INTERNATIONAL RESEARCH JOURNAL OF PHARMACY **ISSN 2230 – 8407**
Available online www.irjponline.com **Review Article**

IN SILICO MODELLING AND DRUG DESIGN – A REVIEW

Gupta Praveen kumar*, Agrawal Pushpa, Shivakumar Neeta and Hiremath Suhasini.B
Department of Biotechnology, R.V College of Engineering, Mysore Road, Bangalore-560 059 India

Article Received on: 17/07/11 Revised on: 20/08/11 Approved for publication: 12/09/11

*E- mail: praveenkgupta@rvce.edu.in

ABSTRACT
Bioinformatics and Computational biology is an interdisciplinary field that applies the techniques of computer science, applied mathematics and statistics to address biological problems. Research in computational biology often overlaps with systems biology. Major research efforts in the field include sequence alignment, gene finding, genome assembly, protein structure alignment, protein structure prediction, prediction of gene expression and protein-protein interactions, and the *in silico* drug modelling. Drug discovery is an intense, lengthy and interdisciplinary endeavour. It is mostly portrayed as a linear, consecutive process that starts with target and lead discovery, followed by lead optimization and pre-clinical *in vivo* and *in vitro* studies to determine if such compounds satisfy a numbers of pre-set criteria for initiating clinical developments. *In silico* methods help in identifying drug targets via bioinformatics tools. They can be used to analyze the target structures for possible binding sites, generate candidate molecules, check for their drug likeness, dock these molecules with the target, rank them according to their binding affinities and further optimize the molecules to improve binding characteristics.

KEY WORDS: Molecular modelling, QSAR, Drug designing, *In silico* modelling,

INTRODUCTION

Bioinformatics is the field of science in which biology, computer science and information science merge into a single discipline. It manages and analyses biological data using advanced computing techniques. *In silico* is an expression used to mean "performed on computer or via computer simulation." The term *'in silico'* was first used by Pedero Miramontes, a mathematician from National Autonomous University of Mexico to characterize biological experiments carried out entirely on a computer. As structures of more and more protein targets become available through crystallography, NMR and bioinformatics methods, there is an increasing demand for computational tools that can identify and analyze active sites and suggest potential drug molecules that can bind to these sites specifically. To combat life-threatening diseases such as AIDS, Tuberculosis, Malaria etc., a global push is essential.[1] Time and cost required for designing a new drug are immense and at an unacceptable level. A new prescription drug costs on an average, $802 million and take up to 15 years of research to develop before it is introduced in the market. Intervention of computers at some plausible steps is imperative to bring down the cost and time required in the drug discovery process.[2] *In silico* modeling is a form of computer based modelling whose technologies are applied in drug target identification or drug discovery processes, cell analysis of prokaryotic and eukaryotic hosts and bioprocess development, optimization, analysis and interpretation of heterogeneous data sets from various sources.

PARAMETERS CONSIDERED FOR THE DRUG DESIGN AND DEVELOPMENT

- Whole genome sequence analysis
- Determining the structure activity relationship by 3DQSAR
- ADME

Whole genome analysis: A paradigm for drug design and development is the human genome project, wherein the entire genome of human is determined.[3] The complete genome of human has been utilized to know the nature and structure of the receptors. This is determined by the genetic code in the genome. Once the structure of the receptor is known, it becomes very easy to develop a molecule which binds to the receptor. This molecule is supposed to be a drug molecule.

Determining the structure activity relationship: This is done using the tool 3DQSAR. It stands for 'Quantitative Structure Activity Relationship'. It is the process by which chemical structure is quantitatively correlated with a well defined process. Drug design is an iterative process which begins with a compound that displays an interesting biological profile and ends with optimizing both the activity profile for the molecule and its chemical synthesis. The process is initiated when the chemist conceives a hypothesis which relates the chemical features of the molecule (or series of molecules) to the biological activity. Without a detailed understanding of the biochemical process(es) responsible for activity, the hypothesis generally is refined by examining structural similarities and differences for active and inactive molecules. Compounds are selected for synthesis which maximizes the presence of functional groups or features believed to be responsible for activity. The combinatorial possibilities of this strategy for even simple systems can be explosive. As an example, the number of compounds required for synthesis in order to place 10 substituents on the four open positions of an asymmetrically disubstituted benzene ring system is approximately 10,000. The alternative to this labor intensive approach to compound optimization is to develop a theory that quantitatively relates variations in biological activity to changes in molecular descriptors which can easily be obtained for each compound. A Quantitative Structure Activity Relationship (QSAR) can then be utilized to help guide chemical synthesis. QSAR is a technique that quantifies the relationship between structure and biological data and is useful for optimizing the groups that modulate the potency of the molecule. It can provide an estimate of the highest potency expected of a molecule in series or can provide information on whether all parts of the molecule are in close contact with the binding site.[4,5,13]

ADME: In pharmacokinetics, the ADME is for absorption, distribution, metabolism, and excretion. It is assumed that the drugs are in a dynamic state within the body as they move between tissues and fluids or bind with plasma or cellular components or are metabolized. The description of drug distribution and elimination is often termed as drug disposition. Characterization of drug disposition is an important prerequisite for determination or modification of dosing regimens for individuals and groups of patients.[2,6,11]

COMPUTER ASSISTED DRUG DESIGNING (CADD)

Computer aided drug design, often called structure based design involves using the biochemical information of ligand-receptor interaction in order to postulate ligand refinements. This technique uses computational chemistry to discover, enhance or study drugs and related biologically active molecules. Methods used include simple molecular modeling. The purpose is to reduce the number of

targets for a good drug that have to be subjected to expensive and time consuming synthesis and trialing. This method comprises of the attempts made to find a ligand (the putative drug) that will interact favorably with a receptor that represents the target site. Binding of the ligand to the receptor may include hydrophobic, electrostatic and hydrogen bonding interactions. The approach used in CADD s dependent upon the amount of information that is available of the ligand and the receptor. The 3-D structural information for the receptor and the ligand-receptor complex from X-ray and NMR should be determined. [5] The approach used in CADD optimizes the fit of the ligand in a receptor site. This approach does not consider the pharmacokinetics of the drug.

The steps involved in the CADD are

- Choice of drug target-Receptor
- Evaluating the structure for the drug design
- Identification of the target site.
- Designing of the drug by homology modeling
- Molecular docking to check the compatibility

Choice of the drug target

For the drug to function, it has to bind to the target molecule and trigger the physiological effect. The target molecule is the receptor. The ideal target molecule is one that is closely linked to human disease and binds a small molecule in order to carry out a function.

Evaluating the structure for the drug design

Once the target has been identified, it is necessary to obtain accurate structural information. The methods for the structure determination that are useful drug design are- X-ray crystallography, NMR. This facilitates the design of the ligand molecule that is the drug by the related computational methods.

Identification of the target site

Structure based design begins with the identification of a potential ligand binding site on the target molecule. Ideally, the target site is a pocket or perturbance with a variety of potential hydrogen bond donors and acceptors and hydrophobic characters.

Drug design by Homology Modeling

Homology modeling involves taking a known sequence with an unknown structure and mapping it against a known structure of one or several similar

Homologous proteins: It would be expected that two proteins of similar origin and function would have reasonable structural similarity. Therefore it is possible to use the known structure as a template for modeling the structure of the unknown structure. [5,6]

Homology modeling, also known as comparative modeling refers to constructing an atomic-resolution model of the "target" protein from its amino acid sequence and an experimental three-dimensional structure of a related homologous protein (the "template"). Homology modeling relies on the identification of one or more known protein structures likely to resemble the structure of the query sequence, and on the production of an alignment that maps residues in the query sequence to residues in the template sequence. The sequence alignment and template structure are then used to produce a structural model of the target. Because protein structures are more conserved than DNA sequences, detectable levels of sequence similarity usually imply significant structural similarity[7]

All homology modeling approaches consists of three steps

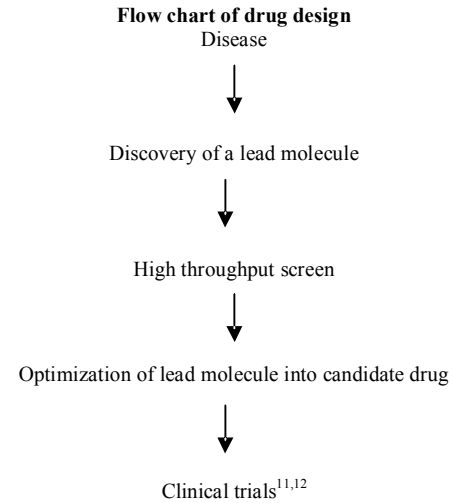
1. Finding homologous PDB files.
2. Creation of the alignment, using single or multiple sequence alignments. (if more than one known is involved, sometimes the known are aligned together, then the unknown sequence aligned with the group; this helps ensure better domain conservation) Analysis of alignments; gap deletions and additions; secondary structure weighting
3. Structure calculation and model refinement .[7,8,9]

Molecular Docking

In the field of molecular modeling, docking is a method which predicts the preferred orientation of one molecule to a second when bound to each other to form a stable complex. Knowledge of the preferred orientation in turn may be used to predict the strength of association or binding affinity between two molecules using for example scoring functions.[10] The associations between biologically relevant molecules such as proteins, nucleic acids, carbohydrates, and lipids play a central role in signal transduction. Furthermore, the relative orientation of the two interacting partners may affect the type of signal produced (e.g., agonism vs antagonism). Therefore docking is useful for predicting both the strength and type of signal produced. Docking is frequently used to predict the binding orientation of small molecule drug candidates to their protein targets in order to in turn predict the affinity and activity of the small molecule. Hence docking plays an important role in the rational design of drugs. [2,9]

The important computational tools for docking are

- DOCK
- AutoDOCK
- FTDOCK
- HEX
- Argus lab
- CHARMM



Benefits of CADD

CADD methods and bioinformatics tools offer significant benefits for drug discovery programs.

- Cost Savings. The Tufts Report suggests that the cost of drug discovery and development has reached $800 million for each drug successfully brought to market. Many biopharmaceutical companies now use computational methods and bioinformatics tools to reduce this cost burden. Virtual screening, lead optimization and predictions of bioavailability and bioactivity can help guide experimental research. Only the most promising experimental lines of inquiry can be followed and experimental dead-ends can be avoided early based on the results of CADD simulations.[2]
- Time-to-Market. The predictive power of CADD can help drug research programs choose only the most promising drug candidates. By focusing drug research on specific lead candidates and avoiding potential “dead-end” compounds, biopharmaceutical companies can get drugs to market more quickly.[11]

- Insight. One of the non-quantifiable benefits of CADD and the use of bioinformatics tools is the deep insight that researchers acquire about drug-receptor interactions. Molecular models of drug compounds can reveal intricate, atomic scale binding properties that are difficult to envision in any other way. When we show researchers new molecular models of their putative drug compounds, their protein targets and how the two bind together, they often come up with new ideas on how to modify the drug compounds for improved fit. This is an intangible benefit that can help design research programs.[12,13]

CADD and bioinformatics together are a powerful combination in drug research and development.

DRUG DESIGN SOFTWARES

1. Binding Affinity Prediction of Protein-Ligand Server (BAPPL): Computes the binding free energy of a protein-ligand complex.
2. Binding Affinity Prediction of Protein-Ligand complex containing Zinc Server (BAPPL-Z): Computes the binding free energy of a metalloprotein-ligand complex containing zinc.
3. Drug-DNA Interaction Energy (PreDDICTA): Calculates the Drug-DNA interaction energy.
4. ParDOCK - Automated Server for Rigid Docking Predicts the binding mode of the ligand in receptor target site.
5. Non Redundant Database of Small Molecules Virtual high throughput screening of small molecules and their optimization into lead- like candidates.
6. Molecular Volume Calculator: Calculates the volume of a molecule
7. DNA Sequence to Structure: Generates double helical secondary structure of DNA using conformational parameters taken from experimental fiber-diffraction studies.
8. Hydrogen Addition to Nucleic Acid: Adds the hydrogen coordinates to the X-ray crystal structures of Nucleic acid
9. Hydrogen Addition to Protein: Adds the hydrogen coordinates to the X-ray crystal structures of Protein
10. DNA Ligand Docking: Rigid Docking predicts the binding mode of the ligand in the minor groove of DNA.

CONCLUSION

The ultimate goal of Bioinformatics is to enable the discovery of new biological insights as well as to create a global perspective from which unifying principles in biology can be discerned. Drug design and development has evolved as a fast emerging field in biotechnology which is conducive to the future of biotechnology which will bring a fortune. It is remarkably an economical process and most importantly time saving. *In silico* method has opened up doors for the pharmaceutical research all over the world. Significant achievements are seen and more are awaited.

REFERENCES

1. Supercomputing Facility for Bioinformatics & Computational Biology, IIT Delhi (Research, drug design) c2004-2011 [cited 2009 Aug 9]. Available from: http://www.scfbioiitd.res.in/tutorial/drugdiscoery.html
2. Soloman AK. Molecular Modeling and Drug Design. 2nd ed (Chennai): MJB Publishers, 2008.
3. Baxevanis AD. Francis Ouellette B. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins; 2nd ed (NY): A John Wiley & Sons, Inc, 2001.
4. Rastogi SC. Bioinformatics: Methods and Applications Genomics, Proteomics and Drug Discovery. 2nd ed (New Delhi): Prentice-hall Of India Pvt.ltd, 2006.
5. Borman S. New QSAR Techniques Eyed for Environmental Assessments. Chem Eng News, 1990; 68: 20-23.
6. Claude Cohen N. The Molecular Modeling Perspective in Drug Design. Academic Press Inc, California, 1996.
7. Searls DB, Kasif S, Salzberg SL. Computational Methods for Molecular Biology. Elsevier Science, Netherlands, 1998.
8. Molecular Modeling of Proteins and Nucleic Acids, Department of Biochemistry, Virginia Tech (Homepage). c1997-2003(cited 2009 Aug 9) Available from: http://www.biochem.vt.edu/modeling/homology.html
9. Lesk AM. Introduction to Bioinformatics. Oxford University Press Inc., New York, 2002.
10. Lengauer T, Rarey M. Computational methods for biomolecular docking. Curr Opin Struct Biol 1996; 6: 402-406.
11. Silverman RB. The Organic Chemistry of Drug Design and Drug Action. 2nd ed (Netherlands): Elsevier Science, 2004.
12. Bourne PE and Weissing H. Structural Bioinformatics.2nd ed (NY): A John Wiley & Sons, Inc, 2009.
13. Brown SM. Essentials of Medical Genomics. 2nd ed (NY): A John Wiley & Sons, Inc, 2003 .