ejpmr, 2021,8(2), 258-262

SJIF Impact Factor 6.222

EUROPEAN JOURNAL OF PHARMACEUTICAL AND MEDICAL RESEARCH

www.ejpmr.com

Review Article

ISSN 2394-3211

EJPMR

A REVIEW ON MATERIALS & METHODS FOR IN SILICO DRUG DESIGNING

Dimpal Rani*, Rajesh Kumari Patil and Hanumanthrao Chandershekar Patil

Adesh Institute of Pharmacy and Biomedical Sciences, Adesh University, Bathinda, Punjab, India.

***Corresponding Author: Dimpal Rani**

Adesh Institute of Pharmacy and Biomedical Sciences, Adesh University, Bathinda, Punjab, India.

Article Received on 10/12/2020 **Article Revised on 30/12/2020** **Article Accepted on 20/01/2021**

ABSTRACT

In silico Drug Designing is the identification of lead drug molecule by applying Bioinformatics. it is playing increasing role in Drug Discovery and Identification of drug candidate. So insilico methodology is important part of drug Discovery process. So in this study we study different software and tools, which are used for searching journals, databases, protein analysis, for retrieving sequence of protein and also comparing the sequences, modeling of protein, validation of model, comparing 3D structure of protein, visualization of 3-D structure, active site finding, protein-ligand interaction, ligand design, ADMET prediction.

KEYWORDS: Drug Designing, Drug Discovery, Protein Modeling.

INTRODUCTION

The goal of insilico drug designing is to find out lead molecule which can be synthesized further and can be used to cure disease by passing through different stages of clinical trial. So here are tools and software which are used for insilico drug designing which involve various steps from protein analysis to visualization.[1,2]

1. FOR PROTEIN ANALYSIS (http://www.expasy.org/)

The ExPASy (the Expert Protein Analysis System) World Wide Web server, used to access to a variety of databases and analytical tools related to proteins and proteomics, genomics, transcriptomics. This portal is operated by Swiss Institute of Bioinformatics .ExPASy databases include resources asSWISS-MODEL, Swiss-orthology, SWISS-PROT, SWISS Drug Design, String, and the SWISS-Lipid repository. Analysis tools are available for specific tasks relevant to proteomics, protein structure homology modelling,similarity searches, tool and data for regulatory genomics, pattern and profile searches, knowledge resource of lipid and protein, post-translational modification prediction, primary, secondary and tertiary structure analysis and sequence alignment. These databases and tools are interlinked: a special emphasis is placed on integration of database entries with related resources developed at the SIB and elsewhere, and the proteomics tools have been designed to read the annotations in SWISS-PROT in order to enhance their predictions. This portal is user friendly for both expert and new users.[3]

2. FOR RETRIEVAL OF SEQUENCES, TEMPLATE SELECTION AND ALIGNMENT OF SEQUENCES

2.1 NCBI (www.ncbi.nlm.nih.gov)

The National Center for Biotechnology Information (NCBI) was created in 1988 at the National Institutes of Health to access biomedical and genomic information. Most popular database are Genebank and Pubchem.In addition to maintaining the GenBank® nucleic acid sequence database, which receives data through the international collaboration with DDBJ and EMBL as well as from the scientific community, NCBI provides data retrieval systems and computational resources for the analysis of GenBank data and many other kinds of other biological data.

There is Popular NCBI resources include Pubmed, Bookshelf, Pubmed central, BLAST, Neuclotide, Genome, SNP, Gene, Protein, Pubchem etc. These resources give information of search and retrieval of biomedical and life sciences literature, free online access to books and documents in life science and healthcare SNP contains human single nucleotide variations, microsatellites, and small-scale insertions and deletions, with main objective of Search, read, and discover. Augmenting many of the web applications are custom implementations of the BLAST program optimized to search specialized datasets.[4]

2.2 CLUSTAL W (www.ebi.ac.uk/clustalw)

The Clustal series of programs are widely used in molecular biology for the multiple alignment of both nucleic acid and protein sequences and for preparing phylogenetic trees. The popularity of the programs depends on a number of factors, including not only the

accuracy of the results, but also the robustness, portability and user-friendliness of the programs. New features include NEXUS and FASTA format output, printing range numbers and faster tree calculation. The Clustal W and Clustal X multiple sequence alignment programs have been completely rewritten in C++. This will facilitate the further development of the alignment algorithms in the future and has allowed proper porting of the programs to the latest versions of Linux, Macintosh and Windows operating systems.[5,6]

2.3 BLAST (www.ncbi.nlm.nih.gov/blast)

Basic local alignment search tool (BLAST), use to find regions of local similarity between sequences. This tool compare the sequences of neucleotide and protein with the sequences of Database and statistically calculate the significance of matches. BLAST offers different search type like BLASTn (Neucleotide BLAST), BLASTx (translated nucleotide sequence searched against protein sequences), tBLASTn (protein sequence searched against translated nucleotide sequences), BLASTp (Protein BLAST). BLAST is an order of magnitude faster than existing sequence comparison tools of comparable sensitivity.[7]

3. FOR PROTEIN MODELING AND VALIDATION OF MODELS

3.1 MODELLAR 9V5

Modeller is a computer program used for homology or comparative protein structure modeling that models three-dimensional structures of proteins by satisfaction of spatial restraints. An alignment of a sequence to be modeled is provided with known related structures then a model with all non-hydrogen atoms will automatically calculate by Modeller.

More generally, the input to the program are restraints on the spatial structure of the amino acid sequence(s) and ligands to be modeled. The output is a 3D structure that satisfies these restraints as well as possible. Restraints can in principle be derived from a number of different sources. These include related protein structures (comparative modeling), NMR experiments (NMR refinement), rules of secondary structure packing (combinatorial modeling),cross-linking experiments, fluorescence spectroscopy, image reconstruction in electron microscopy, site-directed mutagenesis, intuition, residue–residue and atom–atom potentials of mean force, *etc*. The restraints can operate on distances, angles, dihedral angles, pairs of dihedral angles and some other spatial features defined by atoms or pseudo atoms. Presently, Modeller automatically derives the restraints only from the known related structures and their alignment with the target sequence.

Modeller can use for multiple comparison of protein sequences and structures, searching of sequence databases, and clustering of proteins. A scripting language is used by this program and does not include any graphics.

3.2 SWISS MODEL (http:/swismodel.expasy.org)

SWISS-MODEL is a automated server for Homology modeling of three-dimensional (3D) protein structures. Homology Modeling is reliable method used to generate three-dimensional (3D) protein structures models. It is the most widely-used free web-based automated modeling facility today. SWISS-MODEL provides several levels of user interaction through its World Wide Web interface.

SWISS-MODEL build model by different steps which include identification of structural template(s), alignment of target sequence and template structure(s), model-building, and model quality evaluation. To build model these steps can be repeated until the satisfactory modeling results obtained. Complex modeling tasks can be handled with the 'project mode' using DeepView (Swiss-PdbViewer). In this programme the model evaluation can be done by GMQE (Global Model Quality Estimation) and QMEAN estimator based on different geometrical properties . The SWISS-MODEL server is under constant development to improve the successful implementation of expert knowledge into an easy-to-use server.[8]

3.3 DISCOVERY STUDIO

Discovery Studio is a single, powerful, easy-to-use, software for simulating small molecule and macromolecule systems. It is graphical interface for drug design and protein modeling research. Discovery Studio 2.1 includes a set of new functionality and enhancements to existing features that further strengthen its capabilities as a comprehensive modeling and simulation platform for the life sciences. Powerful new analysis tools in the area of Quantum Mechanics-Molecular Mechanics (QM-MM) offer new capabilities for analyzing protein-ligand systems. Enhancements to the ligand profiling protocol enable us to work with commercial curated databases of structure based pharmacophores. Also, CHARMm has been updated to version 34b1.

Experimental structure determination involves difficult methods that require a significant amount of expertise and resources. Protein Modeling on the other hand enables access to sensible structural models in much shorter periods of time. The major aim of protein modeling is the prediction of a protein's tertiary structure from its primary structure. To this end a diverse range of approaches have been developed comparative, de novo or ab initio structure prediction.

The protein modeling methods provided in Discovery Studio provide an extensive set of tools and protocol which are described below for the construction of molecular structures and analysis of their behavior.

- Build and edit proteins or peptides
- Modify protein structures and generate reports of protein structures

- Identify and report abnormalities in a protein structure
- Align sequences according to sequence similarity or structure similarity
- Build homology models for a protein sequence
- Refine side-chain conformation of a protein structure
- Refine conformation of one or more loop regions of a protein structure
- Mutate residues and optimize structure
- Predict protein docking conformations
- Analyze the sequence conservation pattern of a sequence family
- Identify conserved functional patches in a protein structure

3.4 NIH PROTEOMIC SERVER

This is Stuctural Analysis and Verification Server. Depending on how many programs we select to use, this server can take several minutes to run. It also depends on how many residues there are in the protein we submit.

3.5 PROCHECK

PROCHECK programme checks the stereochemistry of a protein structure. It analyze residue-by-residue geometry and overall structure geometry. This programme access the quality of protein structures in the process as well as existing structures. For this it also focus on reporting torsion angle parameters.[9]

4. FOR DESIGNING OF CHEMICAL LIBRARY

4.1 ACD ChemSkech11.0

ACD/ChemSketch is a drawing package that allows draw chemical structures. It also calculate of molecular properties like molecular weight, density, molar refractivity etc. It also include 2Dimensional and 3Dimensional structure cleaning and viewing, functionality for naming structures (fewer than 50 atoms and 3 rings), and prediction of log*P*.

It is easy to use chemical structure drawing application. It can draw chemical structure, Reactions, access a variety of graphical tools and templates. It also Generate names from molecular structure and vice versa, Create professional reports, presentations, and publication-ready figures and Communicate scientific information with clarity and ease.

5. FOR DOCKING

5.1 Gold 4.0

Genetic Optimization for Ligand Docking, (GOLD) is useful in virtual screening, lead optimisation, and identifying the correct binding mode of active molecules**.** It is Reliable, Flexible, and Configurable.

The mechanism for ligand placement is based on fitting points. Fitting points are added to hydrogen-bonding groups on protein and ligand and the program maps acceptor points in the ligand on donor points in the protein and *vice versa*.

GOLD also maps ligand CH groups onto hydrophobic fitting points generated on the protein binding site cavity. GOLD also optimizes flexible ligand dihedrals, ligand ring geometries, dihedrals of protein hydroxyl and amino groups and mappings of the latter. A molecular mechanics-like scoring function that includes terms for hydrogen bond, 4-8 intermolecular van der Waals bonds and 6-12 intramolecular van der Waals bonds (for the internal energy of the ligand) is employed by GOLD to rank the docked poses.[10]

5.2 Ligand fit (Accelrys DS 2.1): The interactions between a receptor and a ligand are fundamental to drug discovery. Discovery Studio provides a set of methods for predicting and analyzing the interactions between protein receptors and ligands. These methods allow we to carry out structure-based design, or even to examine possible interactions with theoretical structures such as homology models. A common technique central to receptor-ligand interactions is docking. Discovery Studio provides several docking methods as well as a rich graphical interface to third-party docking tools such as GOLD. Discovery Studio also includes several methods applicable to fragment-based design such as the De Novo protocols. Analysis of hypothetical poses is also possible via a series of scoring functions, hydrogen bonds and bumps, and high level physics-based scoring methods to predict binding energies.

Receptor-Ligand Interactions provide tools and protocols that allow us to:

- Perform characterization and analysis of receptor binding sites
- Perform detailed analysis of ligand poses, including calculation of RMSD
- Estimate binding free energy for complex formation
- Score ligand poses and find the consensus between a set of scoring functions
- Suggest novel ligands with a De Novo approach
- Dock ligands into an active site using various approaches
- Perform flexible docking by considering both protein and ligand flexibility
- Prepare and minimize a set of ligands

5.3 AutoDock

AutoDock is a suite of automated docking tools. It is designed to predict how small molecules, such as substrates or drug candidates, bind to a receptor of known 3D structure.

AutoDock actually consists of two main programs: AutoDock performs the docking of the ligand to a set of grids describing the target protein; AutoGrid pre-calculates these grids. In addition to using them for docking, the atomic affinity grids can be visualised. This can help, for example, to guide organic synthetic chemists design better binders.[11]

6. FOR ADMET PREDICTION

6.1 DISCOVERY STUDIO (ACCELRYS DS 2.1): Discovery Studio provides methods for assessing the disposition and potential toxicity of a ligand within an organism. The ADMET protocols contain published models that we can use to compute and analyze Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) properties.[12] In addition, we can apply specific rules to remove ligands that are not likely drug-like, unsuitable leads, etc. based on the presence or absence and frequency of certain chemical groups.

We can use the ADMET functionality to:

- Estimate the aqueous solubility of a set of ligands
- Estimate the Blood brain barrier penetration (BBB)
- Estimate Cytochrome P450 (CYP450) 2D6 inhibition.
- Estimate hepatotoxicity
- Estimate human intestinal absorption (HIA)
- Estimate plasma protein binding
- Assess a broad range of toxicity measures for a set of ligands
- Remove ligands that are not likely to be drug-like or lead-like

7. VISUALIZATION TOOL

7.1 RASMOL

RasMol2 is a molecular graphics program intended for the visualisation of proteins, nucleic acids and small molecules. The program is aimed at display, teaching and generation of publication quality images. RasMol runs on Microsoft Windows, Apple Macintosh, UNIX and VMS systems. The UNIX and VMS systems require an 8, 24 or 32 bit colour X Windows display (X11R4 or later). The program reads in a molecule co-ordinate file and interactively displays the molecule on the screen in a variety of colour schemes and molecule representations. Currently available representations include depth-cued wireframes, 'Dreiding' sticks, spacefilling (CPK) spheres, ball and stick, solid and strand biomolecular ribbons, atom labels and dot surfaces.[13]

7.2 PYMOL

PyMOL is an open-source, user-sponsored, molecular visualization system created by Warren Lyford DeLano and commercialized by DeLano Scientific LLC, which is a private software company dedicated to creating useful tools that become universally accessible to scientific and educational communities. It is well suited to producing high quality 3D images of small molecules and biological macromolecules such as proteins. According to the author, almost a quarter of all published images of 3D protein structures in the scientific literature were made using PyMOL.

PyMOL is one of few open source visualization tools available for use in structural biology. The Py portion of the software's name refers to the fact that it extends, and is extensible by the Python programming language.

7.3 SWISS PDB VIEWER: Swiss-PdbViewer is an application that provides a user friendly interface allowing to analyze several proteins at the same time. The proteins can be superimposed in order to deduce structural alignments and compare their active sites or any other relevant parts. Amino acid mutations, H-bonds, angles and distances between atoms are easy to obtain due to the intuitive graphic and menu interface.

Swiss-PdbViewer can also read electron density maps, and provides various tools to build into the density. In addition, various modeling tools are integrated and command files for popular energy minimization packages can be generated.

7.4 DISCOVERY STUDIO VISUALIZER

Discovery Studio® Visualizer 2.5 allows we to view and edit molecular structures, sequences and sequence alignments that created with the Discovery Studio and other applications.

CONCLUSION

In this review, we discussed different software and tools which are contributing great role in the in silico drug designing and drug discovery. This software based approach is useful opportunity for the in vitro identification of biologically active agents. These methods like molecular modeling and docking are helpful for insilico predication of active drugs against multiple disease. Significant advances in this field is going on by developing new software and application of these softwares that is beneficial in drug discovery. These methods are cost effective and also limiting the use of animal models. So, we can conclude that this software based approach is playing major role in Drug Designing.

REFERENCE

1. Contreras-Moreira B., FitzjohnP.W., Bates P.A. Comparative modelling: an essential methodology for protein structure prediction in the post-genomic era. Applied Bioinformatics, 2002; 1(4): 177-190.
2. Wenbo Yu and Alexander D. MacKerell Jr. Computer-Aided Drug Design Methods. Methods Mol Biol, 2017; 1520: 85–106.
3. Gasteiger E., Gattiker A., Hoogland C., Ivanyi I., Appel R.D., Bairoch A**.** ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.*, 2003; 31(13): 3784-3788.
4. Sayers E.W. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, 2010; 38(Database issue): D5-16.
5. Larkin M.A., Blackshields G., etal. Clustal W and Clustal X version 2.0. *Bioinformatics*, 2007; 23(21): 2947-2948.
6. Chenna R., Sugawara H., Koike T., Lopez R., Gibson T.J., Higgins D.G., Thompson J.D. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.*, 2003; 31(13): 3497-500.

7. Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. Basic local alignment search tool. *J Mol Biol.*, 1990; 215(3): 403-10.
8. Schwede T., Kopp J., Guex N., Peitsch M.C. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res*, 2003; 31(13): 3381-3385.
9. Willard L., Ranjan A., Zhang H., Monzavi H., Boyko R.F., Sykes B.D., Wishart D.S. VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res.*, 2003; 31(13): 3316-3319.
10. Joy S., Nair P.S., Hariharan R., Pillai M.R. Detailed comparison of the protein-ligand docking efficiencies of GOLD, a commercial package and ArgusLab, a licensable freeware. *In Silico Biol.*, 2006; 6(6): 601-605.
11. Stefano Forli, Ruth Huey, Michael E Pique, Michel F Sanner, David S Goodsell & Arthur J Olson. Computational protein–ligand docking and virtual drug screening with the AutoDock suite.*Nature Protocols*, 2016; 11(5): 905–919
12. Jie Dong, Ning-Ning Wang, Zhi-Jiang Yao, Lin Zhang, Yan Cheng, Defang Ouyang, Ai-Ping Lu & Dong-Sheng Cao ADMETlab: a platform for systematic ADMET evaluation based on a comprehensively collected ADMET database. *Journal of Cheminformatics*, 2018; 10: 29.
13. J. Tony Pembroke. Bio-molecular modelling utilising RasMol and PDB resources: a tutorial with HEW lysozyme *Biochemistry And Molecular Biology* Education, 2000; 28: 297-300.