

A Review On Advances in Molecular Docking : Methods, Challenges, and Future Direction

Avanti A Yelpale, Sarfaraz M Kazi, Sanjay K Bais

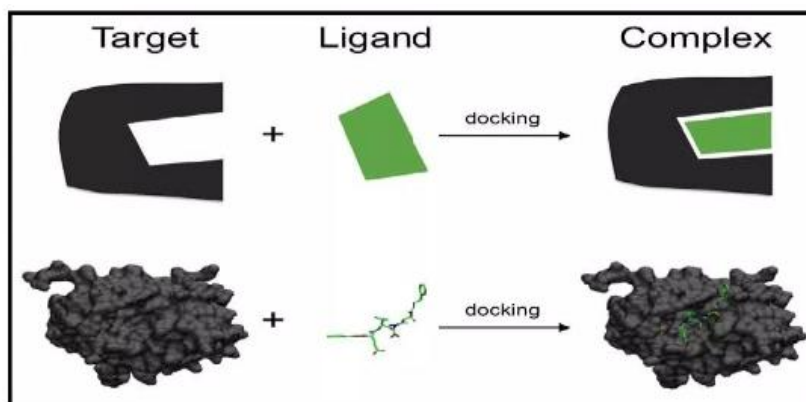
Fabtech College of Pharmacy, Sangola

Abstract: *Molecular docking is a computational technique used to model structural complexes composed of two or more interacting atoms. Its main objective is to predict the three-dimensional structure of a molecule when it interacts with another molecule of interest. This method is widely used in drug development, where access to structural databases has become an essential tool for researchers. Docking software provides valuable, often costly instruments for drug analysis and design, enabling simple molecular predictions and efficient navigation of structural data. These tools have become crucial elements on the modern scientist's desktop. Among the various applications of molecular docking, virtual screening is perhaps the most significant, allowing researchers to examine large libraries of compounds rapidly. Numerous docking applications are available to visualize the three-dimensional structures of molecules, and different computational methods can be employed to analyze docking results and evaluate binding performance.*

Keywords: *Molecular docking.*

I. INTRODUCTION

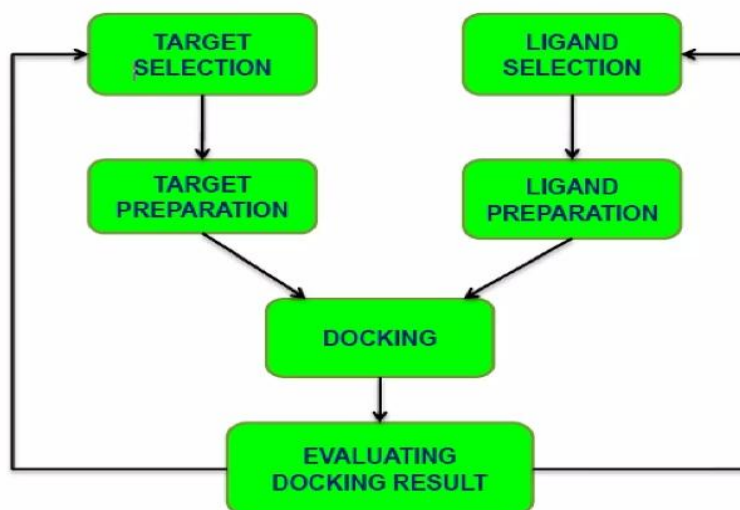
The tiny molecule's experimental binding mechanism and affinity inside the target receptor's binding site. Pose prediction, virtual screening, and binding rate assessment are the small molecules experimental binding procedure and affinity at the intended receptor three principal interrelated Pose prediction, virtual screening, and binding affinity estimation are the three primary interconnected objectives of molecular docking (Jain and Nicholls 2008). The native ligand posture at the receptor binding site (i.e., to determine the real ligand geometry within a specific tolerance limit) and the corresponding physical-chemical molecular interactions must be accurately predicted using a successful docking methodology. Additionally, the method must be able to accurately rank these ligands among the best compounds in the database and reliably identify binding from nonbinding molecules when examining vast compound libraries (Kolb and Irwin 2009).



The core components of a docking methodology for creating and assessing the ligand conformations are an energy scoring function and a search algorithm. The development of predictive docking methodologies that are helpful in future



drug design research depends on the capacity to effectively manage a system's inherent molecular flexibility and accurately represent the energetics of receptor–ligand interactions. The search algorithms and scoring functions most frequently employed in modern molecular docking techniques that concentrate on protein–ligand applications are presented in this review. We try to provide an overview of the key concepts and current methodological and computational developments in protein–ligand docking. In the context of a docking-based study, we additionally take into account methods for incorporating protein flexibility and tactics meant to enhance binding affinity prediction.



COMPUTER-BASED DRUG DESIGN

Computer Aided Drug Design, as well as or CADD, is a technique based on computers used in computational chemistry to find, improve, or analyze drugs and related physiologically engaged molecules. New medication design indicates that it is most helpful. It offers information on the biological and pharmacological characteristics of targets and ligands. It is employed to discover and enhance new drugs. The invention of in-silico filters for predicting undesirable traits of toxic molecules, such as poor pharmacokinetics and toxicity. New drug targets are optimized using it. To locate hits, CADD is utilized. New pharmaceutical substances are examined virtually by employing chemical scaffolds.

DRUG DESIGN BASED ON STRUCTURE

In order to compute interaction strengths for each tested compound, structure-based computer-aided drug design relies on a comprehending of the goal protein structure. There are crystallized target proteins in the structures database. The goal of structure-based design is to create compounds that bind to the target precisely and persistently with the smallest possible amount of energy. A computer-based screening technique allowing the screening of a sizable library of related chemical compounds for a specific biological activity is known more broadly as virtual high-throughput screening. There are several types of virtual high-throughput screening, such as chemical similarity search and quantitative structure-based drug selection based on projected biologic activity.

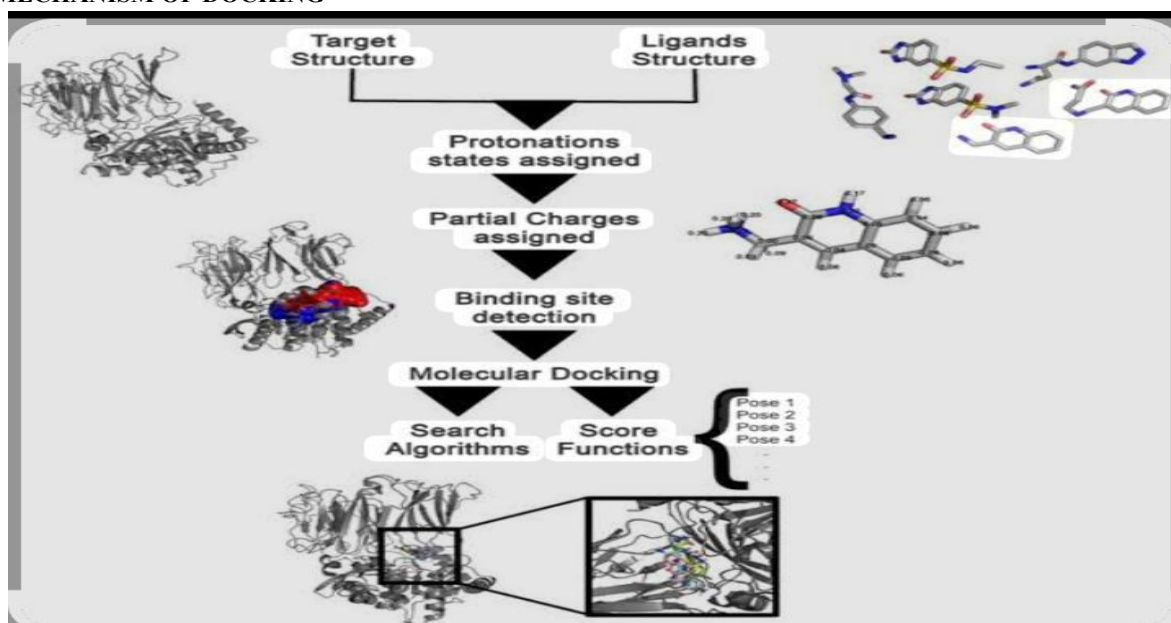
Molecular docking types

Search Algorithm: The experimentation approach establishes the number of possible combinations and binding modes. The Monte Carlo approach, fragment and genetic-based structured searches, and docking assessment are used a Stiff



Docking b. Adaptable Docking. Rigid Docking: Both both the receptor and the ligand molecules are fixed in this docking. Docking is carried out. Flexible Docking: This type of docking allows both the ligands and the receptor to move. It is versatile in terms of conformation. The energy is computed for each rotation. The occupancy of each conformation surface cell is computed. Docking of molecules. Basic molecular docking. The goal of the docking process is to use computer-based techniques to anticipate the ligand-receptor complex²⁷. The two primary processes in the docking procedure are sampling the ligand and applying a scoring function²⁸. By considering the ligand's attachment mode, sampling algorithms assist in determining the ligand's most energetically advantageous conformations within the protein's active region. A score system is then used to rank these confirmations.^{7, 28} algorithms for searching. Finding every conceivable orientation and configuration of the protein in conjunction with the ligand²⁸ is the main goal of the search method.

MECHANISM OF DOCKING



Organizing the protein of interest is the first step towards creating a docking screen. Using a biophysical technique like x-ray crystallography or, less frequently, NMR spectroscopy to identify the structure has usually been established. The protein a docking agenda receives organization and a folder containing ligands as input . The search algorithm and scoring function are two methods that determine a docking program's success. All possible orientations and conformations of the protein paired with the ligand make up the investigation space. It is impossible to fully explore the space that would list every possible distortion of every molecule and every likely rotational and translational orientation of the ligand in relation to the protein at a predetermined level of granularity with current computing capabilities. The majority of docking programs currently in use take into consideration bendable ligands, and several are making an effort to mimic a flexible protein receptor. The process of studying the intermolecular interaction between two molecules in silica is known as molecular docking.

The protein transporter functions as the macromolecule in this enhancement. The Ligand 5 molecule, which can function as an inhibitor, is the tiny particle.



Key Steps in Molecular Docking Mechanics

Thus, the following phases are involve in the docking process

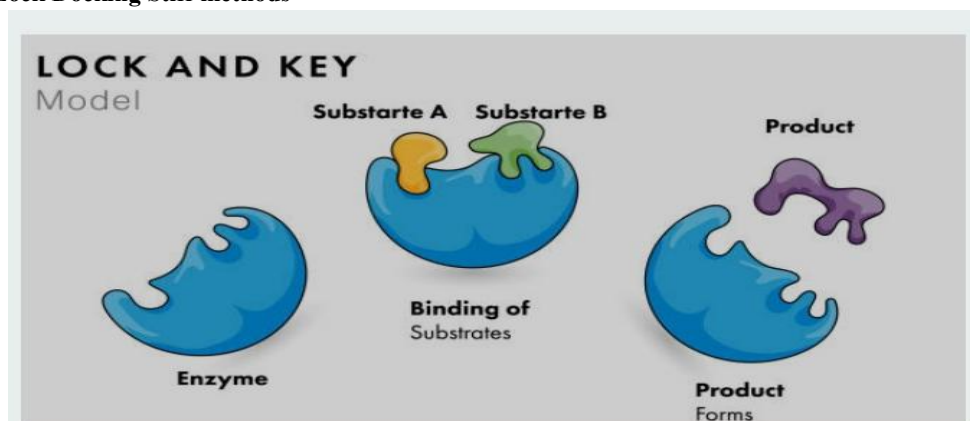
Step I: protein preparation The protein's three-dimensional structure needs to be retrieved from the Protein Data Bank (PDB) and then pre-processed. Depending on the available parameters, this should allow the water molecules to be amputated from the cavity, stabilize the charges, produce side chains, and substantially replace the missing residue.

Active site prediction in Step II

The protein's active site needs to be anticipated after it has been prepared. There are several active sites in the receptor strength; only the one that is of concern should be selected. Heteroatoms and water molecules are typically ignored if they exist.

Step III – preparation of ligand Ligand can be retrieve from numerous databases such as ZINC, Pub Chem. or can be sketched apply Chem. sketch tool. While picking out the ligand, the LIPINSKY'S RULE OF 5 should be utilizing. Lipinski rule of 5 assist in discriminating among non-drug like and drug like. The computer aided drug design and detection (CADD) method. It promises high possibility of achievement or failure due to drug likeness for molecules remaining by with 2 or more than of the complying rules. For choice of a ligand allow to the Lipinsky's rule. A lesser amount of five hydrogen bond donors. A lesser amount of ten hydrogen bond acceptors. Molecular mass less than 500 Da High lipophilicity (expressed as Log not over 5) Molar refractivity should be between 40 130. Step IV Docking: Ligand is docked alongside the protein and the interactions are analyzed.

Key and Lock Docking Stiff methods



For medicinal chemists, the possibility of highly specific interactions during molecular recognition events is nothing new. Fischer (1894) and Ehrlich (1909) first proposed analogies to a "lock-and-key" concept to explain these processes a century ago (see also Lichtentaler, 1994). We may approximate the receptor and the ligand as rigid molecules because this model requires an exact matching of immutable components, which has implications for molecular docking. By lowering the number of degrees of freedom from many thousand to just six, this significantly simplifies the docking problem. The lock-and-key idea is at one extreme of the range of techniques for simulating dynamic processes. For molecular docking studies, what is the value of such a reduced model?

MOLECULAR DOCKING SOFTWARE

Program for molecular docking.design of a molecular docking program. Molecular docking has been essential in many drug development projects, particularly for the virtual screening of phytochemicals or nutraceuticals as potential medicinal compounds⁷. The first docking program was developed in the mid-1980s by Irwin Kuntz of the University of California, and docking computations are continuously being improved. In order to predict an enzyme's capacity, current advancements in docking techniques determine the enzyme's natural substrates³⁶. By



determining that the protein of interest belongs to a certain superfamily, protein complexes can be successfully predicted by limiting the search for probable substrates and reaction types to that region³⁷.rating techniques for docked molecules. The docked molecules are carefully ranked using a variety of approaches and systems.

DOCK 3.5.x. Enzymes catalyze processes by limiting the transition state that is favorable to the substrate, according to the theory underlying this program. Docking molecules that match transition states should produce a larger signal than docking substrates since amidohydrolase superfamily hydrolysis processes the protein to retain its stiffness³⁸.Slide. By fine-tuning and rescoring the docked complex with an increasingly complex material science-based scoring capacity and allowing receptor side chains to move, the program improves positioning precision by identifying the enzymes that belong to a particular subgroup of the enolase superfamily, allowing tapering the arrangement of potential substrates³⁹.Molecular docking software highlights.here are numerous docking applications available; some of the most well-known ones are included in this section.Field-based force empirical role function of scoring consensus based on knowledge.

APPLICATION

An enzyme may be activated or inhibited as a result of a binding communication between a small molecule ligand and an enzyme protein. Ligand binding may have agonist or antagonist effects if the protein is a receptor. Drug design is the primary application of docking. Since the majority of medications are tiny organic compounds, docking can be used for:

Identification of Hits

Large databases of possible medications can be quickly screened in silico using docking collective with a scoring function to identify molecules that are likely to bind to a target protein.

Optimization of Lead

Vina AutoDockQuick, precise, and frequently used with UCSF Chimera due to its intuitive UI. AutoDock is a set of automated docking technologies that predicts binding using a genetic approach. DOCKOne of the earliest programs, it used a shape-based docking technique.

GOLDA commercial program with a reputation for handling flexible proteins and using a genetic algorithm.SchrödingerA whole range of commercial tools for molecular dynamics, virtual screening, and other applications.DockingA commercial toolset that provides a range of applications for certain docking requirements, such as induced-fit posing (IFP) or fast exhaustive docking (FRED) Docking, also known as the binding mode or pose, can be used to determine the location and direction of a ligand's binding to a protein. This could then be utilized to create analogs that are more effective and specific. Bioremediation Pollutants that can be broken down by enzymes can also be predicted via protein ligand docking. A bacterial enzyme called DNA gyrase is being researched as an antibacterial target because it unwinds DNA and adds negative supercoils to bacterial DNA. HTS was unable to identify any new DNA gyrase inhibitors. For this enzyme, Boehm et al. employed de novo design and were able to produce a number of novel inhibitors . First, a shared binding pattern—in which both inhibitors supply one hydrogen bond to Asp73 and accept one hydrogen bond from a conserved water molecule—was obtained by closely examining the three-dimensional complex structures of DNA gyrase with known inhibitors, ciprofloxacin and novobiocin. To have a lipophilic interaction with the receptor, the molecule should also contain some lipophilic fragments.

Potential compounds can be found, ligand-protein interactions can be predicted, and structure-activity correlations can be made clearer with the use of molecular docking.

Additionally, it can be utilized to forecast the binding characteristics of medicines and nucleic acids and explain ligand activity Potential binding geometries between a ligand and a target protein with a known three-dimensional structure can be found via molecular docking. The binding affinities of small compounds to the enzymes that degrade environmental contaminants can be estimated using molecular docking. This could aid in the creation of these enzymes' activators or inhibitors to increase the effectiveness of bioremediation.



The following are some crucial factors for medication development and discovery:

Lead Optimization: By forecasting binding affinities and directing changes to increase efficacy and lessen negative effects, it helps optimize lead compound.

Cost and Time Efficiency: Molecular docking lowers the cost and time needed for drug development by ranking molecules for synthesis and testing.

Comprehending Molecular Interactions: By showing how medications attach to their targets, mechanism of action docking studies shed light on how pharmaceuticals work. **Protein operate:** By showing how proteins interact with possible ligands, it aids in the understanding of how proteins operate. **Personalized Medicine:** Based on a person's genetic composition, molecular docking can be used to anticipate how they could react to a medication, resulting in customized treatment regimens.

Scope of Molecular Docking;

The potential and extent of molecular docking research Molecular docking has been widely used in many aspects of marine drug research since it is a crucial method in computer-aided drug design [CADD] and can foresee the binding energy and mechanism of protein–ligand complexes. The following points help to explain future prospects. **Integration with AI and ML:** Improved predictive models that combine docking with AI to find new compounds and increase the accuracy of binding affinity predictions. **Automated Docking Platforms:** The creation of automated platforms for virtual screening at high throughput.

Improvements in Computational Power: Using quantum computing to manage intricate computations and simulations that are beyond the capabilities of traditional computers. **Cloud computing:** Making large-scale docking research more accessible and scalable by utilizing cloud resources. **Improved Scoring Functions:** Create more precise scoring functions to more accurately forecast the stability and binding affinity of protein-ligand complexes. **Multitarget Drug Design:** Polypharmacology addresses complicated diseases including cancer and neurological disorders by creating medications that can interact with numerous targets. **Improved Structural Biology Methods:** CryoEM and NMR Improved structural biology methods will yield more precise and varied target structures, increasing the dependability of docking studies.

II. CONCLUSION

Docking studies continue to face significant challenges due to receptor flexibility, particularly backbone flexibility and movement of various important secondary parts of the receptor involving ligand binding and the catalyst. Certain approaches to side chain flexibility have been shown to be sufficient and successful in specific situations. An ensemble of proteins is a common option for global flexibility that aligns with the conformer selection perspective. It necessitates an effective method for obtaining and choosing trustworthy protein structures for docking, therefore the ensembles should contain structures that the ligand can fit into.

Additionally, this method's computing expense is another drawback. Since loops tend to be more flexible and challenging to simulate using current methods, particularly because of their potentially dramatic motions, LMMC may be a suitable technique for sampling a ligand within loop-containing active sites. The ability to modify the degree of flexibility is an additional benefit. Users can directly control either the side chain or the loop's entire movement.

One essential element of docking that could be further enhanced is the scoring function. Examples of successful applications demonstrate that computational methods are capable of designing new small compounds and screening hits from a large database. However, experimental technology is still necessary to achieve realistic interactions between tiny compounds and receptors.

REFERENCES

1. Jorgensen WL. The many roles of computation in drug discovery. *Science*. 2004; 303(5665): 1813–1818.
2. Bajorath J. Integration of virtual and high-throughput screening. *Nat Rev Drug Discov*. 2002; 1(11):882–894.



3. Walters WP, Stahl MT, Murcko MA. Virtual screening - an overview. *Drug Discov. Today*. 1998; 3:160–178.
4. Langer T, Hoffmann RD. Virtual screening: an effective tool for lead structure discovery? *Curr Pharm Des.* 2001; 7(7):509–527.
5. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov.* 2004; 3(11):935–949.
6. Gohlke H, Klebe G. Approaches to the description and prediction of the binding affinity of small molecule ligands to macromolecular receptors. *AngewChemInt Ed Engl.* 2002; 41(15):2644–2676.
7. Moitessier N, Englebienne P, Lee D, Lawandi J, Corbeil CR. Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go. *Br J Pharmacol.* 2008; 153(Suppl 1):S7–26.
8. Shoichet, BK.; McGovern, SL.; Wei, B.; Irwin, JJ. Hits, leads and artifacts from virtual and high throughput screening. 2002. *Molecular Informatics: Confronting Complexity.*
9. Bailey D, Brown D. High-throughput chemistry and structure-based design: survival of the smartest. *Drug Discov Today.* 2001; 6(2):57–59.
10. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A geometric approach to macromolecule-ligand interactions. *J Mol Biol.* 1982; 161(2):269–288. [PubMed: 7154081]
11. Melzer, J., Brignoli, R. & Saller, R. Komplementärmedizin, phytotherapie und sojaisoflavonealsphytoöstrogene [complementary medicine: Phytotherapy and soyaisoflavones as phytoestrogens]. *Zentralbl. Gynakol.* 126(3), 138–147 (2004).
12. Zhou, J. et al. Epigallocatechin-3-gallate (EGCG), a green tea polyphenol, stimulates hepatic autophagy and lipid clearance. *PLoS ONE*
13. Nitzan-Kaluski, D., Stern, F., Kachel, J. & Leventhal, A. Soy and phytoestrogens consumption and health policy hesitation or certitude. *Harefuah* 141(1), 61–66 (2002).
14. Dwyer, J. Overview: dietary approaches for reducing cardiovascular disease risks. *J. Nutrition* 125, 656S-665S (1995)
15. Cencic, A. & Chingwaru, W. The role of functional foods, nutraceuticals, and food supplements in intestinal health. *Nutrients* 2(6), 611–625 (2010).
16. Martino, A. et al. Diets and heart disease. Myths and reality. *J. Nutr. Health Food Sci.* 4(2), 1–10 (2016)
17. Ashraf, S. A. et al. Cordycepin for health and wellbeing: A potent bioactive metabolite of an entomopathogenic medicinal fungus *Cordyceps* with its nutraceutical and therapeutic potential. *Molecules* 25, 2735 (2020).
18. American Heart Association News. Consuming about 3 grams of omega-3 fatty acids a day may lower blood pressure. consuming- about-3- grams- of- omega-3- fatty- acids-a- day- may- lower- blood- press ure. (2022).
19. Jain, A. P., Aggarwal, K. K. & Zhang, P. Y. Omega-3 fatty acids and cardiovascular disease. *Eur. Rev. Med. Pharmacol. Sci.* 19(3), 441–445 (2015).
20. Wang, C. Z., Mehendale, S. R. & Yuan, C. S. Commonly used antioxidant botanicals: Active constituents and their potential role in cardiovascular illness. *Am. J. Chin. Med.* 35, 543–558 (2007)
21. Lampe, J. W. Health effects of vegetables and fruit: Assessing mechanisms of action in human experimental studies. *Am. J. Clin. Nutr.* 70, 475S-490S (1999)
22. Ohishi, T., Goto, S., Monira, P., Isemura, M. & Nakamura, Y. Anti-inflammatory action of green tea. *Antiinflamm. Antiallergy Agents Med. Chem.* 15(2), 74–90 (2016).
23. National Cholesterol Education Program (NCEP). Expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel III). Third report of the national cholesterol education program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel III) final report. *Circulation* 106(25), 3143–3421 (2002).
24. Ahmad, A. et al. Introduction and classification of natural polyphenols. In *Polyphenols-Based Nanotherapeutics for Cancer Management* (eds Tabrez, S. & Imran Khan, M.) (Springer, Singapore, 2021).



25. Riccioni, G., Mancini, B., Di Ilio, E., Bucciarelli, T. & D'Orazio, N. Protective effect of lycopene in cardiovascular disease. *Eur. Rev. Med. Pharmacol. Sci.* 12, 183–190 (2008).
26. Chiu, H.-F., Venkatakrishnan, F. & Wang, C.-K. The role of nutraceuticals as a complementary therapy against various neuro degenerative diseases: A mini-review. *J. Tradit. Complement. Med.* 10(5), 434–439 (2020).
27. Ashraf, S. A. et al. Multi-targeted molecular docking, pharmacokinetics, and drug-likeness evaluation of okra-derived ligand abscisic acid targeting signaling proteins involved in the development of diabetes. *Molecules* 26, 5957 (2021).
28. Agu, P. C., Aja, P. M., Ezeh, E. M. & Ekpono, U. E. Neuroprotective potentials of *Ageratum conyzoides* phytoconstituents via inhibition of monoamine oxidase: An in-silico study. *Niger. J. Biochem. Mol. Biol.* 38(1), 43–55 (2023).
29. Limanaqi, F., Biagioni, F. & Busceti, C. L. Phytochemicals bridging autophagy induction and alpha-synuclein degradation in parkinsonism. *Int. J. Mol. Sci.* 20, 3274 (2019).
30. Teter, B., Morihara, T. & Lim, G. P. Curcumin restores innate immune Alzheimer's disease risk gene expression to ameliorate Alzheimer's pathogenesis. *Neurobiol. Disord.* 127, 432–448 (2019).
31. Morowitz, M. J., Carlisle, E. M. & Alverdy, J. C. Contributions of intestinal bacteria to nutrition and metabolism in the critically ill. *Surg. Clin. N. Am.* 91(4), 771–785 (2011).
32. Nissen, L., Chingwaru, W., Sgorbati, B., Biavati, B. & Cencic, A. Gut health-promoting activity of new putative probiotic/protective *Lactobacillus* spp. strains: A functional study in the small intestinal cell model. *Int. J. Foodst. Microbiol.* 135, 288–294 (2009).
33. Siddeeg, A. et al. Recent updates and perspectives of fermented healthy superfood sauerkraut: A review. *Int. J. Food Prop.* 25(1),
34. Subramanian J, Sharma S, C BR. Modeling and selection of flexible proteins for structure-based drug design: backbone and side chain movements in p38 MAPK. *ChemMedChem.* 2008; 3(2): 336–344. [PubMed: 18081134] NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript
35. Brint AT, Willett P. Algorithms for the Identification of Three-Dimensional Maximal Common Substructures. *J. Chem. Inf. Comput. Sci.* 1987; 27:152–158.
36. Fischer D, Norel R, Wolfson H, Nussinov R. Surface motifs by a computer vision technique: searches, detection, and implications for protein-ligand recognition. *Proteins.* 1993; 16(3):278–292.
37. Norel R, Fischer D, Wolfson HJ, Nussinov R. Molecular surface recognition by a computer vision-based technique. *Protein Eng.* 1994; 7(1):39–46.
38. Miller MD, Kearsley SK, Underwood DJ, Sheridan RP. FLOG: a system to select 'quasi-flexible' ligands complementary to a receptor of known three-dimensional structure. *J Comput Aided Mol Des.* 1994; 8(2):153–174.
39. Diller DJ, Merz KM Jr. High throughput docking for library design and library prioritization. *Proteins.* 2001; 43(2):113–124.
40. Burkhard P, Taylor P, Walkinshaw MD. An example of a protein ligand found by database mining: description of the docking method and its verification by a 2.3 Å X-ray structure of a thrombin-ligand complex. *J Mol Biol.* 1998; 277(2):449–466
41. DesJarlais RL, Sheridan RP, Dixon JS, Kuntz ID, Venkataraghavan R. Docking flexible ligands to macromolecular receptors by molecular shape. *J Med Chem.* 1986; 29(11):2149–2153.
42. Kuntz ID, Leach AR. Conformational analysis of flexible ligands in macromolecular receptor sites. *J. Comput. Chem.* 1992; 13:730–748.
43. Ewing TJ, Makino S, Skillman AG, Kuntz ID. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *J Comput Aided Mol Des.* 2001; 15(5):411–428.
44. Welch W, Ruppert J, Jain AN. Hammerhead: fast, fully automated docking of flexible ligands to protein binding sites. *Chem Biol.* 1996; 3(6):449–462.



45. Schnecke V, Kuhn LA. Virtual Screening with Solvation and Ligand-Induced Complementarity. *Perspectives in Drug Discovery and Design*. 2000; 20:171–190.
46. Zsoldos Z, Reid D, Simon A, Sadjad BS, Johnson AP. eHiTS: an innovative approach to the docking and scoring function problems. *Curr Protein Pept Sci*. 2006; 7(5):421–435.
47. Miranker A, Karplus M. Functionality maps of binding sites: a multiple copy simultaneous search method. *Proteins*. 1991; 11(1):29–34.
48. Eisen MB, Wiley DC, Karplus M, Hubbard RE. HOOK: a program for finding novel molecular architectures that satisfy the chemical and steric requirements of a macromolecule binding site. *Proteins*. 1994; 19(3):199–221.
49. Bohm HJ. LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads. *J Comput Aided Mol Des*. 1992; 6(6):593–606.
50. Goodsell DS, Lauble H, Stout CD, Olson AJ. Automated docking in crystallography: analysis of the substrates of aconitase. *Proteins*. 1993; 17(1):1–10.
51. Hart TN, Read RJ. A multiple-start Monte Carlo docking method. *Proteins*. 1992; 13(3):206–222.
52. Goodsell DS, Olson AJ. Automated docking of substrates to proteins by simulated annealing. *Proteins*. 1990; 8(3):195–202.
53. Abagyan R, Totrov M, Kuznetsov D. ICM-A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem*. 1994; 15:488–506

