Available online on 15.06.2025 at http://jddtonline.info

Journal of Drug Delivery and Therapeutics

Open Access to Pharmaceutical and Medical Research

Open Access Full Text Article

Review Article

Computational Approaches to Molecular Docking and Protein Modeling in Drug Discovery

Monali Jagtap *, Ghanshyam Girnar, Vanshika Ahuja

Department of Pharmaceutics, R. C. Patel Institute of Pharmacy, Shirpur, Dhule Maharashtra, India 425405.

Article Info:

Article History:

Received 23 March 2025
Reviewed 09 May 2025
Accepted 02 June 2025
Published 15 June 2025

Cite this article as:

Jagtap M, Girnar G, Ahuja V, Computational Approaches to Molecular Docking and Protein Modeling in Drug Discovery, Journal of Drug Delivery and Therapeutics. 2025; 15(6):278-287 *DOI: http://dx.doi.org/10.22270/jddt.v15i6.7212*

*Address for Correspondence:

Monali Jagtap, Department of Pharmaceutics, R. C. Patel Institute of Pharmacy, Shirpur, Dhule Maharashtra, India 425405.

Abstract

Protein modeling and molecular docking are crucial computational methods in contemporary drug discovery. To identify potential therapeutic possibilities with high affinity and specificity, molecular docking predicts the ideal binding interactions between tiny molecules (ligands). The best binding interactions between target macromolecules, such proteins, and tiny molecules, or ligands, are predicted by molecular docking. When experimental structures are not accessible, protein modeling—including homology modeling and ab initio techniques—allows for the creation of three-dimensional protein structures. By cutting down on time and expense, these methods work together to expedite the drug discovery process. related to experimental techniques. This review explores the principles of molecular docking, emphasizing key algorithms, scoring functions, and software tools like AutoDock Vina and Discovery Studio. Additionally, it highlights advancements in protein modeling approaches, such as AlphaFold and comparative modeling, and their integration with docking workflows. By using these computational approaches, researchers can effectively predict binding mechanisms, find lead compounds, and improve drug design. The growing increases integration between molecular docking, protein modeling, and artificial intelligence holds promise for more accurate predictions and faster drug development processes in the pharmaceutical industry.

Keywords: Molecular Docking; Protein Modeling: AutoDock Vina.

1. Introduction

Molecular docking is an essential technique in structural molecular biology and computer-aided drug creation. The goal of ligand-protein docking is to predict the main binding mechanisms of a ligand with a protein of known three-dimensional structure.[1] is a reputable and well-established computational technique that has been widely used to understand molecular interactions between a natural organic molecule (ideally regarded as a receptor), such as an enzyme, protein, DNA, or RNA, and a synthetic or natural organic/inorganic molecule (referred to as a ligand).[2]

Molecular docking, which entails putting a ligand into the preferred binding site of a target specific region of the DNA or protein (receptor) for both mechanistic research and logical drug design and discovery, is an intriguing paradigm for understanding pharmacological biomolecular-interactions. mostly in a non-covalent way to produce a stable complex with improved efficacy and selectivity. The main objective of molecular docking is to produce a ligand-receptor complex with an ideal shape and a reduced binding free energy.[3]The behavior of small molecules at target protein binding sites is described by molecular docking.

The method aims to predict the ligand-protein affinity and identify the appropriate locations of ligands within a protein's binding pocket.

The technique seeks to determine the proper positions of ligands within a protein's binding pocket and forecast the ligand-protein affinity.

- Protein–small molecule (ligand) docking
- Protein–nucleic acid docking
- Protein-protein docking

1.1. Principle of Molecular Docking:

ligand-receptor binding requires that the geometry, electrostatics, hydrogen bonding, and hydrophobic interactions of the ligand and receptor all be complementary.[4] Docking is the process of placing molecules in the optimal positions to interact with a receptor.[5] Docking is the mechanism that takes place within a cell when molecules bind together to form a sustained complex.

1.2. Types of Molecular Docking:

I. Rigid:

In this docking, the ligand and receptor molecules are both fixed. In order to find a 3D space conversion for one of the molecules that will best match the other molecules according to a score function, docking is done. Both in the presence and absence of receptor binding activity, the ligand's conformation can be produced.[6]

II. Flexible:

Both the ligand and the receptor can be moved. in a flexible docking. It is versatile in terms of conformation. Energy is computed for every spin. Every conformation surface cell's occupancy is calculated. Next, the best binding position is selected.

2. Software of Molecular Docking

2.1. Dock:

The UCSF Chimera team developed a molecular docking tool called Dock. It is a simple tool that can dock small compounds into receptor binding sites. Dock uses a grid-based system to evaluate the affinity of ligand-receptor binding. Furthermore, it possesses characteristics for ranking and scoring. the locations formed during docking. Some of the input files are PDB, MOL2, and SDF.

2.2 Autodock:

The Scripps Research Institute developed Autodock, a well-known molecular docking tool. This open-source, free application can be used for both rigid and flexible docking.

Autodock uses a Lamarckian genetic algorithm to maximize ligand dispersion inside a receptor binding region. approach. It also offers several scoring elements, such as PDB, MOL2, and SDF, to evaluate ligands' affinity for binding to the receptor.

2.3 GOLD

A protein-ligand docking program called GOLD has a number of important characteristics. It employs user-defined scoring functions that can be adjusted based on the computations' spine and side chain flexibility. Both conformational and non-reinforced contact information serve as the foundation for their energy functions. A number of docking methods are also available, such as the removal of crystallographic water molecules from the ligand binding site. Furthermore, if metal atoms are properly configured in the protein data file, GOLD can manage the automatically. Lastly, the companion applications SILVERTM or Gold Mine TM can be effectively used to examine and post-process the results of high-throughput virtual screening. Three elements are included in their most recent edition of GOLD Suite 5.2: Gold Mine 1.5, Hermes 1.6 for thorough protein imaging, and Gold 5.2 for protein-ligand docking.

2.4 MolDock:

This docking program docks tiny compounds into a receptor-binding site quickly and effectively. MolDock assesses the ligand-receptor binding affinity using the fast Fourier transform (FFT) technique. It incorporates a scoring system that takes into account the van der Waals forces, electrostatic interactions, and form complementarity between the ligand and the receptor. PDB, MOL2, and SDF are among the input file formats that MolDock supports[7].

General steps for molecular docking:

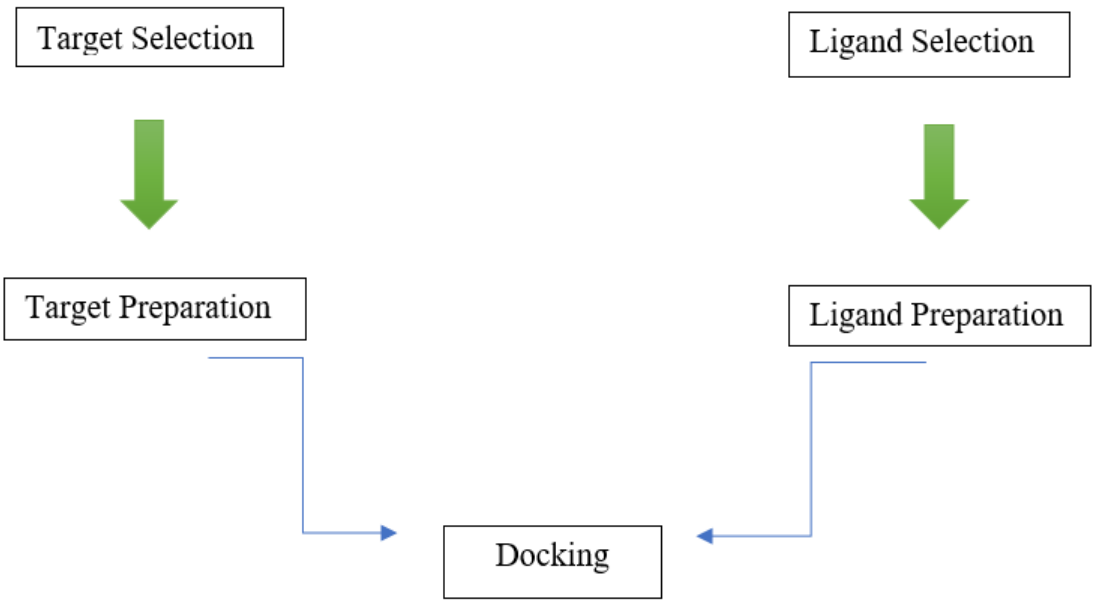

Figure 1: Common work flow chart for docking

3. Method of molecular docking

Molecular Docking: AutoDock Vina was used for the molecular docking experiments, which were aided by a number of computational tools for post-docking analysis, docking, and structure building. The intended target Protein Data Bank (PDB) provided the protein structure, and Open Babel was used to build and optimize the ligand structures. Using ChimeraX, PyMOL, and Discovery Studio, the docking data was visualized and examined.

Table 1: List of Software uses in Molecular Docking

Sr.No	Software's	Application
1	Autodock Vina	Docking
2	Discovery Studio	Visualization tools
3	Open Babel	Files Transferred and Preparation of ligand
4	PyMol	Visualization and Interaction Study
5	ChimeraX	Visualization tools

3.1. AutoDock Vina:

This popular molecular docking software is available for free. It is employed to predict the locations of ligand-protein attachments. It also evaluates how well the ligand and protein interact using the scoring system.

3.2. Open Babel:

This chemical instrument is openly accessible. It allows for smooth format conversion for molecular modeling and docking investigations by supporting a large variety of chemical file formats.

3.3. Discovery Studio Visualizer:

Studio Visualizer is a comprehensive visualization software used to analyze and visualize molecular structures and docking results. It provides an interface to explore molecular interactions, assess binding sites, and understand the structural basis of ligand-protein interactions.

3.4. PyMOL:

PyMOL is a molecular visualization software widely used in structural biology and drug discovery. It allows for detailed visualization of molecular structures, including proteins and ligands, in three-dimensional space. These tools collectively provide essential functionalities for molecular docking

4. Procedure:

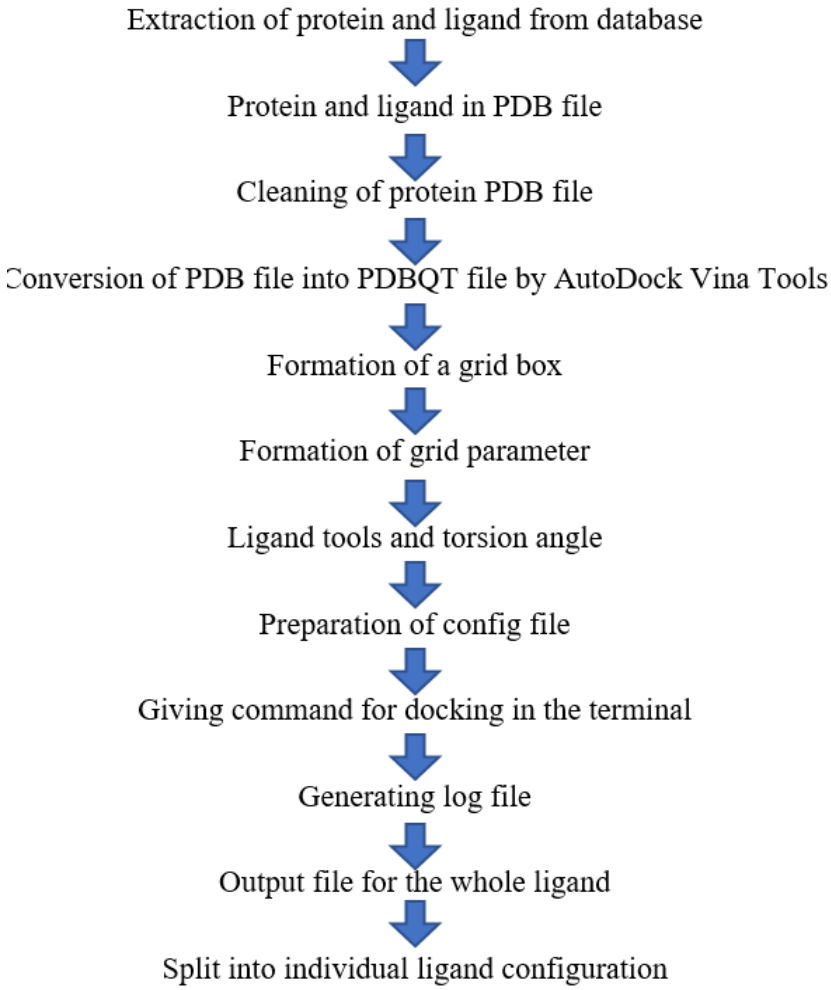

Figure 2: Steps of Molecular Docking

4.1. Extraction of Protein and Ligands:

We can use various freely available databases to extract proteins. Some examples of these databases include UniProt, PDB, and NCBI. And for ligands we can use the zinc database PubChem.

4.2. Conversion ligand of into PDB File:

These are the basic steps for the conversion of a ligand (SDF) File into a PDB file:

1) Ligand extracted from the PubChem or Zinc Database.
2) Ligand can be downloaded in the form of SDF format.
3) The SDF format file was converted into a PDB file using Open babel software.

4.3. Cleaning of Protein

We clean the protein by using Discovery Studio Visualizer Tool Software.

- ✓ Delete H2O molecules
- ✓ Delete heteroatom
- ✓ Add H- Bond
- ✓ Add Charge

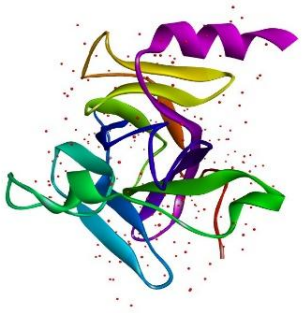

Figure 3: Original 4mzi Protein

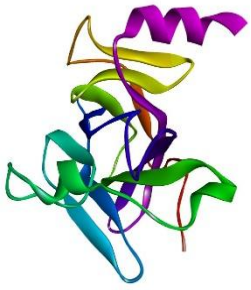

Figure 4: Clean 4mzi Protein

The interface of AutoDock Vina Software

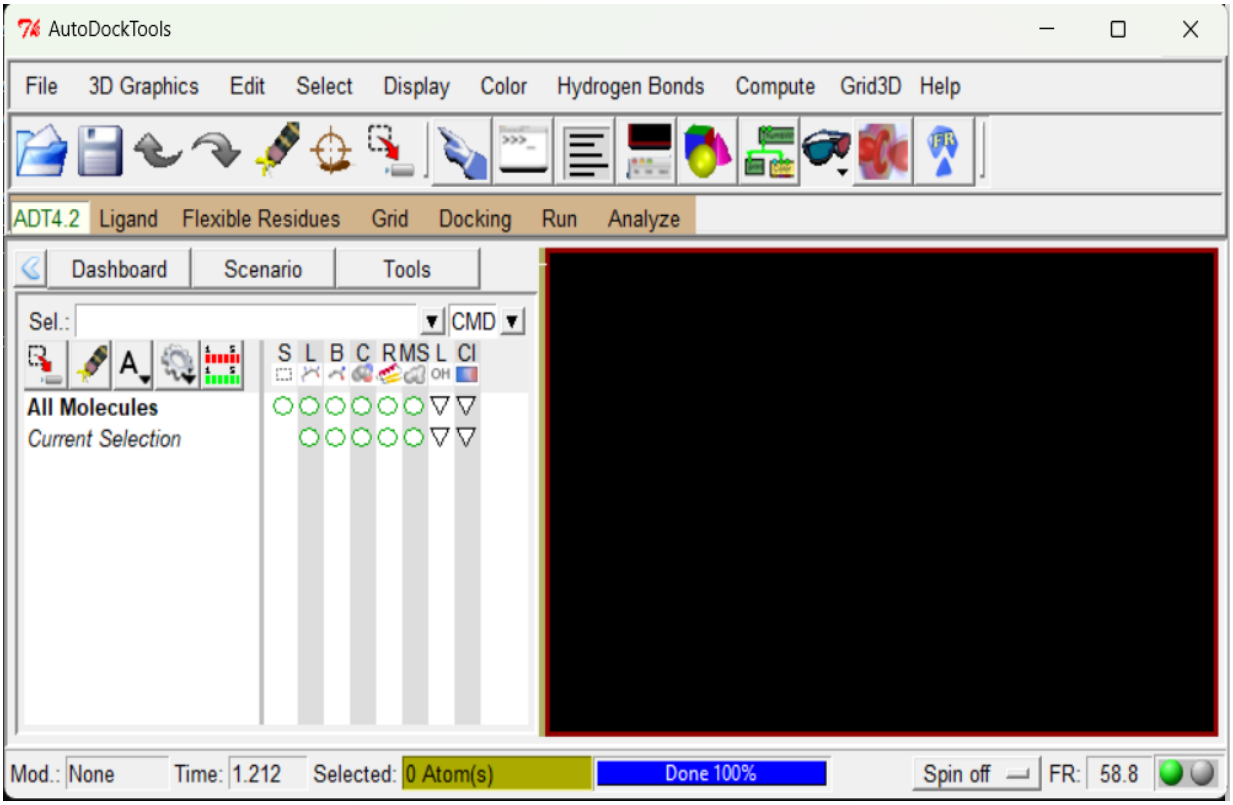

Figure 5: Interface of Autodock Vina

4.4. Conversion of PDB file into PDBQT Format File

Conversion of PDB file into PDBQT format by using **AutoDock Vina tool**

Steps:

Following are the basic steps for conversion of the PDB file into a PDBQT file.

4.5. Formation Grid Box

AutoDock Vina tools adjust the grid box according to its binding site.

Types of docking:

I. **Blind docking**:

In this docking method, the active site is unknown; therefore, the entire protein is adjusted inside the grid box.

II. **Specific docking**:

In specific docking, the active site is known. The active site is highlighted so that the active site is covered under the grid box.

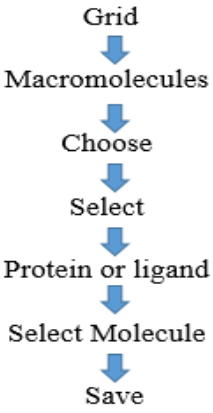

Figure 6: Steps for conversion of PDB to PDBQT O

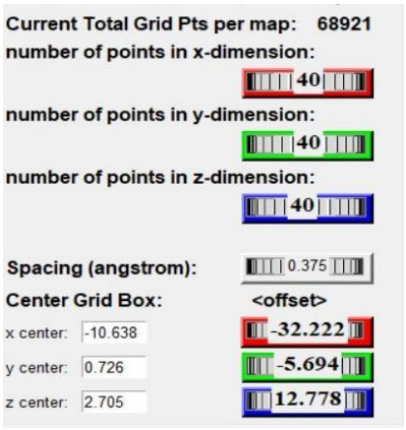

Figure 7: Grid Parameter

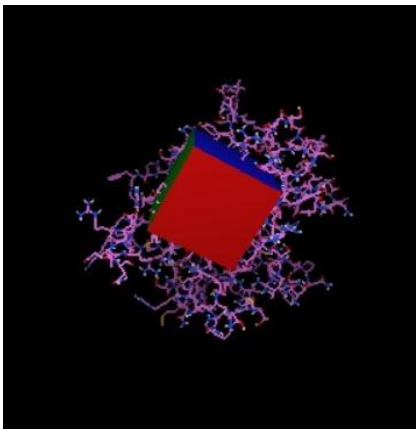

Figure 8: Grid Box

4.6. Ligand tools and torsion angle

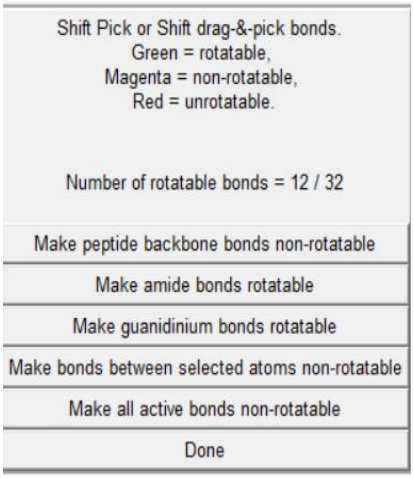

Figure 9: Parameter

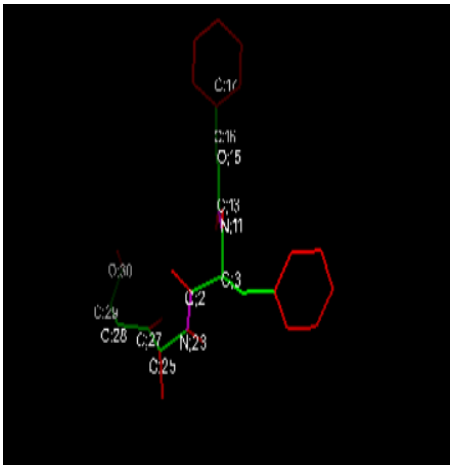

Figure 10: Ligand

4.7. Config File

Example of the config file:

Receptor = protein. pdbqt
ligand = ligand. pdbqt
center_x = -18.961
center_y = -5.167
center_z = 8.644
size_x = 40
size_y = 40
size_z = 40
num_modes = 10

The command for the AutoDock Vina

Docking

.\vina.exe --receptor protein.pdbqt --ligand ligand.pdbqt --config config.txt --log log.txt --out output.pdbqt

each ligand-receptor

.\vina_split.exe --input .\output.pdbqt

4.8. Log file:

Log files are generated by command prompts with the help of vina.exe and vina.split.exe files.

Ex..

mode	affinity | dist from best mode (kcal/mol)	rmsd l.b.| rmsd u.b.	
1	-6.7	0.000	0.000
2	-6.4	14.064	17.209
3	-6.3	26.187	29.737
4	-6.2	1.888	4.989
5	-6.1	16.680	20.337
6	-5.9	25.780	29.613
7	-5.9	16.915	19.713
8	-5.9	26.021	29.863
9	-5.9	13.232	17.293
10	-5.8	15.793	17.747

- RMSD values are inversely proportional to the binding affinity of protein and ligand.
- The more negative values, more the binding affinity.

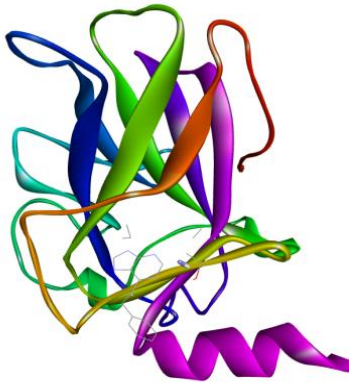

Figure 11: Protein-Ligand

5. Protein-Ligand Complex:

Protein- Severe Acute Respiratory Syndrome Related coronavirus (2OP9)

Ligand-Nalpha-[(benzyloxy)carbonyl]-n-[(1r)-4-hydroxy-1-methyl-2-oxobutyl]-l-phenylalanine amide (WR1)

- The 2OP9 protein was extracted from the PDB database, and the molecular docking steps were followed, as shown earlier for the docking.
- Cleaning of protein and ligand PDB file using Discovery Studio visualizer tool. Identification of the active site of protein via Discovery studio visualizer.
- Generating a sphere for the active site in the protein. Separating the amino acid present in the sphere from the whole protein.
- These separated residues are used to interact with proteins and ligands.
- Highlight the separated amino residue, adjust the grid box, and perform the docking.[8]

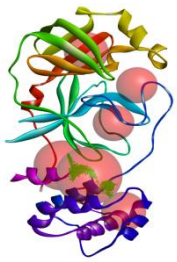

Figure 12: Active sites

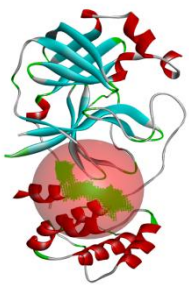

Figure 13: Site 1

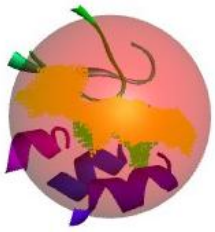

Figure 13: Amino Residue

Highlight the separated amino residue, adjust the grid box, and perform the docking after giving the command, then get the log file.

Log File: mode | affinity | dist from best mode

| (kcal/mol) | rmsd l.b.| rmsd u.b.

-----+------------+----------+----------

1 -6.2 0.000 0.000

2 -6.1 2.389 6.807

3 -5.9 2.135 4.152

4 -5.9 2.682 6.857

5 -5.9 23.381 26.159

6 -5.8 20.748 23.552

7 -5.8 21.455 24.143

8 -5.8 22.328 25.17

9 -5.7 2.066 4.842

10 -5.7 2.288 6.46

RMSD values are inversely proportional to the binding affinity of protein and ligand. The more negative value, the more the binding affinity

6. Application of Molecular Docking:

Molecular docking can demonstrate the viability of biological processes prior to the experimental stage of any inquiry. Several of molecular docking's most important applications

6.1. Optimization of lead:

A ligand's ideal orientation on its target can be predicted using molecular docking. It can forecast several ligand binding mechanisms in the target molecule's groove. This can be applied to create therapeutic candidates that are more effective, selective, and powerful.[9]

6.2. Screening of ligand

Virtual screening is used to find hits and lead compounds in molecular databases using scoring methods. The usage of docking in virtual screening have increased when the method is coupled with other cutting-edge applications. For example, combining free energy with dynamics of molecules.

7. Protein Modeling:

A protein's three-dimensional form can be inferred from its amino acid sequence using protein structure prediction.[10] Two complementary approaches that concentrate on physical interactions or evolutionary history have been used in the development of computer techniques to predict three-dimensional (3D) protein structures from the protein sequence.[11]

A polymeric macromolecule made up of linearly ordered amino acid building blocks connected by peptide bonds is called a protein. The main structure of a protein is its linear polypeptide chain. Usually, a series of letters across a 20-letter alphabet that corresponds to the 20 naturally occurring amino acids represents the fundamental structure.[12] The objectives of protein-structure prediction range from the intellectual challenge of figuring out an intriguing riddle to the variety of uses that can be made possible once precise prediction is achieved. With very few instances up until recently, protein structure prediction was of greater conceptual than practical significance because forecasts were rarely precise enough to support the structure-based design of novel drugs or to infer biological function, for instance.[13] Structural modeling has greatly benefited from advances in computational tools and methods. Methods for predicting protein structures include the ab initio approach, threading, and comparative (homology) modeling. A number of software programs and tools have been created for protein 3D modeling. [14]

7.1. Material and Methods:

The three-dimensional structure was predicted and analyzed using a combination of computational tools. The initial protein sequence was retrieved from the UniProt database, which served as the input for structure prediction.

Table 1: Computer tools used for Structure prediction

Sr No	Computational tools/ software	Application
1	Swiss Model	3D structure prediction
2	AlphaFold	3D structure prediction
3	AlphaFold Colab	Complex structure prediction

7.2. Swiss Model:

Using a comparison method and a database of annotated models for important references, Swiss Model is an automated platform for protein structure homology modeling that creates 3D model.

UniProtKB-based proteomes.

The four phases listed below comprise all homology-modeling techniques:
(i) Choosing a template

(ii) Aligning the target template

(iii) Creating a model

(iv) Assessing.
Until a satisfactory model structure is obtained, these procedures can be repeated.

7.3. AlphaFold:

Amino acid sequences can now be used to reliably predict the three-dimensional structures of proteins. Independent assessments of prediction accuracy show how well AlphaFold, a deep learning-based approach to protein structure prediction, performs.

Interface of AlphaFold:

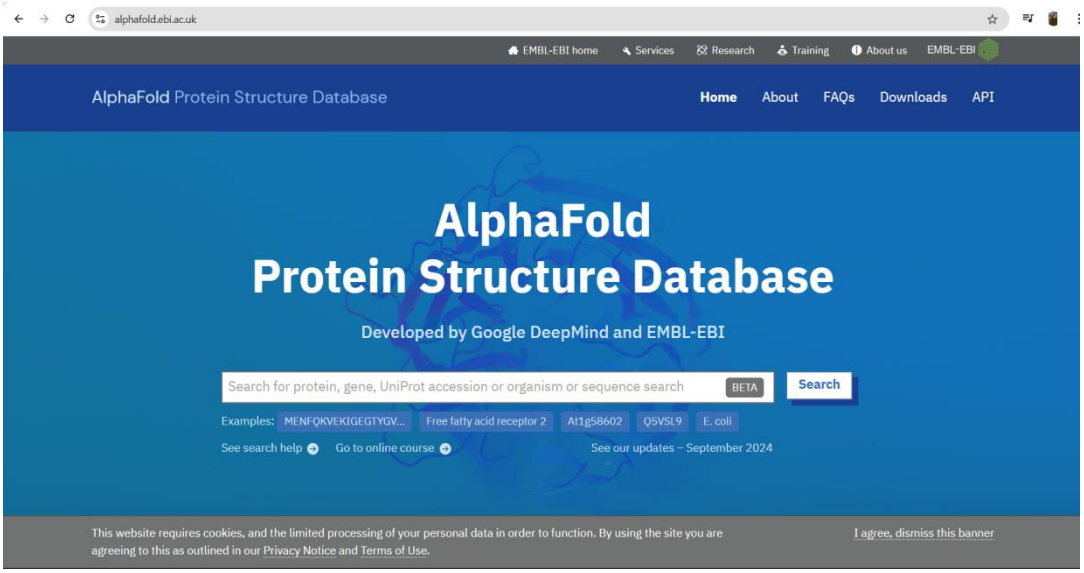

Figure 14: Interface of AlphaFold

How AlphaFold Work

- ✓ Extract the amino sequence in FASTA format from the UniPort database.
- ✓ Download amino sequence.
- ✓ Upload into AlphaFold protein prediction server.
- ✓ Get different templates that are similar to the target protein structure
- ✓ Select the template that contains more than 70% of the identity.
- ✓ Build the Model
- ✓ Analysis of the accuracy of protein structure by confidence score.
- ✓ Download the 3D model structure in PDB format
- ✓ Use this structure in research

7.4 Output of AlphaFold:

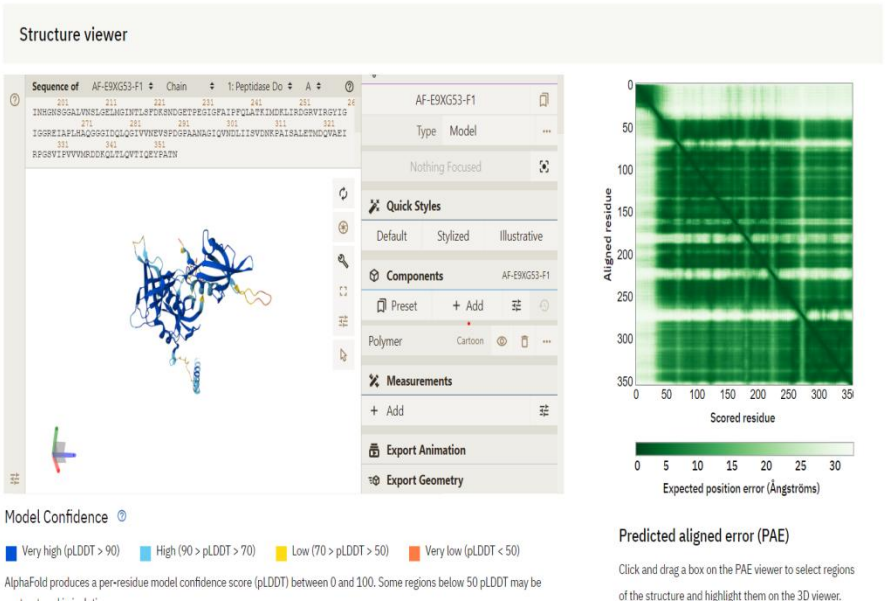

Figure 15: output of AlphaFold

7.5 Protein Complex Structure Prediction:

7.5.1 ChimeraX: AlphaFold Colab

ChimeraX is a powerful software tool for visualizing and analyzing molecular structures, and it includes several features for structure prediction and analysis, like AlphaFold.

AlphaFold Database and Structure Prediction

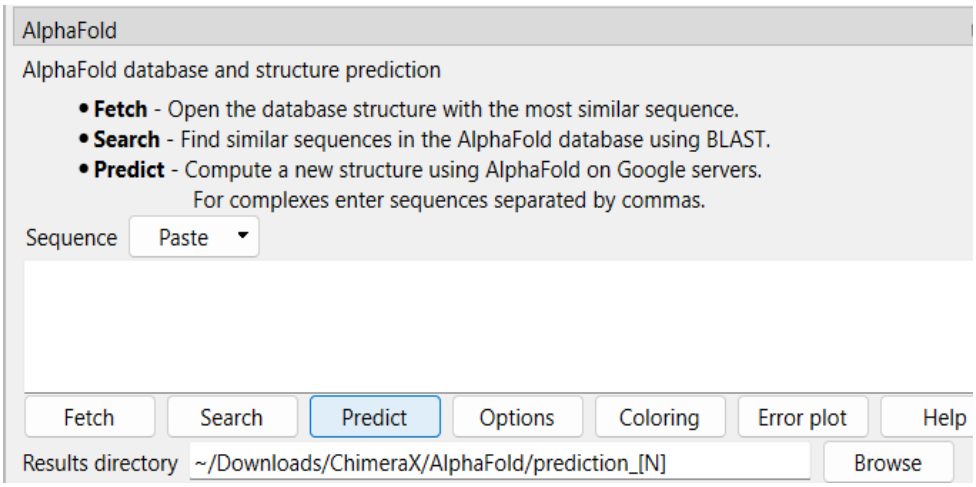

Figure 16: Interface of ChimeraX

7.5.2 Running AlphaFold:

After obtaining the amino sequence of different proteins from UniProtKB or other databases. These sequences paste and predict the structure. Then, run the AlphaFold. The given directory downloads the results.

7.5.3 Results and Discussion:

AutoDock Vina was used to molecularly connect the ligand WR1 with the SARS-CoV-2 major protease (PDB ID: 2OP9). Nine distinct binding modes, arranged by binding affinity, were generated by the docking simulation.

Table 2: Binding Affinity and RMSD of Top Docking Poses of WR1 with SARS-CoV-2 Protease (2OP9)

Mode	Binding Affinity (kcal/mol)	RMSD Lower Bound (Å)	RMSD Upper Bound (Å)
1	-6.2	0.000	0.000
2	-6.1	2.389	6.807
3	-5.9	2.135	4.152

A stable interaction between WR1 and the target protease was indicated by the lowest binding energy of -6.2 kcal/mol that was recorded. Using PyMOL and Discovery Studio, the optimal binding pose (mode 1) was shown, revealing particular molecular interactions. A relatively strong interaction with the target protease is suggested by WR1's binding affinity of -6.2 kcal/mol. A binding energy of less than -6.0 kcal/mol is typically regarded as advantageous for lead optimization in molecular docking.

The idea that WR1 might function as a protease inhibitor is supported by these interactions. WR1's binding mechanism and pattern of interaction are similar to those of peptidomimetic inhibitors that target the SARS-CoV-2 protease's active site cleft. Furthermore, it was shown that WR1 functions identically even when docked with crystal-verified structures by the similar binding scores of current AI-predicted ligands evaluated with AlphaFold-modeled protease structures.

8. Conclusion:

Drug development has been transformed by molecular docking and protein modeling, which make it possible for rational drug design, mechanistic study, and virtual screening. AutoDock Vina, AlphaFold, and Discovery Studio are some of the tools that greatly speed up the identification of lead compounds. Drug development procedures will be streamlined and prediction accuracy increased with increasing AI and machine learning integration.

Acknowledgements: We thank Ghanshyam Girnar for her advice and immense insights while writing this review article.

Authors' contributions: Monali Jagtap – draft writing,

Vanshika Ahuja – draft writing, Ghanshyam Girnar – Supervision.

Funding source: There is no funding source.

Conflict of interest: The authors reported no conflict of interest.

Ethical Approval: Not applicable

References

1. Morris GM, Lim-Wilby M. Molecular docking. In: Kukol A, editor. *Methods in molecular biology*. Clifton NJ: Humana Press; 2008. p. 365–82. https://doi.org/10.1007/978-1-59745-177-2_19
2. Mohanty M, Mohanty PS. Molecular docking in organic, inorganic, and hybrid systems: a tutorial review. *Monatsh Chem*. 2023;1–25. https://doi.org/10.1007/s00706-023-03076-1
3. Dar AM, Mir S. Molecular docking: approaches, types, applications and basic challenges. *J Anal Bioanal Tech*. 2017;8(2):1–3. https://doi.org/10.4172/2155-9872.1000356
4. Creative Proteomics. Principles, processes and types of molecular docking [Internet]. [cited 2025 Feb 3]. Available from: https://www.iaanalysis.com/principles-processes-and-types-of-molecular-docking.html
5. Keval R, Tejas G. Basics, types and applications of molecular docking: a review. *IP Int J Compr Adv Pharmacol*. 2022;7(1):12–6. https://doi.org/10.18231/j.ijcaap.2022.003
6. Rudimentary review on molecular docking: a beginner's guide [Internet]. [cited 2025 Feb 12]. Available from: https://www.researchgate.net/publication/374588198_Rudimentary_Review_on_Molecular_Docking_A_Beginner's_Guide
7. Agu PC, Afiukwa CA, Orji OU, Ezeh EM, Ofoke IH, Ogbu CO, et al. Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management. *Sci Rep*. 2023;13(1):13398. https://doi.org/10.1038/s41598-023-40160-2
8. Goetz DH, Choe Y, Hansell E, Chen YT, McDowell M, Jonsson CB, et al. Substrate specificity profiling and identification of a new class of inhibitor for the major protease of the SARS coronavirus. *Biochemistry*. 2007;46(30):8744–52. https://doi.org/10.1021/bi0621415
9. Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, Green T, et al. Improved protein structure prediction using potentials from deep learning. *Nature*. 2020;577(7792):706–10. https://doi.org/10.1038/s41586-019-1923-7
10. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583–9. https://doi.org/10.1038/s41586-021-03819-2
11. AlQuraishi M. Machine learning in protein structure prediction. *Curr Opin Chem Biol*. 2021;1–8. https://doi.org/10.1016/j.cbpa.2021.04.005
12. Al-Lazikani B, Jung J, Xiang Z, Honig B. Protein structure prediction. *Curr Opin Chem Biol*. 2001;5(1):51–6. https://doi.org/10.1016/S1367-5931(00)00164-2
13. Agnihotry S, Pathak RK, Singh DB, Tiwari A, Hussain I. Protein structure prediction. In: Singh DB, Pathak RK, editors. *Bioinformatics*. Academic Press; 2022. p. 177–88. https://doi.org/10.1016/B978-0-323-89775-4.00023-7
14. https://www.rcsb.org/
15. https://vina.scripps.edu/
16. https://discover.3ds.com/discovery-studio-visualizer-download
17. https://www.pymol.org/
18. https://sourceforge.net/projects/openbabel/
19. https://swissmodel.expasy.org/
20. https://github.com/android/camera-samples
21. https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb
22. https://alphafold.ebi.ac.uk/