

Advancements in Drug Design Technology and Its Impact on COVID-19 Treatment

Mohd. Wasiullah¹, Piyush Yadav², Raj Aryan Prajapati³, Satish Kumar Yadav^{4,*},
Neha Parveen⁵

Abstract

The deadly coronavirus disease 19 (COVID-19) pandemic has recently spread, raising concerns about global health. The search for novel therapeutic compounds is made more necessary by the persistent problem of the absence of licensed medications or vaccinations. By saving money and time, computer-aided drug design has sped up the process of finding and developing new drugs. The structured-based and ligand-based drug discovery subcategories of computer-aided drug design (CADD) are the main topics of this review study. In ligand-based drug design, pharmacophore modeling, quantitative structure-activity relationships (QSARs), and artificial intelligence (AI) are commonly employed molecular modeling techniques, whereas structure-based drug design often involves molecular docking and molecular dynamic simulation. We have touched on the importance of computer-aided drug design in relation to COVID-19 and how scientists are still depending on these computational methods to quickly identify molecules that show promise as potential drugs against different targets connected to the pathophysiology of SARS-CoV-2, or severe acute respiratory syndrome coronavirus. Cross-species viruses known as coronaviruses (CoVs) can quickly move from their original host species into other ones, whereupon they can cause epidemic diseases. This article provides a comprehensive examination of computational methodologies and their utilization in the realm of drug development. Chemogenomics and drug discovery (DR) are highlighted as novel and developing system-based fields that are focused on modeling protein networks versus a library of chemicals in CoV infections. Furthermore, a number of recent successes based on chemogenomics, and molecular docking are given, thoroughly examined, and interpreted to highlight the unique benefits of CADD approaches in quickly developing a treatment for this deadly virus. The review's findings should help researchers who are creating energy-harvesting materials and systems identify new, unanticipated CoV strains or other variations in the future.

*Author for Correspondence

Satish Kumar Yadav
E-mail: slk.pharma@gmail.com

¹Principal, Department of Pharmacy, Prasad Institute of Technology, Jaunpur, (U.P.), India

²Academic Head, Department of Pharmacy, Prasad Institute of Technology, Jaunpur, (U.P.), India

³Scholar, Department of Pharmacy, Prasad Institute of Technology, Jaunpur, (U.P.), India

⁴Associate PROFESSOR, Department of Pharmacy, Prasad Institute of Technology, Jaunpur, (U.P.), India

⁵Lecturer, Department of Pharmacy, Prasad Institute of Technology, Jaunpur, (U.P.), India

Received Date: May 13, 2024

Accepted Date: May 21, 2024

Published Date: May 27, 2024

Citation: Mohd. Wasiullah, Piyush Yadav, Raj Aryan Prajapati, Satish Kumar Yadav, Neha Parveen. Advancements in Drug Design Technology and Its Impact on COVID-19 Treatment. International Journal of Virus Studies. 2024; 1(1): 1–15p.

Keywords: Computer-aided drug design (CADD), COVID-19, Molecular docking, Pharmacophore modelling, Chemoge, omics

INTRODUCTION

Drug discovery is the process of identifying a drug target, validating the target, discovering a new drug, optimizing the lead molecule, and performing clinical and preclinical studies [1]. The path to bringing a drug to market can incur expenses of around \$2.558 billion, with the discovery phase typically spanning between 10 to 15 years [2, 3]. Only 13% of new medications make it through clinical trials successfully, and drug attrition is highly high, despite the significant cost and effort required to create a new drug [4]. Due to a lack of ideal pharmacokinetic properties on absorption,

distribution, metabolism, excretion, and toxicity (ADME / Tox), the majority of medication failure instances (40 to 60%) are reported later in the drug development phase [5]. Drug discovery and development processes have been sped up while final phase expenses and failures have been reduced thanks to the adoption of CADD (Computer Aided Drug Discovery) tools in preclinical investigations by top Pharmaceutical Companies and Research Groups [6]. Rational drug design is an essential component of CADD and provides valuable insights into the binding affinity of the target protein and the molecular interaction of the target protein with the ligand. In pharmaceutical research, lead identification has been made possible by the use of high-performance supercomputers, parallel processing and sophisticated programs, algorithms and tools [7]. The study, comprehension, and elucidation of massive data pertinent to pharmaceuticals in drug discovery have also benefited greatly from AI and machine learning approaches [8]. Various methods used in drug discovery to identify new inhibitors in chemical databases are pharmacophore modelling, QSAR, molecular docking and quantum mechanics as well as statistical learning methods [9].

CADD can broadly be divided into structure based drug design approach and ligand based drug design approach, both of which are widely used in drug discovery for the identification of lead molecules. Target receptors' 3-D structures and their active sites are the foundation of structure-based drug design, which aims to shed light on the molecular interactions that occur between ligand and target. The foundation of ligand-based drug design lies in the interactions between ligands and the target receptor. Computer-aided drug design has yielded numerous success stories and remains an integral step in the drug discovery process. This methodology has been proposed for combating the COVID-19 pandemic by identifying potential treatment candidates. The causative agent of COVID-19, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), belongs to the Betacoronavirus genus and shares significant nucleotide sequence similarity with both SARS-CoV and Middle East respiratory syndrome coronavirus (MERS-CoV). Studies have extensively investigated the epidemiology, genomic composition, pathophysiology, animal models, diagnostics, and vaccine development for MERS-CoV infections, incorporating various computational biology methodologies (Skariyachan et al., 2019) [10]. The genome of SARS-CoV-2 comprises four structural proteins (shell, membrane, spike, and nucleocapsid) along with several accessory proteins. Similar to SARS and MERS, it also codes for 16 non-structural proteins (nsps), including major protease (Mpro), papain-like protease, RNA-dependent RNA polymerase (RdRp), helicase, among others [11–12]. While Spike glycoprotein is essential for viral interaction with host cell receptors, nsps play a key role in the viral life cycle by participating in accessory RNA production [13, 14].

For the purpose of creating and developing antiviral medications against COVID-19, structural and non-structural proteins present intriguing targets [13]. Given the high death rate and lack of effective COVID-19 treatments or vaccines, it is imperative that new medications be discovered quickly [15]. One key tool for this process is computer-aided drug design. Utilizing natural lead molecules acquired through virtual screening and pharmacokinetic prediction, there exists the potential to create efficacious lead compounds against COVID-19 [16]. Repurposing broad-spectrum antivirals is a promising method thanks to the availability of pharmacokinetic data, which could expedite the development of viable therapies for SARS-CoV-2 infection in people. Pharmaceutical pharmacology and pharmacodynamics [17]. The availability of the entire genome sequence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the elucidation of viral protein structures through techniques such as X-ray crystallography, nuclear magnetic resonance (NMR), electron microscopy, and homology modeling have facilitated the identification of potential drug inhibitors targeting key therapeutic targets of COVID-19. This review paper offers helpful details on some of the popular in silico techniques utilized in CADD, as well as how these techniques are being applied and could be helpful for COVID-19 drug discovery.

COMPUTATIONAL METHODS IN DRUG DISCOVERY FOR COVID-19

In 1981, a significant article was published in "Fortune magazine" under the title "The Next Industrial

Revolution: Computerized Drug Design at Merck. This article has raised awareness of the importance of in silico research in drug discovery and CADD disease manifestations (Sliwoski et al, 2014) [16]. Since then, high-throughput screening (HTS) has gained increasing prominence in both the pharmaceutical industry and academic institutions for swiftly identifying key compounds and lead candidates. This automated approach enables the screening of entire compound libraries or multiple chemicals and biological compounds in target-based or cell-based assays to elicit the desired biological response (Bittker& Ross, 2016; Maghsoudi et al., 2021) [17–18]. The rigorous demands for abundant targets and ligands in HTS might suggest disregarding non-ligands due to the cost and time involved, with seemingly little chance of success.

Nonetheless, because of the strength of CADD technology, this issue can be readily resolved through virtual high-throughput screening (vHTS), which makes use of virtual chemical libraries and enables the researcher to concentrate on the ligands that are more likely to provide engaging experiences (Leelananda & Lindert, 2016). During the early days of the COVID-19 outbreak, vHTS was effectively used and developed to research a group of approved drugs to identify agents that could combat SARS-CoV-2 at the Center for Disease Control and Prevention. National Advancement of Science Transfer (NCATS) from States. The National Institutes of Health (NIH) have extensively investigated various aspects of the SARS-CoV-2 life cycle, including viral entry, replication, in vitro infectivity, and live virus infectivity, as outlined by Xu et al. (2021). Different stages of the SARS-CoV-2 life cycle have been identified as potential targets for therapeutic intervention.

- For instance, cephalexin, recognized for its anti-inflammatory properties, has demonstrated efficacy in mitigating the cytopathic effects (CPE) of SARS-CoV-2 (Rogosnitzky et al., 2020), while corilagin, an anti-fibrotic compound, has shown promise in combating the angiotensin receptor binding domain (RBD) of angiotensin-converting enzyme (ACE2) (Yang et al., 2021). Therefore, the development of a robust SARS-CoV-2 infection model with high biosafety and efficient high-throughput screening (HTS) capabilities for potential drug candidates is an urgent priority.

Structure-Based Drug Design for COVID-19

Structure-based drug design (SBDD) is based on the finding of the therapeutic target protein's three-dimensional structure and the investigation of its binding site cavity [18]. The disease can be understood at the molecular level because to this methodical and efficient approach to lead molecule identification and optimization [19]. Structure-based drug design (SBDD) commonly employs techniques such as structure-based virtual screening (SBVS), molecular docking, and molecular dynamics (MD) simulations. These techniques are useful for assessing binding energies, protein-ligand interactions, and receptor conformational changes that occur when a ligand binds [20–21]. The computational method known as SBDD, which is employed by numerous pharmaceutical companies and medicinal chemists, has been instrumental in the discovery of numerous pharmaceuticals that are currently on the market [22].

- For example, the thymidylate synthase inhibitor raltitrexed, utilized against HIV, was discovered using the structure-based drug design (SBDD) approach [23], while topoisomerase II was identified as a target. Additionally, norfloxacin, an antibiotic commonly employed for urinary tract infections, was discovered using structure-based virtual screening (SBVS) [18], leading to the identification of dorzolamide, a carbonic anhydrase inhibitor used for conditions such as glaucoma and cystoid macular edema, through fragment screening [24]. Isoniazid, an anti-tuberculosis medication acting as an inhibitor of enoyl-acyl-ACP reductase (InhA), was found via virtual screening and structure-based pharmacophore modeling [25]. Furthermore, flurbiprofen, a nonsteroidal anti-inflammatory drug (NSAID) used in rheumatoid arthritis treatment, was identified to target cyclooxygenase-2 (COX-2) through molecular docking, among other methods [26, 27]. The basic steps involved in SBDD include target structure preparation, ligand binding site determination, complex library preparation, functional docking and molecular scoring, molecular dynamics simulation, and calculate the binding free energy (Figure 1).

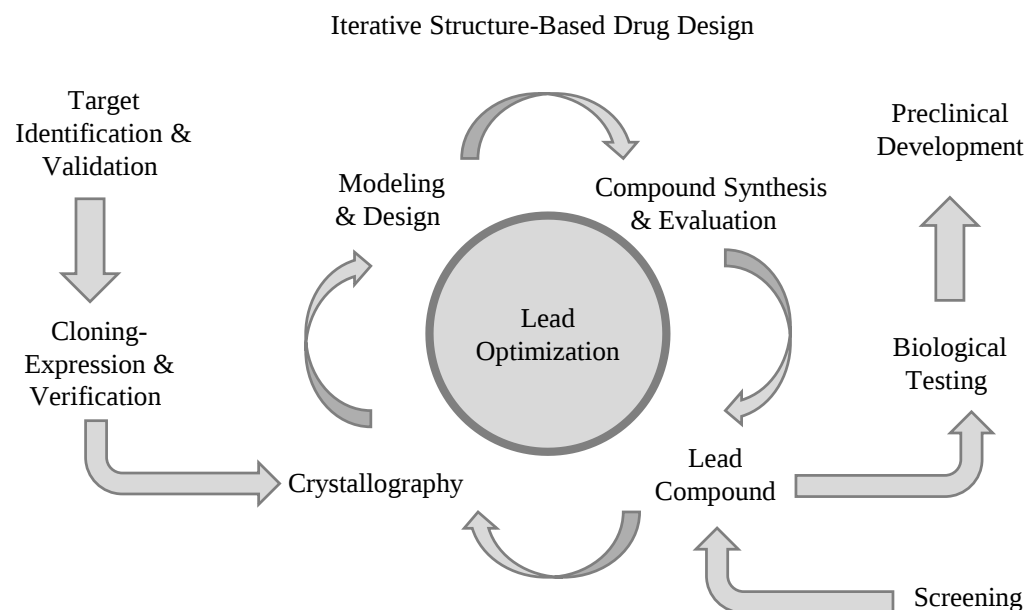


Figure 1.-Basic steps involved in the structure-based drug design approach.

Preparation of the Target Structure

Over the past few decades, the number of structures put into the Protein Data Bank (PDB) has expanded due to the quick development of tools for elucidating structural relationships, such as NMR and X-ray. Many target protein structures are still unknown as a result of shortcomings in experimental methods [28]. It has been possible to decode protein structures from their sequences using computational methods like threading [30], ab initio modeling [31], and comparative homology modeling [29]. The exact determination of a protein's three-dimensional structure from its amino acid sequence can be Enabled by homology modeling, a widely utilized computational method [42], this process involves several steps, including pattern identification, sequencing, target model construction, model refinement, and model validation [29]. When a target protein has low sequence similarity to other proteins in the PDB (as opposed to homology modeling, which just takes similarity into account), protein threading is a substitute protein structure prediction technique that is frequently employed. Protein threading incorporates structural information (secondary structure, solvent accessibility, and pairwise interactions) included in the model to increase prediction accuracy, taking into account sequence similarity between target and template [33]. Ab initio modeling is another computational method favored when the target protein lacks a structural model in current biological databases [31]. To determine the dihedral angle ϕ values for a given protein structure that contribute to the structure's stability (presence of a minimum global or quasi-global potential energy), it takes into account a global optimization problem [34].

Identification of the Ligand Binding Site

Specific docking cannot be carried out without knowledge about the ligand binding location. The X-ray crystal structures of proteins co-crystallized with substrates or inhibitors or the investigation of site-directed mutagenesis can provide information about binding sites [35]. Even in the absence of experimental data for many protein binding sites, we can forecast potential binding sites of target proteins utilizing diverse software and web servers like CASTp [36], DoGSite Scorer [37], NSiteMatch [38], DEPTH [39], MSPocket [40], MetaPocket [41], and Q-SiteFinder [42]. During lead determination, large compounds are rejected if they do not fit securely into the binding site.

Compound Library Preparation

Chemical compounds can be sourced from various chemical databases, including ZINC (with approximately 230 million purchasable compounds) [43], PubChem (providing access to around 111

million pure, characterized chemical compounds) [44], MCULE (offering approximately 122 million accessible synthetic compounds) [source: mcule.com], ChEMBL (housing over 1.6 million unique compounds) [45], DrugBank (containing approximately 14,528 drug molecules) [46], and ChemSpider (providing access to around 25 million unique chemical compounds) [47]. Other potential risks include acute toxicity in rats, carcinogenicity, elevated serum glutamic oxaloacetic transaminase levels, hepatotoxicity, and inhibition of 3A4 oxidation by midazolam. Molecular docking was conducted using drug-like compounds selected based on Lipinski's rule of five and ADMET (absorption, distribution, metabolism, excretion, and toxicity) parameters [48]. ADMET is frequently utilized because of its ability to bind to plasma proteins, inhibit P-glycoprotein (P-gp) and cytochrome P450 (CYP), penetrate the blood-brain barrier (BBB), and human gastrointestinal absorption (HIA) [49]. It is important to take into account these compounds' synthetic accessibility in addition to their pharmacokinetic qualities, druggability, and safety.

Molecular Docking and Scoring Functions

Molecular docking is a computational technique that examines the interactions between molecules, such as ligands and target receptors. It enables the ranking of ligands by determining how well they attach to the receptor using various scoring methods [50]. Favorable binding poses of ligands to the target active site depend on two factors: (a) a large conformational space accounting for different binding poses and (b) well-predicted clarity about the binding affinity of the corresponding ligands for each docking setting (51). Table 1 provides a list of molecular docking applications that are often utilized. Molecular docking is divided into two categories: flexible protein docking and flexible ligand search docking. Flexible protein docking often utilizes Monte Carlo (MC) and molecular dynamics (MD) methods, while flexible ligand search docking employs three algorithms for ligand flexibility: system method, stochastic method, and simulation method [52].

Molecular Dynamic (MD) Simulation

Protein molecular dynamics (MD) simulations were initially conducted in the late 1970s. Based on Newton's laws of motion, which control atomic interactions, this potent physical method is used to forecast the location of individual atoms inside a molecular system over time. Using the appropriate force field that was used to calculate the system's total energy, the force between interacting atoms is computed [62]. There are numerous reasons why MD simulation is so popular. The position and motion of every atom in the system is always recorded, which is quite difficult to use with any experimental technique. It is possible to properly change the simulation circumstances, which are precisely understood. Because MD simulations aid in the large-scale elucidation of several atomic features, including receptor binding and unbinding as well as conformational changes, they are a popular tool in structure-based drug discovery fine resolution that is typically unattainable through experimental research [63]. Additionally, because of it is feasible to measure the thermodynamics, kinetics, and natural energy landscape as well as investigate the dynamics of receptor-ligand interactions (binding and dissociation) using MD simulations. Some examples of MD simulation programs include GROMACS, AMBER, CHARMM, NAMD, and Desmond.

Ligand-based drug design for COVID-19

Another popular approach to computer-aided drug design is ligand-based drug design, which is utilized in situations where the target receptor's three-dimensional structure is unavailable. If structural similarities between a set of molecules active against a particular target receptor correlate to similar biological functions, the physicochemical and structural features responsible for a given biological activity can be used to identify activity. Artificial intelligence (AI), pharmacophore modeling, and quantitative structure-activity relationships (QSAR) are a few of the often employed methods in ligand-based virtual screening [64].

Pharmacophore Modeling

Pharmacophore modeling elucidates the spatial arrangement of a ligand's chemical features essential for interaction with the target receptor. These features encompass hydrogen bond donors, acceptors,

aromatic ring systems, hydrophobic regions, as well as positively and negatively charged ionizable groups. Pharmacophore-based virtual screening identifies ligands with distinct frameworks that share similar spatial arrangements of crucial interacting functional elements. The pharmacokinetic model considers the physiologically active conformation of molecules within the target binding site. Pharmacophore models are integral to molecular binding studies in QSAR investigations [65]. Commonly utilized programs for automatic pharmacophore model construction include Catalyst, PHASE, LigandScout, GALAHAD, and PharmMapper. Steric constraints are also incorporated into well-designed pharmacophore models, which are further refined to reduce model restriction in regions occupied by inactive molecules. The final model was tailored to exclude or include only those pharmacological characteristics that could not be consistently found in the active compounds. A separate external test set must be used for validation, and the developed pharmacokinetic model needs to have the best possible sensitivity and specificity to reduce the possibility of false positive and false negative outcomes. A sequence generated 3D pharmacophore model will be particularly helpful if information regarding the 3D structure of the receptor and the set of known active chemicals is absent. For example, Pharma3D derives common sequence motifs crucial for biomolecular receptor-ligand interactions across protein families using homology modeling and knowledge of 3D crystal structures [66].

Table 1. List of molecular docking applications.

Tools	Key Features	Reference
AUTO DOCK	In AutoDock, shape searching methods include the Lamarckian genetic algorithm, simulated annealing search, and traditional genetic algorithm search. These methods are utilized to estimate the binding free energy of small molecules to protein targets using a semiempirical free energy force field.	[53]
AUTO DOCK VINA	AutoDockVina employs an advanced gradient optimization technique in its calculations, resulting in a significant speed enhancement of about two orders of magnitude and improved accuracy in predicting bond modes compared to AutoDock.	[54]
GOLD	GOLD (Genetic Optimization for Ligand Docking) is an automated program for ligand docking, enabling full ligand conformational flexibility alongside partial protein flexibility. It explores binding conformations through a genetic algorithm.	[55]
CDOCKER	CDOCKER (CHARMm-based DOCKER) is an automated molecular dynamics (MD) docking software utilizing the CHARM 19 force field, offering complete flexibility for CHARMm ligands and receptors while reducing computational time.	[56]
FLEXX	FLEXX is a comprehensive automated docking tool designed for flexible ligands, renowned for its accuracy and precision in delivering reliable results. Its methodology relies on the strategic selection and arrangement of ligand base fragments within the active site, under the premise that optimal base fragments interacting with the site yield superior scores.	[57]
SURFLEX	Surflex is a docking software employing both an empirical Hammerhead scoring function and molecular similarity techniques to generate potential positions of ligand fragments.	[58]
GLIDE	Glide (grid-based ligand docking with energy) conducts an extensive exploration of the position space, orientation, and conformation of the ligand as it binds to the receptor, achieving efficient computational speed. The scoring for linkage conformity relies on the ChemScore function.	[59]
DOCK6	Glide, employing grid-based ligand docking with energy, conducts an exhaustive exploration of ligand position, orientation, and conformation within the receptor, all while maintaining reasonable computational efficiency. The scoring for linkage conformity is derived from the ChemScore function.	[60]
SWISS DOCK	SwissDock is a web server that allows docking of small molecules to target proteins based on the EADock DSS tool.	[61]

Quantitative Structure-Activity Relationships (QSARs)

The foundation of QSAR research is the idea that alterations in a compound's molecular structure might be linked to variations in its biological activity. They are widely utilized during the lead identification and optimization stages of drug discovery. These association studies are utilized to create a statistical model, which may then be used to forecast the biological activity of novel molecules. To develop a reliable QSAR model, it's essential to have an adequate amount of datasets with biological activities obtained through standard experimental procedures, carefully select compounds for both training and test sets, avoid automatic correlations between the physicochemical properties of ligands to prevent data overfitting, and ensure the applicability and validation of the model's predictions through internal and external validation methods [67]. QSAR can be divided into six categories according to the descriptors' methods of derivation: One type of QSAR is 1D, which examines the relationship between bulk molecular properties like logP and pKa and biological activities; another is 2D, which correlates biological properties and activities with structural patterns like 2D pharmacophores and connection indices. (c) 3D-QSAR investigates the relationship between biological activity and non-covalent interaction fields surrounding ligands. Extending from 3D-QSAR, 4D-QSAR includes a range of ligand configurations, while 5D-QSAR integrates different induced fitting models, and 6D-QSAR incorporates additional solvation models. Some notable 3D QSAR programs include Catalyst's HypoGen module, PHASE, Comparative Molecular Field Analysis (CoMFA), and Comparative Similarity Index Analysis (CoMSIA). Various computational tools for molecular descriptors are also available. QSAR procedures can be categorized into linear and nonlinear approaches based on chemical methods, with examples including principal component analysis (PCA), partial least squares (PLS), multiple linear regression (MLR), and principal component regression (PCR) [68].

- Artificial neural networks (ANN), Bayesian neural networks, and closest neighbors (kNN) are a few examples of nonlinear QSAR techniques.

ARTIFICIAL INTELLIGENCE AND DRUG DISCOVERY

Artificial intelligence (AI) relies on computers' ability to learn from existing data and has been utilized in various computational modeling techniques to predict the biological activity and toxicity of pharmacological compounds [69]. AI finds wide-ranging applications in drug discovery, encompassing de novo drug design, QSAR, ADMET characterization, virtual screening, protein-protein interactions, and protein folding prediction. Machine learning (ML) and deep learning (DL) are powerful techniques commonly employed in rational drug design. Support Vector Machines (SVM), Random Forests (RF), and Naive Bayesian (NB) are extensively utilized ML algorithms in drug development [70]. Deep learning techniques include deep neural networks (DNN), autoencoders, recurrent neural networks (RNN), convolutional neural networks (CNN), and restricted Boltzmann machines (RBM) [4]. Simple physicochemical parameters like logP and solubility can be accurately predicted by conventional QSAR approaches. Unfortunately, because the approaches span a narrow chemical space [9] and use short training sets [8], QSAR predictions of complicated biological features like therapeutic efficacy and side effects are frequently inadequate. Machine learning and conventional QSAR approaches face significant hurdles when dealing with large data produced by high-throughput screening (HTS) techniques. In order to accurately forecast the efficacy of medications as well as their negative effects in humans or animals, artificial intelligence (AI) techniques have been created to interpret this large volume, multidimensional big data. Deep learning, which was first applied to drug development in the 2012 QSAR machine learning challenge funded by Merck [10], is now the most promising method in the field of big data. Effective AI learning is hindered by a scarcity of high-quality data from public databases and variability in data sources, particularly those emanating from various biological assays [19].

Case Study of COVID-19

The development of treatments for COVID-19, caused by the SARS-CoV-2 virus, heavily relies on structure-based and ligand-based drug design methods. Thus far, only a handful of potential drug molecules have progressed to clinical trials, primarily consisting of repurposed approved drugs [71].

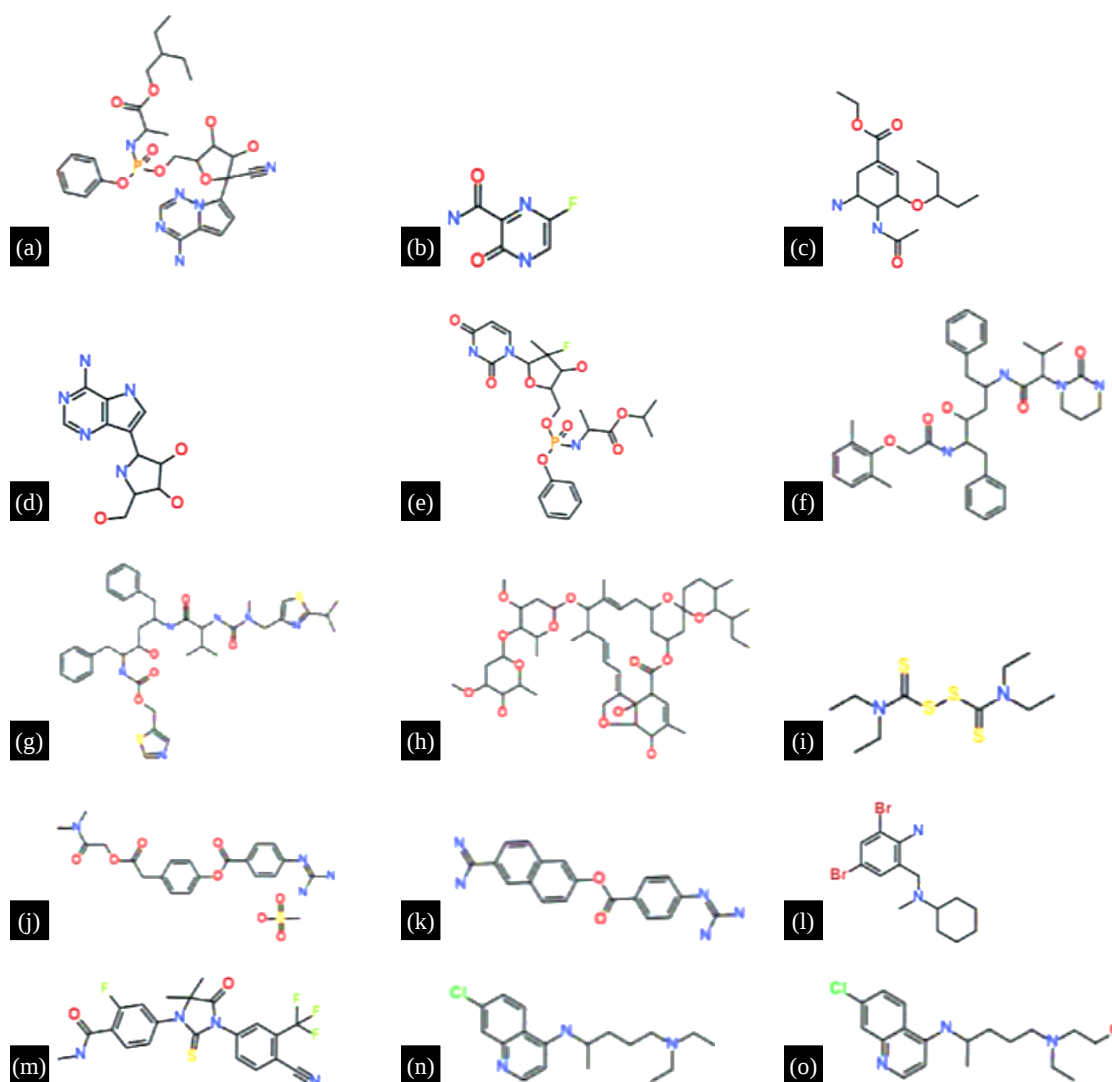


Figure 2. Drug Molecules. (a) Remdesivir (b) Favipiravir (c) Oseltamivir (d) Galidesivir (e) Sofosbuvir (f) Lopinavir (g) Ritonavir (h) Ivermectin (i) Disulfiram (j) Camostat mesylate (k) Nafamostat (l) Bromhexine (m) Enzalutamide (n) Chloroquine (o) Hydroxychloroquine

Figure 2: Drug Molecules

- Currently, molecules under investigation in clinical trials include RNA polymerase inhibitors (molecules 1–5), type 3C protease inhibitors (molecules 6–8), papain-like protease inhibitors (molecules 9–10), TMPRSS2 inhibitors (molecules 10–13 in the blue region), and compounds targeting intracellular acidification (molecules 14–15 in the gray region) shown in figure 2.
- Finding efficient therapies is necessary because to the high death rate of the pandemic, the lack of licensed medications and vaccinations against COVID-19, and other factors. The rapid progress in COVID-19 research is facilitated by the availability of the complete genome sequence of SARS-CoV-2 and the ability to determine the structure of viral proteins through methods like X-ray crystallography, NMR spectroscopy, and electron microscopy. Key drug targets of SARS-CoV-2 include structural proteins such as the spike protein (S), envelope protein (E), membrane protein (M), and nucleocapsid protein (N) [Figure 3]. Additionally, non-structural proteins (Nsps) [Figure 4] serve as targets for host-based pharmacological interventions, including angiotensin-converting enzyme 2 (ACE2), transmembrane serine protease 2 (TMPRSS2), furin, and cathepsin [72]. Furthermore, major protease (3C-like protease 3CLpro, nsp5), papain-like protease (PLpro, nsp3), RNA-dependent RNA polymerase (RdRp, nsp12), nsp15

endoribonuclease, nsp16 2-O-methyltransferase, and nsp13 helicase are among the significant drug targets.

- This section provides a brief discussion of nonstructural protein details. The catalytic pair (cysteine and histidine) in the active pocket of the main protease is a cysteine protease. By cleaving at least 11 locations throughout the C-terminal and central portions of sequenced viral polyproteins, Mpro's catalytic activity on polyproteins releases essential proteins needed for viral replication. L/F /M)-Q ↑(G/A/S)-X correspondence [73].
- The second SARS-CoV-2 protease, papain-like protease (PLpro), has a consensus recognition sequence of "LXGG↓XX" and can be targeted by small compounds that cleave three sites. This target is appealing for drug development because it plays a crucial role in cleaving and maturing viral polyproteins, assembling the replicase-transcriptase complex, and disrupting host immune responses [74].
- RNA-dependent RNA polymerase (RdRp) is derived from the cleavage of ORF1a and ORF1ab polyproteins 1a and 1ab, respectively, and plays a crucial role in SARS-CoV-2 genome replication and transcription. Structurally, the enzyme's catalytic core resembles the human right hand, with distinct regions corresponding to the palm, fingers, and thumb. Targeting RdRp to inhibit viral replication proves to be a promising therapeutic strategy due to the highly conserved and accessible nature of its active site. Nsp15 is an endoribonuclease with specificity for uridine, playing a role in RNA processing and found across various life forms. Its catalytic domain at the C-terminal exhibits both sequence and functional resemblance to enzymes within the EndoUfamily. This hexameric enzyme, with an active size of 234 kDa, cleaves both single- and double-stranded RNA at uridine sites, resulting in the formation of a 2,3-cyclic phosphodiester and 5-hydroxyl termini. SARS CoV-2 2-o-methyltransferase (nsp16) is another important enzyme required for viral replication. This enzyme plays a crucial role in the development of the RNA cap, a particular arrangement that enhances the stability and metabolism of viral RNA and protects it from innate cellular immunity competent interpreter [75].
- Out of the 16 identified CoVNsp proteins, SARS-Cov-2 Nsp13 helicase holds vital importance as it exhibits the highest sequence conservation within the CoV family, underscoring its critical role in viral replication. This enzyme can hydrolyze all forms of NTPs and unwind RNA helicase in an ATP-dependent manner. It also possesses NTPase and RNA helicase activities (. Transmembrane serine protease 2 (TMPRSS2) is an essential host component that primes the viral Spike (S) protein through cleavage at the S1/S2 and S2 locations, hence regulating host cell membrane fusion and cell entrance. Proprotein Convertase (PC) called furin is found in the trans-Golgi complex and is triggered by an acidic pH. The distinctive "RRAR" motif in the SARS-CoV-2 spike protein is recognized and hydrolyzed by this enzyme [72].

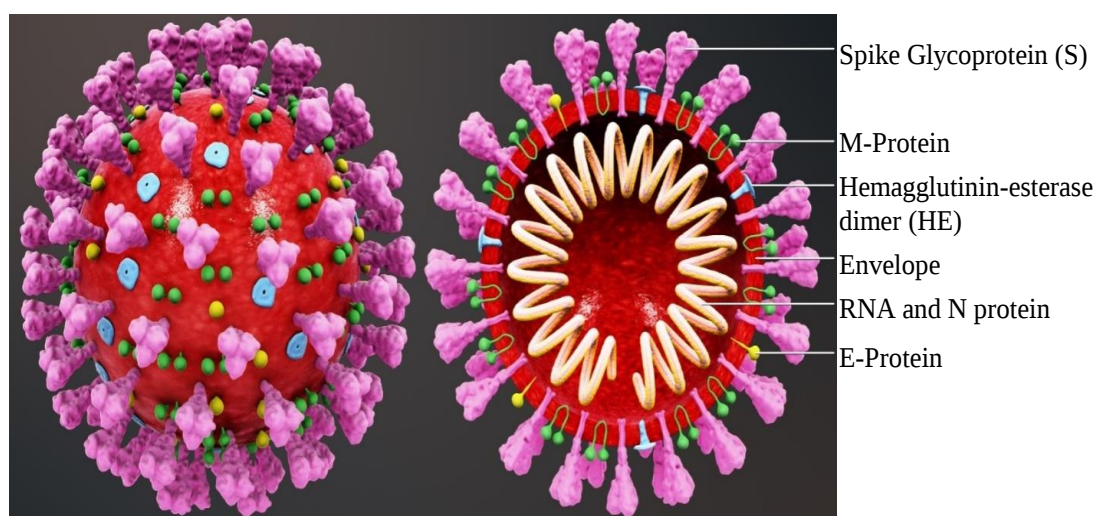


Figure 3. Drug targets of SARS-CoV-2 include structural proteins.

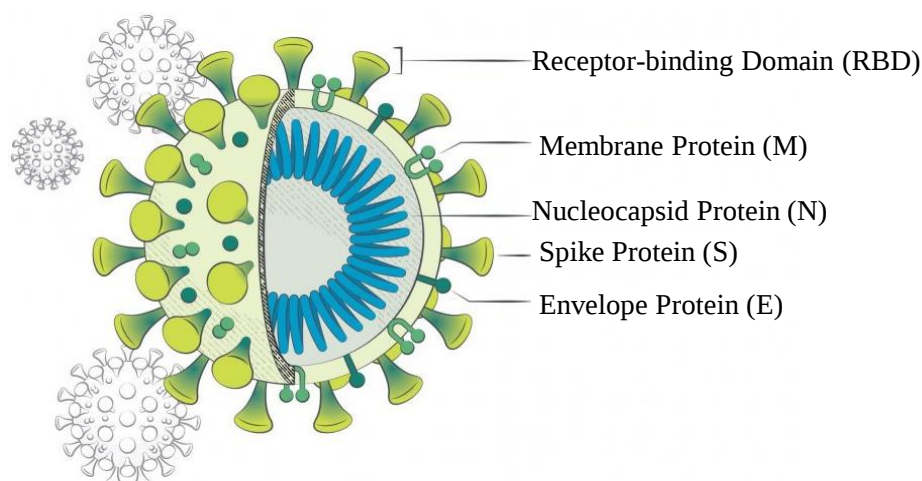


Figure 4. The structural proteins of SARS-CoV-2.

DISCUSSION

Anti-SARS-CoV-2 medications that target Mpro have been reported often since the COVID-19 outbreak, and several of them have started clinical trials. Only Paxlovid, created by Pfizer and Xocova from Shionogi Pharmaceutical Company, has been given the go-ahead to be marketed, nevertheless. Thus, there is still a critical therapeutic need for new, safe, and effective broad-spectrum Mpro inhibitors today and in the future. During this epidemic, researchers quickly conducted virus sequence analysis as well as further studies on function and structure, providing a powerful scientific and technological force to control COVID-19. The development of specific inhibitors and study on Mpro as a therapeutic approach have been fueled by studies on virus-related biology, particularly crystal structure analysis. In this study, we explored the structural features of Mpro, a fully conserved gene sequence with no human homologue, which is highly suitable as a drug target. Compounds having inhibitory action against Mpro that could be turned into medicines were found via virtual screening and pharmacophore design. First, we use pharmacokinetic integration approaches to make sure that the compounds we choose from the small molecule database have inhibitory activity against Mpro. These methods are based on the effective fractions or pharmacodynamic features of the available active compounds. In order to develop a pharmacokinetic model for Mpro, existing medications are utilized to preserve beneficial pharmacodynamic aspects and remove ineffective pharmacodynamic features that have detrimental consequences. We conducted four experiments to verify the pharmacokinetic model's reasonableness, and the outcomes demonstrated that the model's design can reliably identify drugs with Mpro inhibitory activity. Molecular docking of the CDocker program was used to test the activity of the screened compounds, and AutoDock was used to confirm the results. This verified the study's scientific validity and rigor, and the chosen drugs demonstrated good pharmacokinetic characteristics and potent inhibitory efficacy against Mpro. Following analysis of the chosen compounds, we discovered that ENA482732, the first chemical, fulfilled our expectations exactly. Stable binding of compound ENA482732 to the acceptor is ensured by multiple binding sites and distinct intermolecular interaction forces. Research on its composition also reveals that Mpro can interact with the indole group in a number of different ways. According to this, it may be possible to enhance the Mpro binding stability of well-known small molecule inhibitors by adding indole groups at the proper locations. For the development of Mpro inhibitors, this study primarily offers theoretical research. Strong inhibitory efficacy against Mpro was confirmed by molecular dynamics, however this is still speculative. We are still far from developing official inhibitors. Whether or not the chemical has inhibitory activity will be further investigated in future research by using enzyme activity assays to test this notion. Once ENA482732 or modified compounds were identified as potential therapeutic candidates, preclinical safety, pharmacodynamics, and pharmaceutical studies were carried out. In order to enable the start of clinical trials, this is used to monitor the compound's biological activity against the intended ailment and evaluate its safety.

CONCLUSION AND FUTURE PROSPECTS

Substantial advancements have been achieved in the practical implementation and rational utilization of computational drug discovery, resulting in a transformative shift in both industrial and academic sectors. Computer-aided drug design (CADD) proves to be a potent strategy in generating novel medications, particularly when combined with existing chemical methodologies. Although CADD takes advantage of many dilemmas and approximations, this knowledge-based approach plays an important role in the drug discovery process due to its ability to accelerate the drug discovery process. Since CADD has yielded several achievements, there is a promising future to help discover more drugs in the future. Chemogenomics, situated at the intersection of chemistry, biology, and informatics, offers a significant reduction in both cost and time in drug discovery. This interdisciplinary approach enables the translation of genetic sequences into potential biological targets. In recent years, there has been a notable emphasis on drug repositioning (DR) within pharmaceutical and biomedical sectors, underscoring the importance of repurposing existing drugs. This interest has led to the successful introduction of new indications that are more economically viable compared to traditional drug development methods. The global COVID-19 pandemic has highlighted the urgent need for effective treatments, prompting a focus on utilizing DR to expedite research for combating emergent threats like COVID-19. DR is strongly recommended for further exploration in repurposing existing drugs for various diseases. However, the efficacy of computer-aided drug design (CADD), including DR and chemical genomics, in generating viable treatments across different disease areas remains uncertain, as success is not always guaranteed and there are ongoing IT challenges to address.

For instance, high-throughput screening (HTS) has become increasingly prevalent in pharmaceutical and academic settings for swiftly identifying lead compounds. This automated approach enables the screening of entire compound libraries in target-based or cell-based assays to achieve desired biological responses. Leveraging the genomic similarities, particularly chemical genes, is crucial for employing HTS against SARS-CoV-2, serving as a flexible drug screening model with enhanced biosecurity. Presently, CADD represents the primary method for drug screening against COVID-19, either targeting single viral proteins or conducting small-scale screenings of antiviral drugs previously approved by the FDA for SARS-CoV-2. However, this approach may not align well with the biosafety requirements of SARS-CoV-2 HTS. Moreover, drug development faces delays without precise crystal structures of the target virus. Hence, the most effective strategy for promptly identifying COVID-19 treatments involves utilizing the CADD method to identify lead compounds for clinical use. For instance, a program integrating structure-based drug design, HTS, and virtual drug screening was employed to identify new leads targeting Mpro, resulting in the discovery of six compounds with significant IC₅₀s against Mpro and potent antiviral activity in cell-based assays. Furthermore, the crystal structure of Mpro was determined using CADD.

REFERENCES

1. C. M. Song, S. J. Lim, and J. C. Tong, "Recent advances in computer-aided drug design," *Briefings in Bioinformatics*, vol. 10, no. 5, pp. 579–591, 2009.
2. J. A. DiMasi, H. G. Grabowski, and R. W. Hansen, "Innovation in the pharmaceutical industry: New estimates of R&D costs," *Journal of Health Economics*, vol. 47, pp. 20–33, 2016.
3. D. Vohora and G. Singh, *Pharmaceutical Medicine and Translational Clinical Research*, Academic Press, 2018.
4. F. Zhong, J. Xing, X. Li et al., "Artificial intelligence in drug design," *Science China Life Sciences*, vol. 61, no. 10, pp. 1191–1204, 2018.
5. T. Hou and X. Xu, "Recent development and application of virtual screening in drug discovery: an overview," *Current Pharmaceutical Design*, vol. 10, no. 9, pp. 1011–1033, 2004.
6. W. Yu and A. D. MacKerell, "Computer-aided drug design methods," in *Antibiotics*, Humana Press, New York, NY, 2017.
7. S. J. Y. Macalino, V. Gosu, S. Hong, and S. Choi, "Role of computer-aided drug design in modern drug discovery," *Archives of Pharmacal Research*, vol. 38, no. 9, pp. 1686–1701, 2015.

8. W. Duch, K. Swaminathan, and J. Meller, "Artificial intelligence approaches for rational drug design and discovery," *Current Pharmaceutical Design*, vol. 13, no. 14, pp. 1497–1508, 2007.
9. H.-J. Huang, H. W. Yu, C.-Y. Chen et al., "Current developments of computer-aided drug design," *Journal of the Taiwan Institute of Chemical Engineers*, vol. 41, no. 6, pp. 623–635, 2010.
10. M. Hassan Baig, K. Ahmad, S. Roy et al., "Computer aided drug design: success and limitations," *Current Pharmaceutical Design*, vol. 22, no. 5, pp. 572–581, 2016.
11. S. Skariyachan, S. B. Challapilli, S. Packirisamy, S. T. Kumargowda, and V. S. Sridhar, "Recent aspects on the pathogenesis mechanism, animal models and novel therapeutic interventions for Middle East respiratory syndrome coronavirus infections," *Frontiers in Microbiology*, vol. 10, p. 569, 2019.
12. S. A. Amin and T. Jha, "Fight against novel coronavirus: a perspective of medicinal chemists," *European Journal of Medicinal Chemistry*, vol. 201, article 112559, 2020.
13. B. Goyal and D. Goyal, "Targeting the dimerization of the main protease of coronaviruses: A potential broad-spectrum therapeutic strategy," *ACS Combinatorial Science*, vol. 22, no. 6, pp. 297–305, 2020.
14. W. Dai, B. Zhang, X.-M. Jiang et al., "Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease," *Science*, vol. 368, no. 6497, pp. 1331–1335, 2020.
15. Y.-F. Tu, C.-S. Chien, A. A. Yarmishyn et al., "A review of SARS-CoV-2 and the ongoing clinical trials," *International Journal of Molecular Sciences*, vol. 21, no. 7, p. 2657, 2020.
16. D. Gopal and S. Skariyachan, *Recent Perspectives on COVID-19 and Computer-Aided Virtual Screening of Natural Compounds for the Development of Therapeutic Agents Towards SARS-CoV-2*, Methods in Pharmacology and Toxicology, Springer, 2020.
17. T. Pillaiyar, S. Meenakshisundaram, and M. Manickam, "Recent discovery and development of inhibitors targeting coronaviruses," *Drug Discovery Today*, vol. 25, no. 4, pp. 668–688, 2020.
18. M. Batool, B. Ahmad, and S. Choi, "A structure-based drug discovery paradigm," *International Journal of Molecular Sciences*, vol. 20, no. 11, p. 2783, 2019.
19. E. Lionta, G. Spyrou, D. K. Vassilatis, and Z. Cournia, "Structure-based virtual screening for drug discovery: principles, applications and recent advances," *Current Topics in Medicinal Chemistry*, vol. 14, no. 16, pp. 1923–1938, 2014.
20. S. Kalyaanamoorthy and Y.-P. P. Chen, "Structure-based drug design to augment hit discovery," *Drug Discovery Today*, vol. 16, no. 17–18, pp. 831–839, 2011.
21. A. Wlodawer and J. Vondrasek, "Inhibitors of HIV-1 protease: a major success of structure-assisted drug design," *Annual Review of Biophysics and Biomolecular Structure*, vol. 27, no. 1, pp. 249–284, 1998.
22. D. E. Clark, "What has computer-aided molecular design ever done for drug discovery?" *Expert Opinion on Drug Discovery*, vol. 1, no. 2, pp. 103–110, 2006.
23. A. C. Anderson, "The process of structure-based drug design," *Chemistry & Biology*, vol. 10, no. 9, pp. 787–797, 2003.
24. S. Grover, M. A. Apushkin, and G. A. Fishman, "Topical dorzolamide for the treatment of cystoid macular edema in patients with retinitis pigmentosa," *American Journal of Ophthalmology*, vol. 141, no. 5, pp. 850–858, 2006.
25. H. Marrakchi, G. Lanéelle, and A. Quémard, "InhA, a target of the antituberculous drug isoniazid, is involved in a mycobacterial fatty acid elongation system, FAS-II," *Microbiology*, vol. 146, no. 2, pp. 289–296, 2000.
26. S. Dadashpour, T. TuyluKucukkilinc, O. Unsal Tan, K. Ozadali, H. Irannejad, and S. Emami, "Design, synthesis and in vitro study of 5, 6-diaryl-1, 2, 4-triazine-3-ylthioacetate derivatives as COX-2 and β -amyloid aggregation inhibitors," *Archiv der Pharmazie*, vol. 348, no. 3, pp. 179–187, 2015.
27. Z. Miller, K.-S. Kim, D.-M. Lee et al., "Proteasome inhibitors with pyrazole scaffolds from structure-based virtual screening," *Journal of Medicinal Chemistry*, vol. 58, no. 4, pp. 2036–2041, 2015.
28. X. Wang, K. Song, L. Li, and L. Chen, "Structure-based drug design strategies and challenges," *Current Topics in Medicinal Chemistry*, vol. 18, no. 12, pp. 998–1006, 2018.

29. V. K. Vyas, R. D. Ukawala, C. Chintha, and M. Ghate, "Homology modeling a fast tool for drug discovery: current perspectives," *Indian Journal of Pharmaceutical Sciences*, vol. 74, no. 1, pp. 1–17, 2012.
30. C. M.-R. Lemer, M. J. Rومان, and S. J. Wodak, "Protein structure prediction by threading methods: evaluation of current techniques," *Proteins: Structure, Function, and Genetics*, vol. 23, no. 3, pp. 337–355, 1995.
31. J. Lee, P. L. Freddolino, and Y. Zhang, "Ab initio protein structure prediction," in *From protein structure to function with bioinformatics*, Springer, 2017.
32. M. T. Muhammed and E. Aki-Yalcin, "Homology modeling in drug discovery: Overview, current applications, and future perspectives," *Chemical Biology & Drug Design*, vol. 93, no. 1, pp. 12–20, 2019.
33. J. Xu, F. Jiao, and L. Yu, "Protein structure prediction using threading," in *Protein structure prediction*, Springer, 2008.
34. M. Yousef, T. Abdelkader, and K. El-Bahnasy, "Performance comparison of ab initio protein structure prediction methods," *Ain Shams Engineering Journal*, vol. 10, no. 4, pp. 713–719, 2019.
35. L. Pan, C. L. Gardner, F. A. Pagliai, C. F. Gonzalez, and G. L. Lorca, "Identification of the tolfenamic acid binding pocket in PrbP from *Liberibacter asiaticus*," *Frontiers in Microbiology*, vol. 8, 2017.
36. T. A. Binkowski, S. Naghibzadeh, and J. Liang, "CASTp: computed atlas of surface topography of proteins," *Nucleic Acids Research*, vol. 31, no. 13, pp. 3352–3355, 2003.
37. A. Volkamer, D. Kuhn, F. Rippmann, and M. Rarey, "DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment," *Bioinformatics*, vol. 28, no. 15, pp. 2074–2075, 2012.
38. J. Sun and K. Chen, "NSiteMatch: prediction of binding sites of nucleotides by identifying the structure similarity of local surface patches," *Computational and Mathematical Methods in Medicine*, vol. 2017, Article ID 5471607, 16 pages, 2017.
39. K. P. Tan, R. Varadarajan, and M. S. Madhusudhan, "DEPTH: a web server to compute depth and predict small-molecule binding cavities in proteins," *Nucleic Acids Research*, vol. 39, suppl_2, pp. W242–W248, 2011.
40. Zhu H, Pisabarro MT. MSPocket: an orientation-independent algorithm for the detection of ligand binding pockets. *Bioinformatics*. 2011 Feb 1;27(3):351-8.
41. Huang B. MetaPocket: a meta approach to improve protein ligand binding site prediction. *OMICS A Journal of Integrative Biology*. 2009 Aug 1;13(4):325-30.
42. Laurie AT, Jackson RM. Q-SiteFinder: an energy-based method for the prediction of protein–ligand binding sites. *Bioinformatics*. 2005 Feb 8;21(9):1908-16.
43. Hu Y, Bajorath J. Entering the 'big data' era in medicinal chemistry: molecular promiscuity analysis revisited. *Future science OA*. 2017 Mar 6;3(2):FSO179.
44. Cincilla G, Thormann M, Pons M. Structuring Chemical Space: Similarity-Based Characterization of the PubChem Database. *Molecular Informatics*. 2010 Jan 12;29(1-2):37-49.
45. Warr W. Report on an NIH workshop on ultralarge chemistry databases.
46. Oniani D, Hilsman J, Zang C, Wang J, Cai L, Zawala J, Wang Y. Emerging opportunities of using large language models for translation between drug molecules and indications. *Scientific Reports*. 2024 May 10;14(1):10738.
47. Heinemann N, Bub S, Wolfram J, Stehle S, Petschick LL, Schulz R. A Compendium of Chemical Class and Use Type Open Access Databases. *Data*. 2020 Dec 4;5(4):114.
48. Bitew M, Desalegn T, Demissie TB, Belayneh A, Endale M, Eswaramoorthy R. Pharmacokinetics and drug-likeness of antidiabetic flavonoids: Molecular docking and DFT study. *Plos one*. 2021 Dec 10;16(12):e0260853.
49. Zadorozhnyi PV, Kiselev VV, Kharchenko AV. In silico ADME profiling of salubrinal and its analogues. *Future Pharmacology*. 2022 Jun 16;2(2):160-97.
50. Guedes IA, de Magalhães CS, Dardenne LE. Receptor–ligand molecular docking. *Biophysical reviews*. 2014 Mar;6:75-87.

51. Yuriev E, Holien J, Ramsland PA. Improvements, trends, and new ideas in molecular docking: 2012–2013 in review. *Journal of Molecular Recognition*. 2015 Oct;28(10):581-604.
52. Spyraakis F, BidonChanal A, Barril X, Javier Luque F. Protein flexibility and ligand recognition: challenges for molecular modeling. *Current topics in medicinal chemistry*. 2011 Jan 1;11(2):192-210.
53. Morris GM, Goodsell DS, Huey R, Hart WE, Halliday S, Belew R, Olson AJ. AutoDock. Automated docking of flexible ligands to receptor-User Guide. 2001 Nov 20.
54. Huey R, Morris GM, Forli S. Using AutoDock 4 and AutoDockvina with AutoDockTools: a tutorial. The Scripps Research Institute Molecular Graphics Laboratory. 2012 Oct 26;10550(92037):1000.
55. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD. Improved protein–ligand docking using GOLD. *Proteins: Structure, Function, and Bioinformatics*. 2003 Sep;52(4):609-23.
56. Wu G, Robertson DH, Brooks III CL, Vieth M. Detailed analysis of grid-based molecular docking: A case study of CDOCKER—A CHARMM-based MD docking algorithm. *Journal of computational chemistry*. 2003 Oct;24(13):1549-62.
57. Schellhammer I, Rarey M. FlexX-Scan: Fast, structure-based virtual screening. *PROTEINS: Structure, Function, and Bioinformatics*. 2004 Nov 15;57(3):504-17.
58. Jain AN. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *Journal of medicinal chemistry*. 2003 Feb 13;46(4):499-511.
59. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal of medicinal chemistry*. 2004 Mar 25;47(7):1750-9.
60. Allen WJ, Balius TE, Mukherjee S, Brozell SR, Moustakas DT, Lang PT, Case DA, Kuntz ID, Rizzo RC. DOCK 6: Impact of new features and current docking performance. *Journal of computational chemistry*. 2015 Jun 5;36(15):1132-56.
61. Bitencourt-Ferreira G, de Azevedo WF. Docking with swissdock. *Docking screens for drug discovery*. 2019:189-202.
62. Adcock SA, McCammon JA. Molecular dynamics: survey of methods for simulating the activity of proteins. *Chemical reviews*. 2006 May 10;106(5):1589-615.
63. Gubbins KE, Moore JD. Molecular modeling of matter: Impact and prospects in engineering. *Industrial & Engineering Chemistry Research*. 2010 Apr 7;49(7):3026-46.
64. Yang Y, Zhu Z, Wang X, Zhang X, Mu K, Shi Y, Peng C, Xu Z, Zhu W. Ligand-based approach for predicting drug targets and for virtual screening against COVID-19. *Briefings in Bioinformatics*. 2021 Mar;22(2):1053-64.
65. Tyagi R, Singh A, Chaudhary KK, Yadav MK. Pharmacophoremodeling and its applications. In *Bioinformatics 2022* Jan 1 (pp. 269-289). Academic Press.
66. Kutlushina A, Khakimova A, Madzhidov T, Polishchuk P. Ligand-based pharmacophoremodeling using novel 3D pharmacophore signatures. *Molecules*. 2018 Nov 27;23(12):3094.
67. So SS. Quantitative Structure-Activity Relationships. *Evolutionary Algorithms in Molecular Design*. 2000 Jul 27:71-97.
68. Pirhadi S, Shiri F, Ghasemi JB. Multivariate statistical analysis methods in QSAR. *Rsc Advances*. 2015;5(127):104635-65.
69. Staszak M, Staszak K, Wieszczycka K, Bajek A, Roszkowski K, Tylkowski B. Machine learning in drug design: Use of artificial intelligence to explore the chemical structure–biological activity relationship. *Wiley Interdisciplinary Reviews: Computational Molecular Science*. 2022 Mar;12(2):e1568.
70. Nayariseri A, Khandelwal R, Tanwar P, Madhavi M, Sharma D, Thakur G, Speck-Planche A, Singh SK. Artificial intelligence, big data and machine learning approaches in precision medicine & drug discovery. *Current drug targets*. 2021 Apr 1;22(6):631-55.
71. Shanmugam A, Muralidharan N, Velmurugan D, Gromiha MM. Therapeutic targets and computational approaches on drug development for COVID-19. *Current Topics in Medicinal Chemistry*. 2020 Sep 1;20(24):2210-20.

72. Gonsalves RC, Pacharla H, Manohar S, Belliraj SK, Tripathi E, Karyala P, Pakala SB. SARS-CoV-2—host cell interactions and pathways: understanding its physiology, pathology, and targeted drug therapy. In *Pandemic Outbreaks in the 21st Century* 2021 Jan 1 (pp. 185-210). Academic Press.
73. Tušar L, Usenik A, Turk B, Turk D. Mechanisms applied by protein inhibitors to inhibit cysteine proteases. *International Journal of Molecular Sciences*. 2021 Jan 20;22(3):997.
74. Kumar S, Paul P, Yadav P, Kaul R, Maitra SS, Jha SK, Chaari A. A multi-targeted approach to identify potential flavonoids against three targets in the SARS-CoV-2 life cycle. *Computers in biology and medicine*. 2022 Mar 1;142:105231.
75. Wang Y, Anirudhan V, Du R, Cui Q, Rong L. RNA-dependent RNA polymerase of SARS-CoV-2 as a therapeutic target. *Journal of medical virology*. 2021 Jan;93(1):300-10.