International Journal of Research in Engineering, Science and Management
Volume-3, Issue-6, June-2020
www.ijresm.com | ISSN (Online): 2581-5792

A Review in the Use of Chemo-Informatics in The Field of Drug Discovery and Medicine

Divyansh Babel*
Student, Disha Delphi Public School, Kota, India
*Corresponding author: divyanshbabel0102@gmail.com

***Abstract*: There are still a lot of diseases that require medicines. The discovery of medicines has always been a very challenging task for the scientists. This not only involves making the medicine, but also involves the testing of it, and then deciding if it is worth going out in the market. There have been lot of developments in the field of medicine in the past one or two decades and we are slowly getting hold of disease-causing virus, bacteria or germs.**

A lot of us assume that the work of scientists, who discover medicines for various diseases, is very simple. We assume that all they have to do is mix chemicals and see which combination works for that particular disease. But, just so that all of us know, decoding the complex anatomy and structure of the biomolecules is so much harder than we can assume. These biomolecules are responsible for causing fatal and chronic diseases like AIDS, Alzheimer's, Autism, Tumor and Cancer, etc. these illnesses are so fatal in nature that you cannot afford to prescribe any medicines that has other side-effects.

Therefore, finding curing medicines, without side-effects, for chronic illnesses like these is a huge task without a shred of doubt. The lifestyle of humans has already changed a lot, from having unhealthy food to living in an unhealthy environment. Alongside these, if we start having medicines with side-effects, it will only make things worse. Therefore, the entire responsibility falls on the medical researchers and scientists.

Chemo-informatics is very helpful in creating drugs with the help of modern drug discovery techniques. This solves the problems that the traditional drug discovery technique faces. This has turned out to be a very revolutionary change, or discovery, in the field of medicine. In this paper, we will try to present and overview of Chemo-informatics and various other aspects attached to it.

***Keywords*: Chemo-informatics, Drug discovery, Machine Learning.**

1. Introduction

Chemo-informatics is very useful for making drugs and medicines for diseases that are quite advanced. Medicines that are created with traditional drug discovery technique are proving to be not so effective with recent illnesses. Chemo-informatics is also called Chemical Information or Chemical Informatics, and as it literally means, it is the information about various chemicals [1]. For most of the time, Chemo-informatics has been considered to be the interface between a lot of other discipline like biology, chemistry, physics, mathematics, biochemistry, statistics, etc.

The main aim of Chemo-informatics is to study and analyze the chemical information that can represented in 2D and 3D structures, and later used to create various developments in the field of medicine and drugs [2]. Chemo-informatics is majorly used in industrial sectors like agrochemicals, food and beverages, pharmaceuticals, etc.

Chemo-informatics may seem like a heavy and difficult word, but it is simply an umbrella term that researchers use for the creation, design, management, retrieval, analysis, visualization, and use of the information about chemicals [3]. A lot of times, Chemo-informatics is confused with bioinformatics as both deals with chemical compounds in general. But the difference here is that, Chemo-informatics is the analysis and use of small chemical compounds, whereas bioinformatics deals with genes, proteins and other chemical compounds that are of a larger size [4], [5]. Chemo-informatics and bioinformatics work hand in hand when dealing with and analyzing the biomolecular process. Biomolecular processes involve the working of proteins, what they look like, working of enzymes, biochemical reactions, etc.

2. Virtual Chemical Libraries

Virtual chemical libraries are where huge databases related to chemicals compounds are molecules are kept for analysis and development of drugs. They contain imaginary chemical compounds and structures, which are not proven to exist in our natural habitat, but can be created by man. These are a guide for medical researchers towards functions of chemicals that do not exist yet [3], [6]. These libraries also contain methods of synthesis of chemicals and can predict the stability of different objects that undergo a chemical pathway.

Virtual libraries are of huge help to medical researchers and to scientists working in other dimensions of this very field. It takes help of various chemical and physical principles to identify and select the chemical substance that is best suited for a specific reaction or task, from enormous libraries of virtual chemical compounds and substances. They shortlisted molecules and compounds are then taken to the WET lab, where the best candidate is chosen through various tests and experiments [7].

3. Importance of Chemo-informatics

Chemo-informatics has a maintained database from where it accesses and uses the chemical information of millions of chemical compounds for discovery of medicines and other functions too. At present, the database contains detailed information about 4.5 million chemical compounds that have been found till now, and the list doesn't stop here. Thousands of new chemical compounds may get added to the list every year.

In recent times, the discipline of chemistry needs new and advanced techniques to know how to collaborate with bioinformatics to extract data and understand the complex relationship between the structures of various chemical compounds and various biological activities. There are three main aspects of Chemo-informatics:

1) Information acquisition: Information acquisition is the process of generation and then collection of data, either through experiments, or with the help of theory.
2) Information Management: Information Management is the process of storing and retrieving data whenever required.
3) Information use: Information use is the process of using the information that has been acquired and retrieved, like Data Analysis, Correlation, and applying these to the problems available.

4. Applications of Chemo-informatics

Chemo-informatics is the application of the information and technology of chemicals to help chemists and medical researchers in knowing about more chemical compounds, and carry out investigations and researches to develop newer chemical compounds and processes. The primary modules of Chemo-informatics consist of Computer-assisted Synthesis design, Structure representation and Chemometrics [5].

CASD, which stands for Computer-Aided Synthesis Design, is used where Artificial Intelligence is involved. The CASD technique is used in the fields of pharmaceuticals, food and beverage industry, textile industry, and the agro industry.

Chemical representation of different kinds that are machine readable, play an important role to form chemical databases, where researchers store information about various chemicals and analyze and manipulate them in various ways. Medical researchers can design the chemical compounds in a linear manner, in 2D or a 3D format. SMILES, that stands for Simplified Molecular Input Line Entry Specification, is a format for designing a linear chemical compound, as is widely used by chemists and medical researchers [8].

The Reaction Representation is useful in understanding the basic chemical models, analyze chemical reactivity and derive information from the reactions. Structure Searching consists of analysis of features on chemical compounds like bond orders, rings and aromaticity. Its function is to search chemical compounds and determine their entire structure, substructure, the similarity in various structures, and diversity of compounds [9]. The Structure descriptor is used to differentiate between the physical, chemical, and biological characteristics of various chemical compounds and the relationship between two chemical compound structures.

Chemo-informatics fulfils a range of functions and is used is a lot of fields. The knowledge of Chemo-informatics is very rich and one who knows it well can benefit hugely from it. Here are some important applications of Chemo-informatics in various disciplines and professions.

1) It stores data that is created from various experiments or from the simulation of molecules. It is also used to retrieve information about molecules from the Chemo-informatics database.
2) It predicts characteristics of chemical compounds, such as their chemical, physical and biological features.
3) Chemical databases help researchers find the structures and substructures of chemical compounds.
4) It facilitates the interaction between two macromolecules which is called docking.
5) It is used to discover new and effective drugs.

5. Tools that Chemo-informatics Uses

Chemo-informatics has led to the discovery and creation of computer tools that will help researchers to carry out the functions and applications of Chemo-informatics smoothly. We will discuss about a few tools in the following section of this article:

1) ISIS-Draw is used to draw chemical structures and compounds. You can use it in Windows, and this tool is published by the MDL information systems.
2) ChemDraw is our second tool which is developed by the very prestigious company CambridgeSoft. ChemDraw is capable of drawing 3D chemical compounds and molecules.
3) ChemWindow is another tool to draw the structures of chemicals and it comes with a feature of multiple templates. The researchers can create their templates and open them in the dialogue box later.
4) ChemSketch is a good tool as it has predefined templates for drawing chemical structures. Thus, it is a user-friendly tool and cuts off on the time you spend on designing templates.
5) ChemReader is a tool that converts digital raster images of chemical compounds into basic chemical file formats that researchers can analyze and using other Chemo-informatics software as well.
6) ChemMine is another tool that is used by researchers use for analysing small molecules of chemicals. It acts as a mediator between Chemo-informatics and data mining tools for making different analyzes in the field of drug discovery and medicine making.
7) Symyx Draw is a free chemical structure drawing program that researchers use to draw virtual diagrams of chemical compounds and molecules. You can draw the

diagrams manually, or you can also use the IUPAC names present in the program itself. It also allows you to search through databases.

8) JME Molecule Editor is a free drawing software, provided by Java. You can draw and edit diagrams of molecules and also create queries regarding the structure of molecules and their substructure as well.
9) PLSR stands for Partial Least Squares Regression and it is a chemometrics tool that combines your work on two matrixes and analyzes data in the presence of even incomplete variables.

There are a lot of other Chemo-informatics tools that are used to draw and create liner, 2D or 3D images of molecules and chemical structures. Some of them are CAS Draw, DIVA, MOLCAS, WHATIF, XLOGP, ChemTK Lite, CLOGP, Biosoft, and many more [6].

6. What Role Does Chemo-informatics Play in Discovery of Modern Drugs?

The Chemical developers of recent times have invested a lot of time and knowledge on the discovery and creation of new and modern drugs, that will efficiently treat modern and chronic diseases like the AIDS, Alzheimer's, Autism, etc. The developments in chemical compounds that have taken place for the purpose of drug discovery and are leading to a generation of chemical data, also called an information explosion [10], [11]. This has led to researchers being more mindful about collecting, organizing, and analyzing the data about chemicals, and discovers molecules that can be developed into new drugs.

The collaboration between chemistry and information began as early as mid-1970s, and worked effectively in the field of predicting the structure of protein, Fourier transformation of X-ray crystallography, about enzymes and chemical kinetics, analyze different kinds of spectroscopy information and the binding of chemical compounds. In the early 1980s, the only tool scientists used to solve chemical problem was the computer. For example, when you store macromolecular data, it results in the formation of a Protein Data Base [12], [13]. The need in modification and refinement of these solutions arouse, paving the way for more advanced tools to be discovered and worked with.

The traditional technique of drug discovery begins when there is a disease. Firstly, a target is decided, and then medical researchers decide which chemical molecule to use in order to treat that particular disease. Then there is an identification and optimization of the lead. Once all these are dealt with, a Pre-clinical test takes place to check the efficiency of the drug. The identification of the target and molecule synthesis takes a lot of time and is extremely expensive. These functions are carried out in the WET lab [10]. This is where the researchers use the Chemo-informatics technology in the process of the discovery of the drug. The first step is the trial on humans, the authorities them approve the drug and it is then delivered in and around the region. The process of development and delivery of a drug to the market takes almost 10-15 years, which is a long time.

7. Pre-Drug Discovery Process

Before the actual process of discovering a drug starts, the researchers understand the disease that they are trying to treat, how they can alter genes to make the treatment easier, how one protein will reach with the other inside a living body, how the cells that have been hit by the disease change the nature of tissues and how the disease affects the health of the patient.

8. Modern Drug Discovery Process

The drug discovery process has four important processes that are, identification of the target, identification of the lead, optimization of the lead and the pre-clinical trial.

A. Identification of the target

Chemo-informatics is useful is identifying a target molecule that will be further used to treat a certain illness. It can either be a gene or a protein strand that is a potential cure for the disease. The identified target is separated from a human body, crystallized, and bound properly. There are different approaches that are used to productively use the molecule in our bodies [14], [15]. After the target is identified, the next step is to verify whether the change in the target will give the expected result or not.

B. Identification of the Lead

After you have selected your target, you have to go to the "Hit-to-lead" phase. In this phase, you check and see which molecule binds better with the others and works more effectively. Techniques of similarity and diversity are applied to identify the lead.

C. Optimization of the lead

This process involves the identification of the drug molecule from the leads that have been identified. The aim is to refine the structure of chemicals so that the drug characteristic of the chemical is improved. Several techniques of docking are applied to improve the structure of the leads [16]. A lot of techniques are used in the process of lead identification and optimization of the lead, and some of the processes are, Virtual Screening, Molecular database, Data Mining, High-Throughput screening (HTS), QSA R, and many more.

D. Pre-clinical trial

The Pre-clinical stage is a very essential point to verify whether a particular compound can be turned into a potential drug or not, and also whether they have any side-effects or not. Toxicity tests are made to check whether the drug is safe to use or not, and pharmacokinetics testing is done to give report on how a drug is made, absorbed and distributed around the region [17]. Pre-clinical tests can be dome either on animals, or also without them.

9. Development Process

The development process is an important phase in the life cycle of discovering a new medicine. This stage has three important phases that are clinical trial, approval from the authority, and drug in the market.

A. The Clinical Trial

Clinical trial is the first step towards finding a safe and fast way of discovering a treatment for a particular disease. Patients that are suffering from specific diseases undergo the clinical trial and their data will be collected for further evaluation. There are a lot of ways to perform clinical trials[18].

B. Approval from the authority

Based on the rules and regulations formulated for the development of drug in a country, the authorities check whether the drug is safe and other aspects before giving it their acceptance. In India, this body is the Central Drugs Standard Control Organization (CDSCO), and other institutions in different countries.

C. Drug in the market

After the authorities have verified the drug in all the aspects, the drug is ready to get launched in the market. The drug formulation bodies approve the medicine for sales and marketing.

10. Conclusion

The average life expectancy of a human being is decreasing day by day, mainly due to our poor lifestyle that leads to us suffering from chronic diseases. When the amount of diseases and sufferers increase, the medical researchers start looking for a cure to it. For that, they need to understand the structure and function of chemical compounds, and Chemo-informatics helps us to do so. If we do not have a proper database to analyze and create compounds, the discovery of drugs and medicine will become really difficult.

Chemo-informatics has come as a revolution in the world of medicine as it not only helps in the development, creation, trial and distribution of drugs, but also spreads its branch biotechnology, physics and other disciplines. Researchers are trying to develop more advanced technology in the field of drugs and medicine. This paper has primarily focused on Chemo-informatics, its application, the tools it uses, and the stages of drug discovery. Chemo-informatics has been really helpful in making drugs without any side-effects.

We hope this paper was able to give you a detailed overview of Chemo-informatics and the various aspects related to it, and we also hope that this has been useful to you in one or more ways.

References

[1] M. Karthikeyan and R. Vyas, Practical chemo-informatics. 2014.
[2] A. R. Leach and V. J. Gillet, An introduction to chemo-informatics. 2007.
[3] Chemo-informatics Approaches to Virtual Screening. 2008.
[4] T. Engel, “Basic overview of chemo-informatics,” Journal of Chemical Information and Modeling. 2006,
[5] W. L. Chen, “Chemo-informatics: Past, Present, and Future.,” Chem Inform, 2007
[6] N. Brown, “Chemo-informatics—An Introduction for Computer Scientists,” ACM Comput. Surv., 2009
[7] P. Willett, “A bibliometric analysis of the literature of chemo-informatics,” Aslib Proc., 2008.
[8] J. Gasteiger, “Chemo-informatics: Achievements and challenges, a personal view,” Molecules. 2016.
[9] P. Willett, “Chemo-informatics - Similarity and diversity in chemical libraries,” Current Opinion in Biotechnology. 2000.
[10] Y. C. Lo, S. E. Rensi, W. Torng, and R. B. Altman, “Machine learning in chemo-informatics and drug discovery,” Drug Discovery Today. 2018.
[11] P. S. Dittrich and A. Manz, “Lab-on-a-chip: Microfluidics in drug discovery,” Nature Reviews Drug Discovery. 2006.
[12] [T. I. Oprea, Chemo-informatics in Drug Discovery. 2005.
[13] J. P. Hughes, S. S. Rees, S. B. Kalindjian, and K. L. Philpott, “Principles of early drug discovery,” British Journal of Pharmacology. 2011.
[14] T. Cheng, Q. Li, Z. Zhou, Y. Wang, and S. H. Bryant, “Structure-based virtual screening for drug discovery: A problem-centric review,” AAPS Journal. 2012.
[15] N. Chandra, “Drug discovery,” in Systems Biology of Tuberculosis, 2013.
[16] P. Hop, B. Allgood, and J. Yu, “Geometric Deep Learning Autonomously Learns Chemical Features That Outperform Those Engineered by Domain Experts,” Molecular Pharmaceutics, 2017.
[17] X. Q. S. Xie, “Exploiting PubChem for virtual screening,” Expert Opinion on Drug Discovery. 2010.
[18] G. Sliwoski, S. Kothiwale, J. Meiler, and E. W. Lowe, “Computational methods in drug discovery,” Pharmacological Reviews. 2014.