Review literature on Role of Database in CADD

Shyam Narayan Gupta*, Piyush Yadav , Manoj Kumar Yadav,

Shashikant Maury, Shradha Sahu

Abstract:

The objective of this study was to know about use of Database in computer aided drug design. Databases are an important part of modern life. Without them, most computer functions would cease to exist.Chemical database is specifically designed to store chemical information. Drug discovery is the step-by-step process by which new candidate drugs are discovered. Traditionally, pharmaceutical companies follow well-established pharmacology and chemistry-based drug discovery approaches, and face various difficulties in finding new drugs. The increasing pressure to generate more and more drugs in a short period of time with low risk has resulted in remarkable interest in bioinformatics .In fact, now there is an existence of new, separate field, known as computer-aided drug design Various approaches of CADD are evaluated as promising techniques according to their need, in between all these structure-based drug design and ligand-based drug design approaches are known as very efficient and powerful techniques in drug discovery and development.

Keyword: Computer Aided Drug Design, Database, Rational drug design, Cheminformatics, Bioinformatics

Introduction:-

Computer aided drug design (CADD) provides several tools and techniques that helps in various stages of drug design thus reducing the cost of research and development time of the drug. Drug discovery and developing a new medicine is a long, complex, costly and highly risky process that has few peers in the commercial world[1]. This is why computer-aided drug design (CADD) approaches are being widely used in the pharmaceutical industry to accelerate the process. The cost and time invested by the pharmacological research laboratories are heavy during the various phases of drug discovery, starting from therapeutic target identification[2,3]

Drug design often referred to as rational drug design or simply rational design is the inventive process of finding new medications based on the knowledge of a biological target. The drug is most commonly an organic small molecule that activates or inhibits the function of a bimolecular such as a protein, which in turn results in a therapeutic benefit to the patient[3]. In the most basic sense, drug design involves the design of molecules that are complementary in shape and charge to the bimolecular target with which they interact and therefore will bind to it. Drug design frequently but not necessarily relies on computer modeling techniques.

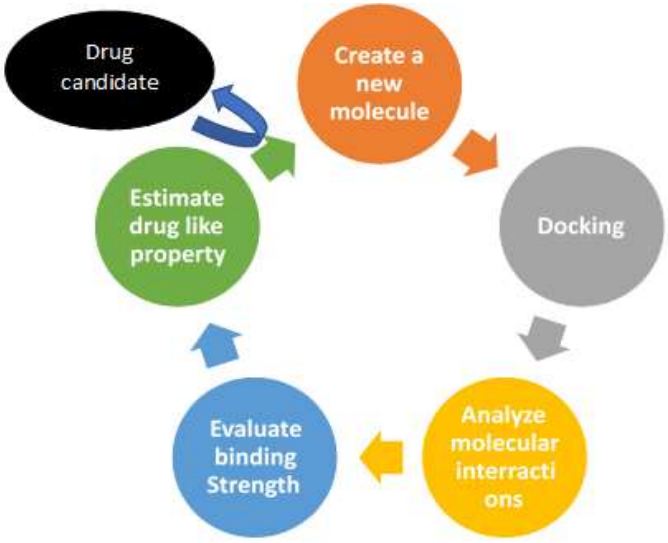

Fig 1: Drug design through CADD

Drug design is the approach of finding drugs by design, based on their biological targets. Typically a drug target is a key molecule involved in a particular metabolic or signaling pathway that is specific to a disease condition or pathology, or to the infectivity or survival of a microbial pathogen. Some approaches attempt to stop the functioning of the pathway in the diseased state by causing a key molecule to stop functioning. Drugs may be designed that bind to the active region and inhibit this key molecule[4]. However these drugs would also have to be designed in such a way as not to affect any other important molecules that may be similar in appearance to the key molecules. Sequence homologies are often used to identify such risks.Other approaches may be to enhance the normal pathway by promoting specific molecules in the normal pathways that may have been affected in the diseased state[5].

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[6].The structure of the drug molecule that can specifically interact with the biomolecules can be modeled using computational tools[2].These tools can allow a drug molecule to be constructed within the biomolecules using knowledge of its structure and the nature of its active site[4,6]. Construction of the drug molecule can be made inside out or outside in depending on whether the core or the residual R-groups are chosen first.

Computational approaches in drug design, discovery and development process gaining very rapid exploration, implementation and admiration.Introducing a new drug in a market is a very complex, risky and costly process in terms of time, money and manpower. Generally it is found that drug discovery and development process takes around 10-14 years and more than 1 billion dollars capital in total[7]. The major pharmaceutical companies have invested heavily in the routine ultra-High Throughput Screening (uHTS) of vast numbers of drug-like‘ molecules In parallel with this, drug design and optimization increasingly uses computers for virtual screening. Recent advancements in DNA microarray experiments explore thousands of genes involved in a disease can be used for gaining in depth knowledge about the disease targets, metabolic pathways and toxicity of the drugs[4,7].

The theoretical tools include empirical molecular mechanics, quantum mechanics and, more recently, statistical mechanics. This latest advance has permitted explicit solvent effects to be Incorporated. All this work is the availability of high quality computer graphics, largely supported on workstations. CADD can be separated into ligand or hit identification and ligand or hit optimization, with both SBDD and LBDD methods useful in the appropriate context. Database screening methods are often

used for hit identification while a number of methods may be used for hit optimization[8,9]. These include the Site-identification by ligand competitive saturation (SILCS) methodology.

Types of Drug Design:-

1. Rational drug design :-

Unlike the historical method of drug discovery, by trial-and-error testing of chemical substances on cultured cells or animals, and matching the apparent effects to treatments, rational drug design begins with a knowledge of specific chemical responses in the body or target organism, and tailoring combinations of these to fit a treatment profile. Due to the complexity of the drug design process two terms of interest are still serendipity and bounded rationality[10]. The first unequivocal example of the application of structure-based drug design leading to an approved drug is the carbonic anhydrate inhibitor dorzolamide which was approved in 1995. The activity of a drug at its binding site is one part of the design. Another to take into account is the molecule's drug likeness, which summarizes the necessary physical properties for effective absorption[7,9].

Fig 2: Process of Rational Based Drug Design

2. Structure based drug design:-

Drug discovery referred to, as 'rational' did not take flight until the first structures of the targets were solved. In 1897, Ehrlich suggested a theory called the side chain theory wherein he proposed that specific groups on the cells combine with the toxin[11]. Ehrlich coined these side chains as receptors. Structure-based drug design of protein ligands has emerged as a new tool in medicinal chemistry.and often proceeds through multiple cycles before an optimized lead goes into clinical trials. The first cycle includes the cloning, purification and structure determination of the target protein or nucleic acid by one of three principal methods: X-ray crystallography, NMR or comparative modeling[18]. In the second cycle, structure determination of the target in complex with a promising lead from the first cycle, one with at least micromolar inhibition in vitro, reveals sites on the compound that can be optimized to increase potency.

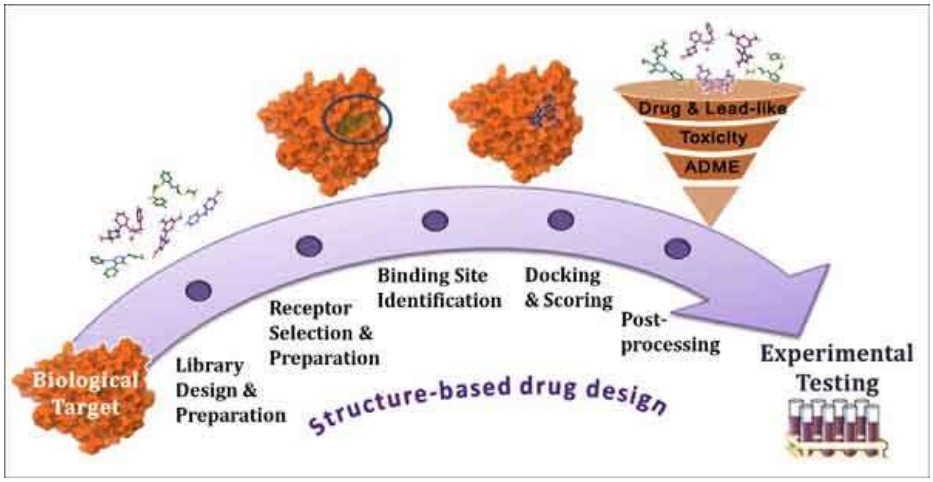

Fig 3: Process of Structure based drug design

3. Ligand Based Drug Design :-

In this method, target is a ligand. Structural and electronic properties of molecules use to predict their biological activity, and this is referred to as a Structure-activity relationship (SAR). In LBDD, 3D structure of the target protein is not known but the knowledge of ligands which binds to the desired target site is known. These ligands can be used to develop a pharmacophore model or molecule which possesses all necessary structural features for bind to a target active site[11,12]. Generally ligand-based techniques are pharmacophore based approach and quantitative-structure activity relationships (QSARs). In LBDD it is assumed that compounds which having similarity in their structure also having the same biological action and interaction with the target protein[11,13].

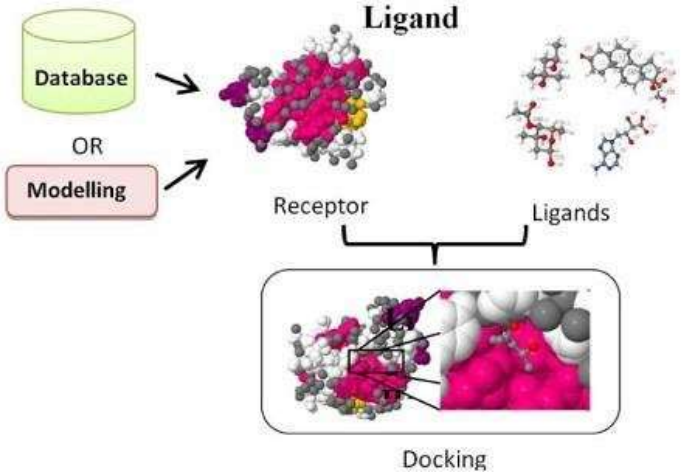

Fig 4 : Process in Ligand Based Drug Desine[13]

Role of Database in Computer Aided Drug Design:-

Database :-

Databases are an important part of modern life. Without them, most computer functions would cease to exist. If you're someone that relies on storing information within a computer, whether as an individual or for your job, then it's important to understand the different types of databases that exist and how you should use them[15]. Researchers in large pharmaceutical companies typically draw on a wide range of data resources and tools to enable decisions regarding target selection, lead identification, optimization

and candidate selection. Much of this information is either generated internally or licensed from commercial vendors[17,14]. As database technology has improved over the years, so too have the different types of databases. There are now many different types of databases, each with its strengths and weaknesses based on how they are designed.

Types of Database

1. **Centralized database**:- A centralized database is one that operates entirely within a single location. Centralized databases are typically used by bigger organizations, such as a business or university. The database itself is located on a central computer or database system. Users can access the database through a computer network, but it is the central computer that runs and maintains the database.

2. **Cloud database**:- A cloud database is one that runs over the Internet. The data is stored on a local hard drive or server, but the information is available online. This makes it easy to access your files from anywhere, as long as you have an Internet connection. To use a cloud database, users can either build one themselves or pay for a service to store their data for them. Encryption is an essential part of any cloud database, as all information needs to be protected as it is transmitted online.

3. **Operational database:-** The purpose of an operational database is to allow users to modify data in real-time. Operational databases are critical in business analytics and data warehousing. They can be set up either as relational databases or NoSQL, depending on needs. Conventional databases rely on batch processing, where commands are carried out in groups. Operational databases, on the other hand, allow you to add, edit and remove data at any moment, in real-time.

4. **Personal database:-** A personal database is one that is designed for a single person. It is typically stored on a personal computer and has a very simple design, consisting of only a few tables. Personal databases are not typically suitable for complex operations, large amounts of data or business operations.

5. **Relational database:-** Relational databases are the other major type of database, opposite of NoSQL. With a relational database, information is stored in a structured way and about other data[15]. A good representation of a relational database would be a person shopping online and their shopping cart. Relational databases are often preferred when you are concerned about the integrity of your data, or when you're not particularly focused on scalability.

6. **Graph database:-** Graph databases are databases that focus equally on the data and the connections between them. In this database, data is not constricted to pre-defined models. Most other databases can find connections between data when you run a search[15]. With a graph database, these connections are stored inside the database right alongside the original data. This makes for a more efficient and faster database when your primary goal is to manage the connections between your data.

Type of Database applicable in Cheminformatics & Bioinformatics:

1. Chemical Database
2. Biochemical Database
3. Pharmaceutical Database
4. Pharmaceutical Database

Chemical Database :

A chemical database is a database specifically designed to store chemical information. This information is about chemical and crystal structures, spectra, reactions and syntheses, and thermo physical data. Chemical structures are traditionally represented using lines indicating chemical bonds between atoms and drawn on paper[18]. While these are ideal visual representations for the chemist, they are unsuitable for computational use and especially for search and storage. Small molecules (also called ligands in drug design applications), are usually represented using lists of atoms and their connections. Large molecules such as proteins are however more compactly represented using the sequences of their amino acid building blocks. Large chemical databases for structures are expected to handle the storage and searching of information on millions of molecules taking terabytes of physical memory.

Chemists can search databases using parts of structures, parts of their IUPAC names as well as based on constraints on properties.Chemical databases are particularly different from other general purpose databases in their support for sub-structure search. All properties of molecules beyond their structure can be split up into either physico-chemical or pharmacological attributes also called descriptors[18,19]. Databases systems for maintaining unique records on chemical compounds are termed as Registration systems. These are often used for chemical indexing, patent systems and industrial databases.

Bioinformatics database :

A biological database is a large, organized body of persistent data, usually associated with computerized software designed to update, query, and retrieve components of the data stored within the system[21]. In biology, bioinformatics is defined as, “the use of computer to store, retrieve, analyse or predict the composition or structure of bio-molecules”.Through it we can reduce the synthetic and biological testing efforts[20]. Bioinformatics is the application of computational techniques and information technology to the organization and management of biological data. Classical bioinformatics deals primarily with sequence analysis

Bioinformatics includes biological studies that use computer programming as part of their methodology, as well as a specific analysis "pipelines" that are repeatedly used, particularly in the field of genomics[8,21]. For instance, bioinformatics tools such as the comparative analysis of genomic and genetic data and signal processing help to interpret and understand the molecular and evolutionary processes[22] and interactions from large volumes of raw data in the field of wet-bench experimental molecular biology. Common uses of bioinformatics include the identification of candidates genes and single nucleotide polymorphisms[21].

Application of bioinformatics:

- Development of database containing all biological information.
- Development of better tools for data designing, annotation and mining.
- Design and development of drugs by using simulation software.
- Design and development of software tools for protein structure prediction function, annotation and docking analysis.
- Creation and development of software to improve tools for analyzing sequences for their function and similarity with other sequences

Conclusion :

The Role of database is very important in CADD.With the help of it we can find specific record of given data. They store special information used to manage the data. CADD database is the file that stores information about the part. In order to make this database an efficient one, a suitable database management system (DBMS) is required. The basic purpose of the CADD system is to create the geometric model of a physical part as close as possible to its true shape and size.

Reference:

1. Narayan, K. Lalit (2008). Computer Aided Design and Manufacturing. New Delhi: Prentice Hall of India.
2. Hoque I, Chatterjee A, Bhattacharya S, Biswas R. An Approach of Computer-Aided Drug Design (CADD) Tools for In Silico Pharmaceutical Drug Design and Development. Int. J. Adv. Res. Biol. Sci. 2017; 4(2) 60-71
3. Madsen U ,Krogsgaard –Larsen P ,Lijefors T (2002). Textbook of Drug Design and Discovery. Washington , DC: Taylor & Francis.
4. Kapetanovic IM. Computer-aided drug discovery and development (CADD): insilico-chemico-biological approach. Chemico-biological interactions. 2008; 171(2):165-76.
5. Greer J, Erickson JW, Baldwin JJ, Varney MD (1994). "Application of the three-dimensional structures of protein target molecules in structure-based drug design". J Med Chem 37 (8): 1035 - 1054.
6. Satyajit D, Sovan S, Kapil S; Computer-aided drug design - a new approach in drug design and discover; 2010; 4(3); 146-151.
7. Daina A, Blatter MC, Baillie Gerritsen V, Palagi PM, Marek D, Xenarios I, et al. Drug Design Workshop: A Web-Based Educational Tool To Introduce Computer-Aided Drug Design to the General Public. Journal of Chemical Education. 2017; 94(3):335-44.
8. Yu W, Guvench O, MacKerell AD. Computational approaches for the design of protein–protein interaction inhibitors. In: Zinzalla G, editor. Understanding and exploiting protein–protein interactions as drug targets. London, UK: Future Science Ltd; 2013. pp. 99–102
9. Sliwoski G, Kothiwale S, Meiler J, Lowe EW. Computational Methods in Drug Discovery. Pharmacol Rev. 2014;66:334–395
10. Macalino SJ, Gosu V, Hong S, Choi S. Role of computer-aided drug design in modern drug discovery. Archives of pharmacal research. 2015; 38(9):1686-701.
11. Tropsha A (2010). "QSAR in Drug Discovery". In Reynolds CH, Merz KM, Ringe D (eds.). Drug Design: Structure- and Ligand-Based Approaches (1 ed.). Cambridge, UK: Cambridge University Press. pp. 151–164.
12. Satyajit D, Sovan S, Kapil S; Computer-aided drug design - a new approach in drug design and discover; 2010; 4(3); 146-151.
13. Hoque I, Chatterjee A, Bhattacharya S, Biswas R. An Approach of Computer-Aided Drug Design (CADD) Tools for In Silico Pharmaceutical Drug Design and Development. Int. J. Adv. Res. Biol. Sci. 2017; 4(2) 60-71.
14. Maithri G, Manasa B, Vani SS, Narendra A, Harshita T. Computational drug design and molecular dynamic studies—a review. Biomedical Data Mining. 2016; 5:1-7.
15. https://www.indeed.com/career-advice/career-development/types-of-databases.

16. Harris, JB (2019). "Post-processing of Large Bioactivity Data". Bioinformatics and Drug Discovery. Methods Mol Biol. 1939. pp. 37–47
17. Austin CP, Brady LS, Insel TR, Collins FS. NIH Molecular Libraries Initiative. Science306(5699),1138-1139 (2004).
18. https://en.wikipedia.org/wiki/Chemical_database
19. Kapetanovic IM. Computer-aided drug discovery and development (CADD): insilico-chemico-biological approach. Chemico-biological interactions. 2008; 171(2):165-76.
20. Zhang S. Computer-aided drug discovery and development. In Drug Design and Discovery. 2011; 716:23-38.
21. https://en.wikipedia.org/wiki/Bioinformatics
22. Saxena A., Soni B.P., Gupta V. A chronological review and comparison of four evolutionary based algorithms. European Journal of Advances in Engineering and Technology, 2015;2:35–41.