico-chemical properties, such as hydrophobicity and polarity that have a profound influence on how drugs bind to proteins [6].

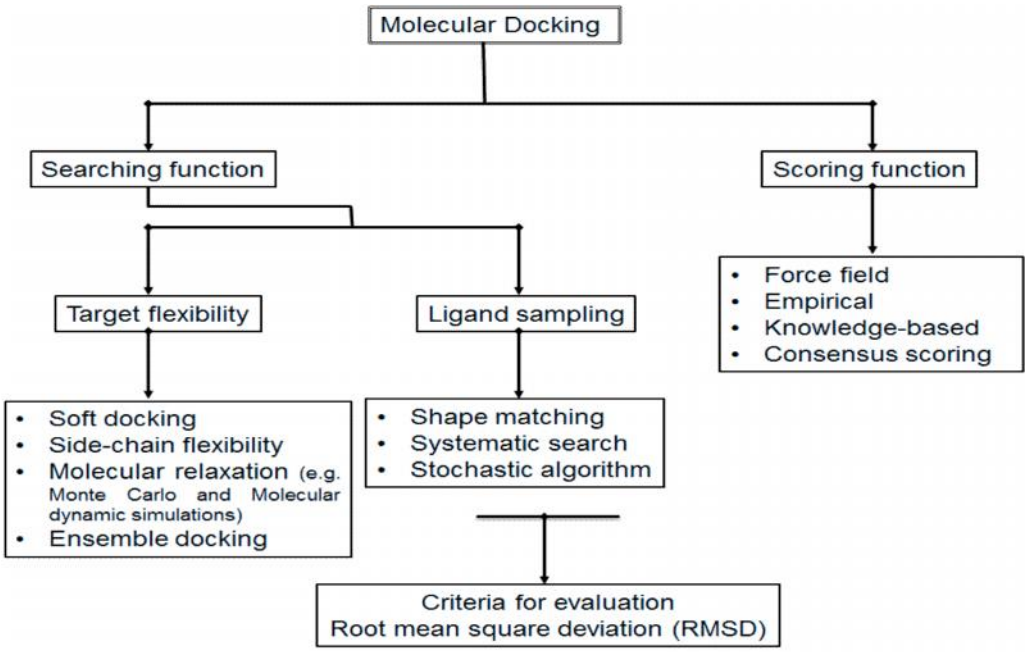


**Figure 8:** Methods used for protein-ligand docking [11]

## Applications and importance of molecular docking:

The uses of docking programmes to indicate the nature of the atoms and functional groups present in the 3D (three-dimensional) structures also enable to examine the binding of a drug to its target site [12].

**(d) DE NOVO DRUG DESIGN**: De novo design is the uses of docking programmes to design new lead structures that fit a particular target site.

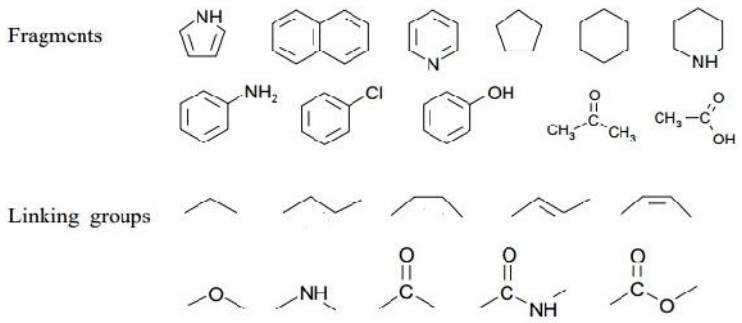


**Figure 9:** Different fragments and other linking groups used in de novo drug design methodology [13]

**Step 1:** Determination of binding pocket on target receptor

↓

**Step 2:** Prediction of interaction sites of target receptor through Ligplot

↓

**Step 3:** Placing the fragments or other linking groups with pharmacophore models at pre-defined interaction site to provide feasible interactions with the residues in the site of the target receptor

↓

**Step 4:** Structurally modification of the fragments to provide possible interactions with the residues in the site of the target receptor

↓

**Step 5:** Joining all fragments together to yield a complete single molecule

**Figure 10:** Steps of de novo drug design methodology [13]

## (e) PHARMACOPHORE-BASED DRUG DESIGN:

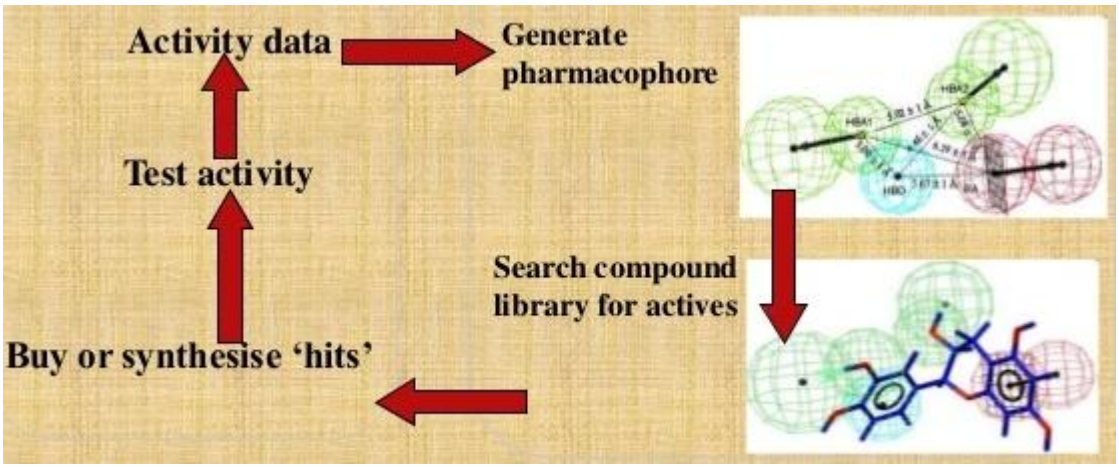


**Figure 11:** Pharmacophore-based drug design [14]

A pharmacophore is an abstract description of molecular features which are essential for molecular identification and recognition of a ligand by a biological macromolecule. Typical pharmacophoric features include hydrophobic centroids, aromatic rings, hydrogen bond acceptors, hydrogen bond donors, positive charge and negative charge. Pharmacophore approaches have become one of the major tools in drug discovery after the past century's development. Various ligand-based and structure-based methods have been developed for improved pharmacophore modeling. A pharmacophore model can be established either in a ligand-based manner, by superposing a set of active molecules and extracting common chemical features that are essential for their bioactivity, or in a structure-based manner, by searching possible interaction points between the macromolecular targets and ligands. Pharmacophore approaches have been used extensively in virtual screening [14].

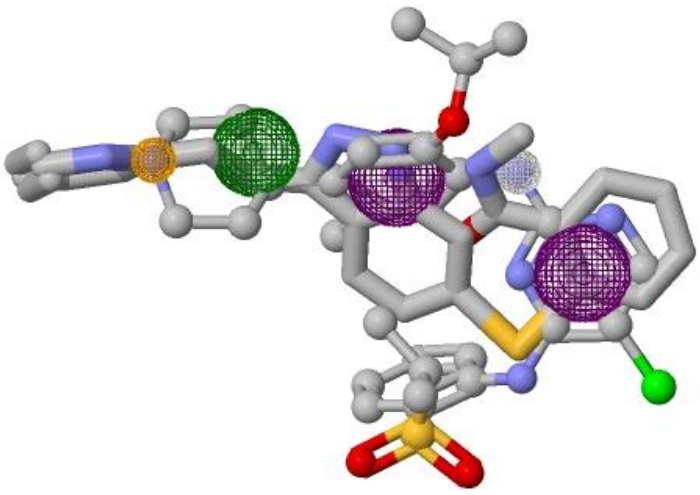

**Figure 12:** Pharmacophore [16]

- **Colours of Pharmacophoric Features** [16]**:**

1. Hydrogen bond acceptor: Orange,
2. Hydrogen bond donor: White,
3. Aromatic ring: Magenta,
4. Hydrophobic centroid: Green.

- **Uses**: Pharmacophores are frequently used as a tool for searching databases for compounds with similar pharmacophores [15].

### (f) VIRTUAL SCREENING (VS):

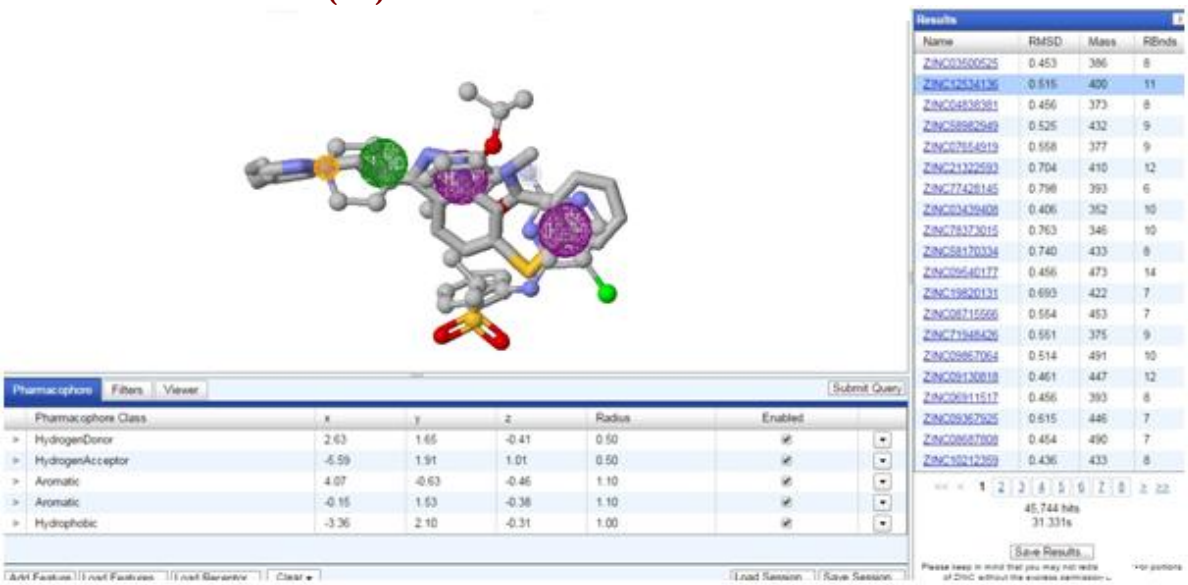


**Figure 13:** Virtual screening from the pharmacophoric model through Zinc pharmer web server [16]

Virtual screening is a computational method where large libraries of compounds are assessed for their potential to bind specific sites on target molecules such as proteins and well-compounds tested. Virtual screening is a computational technique used in drug discovery research. By using computers, it deals with the quick search of large libraries of chemical structure in order to identify those structures which are most likely to bind to a drug target, typically a protein receptor or enzyme. Virtual screening has become an integral part of the drug target, typically a protein receptor or enzyme. Virtual screening has become an integral part of the drug discovery process. Related to the more general and long pursued concept of database searching, the term "virtual screening" is relatively new. Virtual screening has largely been a numbers game focusing on questions like how can we filter down the enormous chemical space of over 1060 conceivable compounds to a manageable number that can be synthesized, purchased and tested. Although filtering the entire chemical universe might be a fascinating question, more practical virtual screening scenarios focus on designing and optimizing targeted combinatorial libraries and enriching libraries of available compounds from in-house compound repositories or vendor offerings. It is less expensive than high-throughput screening, faster than conventional screening, scanning a large number of potential drugs like molecules in very less time [17].

### (g) QUANTATIVE STRUCTURE-ACTIVITY RELATIONSHIPS (QSARs):

Quantitative structure-activity relationship (QSAR) modeling pertains to the construction of predictive models of biological activities as a function of structural and molecular information of a compound library. The concept of QSAR has typically been used for drug discovery and development and has gained wide applicability for correlating molecular information with not only biological activities but also with other physicochemical properties, which has therefore been termed quantitative structure-property relationship (QSPR) [18]. Typical molecular parameters that are used to account for electronic properties, hydrophobicity, steric effects, and topology can be determined empirically through experimentation or theoretically via computational chemistry [19]. A given set of data sets is then subjected to data pre-processing and data modeling through the uses of statistical or machine learning techniques. This review aims to cover the essential concepts and techniques that are relevant for performing QSAR/QSPR studies through the uses of selected examples from our previous work [20].

- **Regression analysis:** Regression analysis is a group of mathematical methods of QSAR used to obtain mathematical equations relating different sets of data that have been obtained from experimental work or calculated using theoretical study. The data are fed into a suitable computer program, which, on execution, produces an equation that represents the line that is the best fit for those data. Regression analysis would calculate the values of m and c that gave the line of best fit to the data [21].

**Importance:** The value of the r (regression coefficient) is a measure of how closely the data match the equation. r (regression coefficient) value greater than 0.60 are usually regarded as representing an adequate degree of accuracy. For example, a value of r > 0.60 or $R^2$ > 0.50 for natural or herbal compounds indicates that 80% of the results can be suitably explained by regression analysis by using the parameters specified [21].

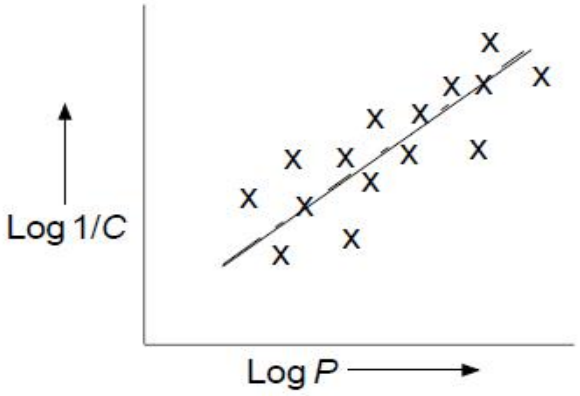


**Figure 14:** A hypothetical plot of the activity (Log1/C) of a series of compounds against the logarithm of their partition coefficients parameters (LogP) [21]

**(h) *IN SILICO* ADMET (ABSORPTION, DISTRIBUTION, METABOLISM, EXCRETION, TOXICITY) AND DRUG SAFETY PREDICTION:** Lipinski's rule is related to ADMET (absorption, distribution, metabolism, excretion and toxicity) which states that, in general, an orally active drug has no more than one violation of the following components [22,23,24,25,26,27,28,29]:

1. Hydrogen bond donor (the total number of nitrogen-hydrogen and oxygen-hydrogen bonds) in a molecule is not more than 5.
2. Hydrogen bond acceptor (all nitrogen or oxygen atoms) in a molecule is not more than 10.
3. Molecular weight (MW) of a molecule is less than 500 daltons or 800 gms.
4. Octanol-water partition coefficient (LogP) of a molecule is not greater than 5.
5. Polar surface area (PSA) of a molecule is not greater than 190 $Å^2$.
6. The range of molar refractivity (MR) of a molecule is in between 40 to 130.
7. The range of total number of atoms in a molecule is in between 20-70.
8. The range of total number of rotatable bonds in a molecule is not greater than 10.

- **Importance:** The rule is important to keep in mind during drug discovery when a pharmacologically active lead structure is optimized step-wise to increase the activity and selectivity of the compound as well as to ensure drug-like physicochemical properties are maintained.

There are various *in silico* tools to predict ADMET (absorption, distribution, metabolism, excretion and toxicity) like 1) ALOGPS, 2) E-dragon, 3) Padel-descriptors etc [30].

## Conclusion

The drug discovery and development process is a long and expensive one. It starts from target identification, after that, validates the targets and identifies the drug candidates before any newly discovered drug is placed on the market. It must undergo extreme preclinical and tests and get the FDA approval**.** Computer-aided drug design (CADD) is a natural outgrowth of theoretical chemistry, the traditional role of which involves the creation and dissemination of a penetrating conceptual infrastructure for the bioinformatics, chemical sciences, particularly at the atomic and molecular levels.

The main aim to decrease the level of manufacturing cost level. In particular, the strong mathematical flavour of CADD links between mathematical and the chemical sciences, and to the past, present and future roles of interdisciplinary research at the interface between these subjects. The issues constitute basis concerns for the present study. The growing number of chemical and biological databases; and explosions in currently available software tools are providing a much improved basis for the design of ligands and inhibitor with desired specificity.

## References


1. Veselovsk AV, Ivanov AS; Strategy of computer-aided drug design; Russian academy of medical science; 2003; 3(1); 33-40.
2. Yun T, Weiliang Z, Kaixian C, Hualiang J*; New technologies in computer-aided drug design*: *Toward target identification* and *new chemical entity discovery;* Drug discovery today: technologies; 2006; 3(3); 307-313.
3. Pranita PK, Madhavi MM, Rishikesh VA, Rajesh JO, Sandip SK; Computer-aided drug design: an innovative tool for modeling; Journal of medicinal chemistry; 2012; 2; 139-148.
4. Pugazhendhi D, Umamaheswari TS; *In silico* methods in drug discovery–a review; International journal of advanced research in computer science and software engineering; 2013; 3(5); 680-683.
5. http://www.slideshare.net/adammbbs/structure-based-drug-design-48389641/12; Structure-based drug design (Docking and de novo drug design); Authored by Adam Shahul Hameed; Anna University Chennai; Available online from 2015.
6. Silvana G; Computer-based methods of inhibitor prediction; Intech open science and open minds; 2013; Chapter 3; 73-90. (http://cdn.intechopen.com/pdfs-wm/40599.pdf)
7. Satyajit D, Sovan S, Kapil S; Computer-aided drug design - a new approach in drug design and discover; 2010; 4(3); 146-151.
8. Choudhary L.K., Shukla A., Zade S., Charde R; (C.A.D.D.)-a new modern software-based approach in drug design and discovery; International journal of pharmaceutical chemistry; 2011; 1(1); 10-20.
9. Vyas VK, Ukawala RD, Ghate M, Chintha C; Homology modeling a fast tool for drug discovery: current perspectives; Indian journal of pharmaceutical sciences; 2012; 74(1); 1–17.
10. Roman AL, Mark BS; LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery; Journal of chemical information and modeling; 2011; 51; 2778-2786.
11. Alejandra HS, Aldo YT, Victor A, Héctor VC, Claudia MB; Protein-protein and protein-ligand docking; Intech open science and open minds; 2013; Chapter 3; 64-81. (http://cdn.intechopen.com/pdfs-wm/44790.pdf)
12. Gaba M, Gaba P, Singh S, Gupta GD; An overview on molecular docking; International Journal of Drug Development and Research; 2010; 2(2); 219-231.
13. Gareth T, Medicinal chemistry an introduction; Computer-aided drug design (Chapter-4); John Wiley and Sons Ltd; 2007; 2nd edition; 128-130.
14. http://www.slideshare.net/8055679610/final-presentation66; Pharmacophore Modeling in Drug Designing; Authored by Vinod Tonde; Rajmata jijau shikshan prasarak mandal's college of pharmacy, Dudulgaon, Pune-412105; Available online from 2014.
15. Udugade BV, Gawade SP; 3D QSAR and pharmacophore modeling on substituted cyanopyrrolidines as type ii anti-diabetic agents potential dipeptidyl peptidase-iv inhibitors; Pharmacophore (An international research journal); 2016; 7(5); 342-348.
16. David RK, Carlos JC; ZINC Pharmer: pharmacophore search of the ZINC database; Nucleic acids research; 2012; 40; 409-414.
17. Reddy AS, Pati SP, Kumar PP, Pradeep HN, Sastry GN; Virtual screening in drug discovery-A computational perspective; Current protein and peptide science; 2007; 8; 329-351.
18. Chanin N, Chartchalerm INA, Thanakorn N, Virapong PA; Practical overview of quantitative structure-activity relationship; EXCLI Journal; 2009; 8; 74-88.
19. Anu G, Dr. Manish G, Dr. Komal S; A Practical overview of quantitative structure-activity relationship; 2016; 5(2); 427-437.
20. Denis F, Eugene M, Alexander T; Trust but verify: On the importance of chemical structure curation in cheminformatics and QSAR modeling research; Journal of Chemical Information and Modeling; 2010; 50(7); 1189–1204.
21. Gareth T, Medicinal chemistry an introduction; Structure–activity and quantitative structure relationships (Chapter-3); John Wiley and Sons Ltd; 2007; 2nd edition; 93.
22. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ; Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings; *Advanced drug delivery* reviews; 2001; 46(1-3); 3–26.

23. Lipinski CA; Lead- and drug-like compounds: the rule-of-five revolution; Drug discovery today: Technologies; 2004; 1(4); 337–341.
24. Leo A, Hansch C, Elkins D; Partition coefficients and their uses; Chemical reviews; 1971; 71(6); 525–616.
25. *Ghose AK, Viswanadhan VN, Wendoloski JJ; A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases;* Journal of combinatorial chemistry*; 1999; 1(1); 55–68.*
26. *Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD; Molecular properties that influence the oral bioavailability of drug candidates; Journal* of *medicinal chemistry; 2002; 45(12); 2615–2623.*
27. Congreve M, Carr R, Murray C, Jhoti H; A 'rule of three' for fragment-based lead discovery?; Drug discovery today; 2003; 8(19); 876–877.
28. Oprea TI, Davis AM, Teague SJ, Leeson PD; Is there a difference between leads and drugs? A historical perspective; Journal of *chemical information* and modeling; 2001; 41(5); 1308–1315.
29. Leeson PD, Springthorpe B; The influence of drug-like concepts on decision-making in medicinal chemistry; Nature reviews drug discovery; 2007; 6(11); 881–890.
30. RamaKrishna K, Rao CVK, Rao RS; ChemoInformatics Part I: Molecular descriptors in omnimetrics research; Journal of applicable chemistry; 2015; 4(4); 1024-1144.

**Access this Article in Online**

Website: www.ijarbs.com

Subject: Pharmaceutical Sciences

Quick Response Code

DOI:10.22192/ijarbs.2017.04.02.009

How to cite this article:
Injamul Hoque, Arkendu Chatterjee, Somenath Bhattacharya, Raj Biswas. (2017). An Approach of Computer-Aided Drug Design (CADD) Tools for *In Silico* Pharmaceutical Drug Design and Development. Int. J. Adv. Res. Biol. Sci. 4(2): 60-71.
DOI: http://dx.doi.org/10.22192/ijarbs.2017.04.02.009