Journal of Science and Technology
ISSN: 2456-5660 Volume 7, Special Issue 03 (MAY 2022)

www.jst.org.in *DOI:https://doi.org/10.46243/jst.2022.v7.i03.pp49-52*

Study of role of Cheminformatics in the Modern Drug Discovery Process

Prashant R. Mahalle*
(Dept. of Chemistry, Late B. S. Arts, Prof. N. G. Science & A. G. Commerce College, Sakharkherda, India)
*Email:*prashantmahalle2014@gmail.com

To Cite this Article
Prashant R. Mahalle *"***Study of role of Cheminformatics in the Modern Drug Discovery Process***", Journal of Science and Technology, Vol. 07, Special Issue 03, May 2022.*

Article Info
Received: 18-04-2022 *Revised: 8-05-2022* *Accepted: 10-05-2022* *Published: 20-05-2022*

Abstract:*Cheminformatics is concerned with the application of computational methods to tackle chemical problems, with particular emphasis on the manipulation of chemical structural information. Many of the techniques used in Cheminformatics are in fact rather well established, being the result of years if not decades of research in academic, government and industrial laboratories. Whilst some of the current interest in Cheminformatics can be ascribed to the natural enthusiasm for things new, the main reason for its emergence can be traced to the need to deal with the vast quantities of data being generated by new approaches to drug discovery such as high-throughput screening and combinatorial chemistry.*

Key Word*:Cheminformatics, computational methods, combinatorial chemistry, drug design.*

I. Introduction

The field of chemistry itself has benefitted greatly from this availability, but the development of many new methods, algorithms, and data sources was necessary to realise the compute power now available to the chemist. The interface science of Cheminformatics has the objective of applying computer science approaches in the representation, analysis, design, and modeling of chemical structures and associated metadata, such as biological activity endpoints and physicochemical properties. The field of Cheminformatics not only draws on expertise in computer science and chemistry, but also mathematics, statistics, biology, physics, and biochemistry. Cheminformatics[1] combines the scientific working fields of Chemistry and computer science especially in the area of chemical Graph theory and mining the chemical space. It is to be expected that the chemical space contains at least 1062 molecules. Cheminformatics is a generic term that encompasses the design, creation, organization, management, retrieval, analysis, dissemination, visualization, and use of chemical information. The transformation of data into information and of information into knowledge is an endeavor needed in any branch of chemistry not only in drug design.

The drug discovery[2] process emphasize the entire pharmaceutical industry, encircling the early stages of research from target discovery and validation, right through to the identification of a drug candidate or lead compound. Initial identification of small remedial candidates comes about via a variety of streams. Research can lead to new insights into disease processes that highlight novel pathways for which drugs can be developed to intercede.

The history of Cheminformatics:

Computers and computational methods[3] applied to chemistry appeared very early: the 1950's for statistical models (now QSAR), and the 1960's for the first computer representations, mainly by curious chemists. The bulk of the foundational work in what we now call Cheminformatics was done in the 70's and 80's, and was strongly supported by the pharmaceutical industry and the need for computational drug discovery research. Cheminformatics has some traditional areas of application (pharmaceutical drug discovery, databases of available chemicals, journal article indexing, patent databases) and some newer ones (pathway databases, probe discovery, polypharmacology, toxicology, etc). In particular, there has recently been a big increase in the amount of chemical information in the

public domain, and a deeper integration with other related areas such as bioinformatics[4] and chemogenomics. The General sources, linkages and other fields related with Cheminformatics can show as follows (Fig. 1)

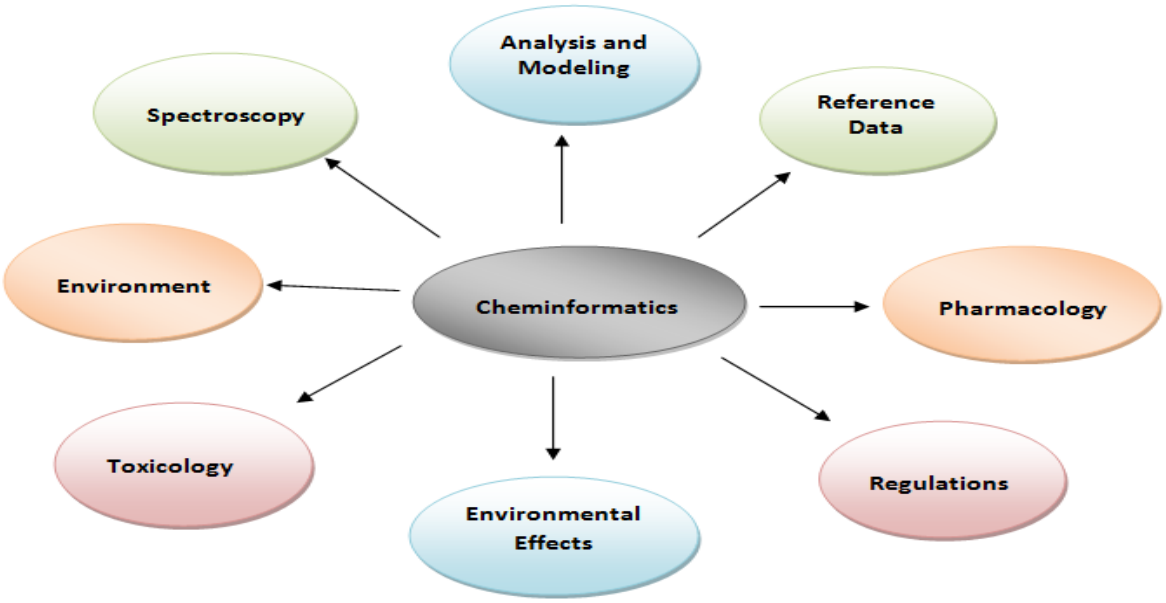

Fig: 1

Drug development:

Drug development[5]is the process of bringing a new pharmaceutical drug to the market once a lead compound has been identified through the process of drug discovery. It includes preclinical research on microorganisms and animals, filing for regulatory status, such as via the United States Food and Drug Administration for an investigational new drug to initiate clinical trials on humans, and may include the step of obtaining regulatory approval with a new drug application to market the drug. The entire process – from concept through preclinical testing in the laboratory to clinical trial development, including Phase I–III trials – to approved vaccine or drug typically takes more than a decade. Broadly, the process of drug development can be divided into preclinical and clinical work.

Step 1: Drug discovery and target validation:

The first step in the drug development process involves discovery work. This is where drug development companies choose a molecule, such as a gene or protein, to target with a drug. This is also where the drug developer will confirm that the molecule is indeed involved with the disease in question. After testing multiple drug molecules, the drug development company will choose those that have promise.

Step 2: Preclinical testing:

The next step in the drug development process is preclinical testing, which in itself is divided into two subcomponents: *in vitro* and *in vivo* testing. *In vitro* testing examines the drug molecules' interactions in test tubes and within the lab setting. *In vivo* testing involves testing the drug molecules on animal models and in other living cell cultures. Although efficacy is beginning to be established here, safety is paramount as the FDA[6] will not let preclinical studies move into human trials without extensive data on safety. This is the stage where researchers will whittle thousands of drug molecule candidates down to between one and five.

Step 3: Investigational New Drug application filing:

The third step involves submitting an Investigational New Drug application to the FDA prior to beginning human clinical trials. This is the point where the FDA will scrutinize the results from the preclinical testing, look at side effects and other safety features of an experimental drug, examine the drugs' chemical structure and how it's believed to work, and take its first look at the manufacturing process of the drug. If the FDA approves a drug

developers' IND, then it can move onto human trials. An IND approval is also the point at which a patented drugs' 20-year exclusivity period begins.

Step 4: Phase 1 clinical studies:

The first phase of human clinical testing involves a relatively small group of healthy people, usually a dozen to a few dozen, and it'll focus entirely on safety. This stage of study involves looking at how a drug is absorbed and eliminated from the body, as well as what side effects it may cause and whether or not it's producing the desired effect. Phase 1 clinical studies are also where maximum tolerated doses are established. It really is all about safety, although it's not uncommon for drug developers to tout early signs of efficacy in phase 1. If everything looks promising the study moves to phase 2, or midstage trials.

Step 5: Phase 2 clinical studies:

The two big changes between early stage and mid-stage trials are that the patient pool widens from a few dozen to perhaps 100 or more patients, and the patients being treated are no longer healthy volunteers but people being afflicted by the disease in question. Safety remains a big focus of phase 2 studies, with short-term side effects being closely monitored, although an increasing emphasis will begin to be placed on whether or not a drug is working as expected and if it's improving the condition or not.

Step 6: Phase 3 clinical studies:

In phase 3 studies, safety remains a priority, but this is where efficacy also plays a big role. Phase 3 studies are designed by drug developers but approved by the FDA with guidelines for a clearly defined primary endpoint to determine the success or failure of a tested drug. Phase 3 trials involve even more patients, perhaps a few hundred to maybe thousands, and they are by far the longest and costliest of all components of the drug development process. This is also the stage where drug developers will begin to think about how they're going to ramp up production if the phase 3 results are promising.

Step 7: New Drug Application filing:

The seventh step in the drug development process is simple: filing a New Drug Application with the FDA. Unfortunately, this isn't just a single page that says "please look at our drug!" An NDA can be tens of thousands or perhaps 100,000 or more pages long, and it contains all research and safety data examined during each of the six prior steps. Still, this stage isn't the point where the FDA has to make a decision to approve or deny the drug; it's merely a stepping stone that says it promises to review the application over the next 10 months. If the NDA is accepted a PDUFA, or Prescription Drug User Fee Act, date is set 10 months down the road (for a standard application) whereby the FDA is expected to make its decision.

Step 8: PDUFA date and decision:

More often than not, the FDA will wait until the PDUFA date to release its decision. Essentially the FDA has three choices: it can approve a drug; it can outright deny a drug (which is pretty rare from what I've witnessed in 15 years), or it can request additional information by sending a complete response letter, or CRL. A CRL simply states what was lacking that prevented the drug from being approved and offers suggestions as to how to remedy the situation. Often times it requires drug developers to run additional studies or perhaps alter their manufacturing process to appease the FDA. If approved by the FDA, the drug becomes immediately available for commercial production.

Step 9: Phase 4 clinical studies:

"Technically" an approved drug can make it to your medicine cabinet after step eight, but that doesn't mean the drug developer is off the hook yet. Even after approval, it's not uncommon for the FDA to request long-term safety studies be undertaken whereby drug developers are required to submit regular reports detailing any adverse events with the drug to the FDA. Even following approval, safety remains the top priority of the FDA.

Programming toolkits for Cheminformatics:

There are fortunately several free and open-source toolkits available that include libraries for many common Cheminformatics functions, such as structure representation and searching. Thus, it is possible to develop new Cheminformatics software quickly without having to reinvent the wheel". The Chemistry Development Kit (CDK)[7], is a widely-used open-source Java toolkit for Cheminformatics. According to the website, it has over 50 developers worldwide. It offers a wide range of functionality including 2D structure input, representation and depiction; file and linear representation conversion; 3D rendering; virtual screening based on simple descriptors and an interface to the R statistics package; simple 3D model building and alignment; substructure searching; NMR[8] prediction; and

structure generation. There is also an interface to the BioJava bioinformatics toolkit. Open Babel[9] is a "toolbox" of ready-made programs and a Cheminformatics C/C++ toolkit (with wrappers for other languages), based on an earlier version of what is now the commercial OEChem toolkit. The ready-made programs include structure format conversion, conformer generation, energy calculation, minimization, and various other functions. The toolkit covers a variety of Cheminformatics 2D and 3D functionality. The Chemistry Descriptors Library (CDL)[10] is a C++ library that provides a wide variety of Cheminformatics functionality, including structure representation and conversion, descriptor and fingerprint generation, pKa prediction and synthetic accessibility estimation.

II. Resultand Discussion

Cheminformatics is the designing tool of informatics methods to solve chemical problems. It has developed over the last 40 years to a mature discipline that has applications in many areas of chemistry. It is an important scientific discipline that stands on the interface between chemistry, biology and Information Technology. Cheminformatics extends a very broad range of problems and approaches which are often inter-related and sometimes difficult to categorize. As high throughput technologies and combinatorial chemistry continue to advance, informatics techniques will become indispensable in managing and analyzing the exploding volumes of data. By organizing, the data, Cheminformatics will further introduce advancements in chemistry and open new possibilities in the field of drug discovery. There are still many problems that await a solution and therefore many new developments in Cheminformatics are foreseen. We believe that this review will be the defining theme and might help to provide much new advancement in the field of Cheminformatics in coming years. Hopefully, the availability of information related to Cheminformatics will catalyze further advancements and would open new advancements in this field.

References

[1]. Brown, F.K. (1998); Cheminformatics: what is it and how does it impact drug discovery. Ann. Rep. Med. Chem., 33, 375-384.
[2]. Abraham M H, A Ibrahim, A M Zissmos, Y H Zhao, J Comer and D P Reynolds (2002). Applications of Hydrogen Bonding Calculations in Property Based Drug Design. Drug Discovery Today 7:1056–1063.
[3]. Brown, N. (2009) ACM Computing Surveys, 41, 2.
[4]. Veber D F, S R Johnson, H-Y Cheng, B R Smith, K W Ward and K D Kopple (2002). Molecular Properties that Influence the Oral Bioavailability of Drug Candidates. Journal of Medicinal Chemistry 45:2615–2623.
[5]. Warmuth M K, J Liao, G Ratsch, M Mathieson, S Putta and C Lemmen (2003). Active Learning with Support Vector Machines in the Drug Discovery Process. Journal of Chemical Information and Computer Sciences 43:667–673.
[6]. Orange Book: Approved Drug Products with Therapeutic Equivalence Evaluations (http://www.accessdata.fda.gov/scripts/cder/ob/default.cfm)
[7]. http://cdk.sourceforge.net
[8]. Kelley L A, S P Gardner and M J Sutcliffe (1996). An Automated Approach for Clustering an Ensemble of NMR-Derived Protein Structures into Conformationally-Related Subfamilies, Protein Engineering 9:1063–1065.
[9]. http://openbabel.org
[10]. http://cdelib.sourceforge.net/doc/index.html