IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A REVIEW ON MOLECULAR MODELING FOR DRUG DISCOVERY: SOFTWARE- BASED APPROACH

Aishwarya Chougule[1], Gajanan Athawale[2], Aarti Kamble[3].

M. Pharmacy

Department of Pharmacology[1], Department of Pharmaceutics[2]

School of Pharmacy, S.R.T.M. University, Nanded

Abstract

To discover and develop new drugs by the traditional method, it requires more time and expensive process. Now a day, these issues are overcome by the use of computer software. There is different software available for molecular modeling and structural activity relationship. By using software the drug discovery process accelerates. The molecular modeling has become a valuable and essential tool in drug discovery but it has the main challenge is to determine structural and chemical information of drug target and ligand binding sites to invent new molecules that have high affinity, selectivity, bioavailability, and less toxicity. This review paper covers the molecular modeling, its application, different types of software such as Maestro, ArgusLab, GRAMM, SYBYL-X Suite, Sanjeevini, and PASS along with its uses.

Keywords

Drug discovery, Molecular modeling, Structural Activity Relationship (SAR), Software-based approach.

1. Introduction

Drug discovery is the process by which potential new drugs/medicines are identified. It includes the identification of drug candidates, characterization, synthesis, screening, and assays for therapeutic efficiency. For drug design, it requires determining lead compound; fit a targeted protein cavity of geometrically and chemically. Drug discovery as well as development is a very expensive and long-lasting process as it requires about one million dollars or more expenditure and about 10-12 years utilized to achieve the desired, successful product. High expenses, high-risk level, more time consuming, highly complex procedures, and then uncertain results are the main challenges. To overcome these challenges, today, methods such as software-based molecular modeling, structure-based drug design, structure-based virtual screening, ligand interaction, and molecular dynamics are developed which playing a principal role in drug discovery and development.

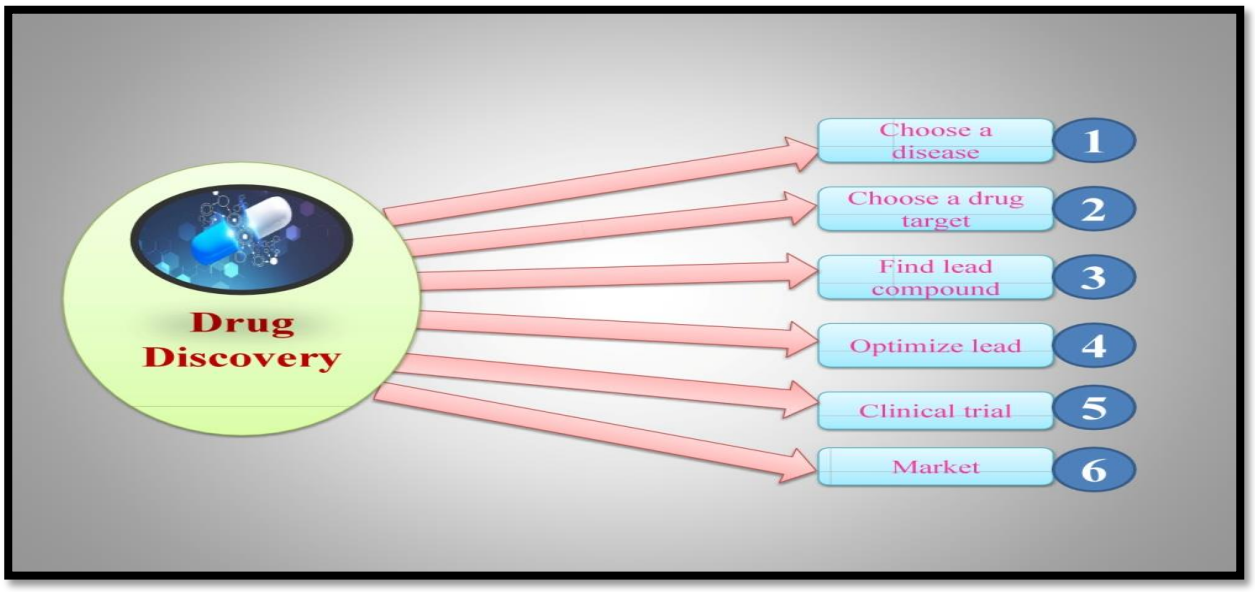

Figure No. 1: Drug discovery cycles

This modern drug discovery approach is specific and target based. The software-based drug design has some features over the traditional method of drug design as-

1. Less time-consuming.
2. Less expensive.
3. Easy identification of the target.
4. Quick experimental finding.
5. Faster and simpler.
6. Trouble-free to handle huge data.
7. Increase the accuracy of the result.
8. Accelerate new drug development.

1.1 Molecular modeling and structural activity relationship (SAR):

Molecular modeling is the most popular and fastest developing field of computational data-based screening methods of drug discovery and development. It is a computer-based method having techniques for deriving, visualizing, manipulating, and optimizing the structures of molecules. Many functions which start from building and visualizing simple protein molecules to performing complex simulation on large protein molecules are carried out by molecular modeling. It is a 'ligand-based screening method' in which a small molecule of the drug (ligand) interacts with the receptor (target protein) through hydrogen or *Van der Waals* bonding. This ligand-target protein interaction determines the binding energy i.e. pose energy of that conformation which depends on the shape or space of the cavity of protein. The negative geometry of the ligand-protein complex will have minimum binding free energy which is predicted as the best pose of that conformation.

The molecular modeling is applying for drug discovery in industry and in many academic laboratories where research is carried out. It is inexpensive, easy to use, and safe which helps to investigate, interpret, explain, and identify molecular characteristics using 3D structures.

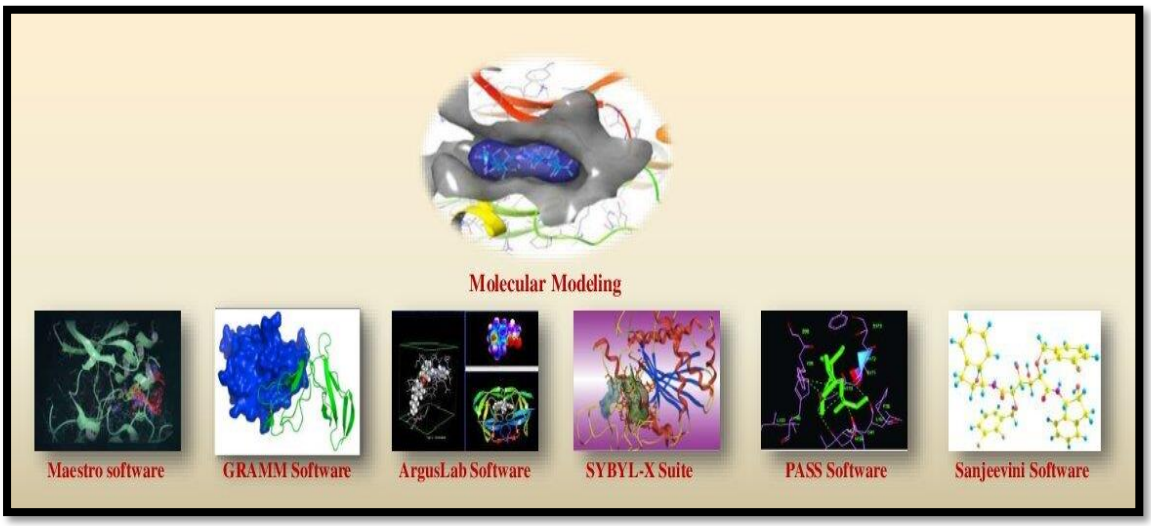

Figure No. 2: Drug design software

1.2 Applications:

1. Useful to investigate the structure, dynamics, surface properties, and thermodynamics of inorganic, biological, and polymeric systems.
2. Biological activities such as enzyme catalysis, protein folding, protein stability, and conformational alert associated with biomolecular function and molecular identification of proteins, DNA, and membrane complexes.

Software and software-based approaches apply to assist inexpensive, complex, and highly challenging drug design and discovery. This paper highlights the commonly used software of molecular modeling for the discovery of new drugs.

2. Molecular modeling software

Table No. 1: Software and computer-based programs used during new drug discovery.

Sr. No	Software name	Developed by	Use
1.	Maestro	Schrodinger Inc.	Molecular modeling analysis.
2.	ArgusLab	Mark A. Thompson, Planaria Software LLC, Washington.	Molecular docking calculations and molecular modeling packages.
3.	GRAMM	A Center for Bioinformatics, University of Kansas, USA.	Protein-protein docking and protein-ligand docking.
4.	SYBYL-X Suite	Tripos Inc.	Molecular modeling and ligand-based design
5.	Sanjeevini	IIT, Delhi.	Predict protein-ligand binding affinity.
6.	PASS	Russian Academy of Medical Sciences	Create an analysis of SAR models.

The commonly used software for drug design and their salient features are as follow:

2.1 Maestro

Maestro is the graphical user interface for all Schrodinger software. It is a powerful, versatile molecular modeling environment. It is built, display, and manipulates structures and organizes, monitor, submit, and visualize results from calculations on structures. It manages organization and analysis of obtained data by making it easy and straightforward.

Maestro is the linchpin of Schrodinger's computational technology. In that the following features:

i. Completely reimagined interface.
ii. Model generation - Maestro provides constructing molecular models of any type, which has an intuitive, full-featured building tool. For Flexible visualization – Maestro supply many viewing options to accommodate the different application of different need. From biomolecular systems to complex materials, it starts clarity to a wider range of modeled systems.
iii. 3D realism – Maestro's superior rendering and stereographic capabilities permit researchers to view complex molecular systems as three-dimensional objects with unmatched realism.
iv. Quantitative structural analysis – Maestro contains versatile determination tools used to quantify a molecule's structural features. Superimposition tools make possible all in detail comparisons between structures.
v. Customization scripts – It can customize and automate difficult work as well as manage workflow via scripting. Other than a proprietary language, Maestro scripts are written in the industry-standard Python language.
vi. Molecular properties – Computed properties like vibrational modes, molecular orbitals, or electron density are visualized in Maestro.
vii. Data management and analysis – It works on a data system that automatically archives structure-related properties. A built-in plotting facility aids elucidate structure-property relationships.
viii. Publication and presentation – Maestro outputs high-resolution, presentation-quality images that easily integrate into documents for publication or for sharing data with colleagues.
ix. Cross-platform support – It hastens natively on Linux, Windows, and Mac.

It is a novel structure-based method that has a similar or better accuracy than competitor methods. Maestro is a multi-agent prediction system based on statistical scoring functions (SSFs) and different machine learning approaches such as artificial neural networks (ANN), support vector machines (SVM), and multiple linear regressions (MLR). It gives as high throughput scanning for multi-point mutations where sites and types of mutations are controlled. Also provides a specific model for the prediction of stabilizing disulfide bonds.

Uses:

a. Pharmacophore modeling.
b. Ligand-receptor docking.
c. Ligand-receptor binding free energy prediction.
d. ADME prediction.
e. Quantitative structural analysis.
f. Prediction of changes in stability upon point mutations in proteins.

2.2 ArgusLab

ArgusLab is a molecular modeling, graphics, and ligand-based drug design program for windows platform by Planaria Software which is very useful, highly-featured.In computational docking, ArgusLab was originally developed as molecular modeling software. It contains two docking engines,

(i) GADock (Genetic algorithm)
(ii) ArgusDock (Shape-based search algorithm)

GADock is superior in terms of accuracy while ArgusLab is advantageous in terms of the low computational time it provides. In a pose construction step, docking poses constructed by GADock were trouble-free to assess than those of ArgusDock.

ArgusLab is a very useful tool used in the research lab, in the pharmaceutical industry and also in academics. The program allows drawing very complex protein configurations, obtaining helical chains of amino acids and folded leaves, etc. It employs a tree system to organize all the elements to add any structure before representing this data as a drawing, allowing analyzing it a visual manner.

The flexible ligand docking is possible with ArgusLab in which ligand is considered as a torsion tree and grids are constructed that cover the binding site. AScore is an empirical scoring function of ArgusLab. The ligand and receptor are required in PBD format. The ligand pose with the lowest energy is considered as the best ligand pose. The protein-ligand interaction, docking of the protein-ligand complex was performed by ArgusDock available under ArgusLab. It helps to determine potential energies, molecular structures, geometry optimization of the structure, vibrational frequencies of coordinates of atoms, bond length, bond angle, and reaction pathway. It is the electronic structure program based on quantum mechanics. The quantum mechanics calculation was accomplished by ArgusLab using Argus Computing Server 16. These docking engines are less accurate than other specialized docking programs (AutoDock) but they are meaningful.

It has the potential to build atom, molecules using templates, and new structure from pre-existing structures. Also, it can change the structure of an atom and bond type.

Advantages:

i. High-speed calculation capability.
ii. Reliable and easy to use program.
iii. A fast and robust method of binding site optimization.
iv. Freely available.
v. Low-cost docking.

Uses:

a. Build a chemical structure in 3D.
b. Calculation of binding energies of protein-ligand complexes.
c. For conformational analysis.
d. Calculation of potential energy.
e. Molecular structure visualization.
f. Molecular docking calculation.

2.3 GRAMM

GRAMM is the Global RAnge Molecular Matching docking method. It is a free program for protein-protein docking and protein-ligand docking. The main goal behind the development of GRAMM is to obtain a method which will give the accurate result of the protein-protein complex from two separate, flexible, and possible inaccurate protein structures. Also to get information about a relative orientation and optimized conformation between two proteins gives a stable structure having minimum potential energy. If it is used at high resolution for molecules that have large conformational changes give an inaccurate prediction of a structure while for molecules that have small conformational changes give an accurate prediction of structure. Structure accuracy identifies prediction quality.

To identify the structure of any complex protein molecule, it requires only atomic coordinates of two molecules. It does not require any information about binding sites. The molecular complex maybe two proteins or a protein and small compound or two transmembrane helices, etc. The surface match between protein molecules is determined by the Fast Fourier Transformation (FFT)-GRAMM method. GRAMM is gathering on SGI R10000, SGI R4000, SGI R4400, SGI R8000, Sun SPARC, IBM RS6000, DEC Alpha, and PC (Windows95 and Linux).

Features:

i. It can smooth protein surface presentation to account for conformational modification upon binding insert the rigid-body docking approach.
ii. It carries out an exhaustive six-dimensional search through its relative translations and rotations of molecules.

Uses:

a. Protein-protein docking.
b. Protein-ligand docking.

2.4 SYBYL-X Suite

It is versatile molecular modeling, simulation suite program which used for homology modeling to improve the process of designing based on its good scoring function. It can perform small molecular modeling and simulation, macromolecular modeling and simulation, chemoinformatics, lead identification, and optimization.

It is commercially developed by Tripos Inc. software for research and available on the Silicon Graphics and Linux workstations. SYBYL-X suite provides tools for molecular modeling, structure building, optimization and comparison, visualization of structures, and associated data. Also has screen capture capacity. It helps to determine molecular properties such as pharmacological, physical, ADME, toxicological properties, and 3D structure of drugs. It is computational informatics software developed to hasten drug design and drug discovery from HTS to lead optimization.

Important features:

i. Molecular editing and building.
ii. Can handle lots of chemical structures.
iii. Easy definition and selection option.
iv. High-quality graphics.
v. Export command scripts.
vi. I/O interface to multiple file formats.
vii. Saving your environment.

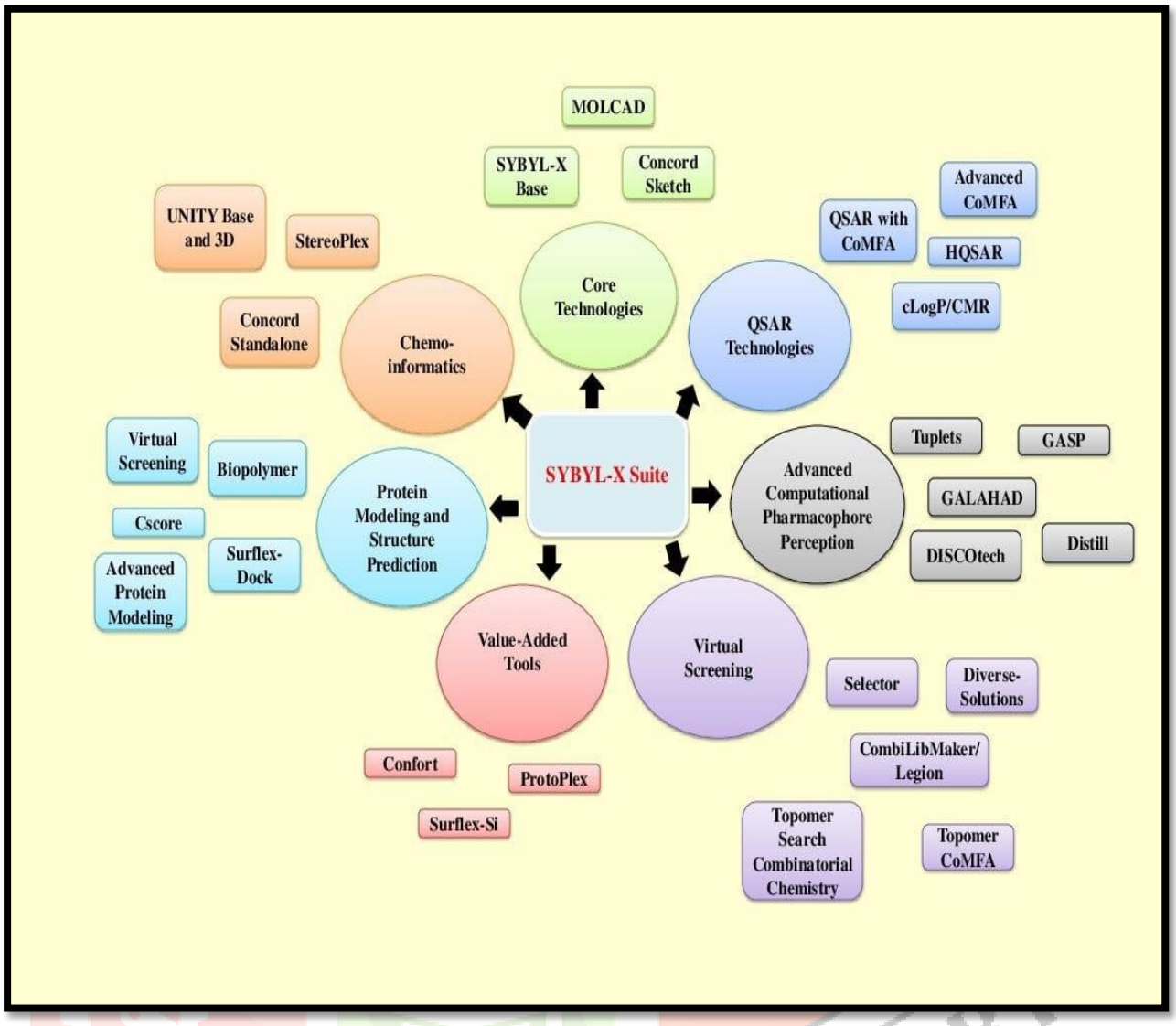

Figure No. 3: SYBYL-X Suite covers the above features.

By applying SYBYL-X, anyone can do the following work:

1. Construct a 3D model or homology model for the receptor of interest.
2. Recognize and evoke the cavities present on the target protein and the properties of the protein/ligand interaction surface.
3. Predict potential drug interaction with receptor by use of Surflex-Dock, docking software.
4. Determine lead compound by use of Surflex-Dock for virtual screening.

Uses:

a. Molecular docking from sequence through lead optimization.
b. Ligand-based design.
c. QSAR modeling.
d. Pharmacophore hypothesis generation.
e. Molecular alignment.
f. Conformational searching.
g. ADME prediction.
h. To build a protein model.

2.5 Sanjeevini

It is a freely accessible drug design software which having computational pathways for protein and DNA targeted lead molecule discovery.

In this software, the user uploads a protein target and a candidate drug. The software starts its work. It identifies the potential active sites, docks, and scores the candidate drug and returns four structures of the candidate drug bound to protein target together with binding free energies. Both protein and candidate drugs should be provided in PBD format. Sanjeevini server developed to overcome issues related to affinity and selectivity of a ligand in the case of a receptor with known structures. It is a composition of a total of six modules with different functions. The modules are as follows:

I. Template Library.
II. Molecule Generation
III. Molecular Descriptors and Drugs like Filters.
IV. Molecular Docking.
V. Energy Minimization of Resultant Complexes.
VI. Binding Affinity Computations on Energy Minimized Complex.

The source codes for all modules are written in a computer language such as C, C++, and FORTRAN. It performs several functions:

i. Automatic determination of active binding sites of ligands on target.
ii. Rapid screening of molecule database
iii. Identification of good candidates for the target protein, geometry optimization, and determination of partial atomic charges by quantum chemical methods.
iv. Docking the candidate in the active site of the target.
v. Determination of binding energy through scoring functions.
vi. Analysis of structure and binding energy for lead optimization.

Now a day, for molecular mechanics and quantum mechanics calculation, Sanjeevini is coupled with AMBER and GAMESS.

Uses:

a. For drug design.
b. Computational automating.
c. Prediction of protein-ligand binding affinity.

2.6 PASS

The acronym PASS stands for Prediction of Activity Spectra of Substance. It is an online tool that predicts more than 4000 types of biological activities such as anti-mycobacterial, anti-tuberculosis, cell wall synthesis inhibitor activity, membrane permeability enhancer ability. It also includes pharmacological effects, mechanism of action, toxic and adverse effects, interaction with metabolic enzymes and transporters, influence on gene expression.

To obtain biological activity profile, the only structural formula is required. Prediction is possible for a virtual structure designed in a computer which not synthesized yet. PASS predicts simultaneously 3678 types of activity with a mean precision of prediction regarding 95% depending on the compound's structural formula. PASS is now actively used in the field of medical chemistry, by both academic organizations and pharma companies. The input compound was submitted in the MOL/ SD file format.

The output has the following predicted activities:

i. A general review of every biological activity assigned to the input compound.
ii. Pharmacotherapeutic effects.
iii. Biochemical mechanisms.
iv. Toxicity i.e. adverse and toxic effects.
v. Metabolism.
vi. Gene regulation expression.
vii. Transporter-related activities.

It reduces the time lag between the designing and the drug discovery. It increases the chances for the discovery of new, safe, potent pharmaceutical agents.

Uses:

a. Create a SAR model.
b. Identification of probable targets and mechanisms of toxicity.
c. Exhibits new effects and mechanisms of action for already identified substances in a corporate and personal database.
d. Select the surest compounds from available samples for high throughput screening.
e. The new substance of lead molecule on the evaluation of existing library structures.

3. Conclusion

In this study, we have briefly mentioned about the different software used in molecular modeling and structural activity relationship that is actively playing role in drug discovery and development. Successfully implementation of molecular modeling and structural activity relationship based technique provide identification of drug candidates, characterization, synthesis, screening, and assays for therapeutic efficiency as well as provides an opportunity for the in vitro identification of biologically active agents. The pharmacokinetics and pharmacodynamics properties of new drug and structural activity relationships between new drug and disease target can determine by this software. Therefore these ligand-based drug design software has advanced application in the pharmaceutical industry and many academic laboratories.

4. Reference:

1. Baldi, A. Computational approaches for drug design and discovery: An overview. *Systematic reviews in Pharmacy* **1**, 99 (2010).
2. Finn, P. W. & Kavraki, L. E. Computational approaches to drug design. *Algorithmica* **25**, 347-371 (1999).
3. Mr. Baheti B. R.and Miss. Patel A. C. An assisting softwares using drug design *World, Journal of Pharmaceutical Research* **8,** 228-235(2019).
4. Senturk, M. Software Used for Drug Design and Development. *Proceedings Book*, 116 (2019).
5. Laimer, J., Hofer, H., Fritz, M., Wegenkittl, S. & Lackner, P. MAESTRO-multi agent stability prediction upon point mutations. *BMC bioinformatics* **16**, 116 (2015).
6. Laimer, J., Hiebl-Flach, J., Lengauer, D. & Lackner, P. MAESTROweb: a web server for structure-based protein stability prediction. *Bioinformatics* **32**, 1414-1416 (2016).
7. Kapoor, G. *et al.* Synthesis, ADME, docking studies and in vivo anti-hyperglycaemic potential estimation of novel Schiff base derivatives from octadec-9-enoic acid. *Bioorganic Chemistry* **84**, 478-492 (2019).
8. Oda, A. *et al.* Evaluation of Docking Accuracy and Investigations of Roles of Parameters and Each Term in Scoring Functions for Protein–Ligand Docking Using ArgusLab Software. *Bulletin of the Chemical Society of Japan* **80**, 1920-1925 (2007).
9. Achutha, A., Pushpa, V. & Manoj, K. Comparative molecular docking studies of phytochemicals as Jak2 inhibitors using Autodock and ArgusLab. *Materials Today: Proceedings* (2020).
10. Tangyuenyongwatana, P. & Jongkon, N. Molecular docking study of tyrosinase inhibitors using ArgusLab 4.0. 1: A comparative study. *Thai Journal of Pharmaceutical Sciences (TJPS)* **40** (2016).
11. Hasan, M. N., Bhuiya, N. & Hossain, M. K. In Silico molecular docking, PASS prediction, and ADME/T analysis for finding novel COX-2 inhibitor from Heliotropium indicum. *J. Comp. Med. Res* **10**, 142-154 (2019).
12. Chikhi, A. & Bensegueni, A. Comparative study of the efficiency of three protein-ligand docking programs. *Journal of Proteomics and Bioinformatics* **1**, 161-165 (2008).
13. Sindhu, T. *et al.* Molecular docking and QSAR studies on plant derived bioactive compounds as potent inhibitors of DEK oncoprotein. *Asian J. Pharm. Clin. Res* **4**, 67-71 (2011).
14. Naz, A., Bano, K., Bano, F., Ghafoor, N. A. & Akhtar, N. Conformational analysis (geometry optimization) of nucleosidic antitumor antibiotic showdomycin by Arguslab 4 software. *Pakistan journal of pharmaceutical sciences* **22** (2009).
15. Sisodiya, D., Pandey, P. & Dashora, K. Drug Designing Softwares and Their Applications in New Drug Discover. *Journal of Pharmacy Research* **5**, 124-126 (2012).

16. Tovchigrechko, A. & Vakser, I. A. GRAMM-X public web server for protein–protein docking. *Nucleic acids research* **34**, W310-W314 (2006).
17. Dubey, A. & Kalra, S. J. S. Computational comparative modeling and visualization for HIV1 and HIV2 proteins via the software SYBYL-X. *International Journal of Scientific and Research Publications*, 108 (2013).
18. Jamkhande, P. G., Ghante, M. H. & Ajgunde, B. R. Software based approaches for drug designing and development: A systematic review on commonly used software and its applications. *Bulletin of Faculty of Pharmacy, Cairo University* **55**, 203-210 (2017).
19. Qin, X. Oxidative and electrophilic structural modification and catalytic regulation of human hydroxysteroid sulfotransferase 2a1 (hsult2a1). (2012).
20. adav, D. K., Rai, R., Pratap, R. & Singh, H. Software and web resources for computer-aided molecular modeling and drug discovery. *Chemometrics Applications and Research: QSAR in Medicinal Chemistry* **33** (2016).
21. DHONGADE, S. APPLICATION OF PASS AS AN EFFECTIVE DRUG DESIGNING TOOL. *Reviews of Literature• Volume* **1** (2013).
22. Morel, P. Gramm: grammar of graphics plotting in Matlab. *Journal of Open Source Software* **3**, 568 (2018.
23. Chowdhury, A. *et al.* International Journal of Pharmaceutical Chemistry.
24. Jayaram, B. *et al.* in *BMC bioinformatics.* S7 (Springer).
25. Heinz, H., Lin, T.-J., Kishore Mishra, R. & Emami, F. S. Thermodynamically consistent force fields for the assembly of inorganic, organic, and biological nanostructures: the INTERFACE force field. *Langmuir* **29**, 1754-1765 (2013).
26. Lee, J. *et al.* CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *Journal of chemical theory and computation* **12**, 405-413 (2016).
27. https://www.schrodinger.com/maestro
28. http://www.arguslab.com/arguslab.com/ArgusLab.html
29. https://arguslab.en.softonic.com/
30. http://vakser.compbio.ku.edu/resources/gramm/grammx/
31. http://www.pharmaexpert.ru/passonline/
32. http://genexplain.com/pass/
33. http://www.scfbio-iitd.res.in/sanjeevini/sanjeevini.jsp