International Journal of Chemistry and Materials Research
2023 Vol. 11, No. 1, pp. 8-22
ISSN(e): 2311-763X
ISSN(p): 2312-0991
DOI: 10.18488/64.v11i1.3484

A short review on recent developments in the computational techniques (2021) to mitigate SARS-CoV-19 disease

Sachin S. Chourasia[1+]
Tikaram D. Kose[2]
Sudhanshu K. Kharkate[3]
Manoj R. Patle[4]

[1,4]*Department of Chemistry, D. B. Science College, Gondia, India.*
[1]*Email: sachinchourasiagondia@gmail.com*
[4]*Email: manojpatle14@gmail.com*
[2]*Department of Chemistry, Dr. Khatri Mahavidyalaya, Tukum, Chandrapur, India.*
[2]*Email: tikukose@gmail.com*
[3]*Department of Chemistry, D. R. B. Sindhu Mahavidyalaya, Nagpur, India.*
[3]*Email: kharkatesudha@gmail.com*


(+ Corresponding author)

Article History
Received: 24 March 2023
Revised: 7 June 2023
Accepted: 10 August 2023
Published: 21 September 2023

Keywords
Computational chemistry
Coronavirus
COVID-19
Drug discovery and development
Drug repurposing
Machine learning
QSAR
SARS-CoV-19.

ABSTRACT

Severe Acute Respiratory Syndrome Corona Virus 2 (SARS-CoV-2) or SARS Corona Virus 2019 (SARS-CoV-19) disease, which was first reported in December 2019 in China, affected nearly all countries of the world with 65,19,18,402 confirmed cases and causing 66,56,601 deaths worldwide till 12th December 2022. No specific drug has been discovered till now because the discovery of effective and reliable drug requires prolonged research and clinical trials. The efforts to discover treatment against Corona Virus Disease 2019 (CoVID-19) have been expedited by the techniques like high-throughput screening, bioinformatics and cheminformatics revealing molecular pathogenesis of coronaviruses. The application of multidisciplinary studies like computational chemistry, drug repurposing, protein-binding, nano-QSAR, fingerprint techniques have offered invaluable information about atomic viral host receptor interaction, biochemical basis of infection, strains evolution, identification of important viral protein for host-ligand binding studies. In the present review, we discuss the role of various techniques like Quantitative Structure Activity Relationship (QSAR), Machine Learning (ML), Deep Learning (DL), Virtual Screening (VS), Drug repurposing (DR), Molecular Dynamics (MD), nano-QSAR, docking study, Artificial Intelligence (AI) and Language Models (ML) in the process of potent drug/vaccine development against coronavirus. This review is an effort to summarize the reported database and tools for computational studies by bringing together resources in the public domain with respect to structure and pathophysiology of the corona virus. The work will offer an insight in prediction of therapeutics of the coronavirus disease.

Contribution/Originality: Various isolated computational studies have been done in the past but the compatibility and the compliance of various studies have not been studied. This study explores the complementary nature of various techniques under computational chemistry which will help the researchers to develop new model to fight COVID-19.

1. INTRODUCTION

Several studies have already reported a lot of information about the testing, diagnosis, transmission, protein structure, disease spectrum, epidemiology of SARS-CoV-2 [1-9]. Although some drug candidates are under trial [10-19] there is no reliable and accurate drug to treat COVID-19, and due to the time and money involved in the drug/vaccine development various in-silico methods and tools are under consideration. The current review is an

effort to compile the in silico studies carried out in the year 2021 providing information related to computational tools like Computer aided drug designing (CADD), Quantitative Structure Activity Relationship (QSAR), 2D QSAR, 3D QSAR, Machine Learning (ML), Deep Learning (DL), Virtual Screening (VS), Drug Repurposing (DR), nano-QSAR, molecular dynamics (MD) and other computational tools and resources for detection of vaccine/drug discovery. The work already reported have shown promising results to mention a few, peptide inhibition study SARS-CoV-2 receptor using Nanoscale Molecular Dynamics and Chemistry at Harvard Macromolecular Mechanics (NAMD and the CHARMM36 protein force field) [20] main protease-drug binding energy study (AutoDock Vina, SMINA) [21] M^{pro}-peptidomimetic inhibitor study using VS-docking- MD through Assisted Model Building with Energy Refinement, Molecular Operating Environment (AMBER, MOE 2019) [22] repurposing of drugs as SARS-CoV-2 main proteinase (M^{pro}) inhibitor by VS employing docking study (Dock 4.2) [23]; Repurposing of DrugBank molecules as M^{pro} inhibitor using Hierarchical VS (GLIDE), MD (AMDER), B. E. calculations using Molecular Mechanics/Poisson–Boltzmann Surface Area/Weighted Solvent-Accessible Surface Area (MM-PBSA-WSAS) [24]. Plant origin terpenoids as M^{pro} inhibitory employing docking (Molegro Virtual Docker) [25]; Drug repurposing against SARS-CoV2 M^{pro} or $3CL^{pro}$ through docking (AutoDock Vina) [26]; M^{pro}-FDA (Food and Drug Administration) approved drug docking studies using (DOCK6), MD using Sire/Open Molecular Dynamics, GROningen MAchine for Chemical Simulations (SOMD, GROMACS) [27]; Antiviral Chinese herb screening study using (Autodock 4.2) [28]; human ACE2-M^{pro} homology modeling studies Internal Coordinate Mechanics (ICM 3.7.3) using docking [29]; MT-DTI (Molecule Transformer Drug Target Interaction), a pre-trained DL based drug-target interaction model development using (AutoDock Vina) [30]; NSP1 protein modeling, docking and MD-VS study for drug repurposing by using tools Homology modeling (Blastp, modeler 9.12), Docking (Maestro, Schrodinger) and MD (DESMOND) [31]; High throughput VS drug repurposing study against ACE2 host cell receptor and S-protein using Homology modeling (SWISS-MODEL) [32]; molecular modeling of 3CLpro –VS study for repurposing of drug using Docking (AutoDock Vina), VS (MTiOpenScreen) [33]; Antagonistic studies against M^{pro} by reducing viral replication utilizing docking tool (AutoDock Vina), VS tool (MTiOpenScreen) and MD tool (AMBER16) [34]; Inhibition studies of garlic oil against Mpro of SARS-CoV-2 by Docking-MD (SYBYL-X 1.1) [35]; interference studies of hydroxychloroquine against Lys353 in human ACE2 through (AutoDock Vina) [36]; structural and molecular modeling for binding of chloroquine (CLQ) and hydroxychloroquine (HCQ) with sialic acid and gangliosides via. MD Molecular Dynamics and Chemistry at Harvard Macromolecular Mechanics (CHARMM36) [11]; combination therapy of HCQ and curcumin docking studies (PatchDock, FireDock) [37]; MD-VS study of approved drugs for binding energy calculation using through (GLIDE-Schrodinger), B. E. calculation (MM-GBSA) [38]; inhibitory study of 50,437 compounds using multiple linear regression based 2D-QSAR model generated by (Autodock tool 1.5.6) and followed by VS [39]; Pharmacophore and docking based VS targeting SARS-CoV-2M protease using tool Pharmacophore (Vrocs (open eye) and VS (Autodock 4,2) [40]; screening studies of 44 million compounds as potential inhibitors of surface glycoprotein using VS employed docking-ML-CNN (Convolutional Neural Network) (Bind-scope) [41]; VS of Dark Chemical Matter (DCM) and food chemicals as potential inhibitors of M^{pro} using similarity searching, docking(Autodock Vina version 1.1.2 and Molecular Operating Environment MOE v.2019) with Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) profiling (SwissADME) [42]; protein reliability analysis of commercially available 9 Mpro proteins inhibitors using (GLIDE-Schrodinger, pkCSM (pharmacokinetic Cutoff Scanning Matrix) and ProTox-II, and MD using DESMOND-Schrodinger) [43].

This study will prove beneficial to the researches that are seeking the correlation and compatibility between the different computational techniques and wants to get various in-silico information under one umbrella. This work will also provide information about the possible drug candidate and vaccine of coronavirus thus expediting the pace of drug discovery.

The review in general, supports the development of drugs, vaccine and in silico aspect which certainly will be supportive for the discovery of the novel drug candidate to fight corona disease.

2. EXPERIMENTAL AND METHOD

An exhaustive literature review is done from the online and offline sources in order to get an idea of the latest developments in the field of computational chemistry for mitigating the SARS-CoV-2 or COVID-19 disease. As discussed earlier in the introduction, a few of the mentioned research work are elaborated here.

2.1. Confluence of QSAR-ML-DL [44]

The work integrates Molecular dynamics simulation (MDS) and ML including DL with the QSAR so as to remove the shortcoming of the QSAR and develop advance model having application in the field of drug development and clinical trials. The generated QSAR learning models still suffer from some problems [45] and are explained in other literature [46]. The latest model uses general representations like self-learning, AutoEncoder, Bidirectional Encoder Representations from Transformers (BERT), Generative Adversarial Network (GAN) and multitasking [47-51]. The proposed framework not only revealed the importance of the features but explained the contribution of each by correlating them. This gave the distributed vectors which helped to understand the correlation between constituent elements and significance of local structure. The author came to the conclusion that the integration of MD with experimental data will lead to better framework to take advantage of merits of ML and DL. The proposed framework can be optimized in future, to acquire low cost high efficiency especially in relation to QSAR in the field of biomedicine.

2.2. Ultra Large Virtual Screening [52]

The authors reported screening of nearly 1 billion molecules using in silico screening platform VirualFlow, in order to search for inhibitors that target SARS-CoV-2. The work studied 40 different target sites on 17 different viral and host targets emphasizing on critical auxiliary sites such as protein-protein interactions. The work specified three targets viz. interaction interface of Receptor-Binding Domain (RBD) and Angiotensin-converting enzyme 2 (ACE2), host protease Transmembrane protease, serine 2 (TMPRSS2) and HR2 hydrophobic binding groove of HR1. This led to potential 161 drugs which were further filtered on the basis of docking score better than -8kcal/mol and number reduced to 80 drugs. Further filtering left to the final 16 selected drug candidate which were then considered for clinical trials for COVID-19. The results are available at https://vf4covid19.hms.harvard.edu/world-approved-drugs. The work also used ZINC is not commercial (ZINC) 15 database taking into consideration overall 3897 drugs out of which 137 were hits and out of 101378 in man 401 compounds were found to be hit. Though the screening gave large number of hits, experimental verification of the target specificity is to be established leading to mechanism of action. The authors call for validation of the results presented for identification of the potent inhibitors of SARS-CoV-2 as it not possible for a single research group to do the task.

2.3. Repurposing the Anti-Covid Drugs using in-Silico Approach [53]

In order to repurpose drugs to destabilize SARS-CoV-2 S-protein angiotensin converting enzyme 2 (SP-ACE2) complex, docking studies were carried out. 147 drugs were selected from the Drug Bank database [54-56] for the possible binding site in the interface region of SP-ACE2 complex (PDB code:6M17) [57]. Out of these 5 best fit candidates were subjected to molecular dynamics studies carried via Auto Dock 4.2.6 [58] and Auto Grid 4.2.6 [59] carrying out 3 screening 25 runs and 25,000,000 evaluations through Lamarckian genetic algorithm and Solis-Wets local search. All atom molecular dynamics simulation (AA-MD) simulations were performed in Desmond via. Schrodinger-Maestro Software [60, 61]. Binding energy (kcal/mol) of SP-ACE2 without ligand, and complexed

with Amprenavir, Enalaprilat, Plerixafor were calculated using sampling molecular dynamics simulation and found to be -29.58, -20.13, -23.84, -19.72 respectively. A reduction in the binding energy of SP-ACE2 complex was observed in presence of these drugs thus preventing the virus from entering the cell. The decreasing order of the potency was in the order Plerixafor > Amprenavir > Enalaprilat. Thus, these three potential candidates can be subjected to further in-silico and in-vivo studies for development as efficient drugs.

2.4. Application of Nano-QSAR for Binding Studies [62]

The work used the principle that the coronavirus can be considered as nanoparticle having outer capsid protein and inner nucleic material. The Engineered Nano Materials (ENMs) therefore, can competitively bind with the receptor cell of the protein envelope of the virus and stabilize it. The major contributor for the mechanism of binding is electrostatic interaction and is responsible for the adsorption energies of organic molecules onto CNP i.e. carbon based nanomaterials. Authors using the QSAR, developed models to describe the three dimensional interactions of 17 different types of Carbon Nano Particles (CNPs) with the SARS-CoV-2 Ribo Nucleic Acid (RNA) fragment. Various molecular descriptors were calculated and predictive model was developed using Multiwfn 3.8 software [63]. For the study, molecular mechanics simulations were performed using Material Studio software package (Ver. 8.0) and the value of the interaction energy (Eint) was calculated. A negative value of the calculated energy indicated stable absorption ion the CNPs. Carbon nanotubes were found to have the highest interaction followed by graphenes and fullerenes respectively with respect to interaction affinity between SARS-CoV-2 RNA fragment and CNPs. The developed model showed high goodness of fit and robustness for the mentioned interactions. The important descriptors deciding these interactions were found to be surface area, molecular weight and sum of degrees of carbon atoms. The outcome could be extended to other applications like nanotechnology based sensor for coronavirus or personal protective equipment (PPE) disinfection. The work creates significant literature with respect or the absorption/separation and inactivation of the coronavirus which could lead to discovery of novel ENMs as potential drug candidate in future.

2.5. Study of ADMET Properties using 3D-QSAR and MD [64]

Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Index Analysis (CoMSIA), based on the usage of the geometric information to discover the best-suited region responsible for biological activity, were employed for the construction of the 3D-QSAR model. As an anti-COVID medication, a group of 18 carboxamide sulfonamide analogues of chloroquine with known antimalarial action was utilized. The structures of 18 carboxamide sulfonamide derivatives were taken from the literature [65] and their MIC values were converted into the negative log of the Minimum Inhibitory Concentration (MIC) or pMIC. Using the sketch module, the structures were designed, and they were then optimized using the Tripos force field and Gasteiger Huckel charges. Using partial least squares (PLS) analysis, a linear relationship between the antimalarial activity and the CoMFA and CoMSIA descriptors was established. The biological activities of four carboxamide sulfonamide derivatives were then predicted using a 3D-QSAR model, which were validated by R^2 test. This led to the CoMFA model (Q2 = 0.579; R2 = 0.989; R2test = 0.791) and COMSIA model (Q2 = 0.542; R2 = 0.975: R2test = 0.964) and could predict the activity. The models created were accurate and showed how several interactions, including hydrophobic, steric, and electrostatic interactions, affect the activity of molecules. The COVID-19 protein's 3D crystal structure was then downloaded from Protein Data Bank (www.rcsb.org) (PDB code: 6LU7), and molecular docking investigations were conducted by identifying the receptor's active site using the Discovery Studio 2016 programme [66] and PyMol's software [67]. The docking studies revealed that the out of the 8 molecules with important inhibitory activity, molecule under consideration (P1) enjoys high stability with efficient binding as compared to chloroquine. The novel carboxamides sulfonamide derivatives can be proclaimed as inhibitors against COVID-19. The work also emphasized on Adsorption, Distribution, Metabolism, Excretion and

Toxicity (ADMET) studies of the molecules so as to assess the oral bioavailability of these potential dug molecules. ADMET and ADME (Absorption, Distribution, Metabolism, Excretion) properties of the 8 proposed molecules were calculated using pkCSM [68] and SwissADME [69] (ADME =Adsorption, Distribution, Metabolism, Excretion) online tools. The carboxamides sulfonamide scaffold showed good pharmacokinetics with moderate absorption, satisfactory metabolic transformation and no toxicity as far as the inhibitory properties against SARS-CoV-2 is concerned.

2.6. QSAR-MD-MM Approach for Anti-Inflammatory Drug Molecules [70]

This study integrates the various computational approaches to discover the lead molecules having predicted anti-inflammatory activity targeting NF-κB complex (Figure 1).

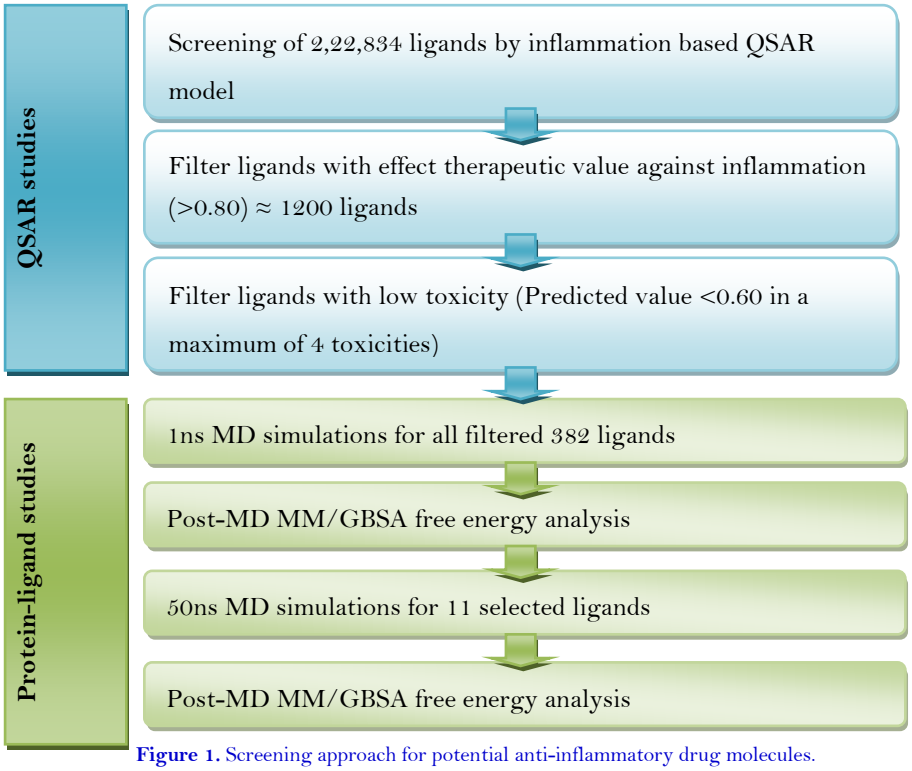

Figure 1. Screening approach for potential anti-inflammatory drug molecules.

It is well known that Nuclear Factor kappa B (NF-κB) mediates and controls the inflammatory response of SARS-CoV-2. 2,22,000 drug like molecules were screened for anti-inflammatory activity against NF-κB/IκBα p50/p65(Re1A) complex. The structure of the molecules were downloaded from https://www.specs.net and protein complexes were prepared using Schrodinger's Maestro suite and optimized using OPLS (Optimized Potentials for Liquid Simulations) force field. A QSAR model was developed by utilizing 598 training set and 93 test set compounds [71] and subjecting them to MetaCore/MetaDrug from Clarivate Analytics platform. A docking study was performed using a grid based docking methodology Glide Standard Precision (Glide/SP) [71-73]. These compounds were then filtered out on the criteria of toxicity less than 0.60 and total 382 compounds showed non-toxicity. Using the Molecular Mechanics-generalized Born Surface Area (MM/GBSA) tool [74-76] and molecular dynamics simulations (MD) via Desmond program [60, 77-79], the filtered out compounds were subjected to free energy calculations. This resulted in the 5 best fit ligands which were having strong protein ligand interaction as compared to known NF-κB inhibitor (procyanidinB2), had conformational stability and no predicted toxicity. The 5 hit ligands were then studied for absorption, distribution, metabolism and excretion (ADME) properties using MetaCore/Metadrug and the outcome suggested they possess drug like profile. To further explore the potency of

the hit molecules as Nuclear Factor Kappa B (NF-κB) inhibitors, they can be tested for chemical optimization and pre-clinical studies. To establish the hit as drug molecules, further in vitro and in-vivo are required.

2.7. Next Generation Computational Tools: From Detection to Vaccine [80]

In addition to virtual screening technologies like phylogenomic, computational structural biology and sequence similarity have proven useful to understand atomic viral-host interaction, biochemical and molecular basis of infection, evolutionary difference across various strains and functional characterization of proteins. Supplementary technique like immune informatics can be used to develop the vaccine by targeting SARS-CoV-2 proteome. This referred work sheds light on the various databases and computational tools for the sequence-structural study of coronavirus by addressing the features of public resources (Figure 2).

Various technologies are used for detection of virus, analysis of genomics of virus, web source for vaccine/drug discovery, docking tools and visualization tools. This will prove beneficial to the researchers to find therapeutics against corona disease. The advances in the multidisciplinary approaches have brought together vast genomic data and findings, which are open sourced for the global scientific community for future development and application.

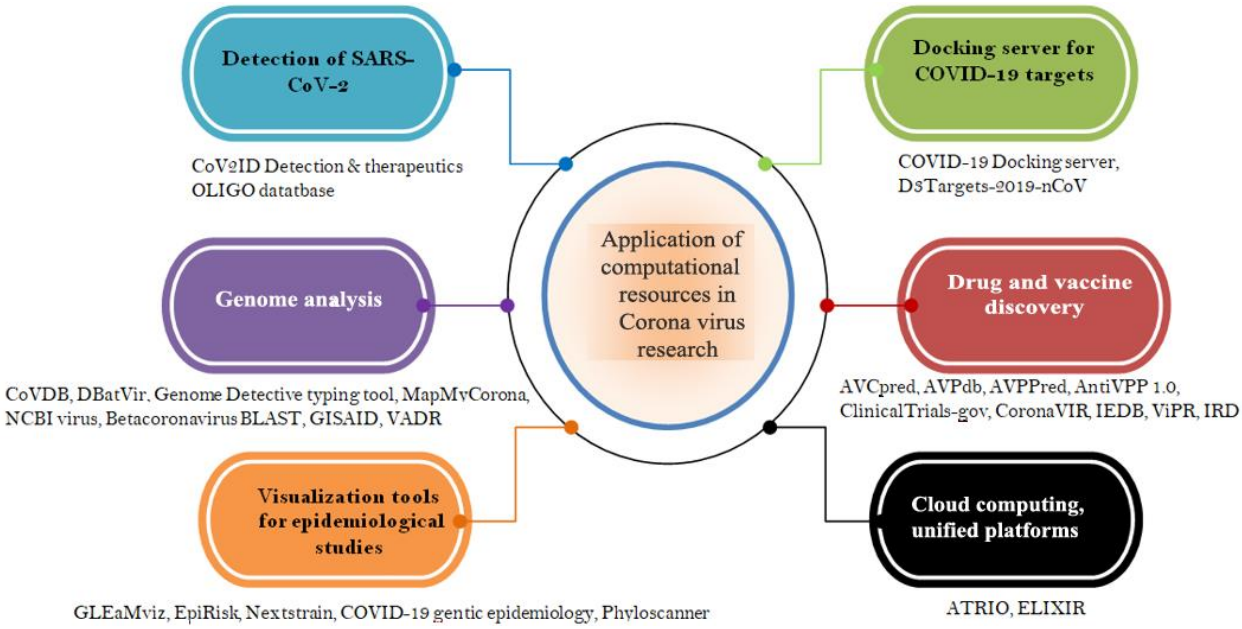

Figure 2. Schematic representation of the different applications of computational resources in coronavirus research.

2.8. CaverDock: Machine Learning Tool for Virtual Screening [81]

CaverDock [82] is an extended tool which used Caver in order to detect tunnel in protein structure and AutoDock Vina algorithm for calculation of binding energies [83] consequently predicting the plausible ligand trajectories along these tunnels. Using the platform, authors performed virtual screening for s-glycoprotein of SARS-CoV-2 in its 6 conformational states with 4358 drug molecules. A total of 26,148 calculations were performed with each calculation taking about 37 minutes to complete proving it to be fast method of virtual screening. Protein data bank RCSB [84] was used to obtain the cryo-EM structure of trimeric spike glycoprotein of SARS-CoV-2 and its trajectory was clustered through CPPTRAJ (main trajectory analysis utility written in C++) [85] and is the module of Amber-tools 16 [86]. HOLE v2.2.005 [81] was used for the characterization of tunnel of timer protein and the output files were converted into CAVER 3 PDB file format [87]. These then were discretized with default setting by the discretiser tool [88] into a series of discs for further CaverDock calculations. For the approved drug dataset, ZINC [89] database was used to download SMILES codes of ligands, stored in CSV (Comma-Separated Values) file, uploaded to Mordred web server [90] so as to get the requisite descriptors. Then the docking studies

of six different protein states with binding energies of 4358 ligands were performed using the Partial Least Square analysis. This resulted into 10 best binders having strong binding capability inside the tunnel and restricting the s-glycoprotein trimer from functional conformational change. The model developed displayed high robustness, with universal applicability to any protein with known tunnel or channel. For the use of scientific community, this developed protocol is planned to be implemented in the next version of CaverWeb [91]. This will allow the scientific community to perform similar work for other proteins and will help to find potential drug molecules from the dataset. The protocol has potential application in the field of drug design and bioengineering e.g. protein and metabolic engineering.

2.9. AI based LM for Drug Discovery and Development [92]

LM which is based on AI resemble human learning process and just like humans, LM knowledge can be transferred to other tasks therefore, have modified the arena of Natural Language Process (NLP). This modification can be effectively applied to the drug as well as treatment development process. The referred work exhaustively signifies the role of AI powered LM in the field of target identification, clinical design, regulatory decision-making and pharmacovigilance in general and is depicted in (Figure 3).

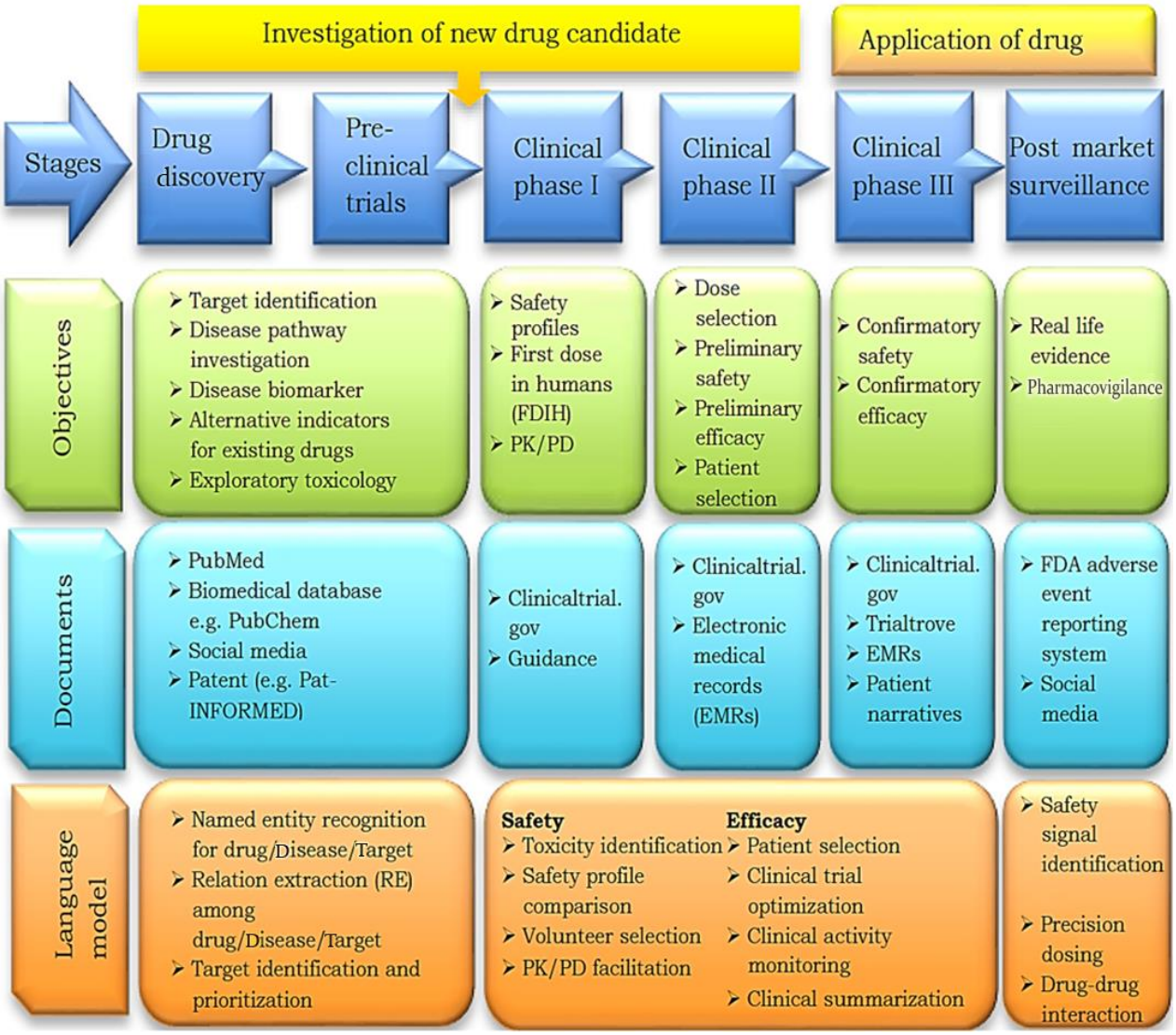

Figure 3. Role of AI based LM with respect to drug discovery and development.

Authors have mentioned a number of transformer based language model viz. BERT, OpenAIGPT, XLNet, ALBERT, RoBERTa, DistillBERT in detail (Table 1 of the article) taking into account various parameters like pertaining corpus, model parameter, training strategies, training time, feature dimension and performance. Authors also have given selected examples of AI based NLP in drug discovery for biomedical named entity recognition, gene based relationship extraction, biomedical text summarization, drug properties prediction, virtual screening, patient trial matching, trial eligibility criteria, biomedical entity normalization, disease coding biomedical mention disambiguation, ADR detection (Table 2 of the article).

Table 1. BLUE tasks.

Corpus	Train	Dev.	Test	Task	Metrics	Domain	Avg.
MedSTS, sentence pairs	675	75	318	Sentence similarity	Pearson	Clinical	25.8
BIOSSES, sentence pair	64	16	20	Sentence similarity	Pearson	Biomedical	22.9
BC5CDR-disease, mentions	4182	4244	4424	NER	F1	Biomedical	22.3
BC5CDR-chemical, mentions	5203	5347	5385	NER	F1	Biomedical	22.3
ShARe/CLEFE, mentions	4628	1075	5195	NER	F1	Clinical	10.6
DDI, relations	2937	1004	979	Relation extraction	micro F1	Biomedical	41.7
ChemProt, relations	4154	2416	3458	Relation extraction	micro F1	Biomedical	34.3
i2b2 2010, relations	3110	11	6293	Relation extraction	F1	Clinical	24.8
HoC, documents	1108	157	315	Document classification	F1	Biomedical	25.3
MedNLI, pairs	11232	1395	1422	Inference	accuracy	Clinical	11.9

Table 2. Corpora.

Corpus	Words	Domain
PubMed abstract	> 4,000M	Biomedical
MIMIC-III	> 500M	Clinical

Table 3. Baseline performance on the BLUE task test sets.

Our BERT								
Task	Metrics	SOTA*	ELMo	BioBERT	Base (P)	Base (P+M)	Large (P)	Large (P+M)
MedSTS	Pearson	83.6	68.6	84.5	84.5	84.8	84.6	83.2
BIOSSES	Pearson	84.8	60.2	82.7	89.3	91.6	86.3	75.1
BC5CDR-disease	F	84.1	83.9	85.9	86.6	85.4	82.9	83.8
BC5CDR-chemical	F	93.3	91.5	93.0	93.5	92.4	91.7	91.1
ShARe/CLEFE	F	70.0	75.6	72.8	75.4	77.1	72.7	74.4
DDI	F	72.9	78.9	78.8	78.1	79.4	79.9	76.3
ChemProt	F	64.1	66.6	71.3	72.5	69.2	74.4	65.1
i2b2	F	73.7	71.2	72.2	74.4	76.4	73.3	73.9
HoC	F	81.5	80.0	82.9	85.3	83.1	87.3	85.3
MedNLI	acc	73.5	71.4	80.5	82.2	84.0	81.5	83.8
Total			78.8	80.5	82.2	82.3	81.5	79.2

Note: * SOTA, state-of-the-art.

The research group, have given a number of publically available FDA documents for AI powered LMs like data labeling FAERS (FDA Adverse Event Reporting System), Orange book, Drugs@FDA, FDA guidance documents, FDA acronyms and abbreviations (Table 3 of the article). In particular, the work discusses the role of LMs strategies in mitigating the COVID-19 emphasizing the drug discovery process. For the assessment of LMs, use of different platforms like Biomedical Language Understanding Evaluation (BLUE) [93] and a question answering dataset, BIOASQ [94] is the can be useful tools. Other models which are based on domain specific corpus and knowledge [95-97] have proven to be beneficial in this regard but they have been place above Bidirectional Encoder Representations from the Transformers (BERT). The work further calls for the training of

the LMs by drug regulatory documents so as to reduce the time for getting the required approvals. Conventionally rule based NLPs have been used in drug discovery and development process and can be complemented by AI based LMs. AI based LMs have shown their application in vast areas of drug discovery especially in clinical drug design. The article concludes that AI based LMs is still in its developmental state and should be allied with other AI based model to enhance the drug discovery and developmental process.

3. DISCUSSION

Various computational techniques like QSAR, ML, VS, and DL have their own pros and cons and if used in combination can deliver better result. Ultra large multi target high throughput virtual screening can provide information about protein binding but has to be consolidated with techniques like Nuclear Magnetic Resonance (NMR) and X-ray crystallography to reveal in depth information about binding site. Drug repurposing/repositioning studies have identified SP (Enalaprilat-Spike) protein in SARS-CoV-2 as possible target for neutralizer, inhibitors drug and vaccine. Techniques like AA-MD and Molecular Dynamics Umbrella Sampling (MD-US) along with QSAR came out as a useful tool for the binding with SP-ACE2 complex. Nanotechnology techniques have proven that the nanoparticles can selectively bind with the SARS-CoV RNA fragment and can stabilize it and the interaction affinity can be used to disrupt the viral replication thus destroying its structure. 3D QSAR techniques like CoMFA and CoMSIA along with pKCSM and SwissADMET tools could easily produce the docking as well as toxicity results for the drug candidates under consideration. New computational tools like CoV2ID and OLIGO can be used for detection; CoVDB, Database of bat-associated viruses (DBatVir), Genome detective typing tool, MapMyCorona, NCBI (National Center for Biotechnology Information), Betacoronavirus BLAST, Global initiative on sharing all influenza data (GISAID), Viral annotation defineR (VADR), D3Targets-2019-nCoV for host pathogen genomic analysis; Antiviral compound prediction(AvCpred), Database for antiviral peptides (AVpdb), AVPpred & AntiVPP 1.0, ClinicalTrials.gov for drug discovery; CoronaVIR, Immune epitope database and analysis resource (IEDB), Viral pathogen resource (ViPR), Influenza research database (IRD) for vaccine discovery, ATRIO (is Composable Cloud computing platform), ELIXIR (Elixir is a functional, dynamically typed language) as unified platform for cloud computing. AI powered LM such as ExBERT can solve the problem of the reproducibility of result and can be applied in various arenas of biomedical research.

4. CONCLUSION

The scientific community all over the world is still trying to find the remedy for the pandemic SARS-CoV-2 or coronavirus. Although no panacea has yet discovered effect are still underway to stop its transmission, mitigate symptoms, find vaccine/drug, and most importantly lower the death rate. We, in the present review have reported the latest computational ways and in-silico tools in an effort for the discovery of the potential drug candidate. Various tools have been discussed to find most suitable candidate with target specific binding capacity and mechanism of action. The discussed resources will certainly expedite the process of drug /vaccine discovery and development process. As a whole the present work provides a strong platform for the drug discovery process related to the COVID-19 mitigation by providing thorough knowledge of potential candidate, supportive data, various in-silico resources and the compatibility of the different computational techniques.

Funding: This study received no specific financial support.
Institutional Review Board Statement: Not applicable.
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] F. Wu *et al.*, "A new coronavirus associated with human respiratory disease in China," *Nature*, vol. 579, no. 7798, pp. 265-269, 2020.

[2] P. Zhou *et al.*, "A pneumonia outbreak associated with a new coronavirus of probable bat origin," *Nature*, vol. 579, no. 7798, pp. 270-273, 2020.

[3] Z.-Q. Zeng *et al.*, "Epidemiology and clinical characteristics of human coronaviruses OC43, 229E, NL63, and HKU1: A study of hospitalized children with acute respiratory tract infection in Guangzhou, China," *European Journal of Clinical Microbiology & Infectious Diseases*, vol. 37, no. 2, pp. 363-369, 2018. https://doi.org/10.1007/s10096-017-3144-z

[4] N. Zhu *et al.*, "A novel coronavirus from patients with pneumonia in China, 2019," *New England Journal of Medicine*, vol. 382, pp. 727–733, 2020.

[5] X. Li, J. Zai, X. Wang, and Y. Li, "Potential of large "first generation" human-to-human transmission of 2019-nCoV," *Journal of Medical Virology*, vol. 92, no. 4, pp. 448-454, 2020. https://doi.org/10.1002/jmv.25693

[6] B. Coutard, C. Valle, X. De Lamballerie, B. Canard, N. G. Seidah, and E. Decroly, "The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade," *Antiviral Research*, vol. 176, p. 104742, 2020. https://doi.org/10.1016/j.antiviral.2020.104742

[7] L. S. Wang, Y. R. Wang, D. W. Ye, and Q. Q. Liu, "A review of the 2019 novel coronavirus (COVID-19) based on current evidence," *International Journal of Antimicrobial Agents*, vol. 55, p. 105948, 2020. https://doi.org/10.1016/j.ijantimicag.2020.105948

[8] D. Wu, T. Wu, Q. Liu, and Z. Yang, "The SARS-CoV-2 outbreak: What we know," *International Journal of Infectious Diseases*, vol. 94, pp. 44-48, 2020. https://doi.org/10.1016/j.ijid.2020.03.004

[9] P. Yang and X. Wang, "COVID-19: A new challenge for human beings," *Cellular & Molecular Immunology*, vol. 17, no. 5, pp. 555-557, 2020. https://doi.org/10.1038/s41423-020-0407-x

[10] X. Yao *et al.*, "In vitro antiviral activity and projection of optimized dosing design of hydroxychloroquine for the treatment of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)," *Clinical Infectious Diseases*, vol. 71, no. 15, pp. 732-739, 2020. https://doi.org/10.1093/cid/ciaa237

[11] J. Fantini, C. Di Scala, H. Chahinian, and N. Yahi, "Structural and molecular modelling studies reveal a new mechanism of action of chloroquine and hydroxychloroquine against SARS-CoV-2 infection," *International Journal of Antimicrobial Agents*, vol. 55, no. 5, p. 105960, 2020. https://doi.org/10.1016/j.ijantimicag.2020.105960

[12] M. Wang *et al.*, "Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro," *Cell Research*, vol. 30, no. 3, pp. 269-271, 2020. https://doi.org/10.1038/s41422-020-0282-0

[13] L. Bouadma, F.-X. Lescure, J.-C. Lucet, Y. Yazdanpanah, and J.-F. Timsit, "Severe SARS-CoV-2 infections: Practical considerations and management strategy for intensivists," *Intensive Care Medicine*, vol. 46, no. 4, pp. 579-582, 2020. https://doi.org/10.1007/s00134-020-05967-x

[14] M. Hoffmann *et al.*, "SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor," *Cell*, vol. 181, no. 2, pp. 271-280. e8, 2020. https://doi.org/10.1016/j.cell.2020.02.052

[15] J. Gao, Z. Tian, and X. Yang, "Breakthrough: Chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies," *Bioscience Trends*, vol. 14, no. 1, pp. 72-73, 2020. https://doi.org/10.5582/bst.2020.01047

[16] X. Tian *et al.*, "Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody," *Emerging Microbes & Infections*, vol. 9, no. 1, pp. 382-385, 2020. https://doi.org/10.1080/22221751.2020.1729069

[17] Recommendations for Investigational COVID-19 Convalescent Plasma, "US FDA ", Retrieved: https://www.fda.gov/. [Accessed 11May 2020], 2020.

[18] T. P. Sheahan *et al.*, "Comparative therapeutic efficacy of remdesivir and combination lopinavir, ritonavir, and interferon beta against MERS-CoV," *Nature Communications*, vol. 11, no. 1, p. 222, 2020. https://doi.org/10.1038/s41467-019-13940-6

[19] D. Guo, "Old weapon for new enemy: Drug repurposing for treatment of newly emerging viral diseases," *Virologica Sinica*, vol. 35, no. 3, pp. 253-255, 2020. https://doi.org/10.1007/s12250-020-00204-7

[20] Y. Han and P. Král, "Computational design of ACE2-based peptide inhibitors of SARS-CoV-2," *ACS Nano*, vol. 14, no. 4, pp. 5143-5147, 2020. https://doi.org/10.1021/acsnano.0c02857

[21] S. Talluri, "Virtual screening based prediction of potential drugs for COVID-19," *Combinatorial Chemistry & High Throughput Screening*, vol. 23, pp. 1-25, 2020. https://doi.org/10.20944/preprints202002.0418.v2

[22] A. Contini, "Virtual screening of an FDA approved drugs database on two COVID-19 coronavirus proteins," 2020. https://doi.org/10.26434/chemrxiv.11847381

[23] F. S. Hosseini and M. Amanlou, "Simeprevir, potential candidate to repurpose for coronavirus infection: Virtual screening and molecular docking study," 2020. https://doi.org/10.20944/preprints202002.0438.v1

[24] J. Wang, "Fast identification of possible drug treatment of coronavirus disease-19 (COVID-19) through computational drug repurposing study," *Journal of Chemical Information and Modeling*, vol. 60, no. 6, pp. 3277-3286, 2020. https://doi.org/10.1021/acs.jcim.0c00179

[25] N. Shaghaghi, "Molecular docking study of novel COVID-19 protease with low risk terpenoides compounds of plants," *Biological and Medicinal Chemistry*, 2020. https://doi.org/10.26434/chemrxiv.11935722.v1

[26] S. Chakraborti, S. Bheemireddy, and N. Srinivasan, "Repurposing drugs against the main protease of SARS-CoV-2: Mechanism-based insights supported by available laboratory and clinical data," *Molecular Omics*, vol. 16, no. 5, pp. 474-491, 2020. https://doi.org/10.1039/d0mo00057d

[27] C. Mendoza-Martinez and A. Rodriguez-Lezama, "Identification of potential inhibitors of SARS-CoV-2 Main protease via a rapid in-silico drug repurposing approach," *Theoretical and Computational Chemistry*, pp. 1-13, 2020. https://doi.org/10.26434/chemrxiv.12085083

[28] D.-h. Zhang, K.-l. Wu, X. Zhang, S.-q. Deng, and B. Peng, "In silico screening of Chinese herbal medicines with the potential to directly inhibit 2019 novel coronavirus," *Journal of Integrative Medicine*, vol. 18, no. 2, pp. 152-158, 2020. https://doi.org/10.1016/j.joim.2020.02.005

[29] C. Wu, Y. Liu, and Y. Yang, "Analysis of therapeutic targets for SARS-CoV2 and discovery of potential drugs by computational methods," *Acta Pharmaceutica Sinica B*, vol. 10, no. 5, pp. 766–788, 2020. https://doi.org/10.1016/j.apsb.2020.02.008

[30] B. R. Beck, B. Shin, Y. Choi, S. Park, and K. Kang, "Predicting commercially available antiviral drugs that may act on the novel coronavirus (SARS-CoV-2) through a drug-target interaction deep learning model," *Computational and Structural Biotechnology Journal*, vol. 18, pp. 784-790, 2020. https://doi.org/10.1016/j.csbj.2020.03.025

[31] A. Sharma, V. Tiwari, and R. Sowdhamini, "Computational search for potential COVID-19 drugs from FDA-approved drugs and small molecules of natural origin identifies several antivirals and plant products," 2020. https://doi.org/10.26434/chemrxiv.12091356.v1

[32] S. Choudhary, Y. S. Malik, and S. Tomar, "Identification of SARS-CoV-2 cell entry inhibitors by drug repurposing using in silico structure-based virtual screening approach," *Frontiers in Immunology*, vol. 11, p. 1664, 2020. https://doi.org/10.3389/fimmu.2020.01664

[33] Y. W. Chen, C.-P. Yiu, and K.-Y. Wong, "Prediction of the 2019-nCoV 3C-like protease (3CLpro) structure: Virtual screening reveals velpatasvir, ledipasvir, and other drug repurposing candidates," *Biological and Medicinal Chemistry*, 2020. https://doi.org/10.26434/chemrxiv.11831103.v1

[34] M. D. L. De Oliveira and K. M. T. De Oliveira, "Comparative computational study of SARS-CoV-2 receptors antagonists from already approved drugs," *Physical Chemistry*, pp. 1-30, 2021. https://doi.org/10.26434/chemrxiv.12044538.v2

[35] B. T. P. Thuy *et al.*, "Investigation into SARS-CoV-2 resistance of compounds in garlic essential oil," *ACS Omega*, vol. 5, no. 14, pp. 8312-8320, 2020. https://doi.org/10.1021/acsomega.0c00772

[36] M. Sehailia and S. Chemat, "Antimalarial-agent artemisinin and derivatives portray more potent binding to Lys353 and Lys31-binding hotspots of SARS-CoV-2 spike protein than hydroxychloroquine: potential repurposing of artenimol for COVID-19," *Journal of Biomolecular Structure and Dynamics*, vol. 39, no. 16, pp. 6184-6194, 2021. https://doi.org/10.1080/07391102.2020.1796809

[37] A. Srivastava and D. Singh, "Destabilizing the structural integrity of SARS-CoV2 receptor proteins by curcumin along with hydroxychloroquine: An insilco approach for a combination therapy," *Biological and Medicinal Chemistry*, 2020. https://doi.org/10.26434/chemrxiv.12090438

[38] L. Mittal, A. Kumari, M. Srivastava, M. Singh, and S. Asthana, "Identification of potential molecules against COVID-19 main protease through structure-guided virtual screening approach," *Journal of Biomolecular Structure and Dynamics*, vol. 39, no. 10, pp. 3662-3680, 2021. https://doi.org/10.1080/07391102.2020.1768151

[39] V. Kumar and K. Roy, "Development of a simple, interpretable and easily transferable QSAR model for quick screening antiviral databases in search of novel 3C-like protease (3CLpro) enzyme inhibitors against SARS-CoV diseases," *SAR and QSAR in Environmental Research*, vol. 31, no. 7, pp. 511-526, 2020. https://doi.org/10.1080/1062936x.2020.1776388

[40] B. S. Andrade *et al.*, "Computational screening for potential drug candidates against the SARS-CoV-2 main protease," *F1000Research*, vol. 9, no. 514, p. 514, 2020. https://doi.org/10.12688/f1000research.23829.2

[41] A. T. Onawole, K. O. Sulaiman, T. U. Kolapo, F. O. Akinde, and R. O. Adegoke, "COVID-19: CADD to the rescue," *Virus Research*, vol. 285, p. 198022, 2020. https://doi.org/10.1016/j.virusres.2020.198022

[42] K. Kapusta *et al.*, "Protein reliability analysis and virtual screening of natural inhibitors for SARS-CoV-2 main protease (Mpro) through docking, molecular mechanic & dynamic, and ADMET profiling," *Journal of Biomolecular Structure and Dynamics*, vol. 39, no. 17, pp. 6810-6827, 2021. https://doi.org/10.1080/07391102.2020.1806930

[43] J. Mao *et al.*, "Comprehensive strategies of machine-learning-based quantitative structure-activity relationship models," *Iscience*, vol. 24, no. 9, p. 103052, 2021. https://doi.org/10.1016/j.isci.2021.103052

[44] J. Dearden, M. T. Cronin, and K. L. Kaiser, "How not to develop a quantitative structure–activity or structure–property relationship (QSAR/QSPR)," *SAR and QSAR in Environmental Research*, vol. 20, no. 3-4, pp. 241-266, 2009. https://doi.org/10.1080/10629360902949567

[45] A. Cherkasov *et al.*, "QSAR modeling: Where have you been? Where are you going to?," *Journal of Medicinal Chemistry*, vol. 57, no. 12, pp. 4977-5010, 2014.

[46] R. Caruana, "Multitask learning," *Machine learning*, vol. 28, pp. 41-75, 1997.

[47] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, pp. 1-16, 2018. https://doi.org/10.48550/arXiv.1810.04805

[48] I. Goodfellow *et al.*, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139-144, 2020.

[49] C.-Y. Liou, W.-C. Cheng, J.-W. Liou, and D.-R. Liou, "Autoencoder for words," *Neurocomputing*, vol. 139, pp. 84-96, 2014. https://doi.org/10.1016/j.neucom.2013.09.055

[50] D. Xu, M. Gopale, J. Zhang, K. Brown, E. Begoli, and S. Bethard, "Unified medical language system resources improve sieve-based generation and bidirectional encoder Representations from transformers (BERT)–based ranking for concept normalization," *Journal of the American Medical Informatics Association*, vol. 27, no. 10, pp. 1510-1519, 2020. https://doi.org/10.1093/jamia/ocaa080

[51] C. Gorgulla *et al.*, "A multi-pronged approach targeting SARS-CoV-2 proteins using ultra-large virtual screening," *Iscience*, vol. 24, no. 2, p. 102021, 2021. https://doi.org/10.1016/j.isci.2020.102021

[52] I. Buitrón-González, G. Aguilera-Durán, and A. Romo-Mancillas, "Potential drugs for destabilizing the SARS-CoV-2 S-protein-angiotensin-converting enzyme 2 complex," *Results in Chemistry*, vol. 3, p. 100094, 2021.

[53] C. Knox *et al.*, "DrugBank 3.0: A comprehensive resource for 'omics' research on drugs," *Nucleic Acids Research*, vol. 39, no. suppl_1, pp. D1035-D1041, 2010. https://doi.org/10.1093/nar/gkq1126

[54] V. Law *et al.*, "DrugBank 4.0: Shedding new light on drug metabolism," *Nucleic Acids Research*, vol. 42, no. D1, pp. D1091-D1097, 2014. https://doi.org/10.1093/nar/gkt1068

[55] D. S. Wishart *et al.*, "DrugBank 5.0: A major update to the DrugBank database for 2018," *Nucleic Acids Research*, vol. 46, no. D1, pp. D1074-D1082, 2018. https://doi.org/10.1093/nar/gkx1037

[56] R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo, and Q. Zhou, "Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2," *Science*, vol. 367, no. 6485, pp. 1444-1448, 2020. https://doi.org/10.1126/science.abb2762

[57] D. Santos-Martins, L. Solis-Vasquez, A. F. Tillack, M. F. Sanner, A. Koch, and S. Forli, "Accelerating AutoDock4 with GPUs and gradient-based local search," *Journal of Chemical Theory and Computation*, vol. 17, no. 2, pp. 1060-1073, 2021. https://doi.org/10.1021/acs.jctc.0c01006

[58] G. M. Morris *et al.*, "AutoDock4 and Autodock tools4: Automated docking with selective receptor flexibility," *Journal of Computational Chemistry*, vol. 30, no. 16, pp. 2785–2791, 2009. https://doi.org/10.1002/jcc.21256

[59] K. J. Bowers *et al.*, "Scalable algorithms for molecular dynamics simulations on commodity clusters," in *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*, 2006, pp. 84-es.

[60] L. Schrodinger, *Maestro-desmond interoperability tools*. New York: Schrödinger, 2021.

[61] F. Zhang, Z. Wang, M. G. Vijver, and W. J. Peijnenburg, "Probing nano-QSAR to assess the interactions between carbon nanoparticles and a SARS-CoV-2 RNA fragment," *Ecotoxicology and Environmental Safety*, vol. 219, p. 112357, 2021. https://doi.org/10.1016/j.ecoenv.2021.112357

[62] T. Lu and F. Chen, "Quantitative analysis of molecular surface based on improved Marching Tetrahedra algorithm," *Journal of Molecular Graphics and Modelling*, vol. 38, pp. 314-323, 2012. https://doi.org/10.1016/j.jmgm.2012.07.004

[63] A. Khaldan *et al.*, "Prediction of potential inhibitors of SARS-CoV-2 using 3D-QSAR, molecular docking modeling and ADMET properties," *Heliyon*, vol. 7, no. 3, p. 06603, 2021. https://doi.org/10.1016/j.heliyon.2021.e06603

[64] D. Ugwu, U. Okoro, P. Ukoha, S. Okafor, A. Ibezim, and N. Kumar, "Synthesis, characterization, molecular docking and in vitro antimalarial properties of new carboxamides bearing sulphonamide," *European Journal of Medicinal Chemistry*, vol. 135, pp. 349-369, 2017. https://doi.org/10.1016/j.ejmech.2017.04.029

[65] D. Systèmes, *Biovia, discovery studio modeling environment*. San Diego, CA, USA: Dassault Systèmes Biovia, 2016.

[66] W. DeLano, "The PyMOL molecular graphics system DeLano Scientific, Palo Alto, CA, USA," Retrieved: http://www.pymol.org. [Accessed 25 February 2017], 2002.

[67] D. E. Pires, T. L. Blundell, and D. B. Ascher, "PkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures," *Journal of Medicinal Chemistry*, vol. 58, no. 9, pp. 4066-4072, 2015. https://doi.org/10.1021/acs.jmedchem.5b00104

[68] A. Daina, O. Michielin, and V. Zoete, "SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules," *Scientific Reports*, vol. 7, no. 1, p. 42717, 2017. https://doi.org/10.1038/srep42717

[69] T. Kanan, D. Kanan, E. J. Al Shardoub, and S. Durdagi, "Transcription factor NF-κB as target for SARS-CoV-2 drug discovery efforts using inflammation-based QSAR screening model," *Journal of Molecular Graphics and Modelling*, vol. 108, p. 107968, 2021. https://doi.org/10.1016/j.jmgm.2021.107968

[70] T. Kanan, D. Kanan, I. Erol, S. Yazdi, M. Stein, and S. Durdagi, "Targeting the NF-κB/IκBα complex via fragment-based E-Pharmacophore virtual screening and binary QSAR models," *Journal of Molecular Graphics and Modelling*, vol. 86, pp. 264-277, 2019. https://doi.org/10.1016/j.jmgm.2018.09.014

[71] T. A. Halgren *et al.*, "Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy," *Journal of Medicinal Chemistry*, vol. 47, pp. 1750-1759, 2004.

[72] L. L. C. Schrödinger, *Glide, version 7.1*. Schrödinger, LLC: New York, USA, 2016.

[73] T. Hou, J. Wang, Y. Li, and W. Wang, "Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations," *Journal of Chemical Information and Modeling*, vol. 51, no. 1, pp. 69-82, 2011.

[74] T. Hou, J. Wang, Y. Li, and W. Wang, "Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking," *Journal of Computational Chemistry*, vol. 32, no. 5, pp. 866-877, 2011. https://doi.org/10.1002/jcc.21666

[75] B. R. Miller III, T. D. McGee Jr, J. M. Swails, N. Homeyer, H. Gohlke, and A. E. Roitberg, "MMPBSA. py: An efficient program for end-state free energy calculations," *Journal of Chemical Theory and Computation*, vol. 8, no. 9, pp. 3314-3321, 2012. https://doi.org/10.1021/ct300418h

[76] D. E. Shaw, *Desmond molecular dynamics system*. New York: LLC, 2009.

[77] W. G. Hoover, "Canonical dynamics: Equilibrium phase-space distributions," *Physical Review A*, vol. 31, no. 3, pp. 1695-1697, 1985. https://doi.org/10.1103/physreva.31.1695

[78] G. J. Martyna, D. J. Tobias, and M. L. Klein, "Constant pressure molecular dynamics algorithms," *The Journal of Chemical Physics*, vol. 101, no. 5, pp. 4177-4189, 1994. https://doi.org/10.1063/1.467468

[79] R. Kangabam *et al.*, "Next-generation computational tools and resources for coronavirus research: From detection to vaccine discovery," *Computers in Biology and Medicine*, vol. 128, p. 104158, 2021. https://doi.org/10.1016/j.compbiomed.2020.104158

[80] G. P. Pinto, O. Vavra, S. M. Marques, J. Filipovic, D. Bednar, and J. Damborsky, "Screening of world approved drugs against highly dynamical spike glycoprotein of SARS-CoV-2 using CaverDock and machine learning," *Computational and Structural Biotechnology Journal*, vol. 19, pp. 3187-3197, 2021. https://doi.org/10.1016/j.csbj.2021.05.043

[81] E. Chovancova, A. Pavelka, P. Benes, O. Strnad, J. Brezovsky, and B. Kozlikova, "3.0: A tool for the analysis of transport pathways in dynamic protein structures," *PLOS Computational Biology*, vol. 8, no. 10, p. e1002708, 2012. https://doi.org/10.1371/journal.pcbi.1002708

[82] O. Trott and A. J. Olson, "AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading," *Journal of Computational Chemistry*, vol. 31, no. 2, pp. 455-461, 2010.

[83] P. W. Rose *et al.*, "The RCSB protein data bank: New resources for research and education," *Nucleic Acids Research*, vol. 4, pp. D475-D482, 2013.

[84] D. Roe, "PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data," *Journal of Chemical Theory and Computation*, vol. 9, no. 7, pp. 3084-3095, 2013. https://doi.org/10.1021/ct400341p

[85] D. Case *et al.*, *Amber 16*. San Francisco: University of California, 2016.

[86] O. S. Smart, J. G. Neduvelil, X. Wang, B. Wallace, and M. S. Sansom, "HOLE: A program for the analysis of the pore dimensions of ion channel structural models," *Journal of Molecular Graphics*, vol. 14, no. 6, pp. 354-360, 1996. https://doi.org/10.1016/s0263-7855(97)00009-x

[87] J. Filipovič *et al.*, "CaverDock: A novel method for the fast analysis of ligand transport," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 17, no. 5, pp. 1625-1638, 2019. https://doi.org/10.1109/tcbb.2019.2907492

[88] T. Sterling and J. J. Irwin, "ZINC 15–ligand discovery for everyone," *Journal of Chemical Information and Modeling*, vol. 55, no. 11, pp. 2324-2337, 2015. https://doi.org/10.1021/acs.jcim.5b00559

[89] H. Moriwaki, Y.-S. Tian, N. Kawashita, and T. Takagi, "Mordred: A molecular descriptor calculator," *Journal of Cheminformatics*, vol. 10, no. 1, pp. 1-14, 2018. https://doi.org/10.1186/s13321-018-0258-y

[90] J. Stourac *et al.*, "Caver Web 1.0: Identification of tunnels and channels in proteins and analysis of ligand transport," *Nucleic Acids Research*, vol. 47, no. W1, pp. W414-W422, 2019. https://doi.org/10.1093/nar/gkz378

[91] Z. Liu, R. A. Roberts, M. Lal-Nag, X. Chen, R. Huang, and W. Tong, "AI-based language models powering drug discovery and development," *Drug Discovery Today*, vol. 26, no. 11, pp. 2593-2607, 2021. https://doi.org/10.1016/j.drudis.2021.06.009

[92] Y. Peng, S. Yan, and Z. Lu, "Transfer learning in biomedical natural language processing: An evaluation of BERT and ELMo on ten benchmarking datasets," in *Proceedings of the 18th BioNLP Workshop and Shared Task Florence, Italy. Association for Computational Linguistics*, 2019, pp. 58–65, doi: https://doi.org/10.18653/v1/w19-5006.

[93] G. Tsatsaronis *et al.*, "An overview of the BIOASQ large-scale biomedical semantic indexing and question answering competition," *BMC Bioinformatics*, vol. 16, no. 1, pp. 1-28, 2015.

[94] J. Lee *et al.*, "BioBERT: A pre-trained biomedical language representation model for biomedical text mining," *Bioinformatics*, vol. 36, no. 4, pp. 1234-1240, 2020. https://doi.org/10.1093/bioinformatics/btz682

[95] K. Huang, J. Altosaar, and R. Ranganath, "Clinicalbert: Modeling clinical notes and predicting hospital readmission," *arXiv preprint arXiv:1904.05342*, 2019. https://doi.org/10.48550/arXiv.1904.05342

[96] I. Beltagy, K. Lo, and A. Cohan, "SciBERT: A pretrained language model for scientific text," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019.

[97] Y. Peng, S. Yan, and Z. Lu, "Transfer learning in biomedical natural language processing: An evaluation of BERT and ELMo on ten benchmarking datasets," in *Proceedings of the 18th BioNLP Workshop and Shared Task, Florence, Italy. Association for Computational Linguistics*, 2019, pp. 58–65, doi: https://doi.org/10.18653/v1/w19-5006.