International Journal of Computational Research and Development (IJCRD)
Impact Factor: 4.775, ISSN (Online): 2456 - 3137
(www.dvpublication.com) Volume 2, Issue 1, 2017

DRUG DESIGN, TYPES AND ITS MODE OF ACTION IN BINDING TARGET PROTEIN WITH LIGAND

Saradha Arumugam*, Dr. S. Saravanan & M. Vasanth*****
CAS in Marine Biology, Faculty of Marine Sciences, Annamalai University, Parangipettai, Chidambaram, Tamilnadu

Cite This Article: Saradha Arumugam, Dr. S. Saravanan & M. Vasanth, "Drug Design, Types and Its Mode of Action in Binding Target Protein with Ligand", International Journal of Computational Research and Development, Volume 2, Issue 1, Page Number 29-36, 2017.

Abstract:

Bioinformatics has become emerging field and has been used in various areas of research. The principal goal of bioinformatics is to enhance the understanding of biological processes. This implicates the establishment and encroachment of databases, algorithms, computational and statistical techniques and theory to solve formal and practical problems originating from the management and analysis of biological data. Molecular interactions including protein-protein, enzyme-substrate, protein-nucleic acid, drug-protein and drug-nucleic acid play important roles in many essential biological processes such as signal transduction, transport, cell regulation, gene expression control, enzyme inhibition, antibody–antigen recognition and even the assembly of multi-domain proteins. These interactions very often lead to the formation of stable protein–protein or protein-ligand complexes that are essential to perform their biological functions. Various organisms from tiny creatures to big one have been put forth in the research field to discover drug molecule for various diseases in the 21st century. Human being has been affected with dangerous viruses for example, ziga virus, HIV virus and so on.

Key Words: Molecular Modelling, Types, Softwares Tools & Binding Activity

Introduction:

The bioinformatics field is growing astoundingly during last decade due to the advancement of biotechnology combined with the development of computational skills and the network effort developed by different countries in sharing information. This field offers specialized analytical technology tools for both laboratory and networking technology fields to achieve the goal of finding natural bioactive compounds from marine organisms as potential drugs, antifouling compounds, biomaterials etc. It has highest implication in marine based ecology, biology, biotechnology and molecular biology, which are integrated for the gene based discovery and development of marine drugs. The marine bioinformatics includes marine genomics and marine proteomics.

As the search for natural bioactive molecules from marine organism as potential drugs, bioinformatics offers many high technology tools to stepping up new drug discovery. Without bioinformatics, new application oriented research in many fields of marine drug discovery and marine biology would come to a standstill. *In silico* molecular docking is one of the most powerful techniques of structure-based drug design (Brooijmans, 2003). Most applications of docking tools focus on the (supposed) primary binding region.

Drug Discovery:

The drug is most commonly an organic small molecule that activates or inhibits the function of a biomolecule such as a protein, which in turn results in a therapeutic benefit to the patient. In the most basic sense, drug design involves the design of small molecules that are complementary in shape and charge to the biomolecular target with which they interact and therefore will bind to it. Drug design frequently but not necessarily relies on computer modeling techniques. Ideally the computational method should be able to predict affinity before a compound is synthesized and hence in theory only one compound needs to be synthesized. The reality however is that present computational methods are imperfect and provide at best only qualitatively accurate estimates of affinity. Therefore in practice it still takes several iterations of design, synthesis and testing before an optimal molecule is discovered. On the other hand, computational methods have accelerated discovery by reducing the number of iterations required and in addition have often provided more novel small molecule structures.

The ultimate goal of modern drug discovery is to identify a therapeutic agent that is effective against a disease. The process of drug discovery is a complex issue. It may be divided into three main steps: *i*) development of relevant biological system for testing of the compounds *in vitro* and *in vivo; ii*) identification of "lead" compounds for concept test in the biological assays; *iii*) optimization of the "lead" structure to enhance the selectivity ratio, toxicity profile or pharmacokinetics and ultimately furnish a candidate drug suitable for appropriate *in vivo* studies and further clinical evaluation. The development of new drugs is an expensive process the toxicology profile must be taken into account at an early stage of the drug development. Thus, a highly potent molecule may be rejected at a late stage of the drug discovery process because of unavoidable and unpredictable side-effects. In order to successfully design safer drugs it has to be tested for its activity towards the particular protein target. (Graffner Nordberg., 2001).

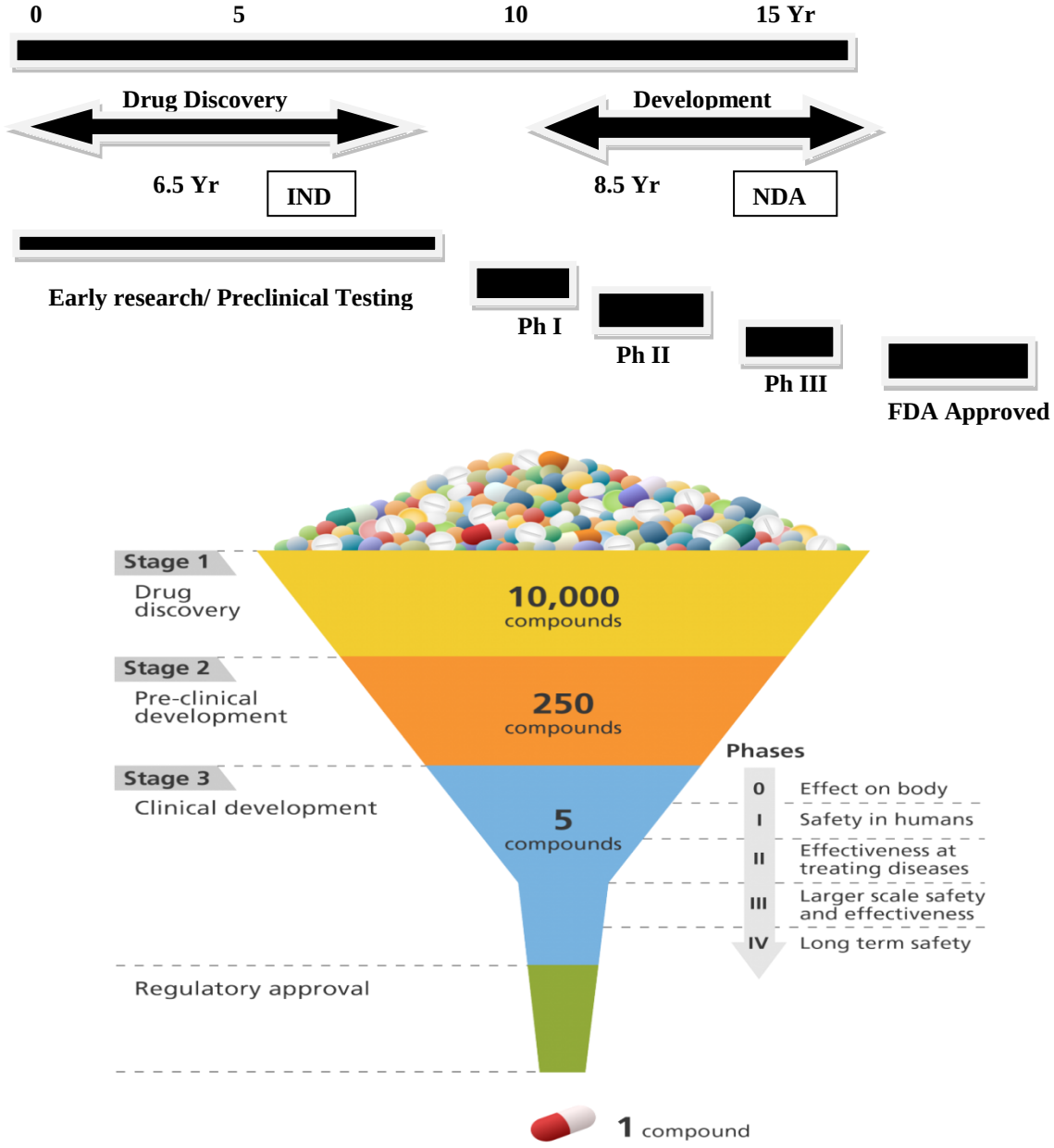

Figure 1: Drug Discovery and Development

In contrast to traditional methods of drug discovery, which rely on 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 hypothesis that modulation of a specific biological target may have therapeutic value. In order for a biomolecule to be selected as a drug target, two essential pieces of information are required. The first is evidence that modulation of the target will have therapeutic value. This knowledge may come from, for example, disease linkage studies that show an association between mutations in the biological target and certain disease states. The second is that the target is "drugable". This means that it is capable of binding to a small molecule and that its activity can be modulated by the small molecule. Once a suitable target has been identified, the target is normally cloned and expressed. The expressed target is then used to establish a screening assay. In addition, the three-dimensional structure of the target may be determined. The search for small molecules that bind to the target is begun by screening libraries of potential drug compounds. This may be done by using the screening assay (a "wet screen"). In addition, if the structure of the target is available, a virtual screen may be performed of candidate drugs. Ideally the candidate drug compounds should be "drug-like", that is they should possess properties that are predicted to lead to oral bioavailability, adequate chemical and metabolic stability and minimal toxic effects. Several methods are available to estimate drug likeness such as Lipinski's Rule of Five and a range of scoring methods such as lipophilic efficiency. Several methods for predicting drug metabolism have been proposed in the scientific literature and a recent example is SPORCalc. Due to the complexity of the drug design process, two terms of interest are

still serendipity and bounded rationality. Those challenges are caused by the large chemical space describing potential new drugs without side-effects.

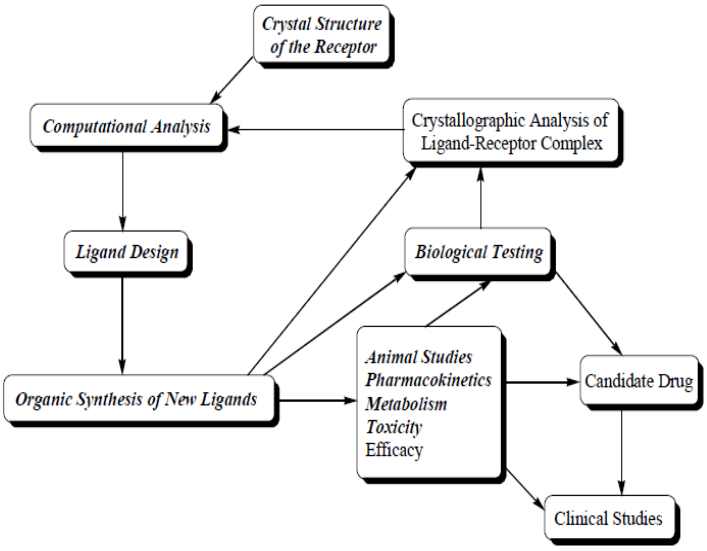

Figure 2: Structure-Based Drug Design Process

Problems in Drug Discovery:

Until recently, the focus in drug discovery, drug development and clinical trials has been on simplifying the complexity of the human being to individual genes, single metabolic pathways, single protein targets and one treatment such as a small molecule drug or a biologic to regulate complex functions. This one gene, one protein, one treatment concept dominated the drug discovery process. To the great disappointment of the pharmaceutical industry the "molecular reductionist" approach to drug discovery has not delivered the promised efficiencies. Instead, the number of hypotheses validated in the clinic as measured by successful product launches has slowed significantly. This failure is due in major part to an over-dependence on the continuity of the one gene, one target and one therapeutic molecule paradigm.

Computer-Aided Drug Design:

Computer-aided drug design uses computational chemistry to discover, enhance or study drugs and related biologically active molecules. The most fundamental goal is to predict whether a given molecule will bind to a target and if so how strongly. Molecular mechanics or molecular dynamics are most often used to predict the conformation of the small molecule and to model conformational changes in the biological target that may occur when the small molecule binds to it. Semi-empirical, ab initio quantum chemistry methods or density functional theory are often used to provide optimized parameters for the molecular mechanics calculations and also provide an estimate of the electronic properties (electrostatic potential, polarizability etc.) of the drug candidate that will influence binding affinity. Molecular mechanics methods may also be used to provide semi-quantitative prediction of the binding affinity. Also, knowledge-based scoring function may be used to provide binding affinity estimates. These methods use linear regression, machine learning, neural nets or other statistical techniques to derive predictive binding affinity equations by fitting experimental affinities to computationally derived interaction energies between the small molecule and the target (Rajamani and Good, 2007).

Drug design with the help of computers may be used at any of the following stages of drug discovery:

- ✓ Hit identification using virtual screening (structure- or ligand-based design).
- ✓ Hit-to-lead optimization of affinity and selectivity (structure-based design, QSAR etc.)
- ✓ Lead optimization of other pharmaceutical properties while maintaining affinity.

Types of Drug Design:

Drug discovery and development is very expensive and time consuming process. Traditional approaches to drug discovery rely on a step-wise synthesis and screening of large number of compounds to identify a potential candidate. Over the past ten to twenty years, there is an increased effort to apply computational power to the combined chemical and biological space in order to streamline drug discovery, design, development and optimization. Computational methods are expected to play an imperative role in understanding the specific molecular recognition events of the target macromolecule with candidate hits leading to the design of improved leads for the target. Computer Aided Drug Design (*in silico*) approaches have been widely employed in Lead Identification and Lead Optimization stages of drug development against various targets over the years. In comparison to traditional drug discovery methods rational drug design methods bring down the time and cost involved in drug development process. It can be used to identify/design new inhibitors

de novo or for optimization of absorption, distribution, metabolism, excretion and toxicity profile of identified molecules from various sources. Advances in computational techniques and hardware have facilitated the application of *in silico* methods in the discovery process. Drug Design can be categorized as two types: Structure based drug design (SBDD) and Ligand based drug design (LBDD) (Kapetanovic *et al*., 2008).

Structure Based Drug Design:

SBDD is the approach where the structural information of the drug target is exploited for the development of its inhibitor. Receptor structure(s) is a prerequisite for this method. Most commonly the structure of the receptor is determined by experimental techniques such as X-ray crystallography or NMR. If the structure of the protein drug target is not available, protein structure can be predicted by computational methods like threading and homology modeling. Threading (also called as fold) is a modeling approach used to model proteins that do not have homologous proteins with known structure. During threading, a given amino acid sequence is searched for compatibility with the structures in a database of known folds. The structure of the query protein is built from these folds. Homology modeling (also called as comparative) is an approach that relies on a clear relationship or homology between the sequence of the target protein and at least one known structure. The process of homology modeling of proteins consists of the following steps: Identification of homologous protein with known 3D structure(s) that can serve as template; sequence alignment of target and template proteins; generation of model for the target based on the 3D structure of the template and the alignment; model refinement and validation. Over the years, homology modeling has become the main alternative to get a 3D representation of the target in the absence of crystal structures.

***De Novo* Drug Design:**

De novo is a Latin expression meaning "from the beginning". Active site of drug targets when characterized from a structural point of view will shed light on its binding features. This information of active site composition and the orientation of various amino acids at the binding site can be used to design ligands specific to that particular target. Computational tools that can analyze protein active site and suggest potential compounds are extensively used for *de novo* design methods. Many promising approaches with the goal of ligand design have been reported. There are six major classes: i. Fragment location methods: To determine desirable locations of atoms or small fragments within the active site. ii. Site point connection methods: To determine locations ("site points") and then place fragments within the active site so that those locations are occupied by suitable atoms. iii. Fragment connection methods: Fragments are positioned and "linkers" or "scaffolds" are used to connect those fragments and hold them in a desirable orientation. iv. Sequential buildup methods: Construct a ligand atom by atom, or fragment by fragment. v. Whole molecule methods: Compounds are placed into active site in various conformations, assessing shape and/or electrostatic complementarity. vi. Random connection methods: A special class of techniques combining some of the features of fragment connection and sequential buildup methods, along with bond disconnection strategies and ways to introduce randomness. Over the years various *de novo* methods especially whole molecule methods like docking have become integrated within disciplines that include chemistry, pharmacology, molecular biology and computer modeling (Schneider and Fechner, 2005). Electrostatic and solvation terms critical for evaluating correct binding energies are difficult and slow to calculate. Advances in algorithm sophistication are providing better and better approximations for these parameters. Finally, it is clear from the recent literature that the drug design process has become an essential part of drug discovery projects.

Structure Based Virtual Screening:

Structure based virtual screening is one of the commonly used approaches in lead identification step and is seen as a complementary approach to experimental high throughout screening (HTS) to improve the speed and efficiency of the drug discovery and development process. This involves explicit molecular docking (process to predict binding mode) of each ligand to the binding site of the target and scoring (process to measure binding affinity). The compounds in the databases screened are ranked with a view to selecting and experimentally testing a small subset for biological activity considered to be appropriate for a given receptor. Many successful applications have been reported in the field of molecular docking based virtual screening. Although the energy calculations involved are crude, the compounds in the library are readily available, making experimental testing easy and false positives tolerable.

Ligand Based Drug Design:

Ligand based drug design is an approach used in the absence of the receptor 3D information and it relies on knowledge of molecules that bind to the biological target of interest. 3D quantitative structure activity relationships (3D QSAR) and pharmacophore modeling are the most important and widely used tools in ligand based drug design. They can provide predictive models suitable for lead identification and optimization (Acharya *et al*., 2011).

Docking:

Docking is a term used for computational schemes that attempt to find the "best" matching between two molecules. A receptor and a ligand. The molecular docking problem can be defined as follows: Given the atomic coordinates of two molecules, predict their "correct" bound association. In its most general form, no

additional data are provided. The simpler problem in docking is referred to as "bound" docking. It relates to computational schemes that attempt to reconstruct a complex using the bound structures of the receptor and the ligand. A "bound" structure is extracted from a structure of more than one molecule, typically a co-crystal of the receptor and the ligand. The goal is, however, the more difficult predictive docking, also referred to as the "unbound" docking. The unbound problem relates to computational schemes that attempt to reconstruct a complex using the unbound structures of the receptor and the ligand. An unbound structure may be a native structure, a pseudo-native structure or a modeled structure. In this terminology, a native structure is the structure of a molecule when it is free in solution, in its uncomplexed state. A pseudo-native structure is the structure of a molecule when complexed with a molecule different from the one used for the docking. There are three key ingredients in the docking: (1) representation of the system (2) conformational space search and (3) ranking of potential solutions. Docking essentially simulates the interaction of the protein surface. Therefore, the first question is how to define a protein surface. The surface can be described by mathematical models such as for example by geometrical shape descriptors or by a grid. Alternatively, it can involve treatment (static or dynamic) of the protein frame such as, for example, rigid vs flexible. Docking involves two separate molecules. It initiates from folded protein chains and ligand conformations. In contrast, protein folding initiates from some non-native protein conformations. Hence, docking is often viewed as distinct from folding. Yet, while currently computational prediction of protein structures largely addresses relatively small, single domain proteins for large multi domain proteins, one faces the problem of domain docking. Such an approach is consistent with experiment. Experimentally, complementary fragments provide a system for studying protein folding, consistent with intermolecular binding resembling intra molecular folding events. Intra molecular domain docking appears simpler owing to chain linkage. Just like in protein folding, solving the docking problem also involves two components: an efficient search procedure and a good scoring function.

The two critical elements in a search procedure are speed and effectiveness in covering the relevant conformational space. On the other hand, the scoring function should be fast enough to allow its application to a large number of potential solutions and in principle, effectively discriminate between native and non-native docked conformations. The scoring function should include and appropriately weigh all the energetic ingredients. Hence, as in folding the performance of a particular docking program should not be viewed as representing one complete piece. To solve the docking problem, ideally, the best matching algorithms and scoring schemes should be combined (Inbal Halperin *et al*., 2002).

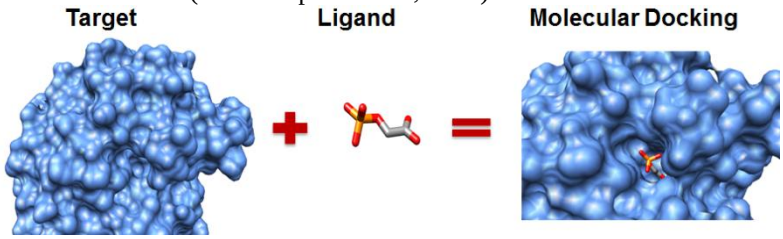

Figure 5: Example of Docking

Search Algorithm:

The search space consists of all possible orientations and conformations of the protein paired with the ligand. With present computing resources, it is impossible to exhaustively explore the search space. This would involve enumerating all possible distortions of each molecule since molecules are dynamic and exist in an ensemble of conformational states and all possible rotational and translational orientations of the ligand relative to the protein at a given level of granularity. Most docking programs in use account for a flexible ligand and several are attempting to model a flexible protein receptor. Each snapshot of the pair is referred to as a pose.

Scoring Function:

The scoring function takes a pose as input and returns a number indicating the likelihood that the pose represents a favorable binding interaction. Most scoring functions are physics-based molecular mechanics force fields that estimate the energy of the pose where a low or negative energy indicates a stable system and thus a likely binding interaction. An alternative approach is to derive a statistical potential for interactions from a large database of protein-ligand complexes, such as the Protein Data Bank and evaluate the fit of the pose according to this inferred potential**.**

Evaluation:

One of the major uses of docking is in the ranking of ligands in order of their relative binding affinities. Evaluation of the results may be done by analyzing drugs based on these scores. Due to uncertainty in the approximation in the scoring function it is better to consider a range of top scoring DOCK prediction. There can be minimum energy positions that are redundantly reported, resulting in many nearly identical orientations with slight differences in their scores. These redundant orientations can overwhelm the ranked list of all the results, reducing their spatial diversity. Clustering provides a useful tool for pruning redundant results from this top scoring list. The most appropriate method for analyzing the dock results would be correlating the result obtained by docking with the experimentally validated results.

Softwares Involved in Drug Discovery:

The number of protein-ligand docking softwares are available in the internet resource and abundantly used by research scholars and students worldwide. This programs are free, open source and commercial packages to carry out various tasks in resolving several bioactive compounds in relation to study molecular modelling and ADME properties.

Docking Program	Developed by	Countries Name
Smooth dock	Carlos J. Camacho and P. Christoph Champ at University of Pittsburgh	USA
1- Click Docking	Mcule	USA
AADS	Indian Institute of Technology	India
ADAM	IMMD Inc.	Toyko
Auto Dock	Olson group at The Scripps Research Institute	USA
Auto Dock Vina	The Scripps Research Institute	USA
Beta Dock	Hanyang University	South Korea
Blaster	University of California	San Francisco
BSP-SLIM	University of Michigan	USA
Clus Pro	Camacho group at Boston University	USA
EhiTS	SimBioSys in Toronto	Canada
DARWIN	The Wistar Institute	USA
DIVALI	University of California	San Francisco
Dock	Kuntz and Shoichet Groups at the University of California	San Francisco
Docking Server	Virtua Drug Ltd	Hungary
Dock Vision	Dock Vision	Canada
EA Dock	Swiss Institute of Bioinformatics	Australia
eHiTS	Sym Bio Sys Inc	Switzerland
EUDOC	Mayo Clinic Cancer Center	USA
FDS	University of Southampton	UK
Flex E	University of Halle-Wittenberg,	*Germany*.
Flex X	Bio Solve IT	Germany
Flex AID	University of Sherbrooke	
Flex Pep Dock	The Hebrew University	
FLIP Dock	Scripps Research Institute	USA
FLOG	Open Eye Scientific	USA
FT DOCK	Biomolecular Modelling Laboratory	UK
GEMDOCK	National Chiao Tung University	Taiwan
Glide	Schrodinger, USA	USA
GOLD	Collaboration between the University of Sheffield, Glaxo Smith Kline plc and CCDC	UK
GPCR automodel	INRA	
HADDOCK	Centre Bijvoet Center for Biomolecular Research	
Hammer Head	Arris Pharmaceutical Corporation	USA
ICM-Dock	Mol Soft	USA
Id Target	National Taiwan University	
I Screen	China Medical University	
Lead finder	Mol Tech	Russia /Canada
Ligand Fit	Bio Via	USA
Lig Dock CSA	Seoul National University	South Korea
LIGIN	Weizmann Institute of Science	Israel/ Germany
LPCCSU	Weizmann Institute of Science	
MCDOCK	Georgetown University Medical Center	USA
MDock	University of Missouri	USA
ME Dock	SIGMBI	
Mol Dock	Molegro ApS	Denmark
MS-DOCK	INSERM	France
Par DOCK	Indian Institute of Technology	India
Patch Dock	Tel Aviv University	
PLANTS	University of Konstanz	Belgium /

		Germany
PLATINUM	Moscow Institute of Physics and Technology (State University)	
PRO DOCK	Cornell University	USA
PSI-DOCK	Peking University	China
PSO@AUTODOCK	University of Leipzig	Germany
Pyth Dock	Hanyang University	South Korea
Q-Dock	Georgia Institute of Technology	USA
QXP	Novartis Pharmaceuticals Corporation	USA
R Dock	University of York/ Open Source Project	
SANDOCK	University of Edinburg	UK
Score	Alessandro Pedretti & Giulio Vistoli	
Smooth Dock	Carlos J. Camacho and P.Christoph Champ at University of Pittsburg	
SODOCK	Feng Chia University	Taiwan
SOFT Docking	University of California, Berkeley	USA
Surflex-Dock	Tripos	USA
Swiss Dock	Swiss Institute of Bioinformatics	
Vote Dock	University of Warsaw	Poland
YUCCA	Virginia Tech	USA

Table 1: Freeware and Commercial Software Packages for docking process

The ultimate goal of modern drug discovery is to identify a therapeutic agent that is effective against a disease. The process of drug discovery is a complex issue; it may be divided into three main steps:

- ✓ Development of relevant biological system for testing of the compounds *in vitro* and *in vivo;*
- ✓ Identification of "lead" compounds for concept test in the biological assays and
- ✓ Optimization of the "lead" structure to enhance the selectivity ratio, toxicity profile or pharmacokinetics, and ultimately furnish a candidate drug suitable for appropriate *in vivo* studies and further clinical evaluation.

To Analysis of Protein-ligand Binding Interactions Using Ligplot:

Despite well-known limitations, the structure determination of a ligand in complex with its target protein is generally regarded as a highly valuable piece of information for helping to understand ligand target complementarity and possibly direct subsequent design strategy. Detailed inspection of available 3D structures will provide the fullest interpretation, but these are often difficult to investigate quickly, even for experts; while potential users from other disciplines, such as medicinal chemistry, need time to first acquaint themselves with 3D coordinate visualization software. The LIGPLOT program automatically generates schematic 2-D representations of protein-ligand complexes from standard Protein Data Bank file input. The output is a colour or black-and-white. PostScript file giving a simple and informative representation of the intermolecular interactions and their strengths including hydrogen bonds, hydrophobic interactions and atom accessibilities. The program is completely general for any ligand and can also be used to show other types of interaction in proteins and nucleic acids. It was designed to facilitate the rapid inspection of many enzyme complexes. But, has found many other applications. The interactions shown by LIGPLOT are those mediated by hydrogen bonds and by hydrophobic contacts. Hydrogen bonds are indicated by dashed lines between the atoms involved; each hydrogen bonded residue from the protein is shown in full, although there is an option to include/exclude its main-chain atoms. Hydrophobic contacts are indicated more schematically; residues from the protein involved in these contacts are represented by an arc, with spokes radiating towards the ligand atoms they contact. The contacted atoms are shown with spokes radiating back (Andrew Wallace *et al*.,1995).

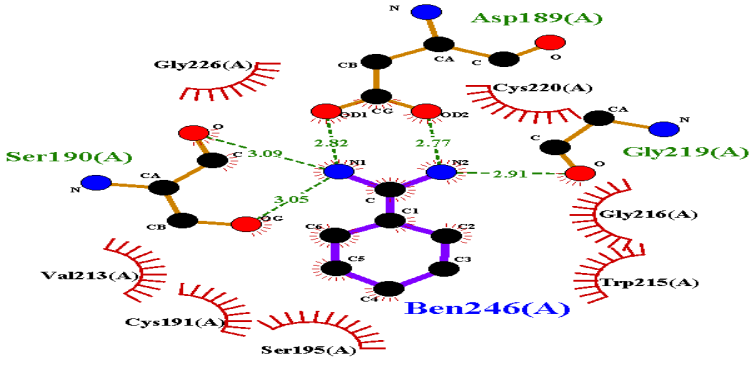

Figure 3: Example of Ligplot

Conclusion:

Current drug discovery efforts are often based on increasing the throughput of experiments to increase the likelihood to find a promising lead compound (e.g. HTS). Systems Biology harbors the potential to learn more about the causal mechanisms of disease and therapeutic targets that ameliorate disease phenotypes. The computer helps in both the steps to find out possible biological functions of a protein by comparing its amino acid sequences to database of proteins with known functions and to understand the mole. Bioinformatics techniques are also used in the process of drug discovery, these speeds up the process. The ligplot diagrams portray the hydrogen-bond interaction patterns and hydrophobic contacts between the ligand(s) and the main-chain or side-chain elements of the protein. The system is able to plot, in the same orientation, related sets of ligand protein interactions. This facilitates popular research tasks, such as analyzing a series of small molecules binding to the same protein target, a single ligand binding to homologous proteins, or the completely general case where both protein and ligand change.

Acknowledgement:

We, the authors are thankful to our Dean and Director, CAS in Marine Biology, Faculty of Marine Sciences and Authorities of Annamalai University for providing facilities and encouragement.

References:

1. Andrew C.Wallace1, Roman A.Laskowski and Janet M.Thornton. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Engineering vol.8 no.2 pp.127-134, 1995.
2. Brooijmans, N. and Kuntz, I.D. Molecular recognition and docking algorithms. Annu. Rev. Biophys. Biomol. Struct. 32, 335–373, 2003.
3. Graffner Nordberg M, Kolmodin K, Aqvist J, Queener S.F,Hallber A. Design, Synthesis, computationals predictions and biological evaluation of ester soft drugs as inhibitors of dihydrofolate reductase from pneumocystis carinii, J.Med.Chem., 44, 2391-2402, 2001.
4. Rajamani R, Good A.C. Ranking poses in structure based lead discovery and optimization: current trends in scoring function development. Curr. opin. Drug Discov Devel 10 (3), 308- 315, 2007.
5. Schneider G, Fechner U. Compurt based *de novo* design of drug molecules. Nat Rev Drug Discov, 4 (8) 649-663, 2005.
6. Kapetanovic, I.M. Computer-aided drug discovery and development (CADDO): in silico-chemico-biological approach. Chem Bioi Interact, 171, 165-176, 2008.
7. Acharya C, Coop A, Polli J E and Mackerell A D. Recent advances in ligand based drug design: relevance and utility of the conformationally sampled pharmacophore approach. Curr. Comp-Aided Drug Design 7(1): 10-22, 2011.
8. Inbal Halperin, Buyong Ma, Haim Wolfson, and Ruth Nussinov, Principles of docking: an overview of search algorithms and a guide to scoring functions, PROTEINS: Structure, Function, and Genetics 47:409-443, 2002.