ISSN: 0975-766X
Review Article

Available Online through
www.ijptonline.com

OVERVIEW OF DIFFERENT SOFTWARE USED IN DRUG DESIGN AND DRUG DISCOVERY: A REVIEW

Suman Bala*, Goldie Uppal, Sunil Kamboj and Minaxi Saini
M. M. College of Pharmacy, Maharishi Markandeshwar University, Mullana, Ambala, Haryana.
E-mail: *sumankmj7@gmail.com*

Received on 17-08-2010 ***Accepted on 28-08-2010***

ABSTRACT

The new drug discovery is found to be a time consuming and costly process. Recently, a trend towards the use of in-silico computational chemistry and molecular modeling for computer aided drug design has gained significant momentum. This review is an investigation of the applications of different software in drug design. The modern drug discovery process is steadily becoming more information driven. Structural, physicochemical and ADME information about property profiles of reference ligands, along with structural information of their target proteins, have been extremely useful for early-stage drug discovery. A number of examples have recently been reported for the successful applications of structure based drug design to the discovery of compounds with a potential to become more useful therapeutic agents. Among the reviewed software programs are applications programmed in Grid computing, window based general PBPK/PD modeling software, PKUDDS for structure based drug design, APIS, JAVA, Perl and Python, CADD as well as software including software libraries. These all programs are useful for cheminformatics approaches to drug design and discovery including QSAR studies, energy minimization as well as docking studies in drug design. Furthermore this review explains options for using different computer modeling software programs in drug discovery and design.

Keywords: Drug design, drug discovery, molecular modeling software.

INTRODUCTION

The greatest source of inefficiency in traditional discovery and development arises from the high percentage of evaluated compounds that have a low probability of ultimate success. To bring a new drug to the market is very costly, with the current price tag approximating US$ 800 million, according to recent studies. This is why various pharmaceutical companies are seeking various ways to reduce the cost associated with drug design. Now a day's various computer based software such as combinatorial and HTS (High-throughput screening) adopted by pharmaceutical companies in order to save time and money. In HTS large libraries of compounds are screened against drug target to identify a lead compound that can leads to a particular outcome. Although these methods are costly and most of the time they are not able to represents the basic needs of many biological systems. Although it increases the rate at which lead compounds are identified [1].

Several commercial tools to support such activities are available in the market. However, these programs usually run on UNIX workstations and their uses require specialist knowledge and experts in the fields of cheminformatics and molecular modeling and are therefore used mainly by molecular modelers. However, the success of a given project is dependent on synthetic bench-chemists becoming much more involved in the direct cheminformatics work because their project-specific knowledge is important for the projects to progress well. De novo ligand design methods are computational methods for designing molecules that complement a receptor or binding site structurally and energetically. Successful de novo design results in proposed ligand structures that have high binding affinity for their desired binding sites. De novo design has been most successful where biological and experimental knowledge of the ligands and substrates exists. Two major approaches have been applied to the development of de novo ligand design software: molecular fragment approaches and sequential growth approaches (fig. 1).

The molecular fragment approaches dock molecular fragments to determine various energetically favorable positions on the active (binding) site that are then "linked" together. The first step is to identify key

locations in the binding pocket and then bind small seeds or fragments to these locations. Once the seed fragments/functional groups are positioned, the next step is to link them together with scaffolds. Current de novo ligand design methodologies suffer from one or more of three major deficiencies that have severely limited their use and acceptance in drug discovery programs. The first and most important deficiency is the fact that a large number of the generated structures are synthetically unfeasible.

This is particularly evident for the fragment methods where in many cases it is chemically infeasible to bridge the bound functional groups and join all fragments in their most favorable locations [2]. The second major deficiency arises from the commonly observed differences in experimental and calculated binding affinities.

De novo design methods utilize a scoring function to evaluate each step of the process. This scoring function is really a calculated measure of receptor-ligand binding affinity. Unfortunately, available scoring functions are limited in their abilities to accurately predict experimentally determined binding affinities and activities.

The third deficiency arises from the large combinatorial problem of quickly and efficiently searching diversity space for good solutions (e.g., structures with reasonable binding affinity).

As an attractive alternative, in-silico methods show promise in identifying new lead compounds faster and at fraction of the cost of combinatorial approaches and other software. The addition of computer aided drug design technologies to the R&D approaches of company, could lead to a reduction in the cost of drug design and development by up to 50%. These in silico approaches may include:

a) Docking studies: where a ligand is studied during the binding to particular target protein.

b) Cheminformatics: where activity and structure are correlated by structural means.

c) Bioinformatics: where drug target obtained from genomic data. Use of these approaches has led to the discovery of indinavir and HIV-protease inhibitors and identification of haloperidol as a lead compound in a structure-based design study of non-peptide inhibitor of HIV [3, 4, 5].

The software used may originate from different sources. It may include commercial companies, academic institute's open-source software and in-house development. Each of these softwares has advantages and disadvantages and their selection depends upon institutions that use the software [6].

GaussDal: An open source database management system for quantum chemical computations

Open source software is known as GaussDal. This is used for management of results obtained from quantum chemical computations. Open source software refers to any program that has the source available for modification as user or for other developers. They are mostly developed in public collaboration and are made available freely, although 'free' can sometimes refers only to academic institutes where as commercial enterprises required to pay a fee.

Recently Campbell McInnes [7] discussed the use of open source software in the strategies used in drug discovery. The open source software has large number of advantages which make it suitable and attractive for drug designing, especially to academic scientist. Usually commercially available drug designing software packages come with an expensive license fee and they have to renewable every year. Recently DeLano discussed the use and advantages of open source software. It is easy to use open source and free drug discovery software as they include the ability to download a program directly and immediately from internet, no license fees and flexibility at lower cost with option to customize the software for a particular project. These features make open source software most attractive and useful software [8].

Although open source software are very attractive and useful but they too have some disadvantages. Most of the times they lack the proper information about the use which may be inconvenient for new user. Also in the field of chemistry it is difficult to convince experts to contribute their spare time to those open source software projects. Therefore new contributors to open source software can be increased by graduating students in computational or computer sciences as a part of their work. Most of the times the installation of these programs is not easy and need to be compiled by programming language-specific compiler such as C^{++}, FORTRAN,

JAVA or run from the command line and most of these programs are written for Linux or SGI platform [9]. This is the reason why most of cheminformatics and bioinformatics software have poorly written graphical user interfaces that are not user friendly. This results in a bench scientist spending more time trying to install and program it. This is the reason why in-silico methods are used in first place – that is to save the time and thereby saving money. Further some software is set to run by command line by imply typing to a command line editor. These interactions are preferred by the users because it speeds up the process [10, 11, 12].

An open source software system called GaussDal for management of results from quantum chemical computations is presented. Chemical data contained in output files from different quantum chemical programs are automatically extracted and incorporated into a relational database (PostgreSQL). The Structural Query Language (SQL) is used to extract combinations of chemical properties (e.g., molecules, orbitals, thermo-chemical properties, basis sets etc.) into data tables for further data analysis, processing and visualization (fig. 2). This type of data management is particularly suited for projects involving a large number of molecules. In the current version of GaussDal, parsers for Gaussian and Dalton output files are supported, however future versions may also include parsers for other quantum chemical programs. Catalog identifier for gaussDal program is ADVT, Program summary [13], Operating system under which the system can be tested is Linux and Programming language used is Python [14].

There are existing software implementations available that try to solve some of the management problems from Daylight Chemical Information Systems [15, 16]. However, all these systems are commercial software and subjected to strict licensing requirements. There are two open source project related to GaussDal: OpenChem Workbench [17] and GaussSum [18]. The OpenChem Work- Bench software appears to contain some of the desired properties of GaussDal, however inspection of the web site indicates that the project has had little activity in the last years and has not progressed to a sufficiently mature state. The GaussSum program is basically a parser for Gaussian output files which does not contain a database for handling the results which

is so central to GaussDal. Within the field of molecular mechanics simulations, there is software which follows several of the ideas used for GaussDal. This is the BioSimGrid [19] database which allows contributors in a GRID computation network to share results from their computations. This is a centralized approach to data management and is in contrast to GaussDal which can run on the local machine of each user. This is to ensure security and flexibility in the use of the program. In addition, GaussDal is also able to work over the Internet in a centralized mode [20].

The World Wide Web as Graphical User Interface to Molecular Modeling and Structural Based Drug Design

World Wide Web (WWW or Web) has been employed to provide access to computational chemistry software and protein structure data by program macros. We show how the combination of web technology and macros can automate both the running of chemistry software and the execution of complex operations on protein structure. Historically the use of molecular graphics, molecular modeling and chemical information software has been limited to the experienced persons with the background concepts, specific user interfaces and overall computing environment. Furthermore the use of most of software packages has been labor intensive with considerable time and efforts often been required for computer based approaches to molecular problems. One solution to this problem has been the use of program macros also known as scripts. These are program consisting of program for running and controlling software packages. Program macros can be used as end user tool so that the user at any level may safely and correctly apply the most advanced features of sophisticated modeling software.

The rise in popularity of World Wide Web has resulted in revolution in the way computers are used and the way information is accessed. The current version of the system supports the molecular visualization packages GRASP, RASMOL, MOLVIEWER-OGL and INSIGHTS95 and the ligand design tool GRID.

The modeling methods described herein focuses on the interpretation of protein structure information and its use in receptor based drug design. We have sought to obtain maximum benefit from specialized protein structure visualization software, particularly GRASP, which is particularly suited to mapping chemical properties on to molecular surfaces, RASMOL a widely used package for protein structure viewing, MOLVIEWER-OLG which we have found to be highlighting individual residues in protein ligand complexes and INSIGHTS95 a comprehensive molecular modeling method package that is used for superimposition of multiple structures. Neil R. Taylor and Ryan Smith put forwards a version of interface which provides easy access to ligand design tool GRID [21]. They attempted to create a system that offers the user simple, intuitive set of options that rapidly delivers meaningful results. More of their studies are a set of macros that perform specific operations on protein structure. The use of macros is that they enable complex operations that normally require many steps to be performed with single command. Each macro is written in command language that is specific to each modeling package. In some cases macro may be only two commands e.g. in case to highlight the two amino acid residue in active site of a protein, while some other cases it may be of hundred commands, e.g., reading a protein-ligand complex and in displaying the molecular surface of the protein in vicinity of the active site.

A Web browser can be used as Graphical user interface via the common gateway interface and controlled by a scripting language (e.g. TCL: Tool Command Language). An HTML document that is used to pass parameter to a CGI program is known as a form. The basic operation of the macros is their use in rational drug design, in the protein structure data and receptor based drug design. There are many advantages of using Web as a GUI to program macros and computational chemistry software. These advantages include following:

(a) The process of viewing structure and their properties is fast.

(b) Access to the protein data and software tools is opened to a wider audience.

(c) Data are managed efficiently owing to reduced duplication of files.

(d) The time required to gain experience with software is reduced thereby lowering training costs.

(e) Many changes associated with running modified upgraded software packages are hidden from inexperienced end user.

(f) The implementation of the state of the art techniques is usually straight forward.

(g) Documentation and manual pages can be easily embedded into the GUI (Graphical User Interface) in the form of hyperlinks.

Despite of so many advantages they have some disadvantages too, e.g. machine dependency. That is the software packages implemented run only on Silicon Graphics work-stations. But this is not a serious issue as work stations are standard for molecular modeling. Another limitation is that the techniques reported are suited to local internet systems and not the whole World Wide Web. Despite of these minor drawbacks the tremendous gain described here ensure the future success of using the Web as an interface to the molecular modeling software [21].

APIS: Application Programming Interface Software

This software used for model identification, simulation and dosage regimen calculations in clinical and experimental pharmacokinetics. Basically pharmacokinetics consists of analyzing the kinetic profile of a given drug during its absorption, distribution, metabolism and excretion process in animal or human body. The drug dynamic behavior can be monitored through the concentration-time curve, obtained by sampling biological fluids and after, by assaying drug concentrations. The basic goal is to optimize drug use on patient by computing adapted dosage regimens. The wide inter-individual variability in the pharmacokinetic process, a given dose can be efficacious without side effects to one patient but toxic to another patient. Thus to design adequate dosage regimens for individual patients, it is important to obtain information on individual pharmacokinetic i.e. to be able to relate concentration to doses [22].

In the past a great attention is given to mathematical modeling of drug distribution in human body in order to perform pharmacokinetic process identification and to compute appropriate dosage regimen. Computer may be used to perform individual pharmacokinetic modeling. But the difficulty arises that the amount of kinetic information obtained from patient is restricted by clinical constraints such as ethics, time-duration and cost of analysis. Further to reduce the number of samples required to produce a mathematical a Bayesian estimation criterion may be used. A Bayesian approach require knowledge of the probability density, p(x), of unknown pharmacokinetic parameters, x, and of conditional probability density function p(y/x), called likelihood function and involve the statical distribution of individual observations, y. The main difficulty for Bayesian estimator is the deviation of p(x). The best conditions under which p(x) and p(y/x) appear in Bayesian criterion is critical for a reliable estimation of pharmacokinetic parameters and therefore a successful dosage regimen adjustment [23].

APIS is a software package based on mathematical modeling which provides a reliable approach in optimizing the drug therapy. It was designed to assist clinician in interpreting blood drug levels so that drug therapy may be better and more cost effective. This method can be used to describe, predict and control the kinetic behavior of the drug. This software includes the principle of Bayesian procedures i.e. one can use all patient available information to determine patient specific parameter estimates. These estimates can then be used to design an optimal and individualized drug regimen. APIS is an attractive and useful tool for clinical and experimental pharmacokinetic. APIS may be used on any IBM compatible computer using Microsoft-windows environment. The software is menu driven to provide a very user friendly tool or analyzing pharmacokinetic data and for designing dosage regimen [24].

A new molecular simulation software package – Peking University Drug Design System (PKUDDS) for structure-based drug design

It is a comprehensive molecular simulation program package, the Peking University Drug Design System (PKUDDS), which runs on personal computers. PKUDDS has been developed mainly for computer-aided drug design using the methods of two-dimensional quantitative structure activity relationships, three-dimensional quantitative structure activity relationships, molecular docking and database screening. Tingjun Hou and Xiaojie Xu presented an overview of its functionality, especially of methods developed. They studied PKUDDS uses genetic algorithms in molecular docking, conformational analysis, and quantitative structural-activity relationships as the most useful optimization technique. A user-friendly graphical interface provides easy access to many functions of PKUDDS [25].

The development of new drugs is a lengthy and expensive process. The first step is to find potential lead compounds with desired biological activity. Computer-aided drug design (CADD) techniques can increase the pool of interesting structures that can be evaluated. The rapid increase in computer speed, memory and the decreased cost of personal computers and workstations have brought significant computational resources within the reach of most researchers. Inexpensive computer graphics programs offer improved methods of organizing and visualizing molecular information. The fundamental assumption of most CADD procedures is that the key biological event, at the molecular level, is the recognition and non-covalent binding of small molecules to specific sites on target biological macromolecules (receptors). Generally, CADD procedures can be divided into two categories: ligand structure-based methods; and receptor structure based methods.

The Peking University Drug Design System (PKUDDS) provide a convenient method of accessing methods for drug discovery (fig. 3). For ease of maintenance and future extensions the system was developed on personal computers to function with Windows 95, Windows 98, or Windows NT operating systems.

PKUDDS provides a powerful simulation capability and a friendly graphical user interface. The computational code and graphical user interface are written with visual C^{++}.

Ligand structure based methods, including quantitative structure activity relationship (QSAR) methodologies and pharmacophore searches, share the goal of predicting biological activities and devising common pharmacophore models from the physicochemical properties of ligand structures. In most cases, the structure of the receptor is unavailable and the only way to study the SAR and pharmacophores is by using ligand structures. Receptor-based methods including molecular docking and de NOVO design. They seek to find lead compounds by modeling the molecular details of drug action or receptor ligand interactions. With recent developments in X-ray and NMR techniques, many protein structures have been solved, providing better information about receptor ligand interactions [26].

With a receptor model in hand, the next step is normally to build or find potential ligands that will fit into the active site model. The key to this step is using 3D information to find or build complementary structures. A crystal structure of the receptor of interest having a ligand bound in the active site offers an ideal place to begin, providing valuable information about the location of important contacts and the conformation of the bound ligand. The goal of designing PKUDDS was to develop an integrated system on a personal computer that contains all of the functionality necessary for structure-based drug design. Major modules include 2D-QSAR, enhanced CoMFA (Comparative Molecular Field Analysis) and molecular docking. The source code of the modules in PKUDDS is mainly written in C and C^{++} languages. By using PKUDDS all calculations were performed on a personal computer. Source code and corresponding parameter files in this study, including the 2D-QSAR based on GA (genetic algorithm), the soft-docking procedure and the conformational analysis based GA, can be obtained upon request. The structural analysis of ligands and receptors by PKUDDS provides useful information for drug design. Moreover it can also be used to study important processes in the design of new functional materials [27].

Various researchers developed new methods for measurement and evaluation of various physicochemical properties of drug like compounds. For example: Pablo R. Duchowicz and et al developed new QSPR (Quantitative Structure Property Relationship) study for the prediction of aqueous solubility of drug like compounds [28].

Web based Cheminformatics tool: Drug Design, Internet, World Wide Web and Java

Web-based molecular processing tools installed on corporate Internets bring easy to use cheminformatics and molecular modeling capabilities directly to the desks of synthetic chemists, giving them comfortable access to data and their visualization and analysis, considerably improving efficiency of the drug design and development process. User friendly tools that use a standard Web browser as an interface allow users access to a broad range of expert molecular processing tools and techniques, without the need for extensive expertise in their use. Computational methods are becoming more involved in the modern medicinal chemist's work. This is mainly because of the exponential growth of the amount of data that need to be processed and analyzed in the modern drug discovery process. Typical tasks involved in this process include fast access to structural and bioactivity data from the corporate database, calculation of molecular properties and structure-activity correlations for large datasets, analysis of molecular diversity and the design of combinatorial libraries [29, 30].

Several commercial tools to support such activities are available on the market. However, these programs usually run on UNIX workstations. Their use requires specialist knowledge and expertise in the fields of cheminformatics and molecular modeling. Therefore these are used mainly by molecular modelers. However, the success of a given project is dependent on synthetic bench-chemists becoming much more involved in the direct cheminformatics work because their project-specific knowledge is important for the projects to progress well. Although most chemists are interested in doing so, they are discouraged by factors

such as the necessity to master the complicated interface and command set of commercial applications or to remember UNIX commands [31, 32].

However, there is a solution to this dilemma - the World Wide Web. The enormous, and still increasing, popularity of Web technology is due to its ease of use and high degree of interactivity. In addition, various emerging technologies, such as Java, sophisticated Web scripting, VRML (Virtual Reality Modeling Language) or chemical markup language have added new functionality to the Web and made it a true dynamic environment that is ideal for the development of user- friendly cheminformatics applications. Free cheminformatics tools on the Internet can be useful in the drug design and discovery process is available on the Internet for free. These services are not only offered by academic institutions and non-profit organizations, but also by providers of commercial software. Such free services are important resources for students or academic researchers [33].

Ciba-Geigy (which merged in 1996 with Sandoz to form Novartis) [33, 34] was one of the first large pharmaceutical companies to recognize the advantages of Web technology for processing molecular information. By 1995, its chemists were already able to use a Web-based cheminformatics system through the company Internet. The Novartis Web-based cheminformatics system has been continually updated and currently comprises >20 modules, supporting a broad range of molecular modeling and molecular processing tasks including:

- Calculation of molecular properties.
- Sophisticated toxicology alerting.
- Retrieval of information about various molecules from the company databases (screening results, experimental properties, sample availability).
- Molecular visualization, including also visualization of surface molecular properties.
- Support for diversity analysis and enumeration of combinatorial libraries.

- Drug design based on bioisosteric principle.
- Various helper tools (e.g. file format conversion, molecule depiction and 2D to 3D conversion).

In addition to the general advantages of Web-based tools for processing chemical information, the chemical industry can benefit from several other specific features of Web technology. On a company network comprising various types of computers with different operating systems (e.g. Windows based PCs, Macs, various types of UNIX computers including LINUX servers), the ability to connect all these computers, as well as integrating useful older legacy applications (which usually run on mutually incompatible systems) is important. Maintenance of software on a single Web server is a viable alternative to updating and maintaining applications on several hundreds of desktop PCs. These software are user friendly and do not require any special training. Another benefit of in-house software development is that the time lag between the appearance of a new methodology or algorithm in the literature and the availability of this solution to the users can be kept to a minimum. Pricing and licensing issues are also encouraging large pharmaceutical and agrochemical companies to move more resources towards in-house development of Web-based cheminformatics tools running on a central server, rather than licensing software for hundreds of desktop computers [35].

The advantages of Web technology in the chemical industry is that such systems provide easy access to relevant data, allow visualization, processing and analysis of the data and development of models that help users to understand complex relationships within the data. The Web tools support medicinal chemists in their daily tasks as well as more specialized activities traditionally reserved for molecular modeling specialists, such as the design of targeted combinatorial libraries, bioisosteric design or virtual screening, including sophisticated virtual docking applications.

There is a limit, determined by the complexity of the required user interface. The interface of Web tools must be simple, ideally offering only structure input and a minimum number of parameters to be set. Web tools that

require complex multi-step input, with numerous input parameters and options have no chance of success [36]. Examples of Web-based cheminformatics applications used in the pharmaceutical industry include:

a) The extensible cheminformatics system VERDI - the Vertex Research Database Interface - is available on the Intranet of Vertex Pharmaceuticals. The system helps Vertex scientists to handle large amounts of data and organizes the chemical, biological and intellectual property information into a format that is clear and easily accessible. The Web system comprises a multi-tier client server architecture that simplifies the integration of a variety of commercial, academic and homegrown tools and can be used to perform a variety of functions including database searching, reagent selection and combinatorial library design. A tool called REOS (Rapid Elimination of Swill) provides a simple means for scientists to select screening compounds and reagents.

b) A Web system for combinatorial library design has been developed at Scynexis. The system offers synthetic chemists access to expert tools without the need for extensive expertise in their use. The drug - likeness, ADME and toxicity filters, as well as scoring functions for library diversity and reagent suitability are available.

c) At Ionix Pharmaceuticals, a set of intercommunicating Intranet packages was created to assist Ionix chemists in the design of combinatorial libraries. The core component of the system is a database of available reagents coupled with a combinatorial library design package to create and filter virtual parallel arrays.

d) Celltech Chiroscience has implemented a Web-based tool for hydrogen bond strength prediction. The method is based on the correlation of experimentally observed hydrogen bonding strength with quantum-chemically derived parameters [37].

Python: A Programming Language for Software Integration and Development

Python is an interactive, object-oriented programming language. It provides high level data structures such as list and associative arrays (called dictionaries), dynamic typing dynamic binding, modules, classes, exceptions and automatic memory management etc. It has a remarkably simple and elegant syntax and yet is a

powerful and general purpose programming language. It was designed in 1990 by Guido van Rossum. Like many other scripting languages, it is free, even for commercial purposes and it can be run on practically any modern computer. A Python program is compiled automatically by the interpreter into platform independent byte code that is then interpreted. The Scripps Research Institute running various unmodified components written in Python under Linux, Windows NT, 98, 95, IRIX, SunOS, and OSF. Python is modular by nature. The Python distribution includes a diverse library of standard extensions (written in Python, C, or C^{++}) for operations ranging from string manipulations and Perl like regular expressions, to Graphical User Interface (GUI) generators and including web related utilities, operating system services, debugging and profiling tools, etc. New extension modules can be created to extend the language with new or legacy code. The extension modules, sometimes referred to as "packages" or components, include GADFLY, an SQL database manager written in Python, PIL, the Python imaging library, FNORB, OmniBorker, CORBA compliant Object Request Brokers (ORB) written in Python, Gendoc, (an automated documentation tool) and Numeric Python. The best resource for Python, along with the books that are available, is probably the Python web site. It provides access to code, documentation, packages, articles, mailing lists etc. It is also worth mentioning the recent creation of the biopython.org web site, a collaborative software effort for computational biology and chemistry very much like bio Perl. Finally, besides the C implementation of the Python interpreter, there is also a 100% pure Java implementation called JPython. JPython allows Python use as an interpreted language for programming in the Java world. This interpreter allows initiation of a Java class and Java code can call Python code. The native extensions first need to be made available in the Java world before they become available in JPython [38, 39].

Computer Assisted Drug Development (CADD): An emerging technology for drug design

Computer Assisted Drug Development (CADD) is an emerging technology for accelerating drug development based on the integration of mathematical modeling and simulation. This methodology provides knowledge based decisional tool on alternative development strategies based on the evaluation of potential risks

on drug safety and the definition of experimental design of new trials with expected power and probability of success. Research and development in pharmaceutical industries are under constant pressure to improve operational efficiency and to continuously simplify and automate processes. This pushes toward a re-engineering of the decisional processes suitable to account for the rapid evolution of the information systems and information technologies [40].

Today, a novel technology is available to combine the mathematical modeling and simulations tools in an integrated environment: the computer-assisted drug development (CADD) system. The CADD system jointly with the development of an effective information system suitable to insure connectivity, workflow management and access to comprehensive information repository (Gomeni, 2000) can supply an effective framework for modern drug development. Using these technologies, researchers can now test in a real and virtual environment drug compounds against protein targets, study the PK and PK/PD of optimized lead, study drug effects on isolated organs or animals, design a clinical trial to test alternative assumptions, and even answer some regulatory questions through simulation. The need to make drug development more efficient and informative pushed toward the extensive use of simulation in clinical development over the past decade. However, the computer simulation has been predominantly used in clinical trials of phase II to phase IV while this methodology has been only occasionally applied to drug development from pre-clinical studies up to the first-time-in-man (FTIM) and the proof-of-concept (PoC) studies. An example of the use the CADD technology to integrate the knowledge about the pharmacokinetics and pharmacodynamics of a new CNS compound in different animal species and to build predictive models to design the FTIM and the PoC study was presented by Roberto Gomeni, Massimo Bani, Carla D. Angeli, Mauro Corsi, Alan Bye of GlaxoSmithKline Group [40]. CADD is knowledge based iterative process where the newly collected information are integrated in the existing drug and disease-specific knowledge frame and used to refine and update the knowledge on the drug properties. This process extensively uses mathematical models and simulation tools to describe and predict the

behavior of dynamic system with the objective to better understand the system itself and explore and improve the system characteristic. The ultimate goal of the CADD approach is to provide a decisional tool based on the exploration of alternative development scenarios, the evaluation of potential risks on drug safety, the definition of experimental design of new trials with expected power and probability of success. CADD can be effectively applied in all the phases of a new drug research: from lead optimization up to the post marketing studies [41].

Grid computing and the drug discovery process

Grid computing has emerged as an important new field. It differs from conventional distributed computing in its focus on large-scale resource sharing, innovative applications and in some cases high performance and high throughput orientation. Grid computing enables the development of large scientific applications on an unprecedented scale. Grid-aware applications, also called meta-applications or multi-disciplinary applications and make use of coupled computational resources that are not available at a single site. In this light, Grids let scientists solve larger or new problems by pooling together resources that could not be coupled easily before. Designing and implementing Grid-aware applications often requires interdisciplinary collaborations involving aspects of scientific computing, visualization and data management. It is well known that the programmer's productivity in designing and implementing efficient parallel applications on high-performance computers is a very time consuming task. Grid computing makes the situation worse. Consequently the development of Grid programming environments that would enable programmers to efficiently exploit this technology is an important and hot research issue [42].

The myGrid is proposed as the next-generation Web. It will provide the computational power and data management infrastructure necessary to support the collaboration of people, together with data, tools and computational resources. The scientific process is an example of such a collaborative process where Grid technology can facilitate virtual organizations of people, machines instruments, data and computational resources. Robert Stevens, Robin McEntire, Carole Goble, Mark Greenwood, Jun Zhao, Anil Wipat and Peter

Li use the drug discovery process as an example of a knowledge rich application domain that can be facilitated by technology such as myGrid [43]. Several other Grid projects orientated towards the life sciences are underway, such as the Asia Pacific BioGrid Initiative [44], the North Carolina BioGrid [45], the Canadian BioGrid [46], the EUROGRID project [47] and the Biomedical Informatics Research Network [48]. All these projects have primarily focused on the sharing of computational resources and the large-scale movement of data for simulations, remote instrumentation steerage or high throughput sequence analysis. However, the life sciences require support for a scientific process with more modest computational needs, but with a high level of semantic complexity, of which the drug discovery process provides many examples. In its early development, Grid computing has focused on providing the computational power necessary for solving computationally intensive scientific problems. However, the scientific process in the life sciences is less demanding on computational power but contains a high degree of inherent heterogeneity, semantic and task complexity. The myGrid project has developed a Grid-enabled middleware framework to manage this complexity associated with the scientific process within the bioinformatics domain. The drug discovery process is an example of a complex scientific problem that involves managing vast amounts of information. The technology developed by the myGrid project is applicable for managing many aspects of drug discovery and development by leveraging its technology for data storage, workflow enactment, change event notification, resource discovery and provenance management [49, 50].

Conclusion

Various computer based molecular modeling software has a significant impact on various areas of drug designing and development. Although a number of free and open source software packages such as QSAR molecular modeling descriptors or docking software are available for drug discovery, but sometimes they might be inaccessible to the bench chemist because of either a poorly programmed GUI or insufficient information is available about how to run the software. However irrespective of these disadvantages and difficulties these

software are useful to speed up the process of new drug development. Some software can also be helpful to study the pharmacokinetic and pharmacodynamic parameters of the drugs in human and animal models e.g. APIS. Some software may be used to study the structural activity relationship of different drugs and their analogues e.g. PKUDDS and CADD. Therefore the use of these softwares is useful in drug designing. When more use of such successful software is made, especially in the academic community, many more drug discovery projects will be benefit and programs with added functionality and user friendliness will, as a result, likely become available to assist such endeavors even further.

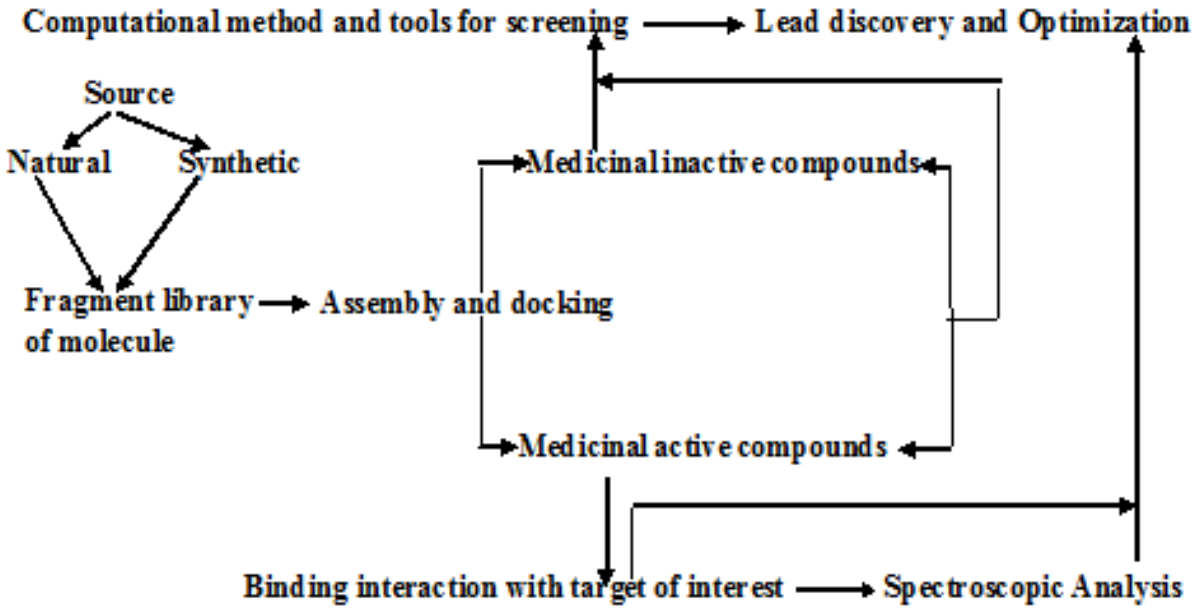

Fig. 1: Basic layout for drug discovery

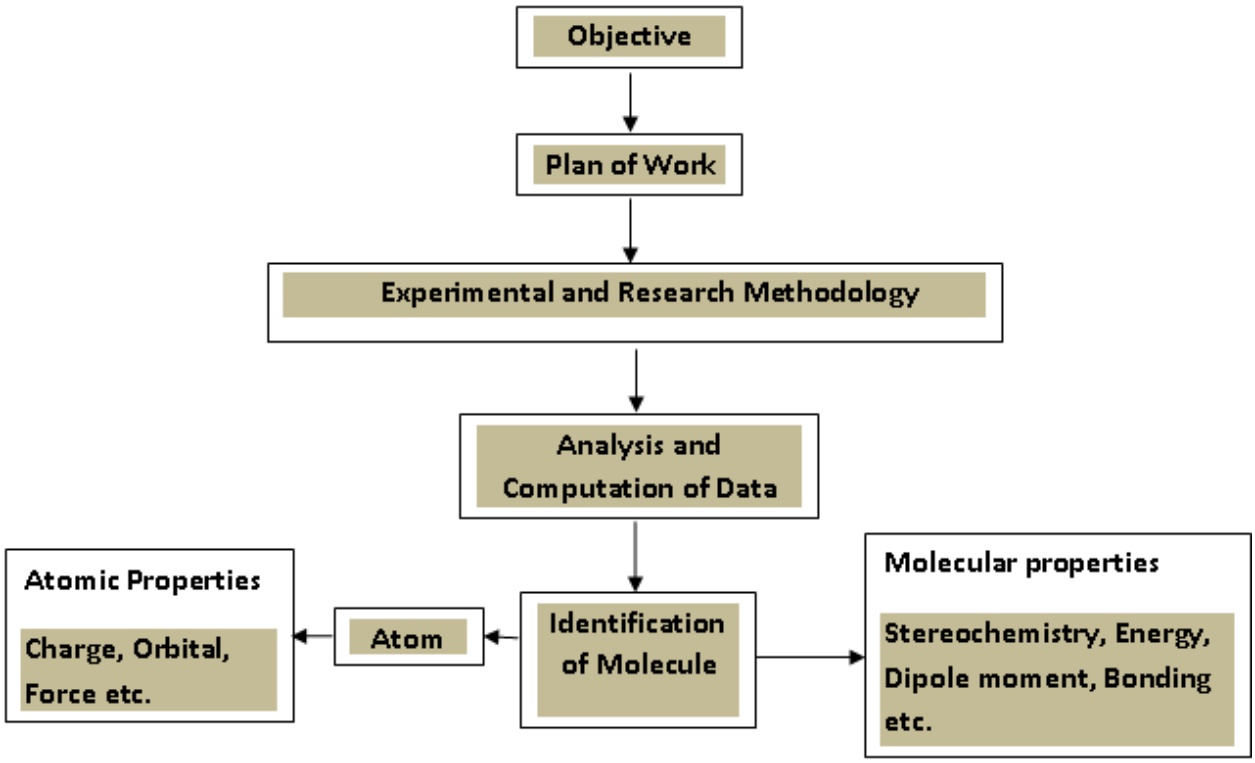

Fig. 2: An overview of the GaussDal database structure for storing molecular properties

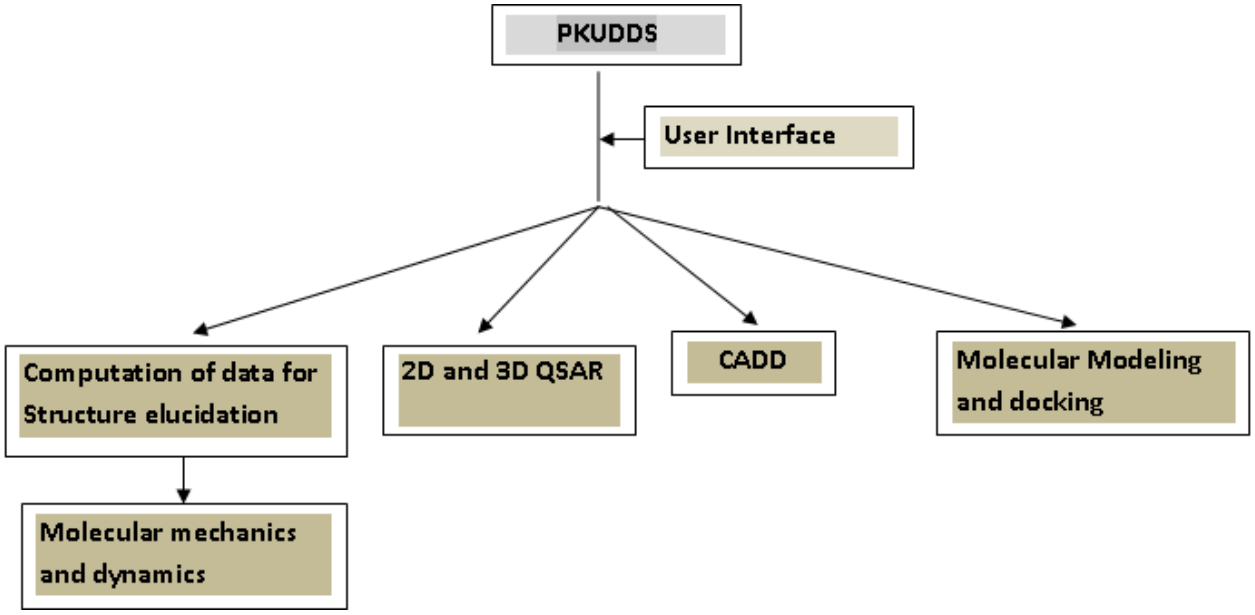

Fig. 3: A flow chart for Peking University Drug Design System (PKUDDS)

References

1. Werner J.Geldenhuys, Kevin E.Gaasch, Mark Watson, David D.Allen and Cornelis J.Vander Schyf; Optimizing the use of open-source software applications in drug discovery; Elsevier DDT.volume 11, February 2006, 127-131.
2. Hecht David and Fogel Gary B.; A Novel In Silico Approach to Drug Discovery via Computational Intelligence. J. Chem. Inf. Model. 2009, 49, 1105–1121.
3. Hou, T.J., Wang, J.M., and Xu, X.J. Applications of genetic algorithms on the structure–activity correlation study of a group of non-nucleoside HIV-1 inhibitors. Chemometr. Intell. Lab. 1999, 1–2, 303–310.
4. Hohman Moses, Gregory Kellan, Chibale Kelly, Smith Peter J. Ekins Sean and Bunin Barry. Novel Web based tool combining chemistry informatics, biology and social networksfor drug discovery. 2009 Elsevier Ltd. 261-270.
5. Wayne C. Guida, Software for structure based drug design, Current Opinion in Structural Biology 1994, 4: 777-781.
6. DiMasi, J.A. et al. The price of innovation: new estimates of drug development costs. J. Health Econ. 2003, 22, 151–185.
7. Campbell McInnes, Virtual screening strategies in drug discovery, Current Opinion in Chemical Biology, 2007, 11:494–502.
8. DeLano, W.L. The case for open-source software in drug discovery. Drug Discov. Today, 2005, 10, 213–217.
9. Product review, JATOON: Java tools for neural networks, Chemometrics and Intelligent Laboratory Systems, 2002, 61, 167– 173.
10. Wolohan P.R. et al. Predicting drug pharmacokinetic properties using molecular interaction fields and SIMCA. J., Comput. Aided Mol. Des. 2003, 17, 65–76.

11. William R. Bush, Software regulation and domain specificity, Information and Software Technology, 2007, 49, 44–54.
12. Curcin, V. et al. Web services in the life sciences. Drug Discov. Today, 2005, 10, 865–871.
13. http://cpc.cs.qub.ac.uk/summaries/ADVT.
14. http://www.molworks.com
15. Bures, M. et al Physical property calculation: a Web browser application, Daylight User Meeting MUG'98, Santa Fe, NM, USA. 1998, 85-89.
16. http://www.daylight.com/
17. http://www.mines.edu/research/ccre/
18. http://gausssum.sourceforge.net
19. http://www.biosimgrid.org/
20. Alsberg Bjorn K, Bjerke Håvard, Navestad Gunn M, GaussDal: An open source database management system for quantum chemical computations, Computer Physics Communications, 2005, 171, 133–153.
21. Taylor Neil R. and Smith Ryan. The World Wide Web as Graphical User Interface to program macros for the Molecular Graphics, Molecular Modeling and Structural Based Drug Design. Journal of Molecular graphics, 1996, 14, 291-296.
22. J.R. Koup, Therapeutic drug monitoring and other clinical applications, Computer Applications in Pharmacy, eds W.E.Fasett and D.B.Christensen, 1986, 119-139.
23. L.B. Sheiner, S.L.Beal, B.Rosenberg and V.V. Marathe, Forecasting individual pharmacokinetics, Clinical Pharmacology Ther.1979, 20 (3), 294-305.
24. Iliadis Athanassios; Brown Allan C.; Huggins Mary L.; APIS: a software for model identification, simulation and dosage regimen calculations in clinical and experimental pharmacokinetics.1992 Elsevier science publishers B.V. Computer methods and program in Biomedicine, 1992, 227-239.

25. Hou Tingjun and Xu Xiaojie, A new molecular simulation software package – Peking University Drug Design System (PKUDDS) for structure-based drug design. Journal of Molecular Graphics and Modelling, 2001, 19, 455–465.
26. Allinger, N.L., Yuh, Y.H., and Lii, J.H. Molecular Mechanics. The MM3 force field for hydrocarbon. 1. J. Am. Chem. Soc. 1989, 111, 8551–8565.
27. Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S. Jr., and Weiner, P. A new force field for molecular mechanical simulation of nucleic acids and proteins. J. Am. Chem. Soc. 1984, 106, 765–784.
28. Pablo R. Duchowicz a, Alan Talevi a, Luis E. Bruno-Blanch , Eduardo A. Castro, New QSPR study for the prediction of aqueous solubility of drug-like compounds, Bioorganic & Medicinal Chemistry, 2008, 16, 7944–7955.
29. Yamashita Fumiyoshi, Itoh Takayuki, Yoshida Shuya, K. Haidar Mohammad, Hashida Mitsuru. Computer and Chemical Engineering -A novel multi-dimensional visualization technique for understanding the design parameters of drug formulations 2003, 28(2), 222-239.
30. Gieschke, R., Reigner, B.G., Steimer, J.-L., 1997. Exploring clinical study design by computer simulation based on pharmacokinetic/pharmacodynamic modelling. Int. J. Clin. Pharmacol. Ther. 35, 469–474.
31. Arup K. Ghose, Torsten Herbertz, Joseph M. Salvino and John P. Mallamo, Knowledge based chemoinformatic approaches to drug discovery, Drug Discovery Today, 2006, 11, 23-24,
32. Moses Hohman, Kellan Gregory, Kelly Chibale, Peter J. Smith, Sean Ekins and Barry Bunin, Novel web based tool combining chemistry informatics, biology and social networks for drug discovery, Drug Discovery Today, 2009, 14, 5-6.

33. Ertl Peter, Selzer Paul and Muhlbacher Jorg, Web-based cheminformatics tools deployed via corporate Intranets. Novartis Institutes for BioMedical Research, In Silico Sciences/Cheminformatics, CH-4002, 2003, 14, 321-328.
34. P. Ertl, World Wide Web-based system for the calculation of substituent parameters and substituent similarity searches. *J. Mol. Graph. Model.* 16 (1998), pp. 11–13.
35. I.V. Tetko, The WWW as a tool to obtain molecular parameters. *Mini Rev. Med. Chem.* 2003,3, 809–820.
36. Bonate, P.L., 2000. Clinical trial simulation in drug development. Pharm. Res. 17, 252–256.
37. Huang SY, Zou X: Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking. Proteins 2007, 66:399-421.
38. David S. Goodsell, The Art of Molecular Graphics The 1998 Molecular Graphics Art Show. The Scripps Research Institute, Journal of Molecular Graphics and Modelling, Science Inc.1999, 55-84.
39. Wlodawer, A. Rational approach to AIDS drug design through structural biology. Annu.Rev. Med. 2002, 53, 595–614.
40. Gomeni, Roberto, Bani Massimo, D'Angeli Carla, Corsi Mauro , Bye Alan, Computer-assisted drug development (CADD): an emerging technology for designing first-time-in-man and proof-of-concept studies from preclinical experiments. European Journal of Pharmaceutical Sciences, 2001, 13, 261–270.
41. Yamashita, F., Hara, H., Itoh, T., & Hashida, M. Novel hierarchical classification and visualization method for multiobjective optimization of drug properties: Application to structure–activity relationship analysis of cytochrome P_{450} metabolism. Journal of Chemical Information and Modeling, 2008, 48, 364.
42. Laforenza, D.et.al, Grid programming: some indications where we are headed. In Elsevier Parallel Computing, 2002 28, 1733–1752.
43. Stevens Robert, McEntire Robin, Goble Carole, Greenwood Mark, Zhao Jun, Wipat Anil and Li Peter; myGrid and the drug discovery process Reviews research focus; DDT: BIOSILICO, 2004, 2, 140-148.

44. http://www.apbionet.org/apbiogrid.

45. http://www. ncbiogrid.org.

46. http:// www.cbr.nrc.ca.

47. http://www.eurogrid.org.

48. http://www. nbirn.net

49. Stevens, R.D. *et al*. myGrid: personalised bioinformatics on the information grid. *Bioinformatics,2003*, 19, 1302–1304

50. Goble, C. et al. The myGrid Project: services, architecture and demonstrator. In Proceedings UK e-Science All Hands Meeting, 2001, 15(3), 200-222.

* Corresponding Author
Suman Bala,
MM College of Pharmacy,
MM University, Mullana,
Ambala, Haryana, INDIA -133203
Phone No: +91-99961-21027
E-mail: sumankmj7@gmail.com