

Identification of Potential Inhibitors of SARS-CoV-2 Using Machine Learning, Molecular Docking and MD Simulation

Anuraj Nayariseri

anuraj@eminentbio.com

Eminent Biosciences

Anushka Bhirdwaj

Eminent Biosciences

Arshiya Khan

Eminent Biosciences

Khushboo Sharma

Eminent Biosciences

Uzma Shaheen

Eminent Biosciences

Umesh Panwar

Eminent Biosciences

V. Natchimuthu

M.Kumarasamy College of Engineering

Rinku chaudhary

Eminent Biosciences

Abhishek Kumar

Eminent Biosciences

Taniya Dey

Eminent Biosciences

Aravind Panicker

Eminent Biosciences

Leena Prajapati

Eminent Biosciences

Francisco Jaime Bezerra

State University of Paraiba

Sanjeev Kumar Singh



Alagappa University

Article

Keywords: SARS-CoV-2 inhibitors, Machine-Learning, Deep Learning, Molecular Docking, Molecular Dynamics Simulation, R Programming, Python, ADMET studies

Posted Date: August 20th, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-4323991/v2>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Identification of Potential Inhibitors of SARS-CoV-2 Using Machine Learning, Molecular Docking and MD Simulation

Anuraj Nayarissari^{*1,2}, Anushka Bhirdwaj^{1,2,3}, Arshiya Khan^{1,2,3}, Khushboo Sharma^{1,2,3}, Uzma Shaheen¹, Umesh Panwar^{1,3}, V. Natchimuthu⁴, Rinku Chaudhary¹, Abhishek Kumar¹, Taniya Dey¹, Aravind Panicker¹, Leena Prajapati¹, Francisco Jaime Bezerra Mendonça Junior⁵, Sanjeev Kumar Singh^{*3}

1. In silico Research Laboratory, Eminent Biosciences, 91, Sector-A, Mahalakshmi Nagar, Indore – 452010, Madhya Pradesh, India.
2. Bioinformatics Research Laboratory, LeGene Biosciences Pvt Ltd, 91, Sector-A, Mahalakshmi Nagar, Indore – 452010, Madhya Pradesh, India.
3. Computer Aided Drug Designing and Molecular Modelling Lab, Department of Bioinformatics, Alagappa University, Karaikudi - 630003, Tamil Nadu, India.
4. Department of Physics, M.Kumarasamy College of Engineering, Karur - 639113, Tamil Nadu, India.
5. Laboratory of Synthesis and Drug Delivery, Department of Biological Sciences, State University of Paraíba, João Pessoa 58429-500, Brazil.

* Corresponding authors:

Dr. Anuraj Nayarissari

Email: anuraj@eminentbio.com

In silico Research Laboratory, Eminent Biosciences, Mahalakshmi Nagar, Indore – 452010, Madhya Pradesh, India.

Bioinformatics Research Laboratory, LeGene Biosciences Pvt Ltd, 91, Sector-A, Mahalakshmi Nagar, Indore – 452010, Madhya Pradesh, India.

Tel: +91 9752295342

ORCID: <https://orcid.org/0000-0003-2567-9630>

Dr. Sanjeev Kumar Singh

Email: skysanjeev@gmail.com

Affiliation: Department of Data Sciences, Centre of Biomedical Research, SGPGIMS Campus, Raebareli Rd, Lucknow, Uttar Pradesh - 226014, India

ORCID: <https://orcid.org/0000-0003-4153-6437>

Abstract

The advent of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the etiological agent of the coronavirus disease 2019 (COVID-19) pandemic, has promoted physical and mental health worldwide. Due to the unavailability of effective antiviral drugs, there is an unmet demand for a robust therapeutic approach for the development of anti-SARS-CoV-2 drugs. Myriad investigations have recognized ACE2 as the primary receptor of SARS-CoV-2, and this amalgamation of ACE2 with the spike protein of the subsequent coronavirus is paramount for viral entry into host cells and inducing infection. Consequently, limiting or restricting the accessibility of the causal virus to ACE2 offers an alternative

therapeutic approach for averting this illness. Thus, the objective of the study was to determine the highly efficacious inhibitors exhibiting an augmented affinity for ACE2 protein and assess their pharmacological efficacy using molecular docking analysis. Machine learning algorithms were employed to govern the novel compounds by taking the ACE2-inhibiting compounds, Quinapril, Moexipril, etc, and pre-established repurposed viral compounds, Birinapant, Remdesivir, etc., as test datasets. Structural stability was further confirmed via MD simulation approach which comparatively assessed the novel machine-learning, and pre-established compounds, followed by toxicity and pharmacophore studies. The study therefore concludes that the novel machine-learning compound (PubChem ID: 23658468) can be a potent therapeutic agent for combatting SARS-CoV-2.

Keywords: SARS-CoV-2 inhibitors, Machine-Learning, Deep Learning, Molecular Docking, Molecular Dynamics Simulation, R Programming, Python, ADMET studies.

1. Introduction

The spread of catastrophic coronavirus 2019 (COVID-19) or severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has affected >2.3 million people and >1,60,000 deaths globally, causing massive disruptions to social, economic, and public healthcare paradigms [1-2]. Studies have documented seven endemic human coronaviruses, such as, (HCoV)-OC43, 229E, HKU1 and NL63 strains which cause mild respiratory tract infections. Over the past two decades, the world has witnessed viral epidemics, such as severe acute respiratory syndrome (SARS-CoV), H5N1 influenza A, H1N1, and Middle East respiratory syndrome (MERS-CoV), which cause acute lung injury (ALI) and acute respiratory distress syndrome (ARDS) [3-5].

SARS-CoV-2, the novel beta-coronavirus responsible agent for COVID-19 pandemic, originated in Wuhan, China has emerged as a lethal infection. The genomic sequence of SARS-CoV-2 has 96% similarity with that of the bat coronavirus RaTG13, which suggests that bats are the most likely source of zoonotic infection [6-7]. SARS-CoV-2 exhibits a genome of 30 kb in size, with an enveloped, non-segmented, positive single-stranded RNA, encoding two open reading frames (ORFs); ORF 1a and 1b. The largest coding gene in the downstream region is of nonstructural proteins (Nsp), such as proteases (3CLpro and PLpro), RNA-dependent RNA polymerase (RdRp), helicases and other accessory proteins involved in viral replication and assembly in the host cell. Moreover, the genome also encodes four

structural proteins; envelope (E), membrane (M), spike (S), and nucleocapsid (N) glycoproteins (GPS)[8-11].

The transmembrane S protein SARS-CoV-2 has been identified to exploit the host entry receptor - angiotensin-converting enzyme 2 (ACE2), and the host cellular serine protease TMPRSS2 type-2 enzyme which cleaves this S protein into S1 and S2 subunits. S1 exhibits the receptor-binding domain (RBD) that binds to the extracellular N-terminus or protease domain (PD) of the full-length ACE2 receptor. Consequently, S2 endo C-terminus region facilitates the fusion of cellular and viral membranes, thus allowing the entry of SARS-CoV-2, and ultimately its replication in target cells, as illustrated in Figure 1 [12-13]. The 3D structure of S protein in SARS-CoV-2 demonstrates a similar structure to SARS-CoV except for a few amino acid variations at key residues, and the S protein of SARS-CoV-2 exhibits stronger interactions and greater binding affinity with ACE2 as compared to SARS-CoV in humans and bats [14-17].

The ACE2 receptor is identified as a carboxy peptidase enzyme. In humans, it is broadly expressed in myriad organs/tissues, such as the kidney, heart, intestines, and extensively expressed in lung alveolar type 2 (AT2) cells [18]. Physiologically, ACE2 regulates the renin-angiotensin system (RAS) signaling pathway and converts angiotensin I and II into angiotensin 1-9 and angiotensin 1-7, respectively [19]. Primarily, SARS-CoV-2 infection occurs in mucosal cells and spreads to AT2 cells, inhibiting ACE2 expression and disrupting the RAS pathway, which is likely involved in the pathogenesis of ARDS and acute lung injury (ALI), further leading to inflammation [20]. Knockout of the ACE2 receptor in mice after experimental SARS-CoV infection significantly mitigated its replication and infection. Hence, the interaction of the SARS-CoV-S protein with ACE2 is critical for the pathogenesis of SARS-CoV infection [21]. Diacritic classes of *In-vitro* tested drugs, such as remdesivir, favipiravir (FPV), darunavir, ritonavir, hydroxychloroquine, teicoplanin, nitazoxanide, chloroquine, and tocilizumab, have demonstrated activity against COVID-19 infection.

To date, there are no suitable drugs available for treating SARS-CoV-2 infection, thus there is an urgent demand for the development of new anti-COVID-19 drugs. Drug repurposing has emerged as a promising strategy and is defined as the process of generating an innovative pharmacological approach for existing, and approved drugs. Similarly, repurposing drugs not only mitigates the duration of chemical compound optimization but also eliminates the cost

of testing the toxicity of the compounds compared to the discovery of *de novo* drugs [22-26]. Furthermore, non-structural proteins of SARS-CoV-2, such as RNA-dependent RNA polymerase (RdRp) and the main protease Mpro, are being targeted to discover vaccines against COVID-19 [27].

This study emphasizes the role of ACE2 as a potential therapeutic target for identifying new treatment strategies to prevent COVID-19 infection and transmission. Provisionally, the use of full-length ACE2 inhibitors as small molecules or compounds could lead to the development of novel therapeutic approaches that block SARS-CoV-2-S protein-mediated cell fusion to prevent entry into target cells.

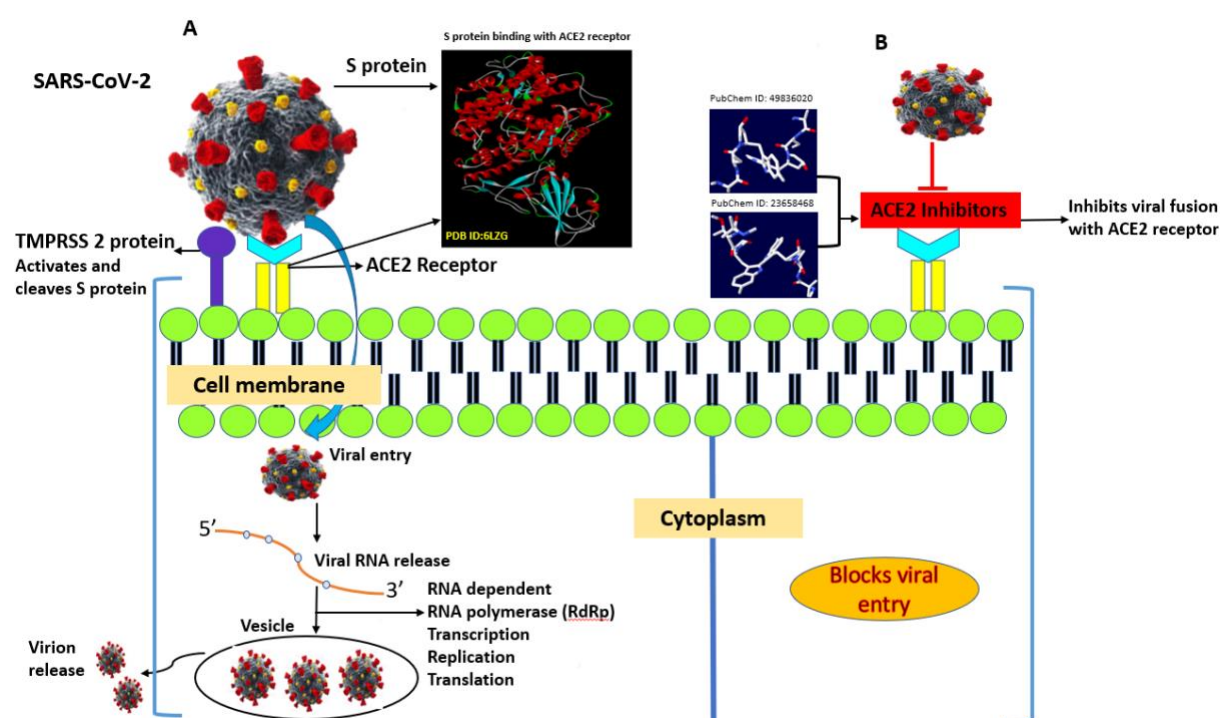


Figure1. Pictorial representation of the binding of the SARS-CoV-2 S protein to the ACE2 receptor. (A) The SARS-CoV-2 S protein interacts with the ACE2 receptor, and the virus enters the host cell. (B) ACE2 inhibitors prevent the binding of the SARS-CoV-2 S protein to ACE2 receptors and block viral entry into the host cell.

2. Methodology

2.1 Selection of Inhibitors

A comprehensive literary study was conducted to assess the effectiveness of selected inhibitors against the target protein ACE2 (PDB ID: 6LZG) [28]. In total, 201 inhibitors were

employed for this study. There are two categories of inhibitors - type I; constitutes repurposed antiviral compounds such as birinapant, remdesivir, lopinavir, and ivermectin, and type II evinces particular ACE2 inhibitors; Quinapril, moexipril, benazepril, elanapril, etc. [Supplementary Table 1] displays the PubChem ID, molecular weight in g/mol, H-bond donor and acceptor, and log-*P* values. The structure of each retrieved compound was stored in 3D SDF Format [29-92].

2.2 Protein and Ligand Preparation

The structural coordinates of protein (HACE-2) human angiotensin-converting enzyme-2 with bound co-crystallized molecules (PDB ID: 6LZG) [28] were obtained from the Protein Data Bank with a resolution of 2.5 Å and R value-free parameters (0.218). The 3-dimensional structure of the target protein was retrieved and pre-processed using the Protein Preparation Wizard Module in Schrödinger 2020, which modifies the structural coordinates using finite gamuts, includes tasks like assigning accurate bond orders and zero-order bonds to metal atoms, addressing the absence of hydrogen cavities, side chains, and loops [93-97]. Moreover, for ligand binding, numerous tautomerization and protonation states were envisioned at pH 7 particularly. Finally, the target protein, ACE2, was refined and minimized in the force field of OPLS3e, which has an RMSD value of 0.3 Å. All pre-established inhibitors were retrieved from the PubChem database and subjected to analysis using the LigPrep Module within the Schrödinger Suite 2020. The process entailed optimization, ring conformation, conversion from 2D to 3D states, resolving stereoisomers, tautomers, and stabilizing ionization states at a pH of 7.00. Additionally, partial atomic charges were assigned using the OPLS3e force field [98-103].

2.3 Molecular Docking

The Molecular Docking Program Molegro Virtual Docker (MVD) was employed which utilizes the piecewise linear potential (PLP), a heuristic search function, and the MolDock scoring function with an appropriate docking paradigm [104-108]. During ligand binding, a large cavity was generated by orchestrating the chemical-tailored complex structure by removing the drug and pre-existing ligand from the hACE2 protein structure. Consequently, the primal cavity, exhibiting a volume of 4825.09 Å³, was assigned as an active site by the cavity prediction algorithm. The governing docking parameters were maximum population size (50), maximum number of iterations (1500), and grid resolution (0.3) [109-114].

The post-docking procedure required hydrogen bond optimization and energy reduction, which required the MolDock simplex evolution to be set at 300 maximum steps with a neighbor distance factor of 1.00. To minimize the complex energy of the ligand-receptor interaction, Nelder–Mead simplex minimization algorithm was employed using hydrogen bond directionality, and a non-grid force field. The best procured compound from docking was screened based on Gibb’s free energy for executing machine-learning models [115-121].

2.4 Machine Learning

The two most effective compounds obtained from molecular docking of the repurposed compound and ACE2 inhibitors were further tested for the identification of better compounds that can bind to ACE2 via machine learning. The present investigation describes training dataset preparation and preprocessing, molecular prediction using an internally supervised machine learning technique, and molecular refinement using docking and molecular dynamics [Figure 2][80][98-99].

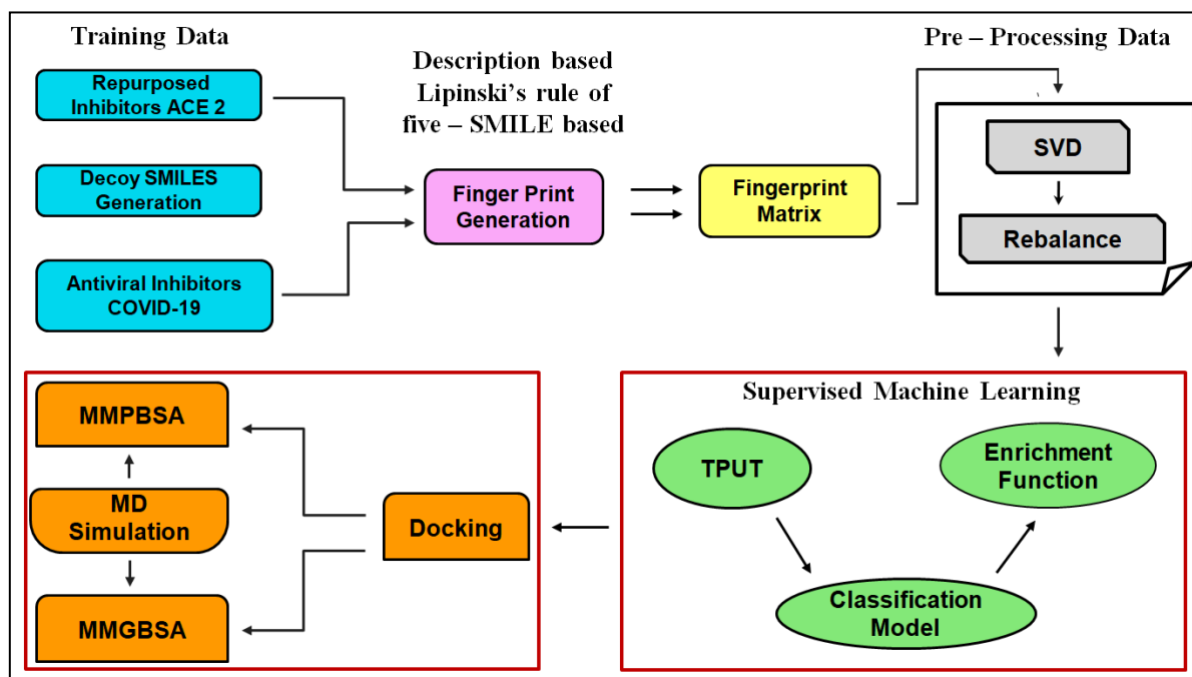


Figure 2. Current Investigation Strategy: Data preparation, preprocessing of the data, supervised learning, and refining of simulations make up its four building blocks. The ellipsoids represent the datasets, while the rectangles denote the processing stages.

2.4.1 Data preparation

The present investigation employed the “Database of Useful Decoys: Enhanced (DUD-E)” server for generating 700 molecular decoys for both the repurposed compounds and ACE2 inhibitors [122-124]. For feature selection in machine learning, chemical data are transformed into an instructive numerical feature that predictive models can make sense of. This was accomplished by applying molecular fingerprints, which are numerical properties taken from chemical structures; furthermore, these fingerprints were used to compare two chemical substances. By determining the Tanimoto coefficient based on the chemical fingerprints of the two chemicals, the structural similarity between them was assessed by the following equation [98-99], [124-125]:

$$T(a, b) = \frac{N_c}{N_a + N_b - N_c}$$

Where N indicates the number of attributes in each object (a,b). C is the intersection set. The Tanimoto coefficient for both the repurposed and ACE2 inhibitor datasets was generated by using the following Python script.

```
import pandas as pd
lista=pd.read_csv('repurposed.csv', sep='\t')
listb=pd.read_csv('aceII.csv', sep='\t')
def tanimoto (lista, listb):
    intersection = [common_item for common_item in lista if common_item in listb]
    return float(len(c))/(len(a) + len(b) - len(c))
```

The two sets of compounds were compared when their values were higher, but they did not reveal the chemical groups that they specifically shared. Various strategies were employed to improve the overall quality of the data and aid machine learning strategies in capturing better correlations between features outside of the Tanimoto coefficient to avoid such problems. The desired dataset was composed of high-affinity ligands, ACE2 inhibitors, and decoys and was derived from the repurposed dataset. The molecular characteristics were defined by using a mixture of two molecular fingerprints that were applied by RDKit [125-127].

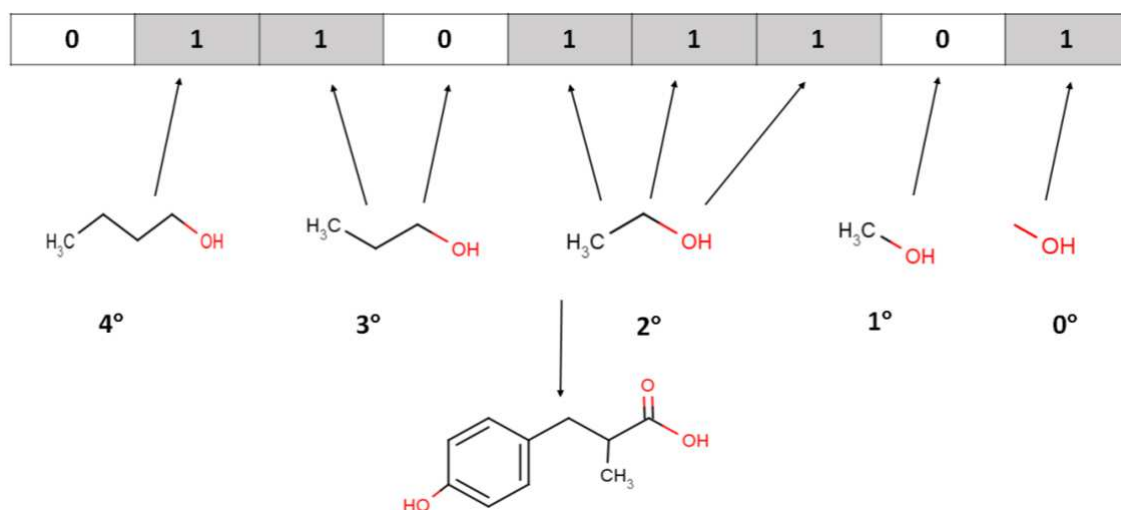


Figure 3. An illustration of a hypothetical 10-bit topological fingerprint, in this case, a 5-fragment linear path-based fingerprint.

The first method involved a topological-based fingerprint search and collection of the chemical patterns present in the obtained chemical graph, and the second method involved structure key-based fingerprint enactment of the public MACCS keys, indicating the presence or absence of specific chemical features or chemical substructures from a list of structural keys assigned to specific compounds [128-132]. The present research created a data matrix exhibiting the numerical chemical information of high-affinity compounds from repurposed and ACE2 datasets and fingerprint features, which were stored as column decoy information in the rows of the matrix, as shown in Figure 3. This matrix was further used for data processing.

2.4.2 Preprocessing data

The data were preprocessed to improve the predictive performance. Because molecules vary in size and functional groups, the matrix data generated by fingerprints are sparse. As a result, many cells were filled with zeros. Furthermore, the assumption that similar compounds in the datasets will have similar activities may not always be correct because minor changes in the compound's functional groups can cause an abrupt change in chemical biological activity. Even a small change in the descriptor values of the compounds will create significant changes in their molecular properties. As a result, the current study did not use fingerprint data directly but rather singular value decomposition (SVD), which provides lower-dimensional projections of the data. SVD reduces the dimensionality of the raw data matrix while preserving essential properties and detecting meaningful patterns while discarding weak signals and noise in both the repurposed and ACE2 inhibitor datasets. The current study also

employed a resampling technique to re-energize the sample space to mitigate the effects of skewed class distributions during the machine learning process. Most of the compounds in the dataset were subdivided into similar sizes and positive classes. All the negative chemical datasets were combined with the positive chemical datasets, which resulted in multiple symmetrical sub-datasets that were further used as inputs for ensemble-based machine learning classifiers [133-138].

2.4.3 Supervised Machine Learning

After preprocessing the data, the matrix of active substances and their corresponding decoys for each target served as training data to create a prediction model. The present study employed the Tree-based Pipeline Optimization Tool (TPOT), a Python-based automated machine learning system, to choose the best learning model for each dataset and optimize its parameters. Python's sci-kit-learn module was used to train and assess the pipelines under fivefold cross-validation. Using the best pipeline determined by TPOT, approximately 4000 meaningful compounds were retrieved from the PubChem database and evaluated for each target. This set of compounds was used as a validation set to determine whether the most effective machine-learning pipeline could successfully retrieve active ingredients from PubChem's list of FDA-approved drugs. PubChem-approved compounds were screened and then ordered according to the class probabilities returned by the relevant pipeline. The top-ranked compounds can be chosen as potential research subjects for more thorough investigations, such as molecular docking and molecular dynamics simulation [139-140].

2.5 ADME/T Prediction and Validation

The ADMET Structure-Activity Relationship database (AdmetSAR) holds immense value within the realm of pharmaceutical research and drug development. It serves as an indispensable tool for the thorough evaluation of absorption, distribution, metabolism, excretion, and toxicity (ADMET) characteristics of compounds. AdmetSAR leverages a diverse array of prediction models, encompassing 22 qualitative classifications and 5 quantitative regression models, to accomplish its intended purpose. These models facilitate the prediction of crucial aspects including AMES toxicity, carcinogenic potential, Caco-2 permeability, propensity for P-glycoprotein interaction as a substrate or inhibitor, hERG inhibition, identification of cytochrome P450 (CYP) metabolism sites, and the ability to penetrate the formidable blood-brain barrier. These variables hold significant implications for

the safety and efficacy of a compound, thereby offering requisites that can drive the intricate process of drug design [141-145].

2.6 Drug likeliness studies

To assess the Drug likeliness, studies were carried out using SwissADME. The versatile tool is tailored to provide a comprehensive evaluation of the suitability of the compounds for oral bioavailability, thereby augmenting the drug discovery process [146-147]. The software evaluated six primary physicochemical properties: lipophilicity often denoted as LIPO or XLOGP3, molecular weight (MW), solubility (INSOLU), polarity (POLAR), saturation (INSATU), and flexibility (FLEX). The depiction of an ideal physiochemical space was demonstrated through a distinctive pink zone, wherein compounds exhibit crucial attributes necessary for oral bioavailability. Additionally, SwissADME offered an extensive overview of the molecular and physiochemical characteristics, encompassing molecular formula, molecular weight, heavy atom count, specific atom count, molar refractivity, and polar surface area. Furthermore, SwissADME utilized five distinct prediction models to forecast lipophilicity, a critical parameter defined as the logarithm of the partition coefficient between n-octanol and water ($\log P_{O/W}$). The multimodal approach significantly improved the credibility of pivotal analysis [148-149].

2.7 Drug-Drug Comparative Study

A comparative graphical analysis was performed using R programming to juxtapose the well-established compounds and machine learning-generated compounds. This analysis concentrated on key pharmacological properties including human intestinal absorption (HIA), blood-brain barrier (BBB) penetration, neurotoxicity, and the LD50. The comparison between top-established and machine learning-derived compounds was executed in the GGLOT package [148-149].

2.8 MD Simulation

The GROMACS 2022 version (<http://www.gromacs.org>). was utilized to conduct comprehensive simulation experiments on molecular complexes of pre-established and ML-screened hits, aiming to investigate the dynamic properties of binding stability, protein compactness, and interactions, version (<http://www.gromacs.org>). The ligand topologies were determined by employing the GROMOS force field through the PRODRG tool

(<http://prodrgr1.dyndns.org/>). To achieve adequate solvation, a cubicle periodic box was used to enclose all complexes. The box was filled with an SPC (single point charge) water model and a width of 1.0 was maintained between the walls of protein complex components. Additionally, counterions with an ionic strength of 0.1 M, specifically sodium ions (Na⁺) and chloride ions (Cl⁻), were incorporated into the solvated systems to attain neutralization. Energy minimization was then performed using the steepest descent energy method for 100 pf to mitigate potential steric clashes within the systems. Following this phase, two equilibrium processes, the NVT (constant volume and temperature) and NPT (constant pressure and temperature) ensembles, were performed at 300 K utilizing a Berendsen thermostat and Leap-Frog algorithm for 1000 ps. To gain insights into the complex behavior over an extended duration, a molecular dynamics (MD) production run was executed. The simulation was conducted for 100 ns, employing an integration time of 2 femtoseconds. Subsequently, the GROMACS Toolkit was used to analyze the trajectories for statistical analyses, including the root mean square deviation (RMSD), radius of gyration (Rg), and number of hydrogen bonds. The results were then visualized using Origin Software [150-154].

2.9 MM-PBSA – free binding energy calculation

The `g_mmpbsa` tool (Molecular Mechanics Poisson–Boltzmann Surface Area) was used to predict the Free energy of simulated protein-ligand complexes. The molecular mechanics potential energy, comprising the free energy of solvation of individual complexes, was calculated by the following equations.

$$\Delta G_{\text{binding}} = G_{\text{complex}} - (G_{\text{protein}} + G_{\text{ligand}})$$

where G_{complex} denotes the total free energy of the protein-ligand complexes and $G_{\text{protein}} + G_{\text{ligand}}$ represents the total free energies of the separated form of protein and ligand in the solvent, respectively.

$$\Delta G_{\text{binding}} = E_{\text{gas}} + G_{\text{sol}} - T\Delta S$$

$$E_{\text{gas}} = E_{\text{int}} + E_{\text{vdw}} + E_{\text{ele}}$$

$$G_{\text{sol}} = G_{\text{pol}} + G_{\text{non-pol}}$$

$$G_{\text{nonpol}} = \gamma \text{SASA} + \beta$$

In this particular scenario, the total $\Delta G_{\text{binding}}$ energy for each complex, ligand, or receptor can be dissected into three distinct components: gas phase energy (E_{gas}), solvation energy (G_{sol}), and an entropy term ($T\Delta S$). The E_{gas} is the summation of the internal energy of bonds (E_{int}) and nonbonds (E_{vdw} & E_{ele}), while the G_{sol} is further divided into polar (G_{pol}) and nonpolar ($G_{\text{non-pol}}$) energies. The calculation of G_{pol} and $G_{\text{non-pol}}$ is accomplished through the application of the Born (GB) equation and solvent-accessible surface area. The entropy term (referred to as $T\Delta S$) encompasses various translational, rotational, and vibrational terms associated with solute molecules [155-162].

3. Results and Discussion

3.1 Protein and Ligand Preparation

The protein (PDB ID: 6LZG) employed in the current study was a chemically tailor-made structure of a novel coronavirus spike receptor-binding domain complexed along its receptor ACE2. The structure of SARS-CoV-2/hACE2 exhibits a resolution of 2.5 Å and is located in an asymmetrical unit in a way that the SARS-CoV-2-CTD uses its external sub-domain to recognize with sub-domain I of hACE2 N-terminal. The structure of the protein complex needs to be altered to perform docking. The SARS-CoV-2 at the N-terminal domain of ACE2 and the existing drug were precisely removed, leaving a vacant site in the ACE2 protein. ACE2, encoded by Chr Xp22 is a type I transmembrane carboxy peptidase protein. The 3D structure of this protein is visualized through Accelrys Discovery Studio Visualizer 3.0 [163-168] and comprises 791 groups with an aggregate of 6411 atoms, 6653 bonds, and 7 bridges. The crystal structure constitutes 537 bonds consisting of hydrogen bonds and zero disulfide bonds. The protein bears 41 helices, 15 strands, zero loops, and turns. Further, for ligand preparation, all structures were retrieved from the PubChem database and were saved in a single SDF file, employing the LigPrep module of the Schrödinger Suite 2020.

3.2 Molecular Docking of Established Compounds

By utilizing the piecewise linear potential (PLP) and MolDock scoring functions, the Molecular Docking Programme Molegro Virtual Docker (MVD) was employed in association with an appropriate docking platform. NCBI PubChem Database was employed for retrieving all the 201 pre-established inhibitors corresponding to our target protein. The optimized structures were docked into the binding cleft of hACE2. The best-established compounds; berginapant (PubChem CID: 49836020) and elbasvir (PubChem CID: 7166125), were procured which exhibited re-ranks of -151.359 KJ/mol and -138.115 KJ/mol, respectively. A mitigated score dictates strong energy interaction between the ligand and the protein, therefore indicating an augmented affinity score for the hACE2 protein. Birinapant and elbasvir comprise molecular weights of 806.9 g/mol and 882 g/mol, logP values of 3.3 and 6.7, hydrogen bond counts of 8 and 4, and hydrogen bond acceptor counts of 10 and 9, respectively. The structure is illustrated in Table 1.

Table 1. Molecular Docking Results of established compounds.

Name	Ligand	MolDock Score	Re-rank Score	H-Bond	MW
[00] 49836020	49836020	-255.307	-151.359	-5.58032	806.941
[03] 71661251	71661251	-216.161	-138.115	-1.96727	882.017
[02] 71661251	71661251	-211.741	-133.333	-0.38188	882.017
[02] 5745214	5745214	-194.888	-130.409	-10.5118	744.896
[00] 44264212	44264212	-179.724	-129.355	-2.05342	895.082
[00] 5745214	5745214	-202.442	-128.248	-11.123	744.896
[03] 37723	37723	-275.438	-128.205	-16.8444	1153
[01] 37723	37723	-291.337	-127.684	-24.4228	1153
[03] 49836020	49836020	-237.241	-123.28	-2.5	806.941
[00] 9854012	9854012	-176.543	-123.178	-10.7525	776.359

3.3 Molecular Docking of Machine Learning-Screened Compounds

The Tree-Based Pipeline Optimization Tool (TPOT), a Python machine learning system, was employed to select the best learning model for each dataset by optimizing its parameters. A fivefold cross-validation method was also employed from Python's sci-kit-learn module, which aids in training and assessing the pipelines. The best pipeline of TPOT analysis aided in the retrieval of ~4000 compounds from the PubChem database, and their interactions were

also evaluated. Among the active procured compounds, the top 10 have been listed in Table 2. The best-established machine learning compounds, PubChem ID: 23658468 and 117637105, had re-rank scores of -168.435 KJ/mol and -165.99 KJ/mol, respectively, clearly demonstrating more negative energies than those of the pre-established compounds. The compounds exhibit molecular weights of 895.046 g/mol and 820.967 g/mol, hydrogen bond values of -15.1118 and -14.6157, and Mol Dock scores of -290.047 and -270.528, respectively.

Table 2. Molecular Docking results of Machine Learning screened compounds.

Name	Ligand	MolDock Score	Re-rank Score	H-Bond	MW
[00] 23658468	23658468	-290.047	-168.435	-15.1118	895.046
[00] 117637105	1.18E+08	-270.528	-165.99	-14.6157	820.967
[00] 139403642	1.39E+08	-254.574	-157.991	-4.11129	808.932
[00] 139403597	1.39E+08	-261.89	-157.979	-8.19184	806.941
[01] 139403597	1.39E+08	-258.658	-156.929	-0.17316	806.941
[00] 50902610	50902610	-280.444	-154.749	-3.89002	887.069
[01] 50902609	50902609	-240.155	-152.595	-3.01409	806.941
[00] 54765240	54765240	-254.187	-150.857	-12.9112	863.047
[00] 102358917	1.02E+08	-236.623	-149.84	-4.31412	806.941
[00] 159308911	1.59E+08	-230.62	-149.23	-8.96039	804.965
[01] 154814211	1.55E+08	-247.125	-148.284	-5.76766	806.941
[00] 125130383	1.25E+08	-237.809	-148.096	-3.4751	806.941
[01] 125130384	1.25E+08	-261.086	-147.806	-8.03618	806.941

3.4 Pharmacophore studies of the most effective compounds obtained from machine learning

3.4.1 H-Bond Interactions

A pharmacophore is a conceptual explanation that elucidates how structurally different ligands bind to a prevalent receptor which is probably useful for identifying and validating a novel drug. The strong interaction between the chemical and the target protein is confirmed by various poses of the molecule. The pharmacophore study aims to determine the target and the compound in its cavity, the machine learning compound (**PubChem ID: 23658468**) binds efficiently in the first cavity, exhibiting a volume of 4820.99 Å³ and a higher binding affinity.

Figure 4 depicts the H-bond interaction within the machine learning compound and the target protein (ACE2) depicted by green dotted lines. In total, 11 hydrogen bond interactions were perceived: 2 H bonds with SER47, ASN349, ASP51 and GLU22 each and 1 H-bond with HIS401, ASP350, and AA348. Thus, in comparison with any other repurposed drug, the machine-learning compound exhibits high affinity and demonstrates a notably increased hydrogen bond interaction with the ACE2 protein.

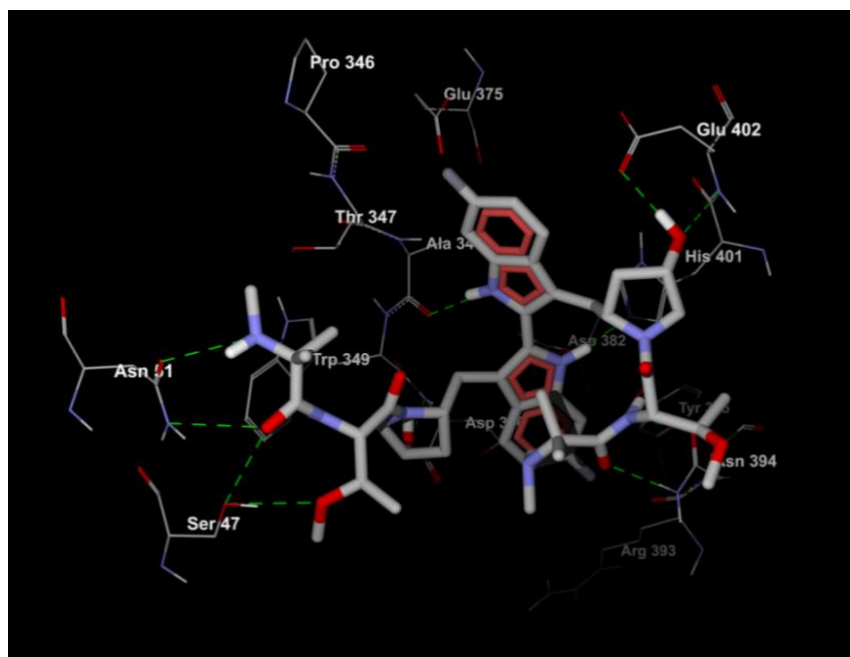


Figure 4. The most effective compound (PubChem ID: 23658468) obtained from machine learning showed H-bond interactions with ACE2.

3.4.2 Van Der Waals Interaction

Numerous 2D interactions occurring between the ligand and the binding pocket are shown in Figure 5. The van der Waals interaction is depicted in green, namely, PHE A:40, SER A:43, SER A:44, VAL A:343, HIS A:345, HIS A:378, GLY A:395, GLU A:398, THRA:347, and LEU A:359. The electrostatic interactions are depicted as pink in the complex TRP A:349, SER A:47, LEU A:351, TYR A:385, ASP A:350, ASN A:394, GLU A:402, PRO A:346, ARG A:514, GLU A:375, and ASN A:51.

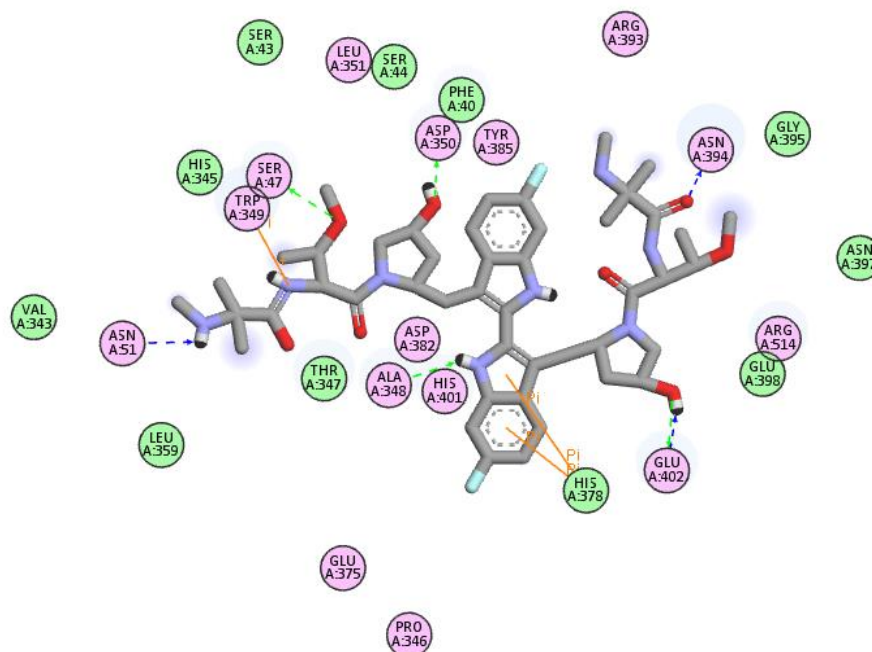


Figure 5. The most effective compound (PubChem ID: 23658468) obtained from machine learning showed 2D interactions with ACE2.

3.4.3 Receptor–Ligand Interaction

This is the most effective form of target ligand interaction. The target was imaged in a blue wireframe with the machine learning compound (PubChem ID: **23658468**). The H-bond interactions are indicated by green dotted lines between SER47 and GLU 402. The ligand interacts with the target at multiple discrete positions, TRP 48, ASN 51, TRP 349, ARG 357, TYR 385, GLY 395, and PHE400, as depicted in Figure 6.

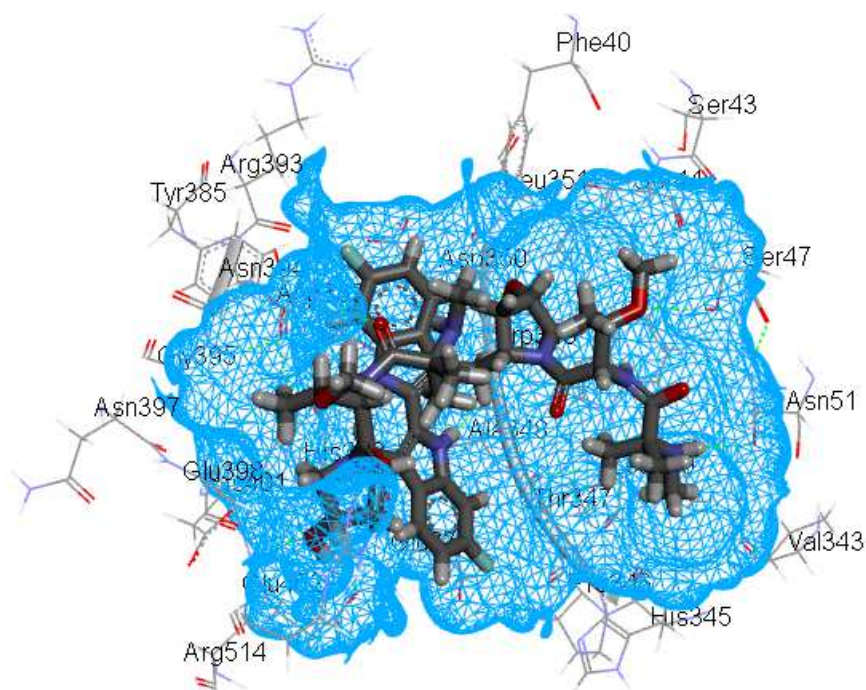


Figure 6. The most effective compound (PubChem ID: 23658468) obtained from machine learning showed receptor-ligand interactions with ACE2.

3.4.4 Aromatic interactions

The aromatic interaction of the machine learning compound (PubChem ID: **23658468**) is illustrated in Figure 7. The blue-colored edge surface indicates the presence of a higher number of aromatic residues, while the brown color represents fewer aromatic residues, characterizing the face surface. While aromatic interactions may not be as prevalent as other types of interactions, they nevertheless play a crucial role in stabilizing ligand complexes, facilitating biological recognition, and biomolecular structural organization.

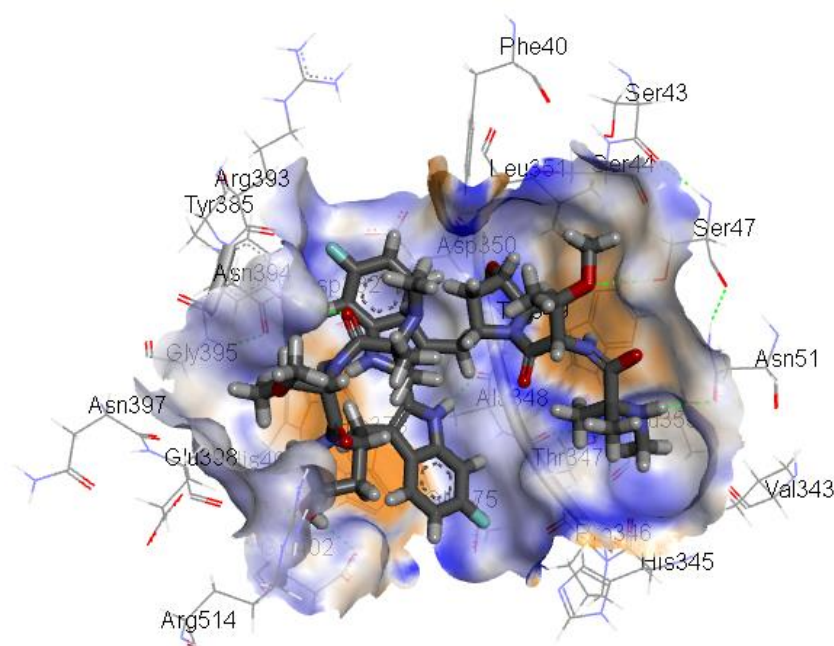


Figure 7. The most effective compound (PubChem ID: 23658468) obtained from machine learning showed aromatic interactions with ACE2.

3.5 Comparative study of ADME/T prediction between the best-established compound and best ML compound

3.5.1 Prediction by AdmetSAR

The ADME/T values for the established compound (PubChem ID: 49836020) and the ML compound (PubChem ID: 23658468) are presented in Table 3. The absorption parameters of the ML compound (PubChem ID: 23658468) are analogous to the established compound. The BBB value of the established compound was 0.9059, while that of the ML compound was 0.9282, indicating a negative blood-brain barrier permeability. The intestinal absorption values are 0.9883 and 0.9662, respectively, facilitating their absorption into the bloodstream. The P-glycoprotein substrate value of the established compound was lower than the Machine learning compound (PubChem ID: 23658468) suggesting that according to the AdmetSAR criteria, the software predicts with approximately 87% probability that the ML compound (PubChem ID: 23658468), serves as a substrate for the P-glycoprotein transporter. The ML compound (PubChem CID: 23658468) has low CYP inhibitor promiscuity, exhibiting a value of 0.5238, which infers that the compound lacks significant CYP450 inhibitory promiscuity. However, its CYP450 2D6 substrate value is significantly lower than that of the established

compound. Both compounds are noncarcinogens and show no AMES toxicity, implying their non-mutagenic nature.

Table 3. ADME/T classification prediction using AdmetSAR

Model	Result	Probability for Established Compound PubChem ID: 49836020	Probability for ML Compound PubChem ID: 23658468
Absorption			
Blood–Brain Barrier	BBB-	0.9059	0.9282
Human Intestinal Absorption	HIA+	0.9883	0.9662
Caco-2 Permeability	Caco2-	0.7234	0.7201
P-glycoprotein Substrate	Substrate	0.8238	0.8784
P-glycoprotein Inhibitor	Non-inhibitor	0.6475	0.5885
	Non-inhibitor	0.7489	0.5197
Renal Organic Cation Transporter	Non-inhibitor	0.8794	0.8622
Distribution			
Subcellular localization	Mitochondria	0.7236	0.6906
Metabolism			
CYP450 2C9 Substrate	Non-substrate	0.8547	0.8524
CYP450 2D6 Substrate	Non-substrate	0.7559	0.709
CYP450 3A4 Substrate	Substrate	0.6661	0.7196
CYP450 1A2 Inhibitor	Non-inhibitor	0.8262	0.7249
CYP450 2C9 Inhibitor	Non-inhibitor	0.6282	0.5235
CYP450 2D6 Inhibitor	Non-inhibitor	0.8664	0.7851
CYP450 2C19 Inhibitor	Non-inhibitor	0.6208	0.576
CYP450 3A4 Inhibitor	Non-inhibitor	0.6909	0.5753
CYP Inhibitory Promiscuity	Low CYP Inhibitory Promiscuity	0.6918	0.5238
Excretion			
Toxicity			
Human Ether-a-go-go-Related Gene Inhibition	Weak inhibitor	0.9788	0.9922
	Inhibitor	0.8402	0.7991
AMES Toxicity	Non-AMES toxic	0.8012	0.7855
Carcinogens	Non-carcinogens	0.7599	0.7561
Fish Toxicity	High FHMT	0.9898	0.9948
Tetrahymena Pyriformis Toxicity	High TPT	0.9851	0.9899
Honeybee Toxicity	Low HBT	0.8399	0.7935
Biodegradation	Not readily biodegradable	1	1
Acute Oral Toxicity	III	0.6016	0.6184
Carcinogenicity (Three-class)	Non-required	0.5791	0.5454

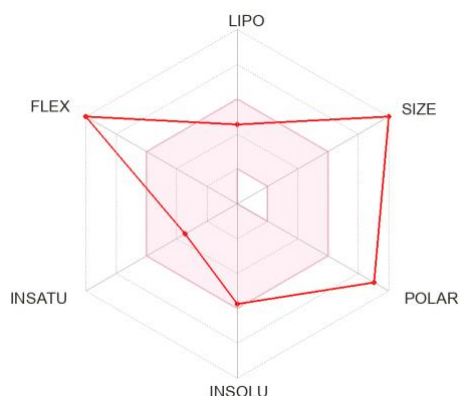
The ML compound (PubChem ID: 23658468) displayed a higher regression value for absorption and toxicity (aqueous solubility and rat acute toxicity) than the established compound, as shown in Table 4.

Table 4. ADME/T predictive regression profile of Pre-Established and ML Screened Compound

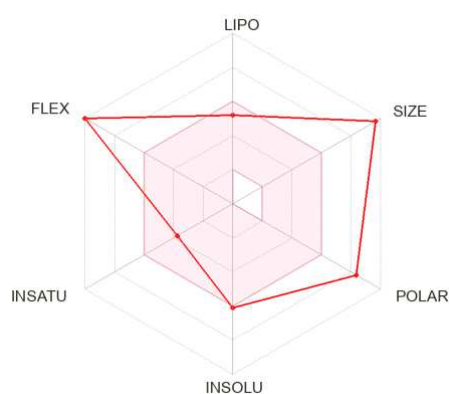
Model	Value	Probability for Established Compound PubChem ID: 49836020	Probability for ML Compound PubChem ID: 23658468	Units
Absorption				
Aqueous solubility		-3.5418	-3.4754	LogS
Caco-2 Permeability		0.3421	0.4847	LogPapp, cm/s
Toxicity				Unit
Rat Acute Toxicity		2.7235	2.6888	LD50, mol/kg
Fish Toxicity		1.3322	1.1937	pLC50, mg/L
Tetrahymena Pyriformis Toxicity		0.4736	0.5893	pIGC50, ug/L

3.5.2 Drug likeness using Swiss-ADME

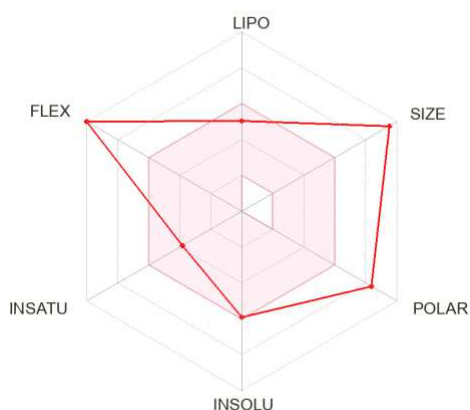
The assessment of compound activity within the oral bioavailability radar for the machine learning compound and established compounds incurred from the SwissADME web server is presented in Figure 8. The pink area highlights the ideal range of characteristics for XlogP3 should include lipophilicity (LIPO), which should lie within the range of -0.7 to +0.5, a molecular weight (SIZE) spanning from 150 g/mol - 500g/mol, a topological polar surface area (TPSA: POLAR) ranging from 20 Å² to 130 Å², solubility (INSOLU) in logS should be notably lower than 0, a saturation fraction of carbons with sp³ hybridization (INSATU) lower than 1 and a pliability of rotatable bonds (FLEX) not beyond 9. The machine learning compound (PubChem ID: 23658468) demonstrates preferable lipophilicity and saturation fractions of carbons with sp³ hybridization when compared to the established compound (PubChem ID: 49836020).



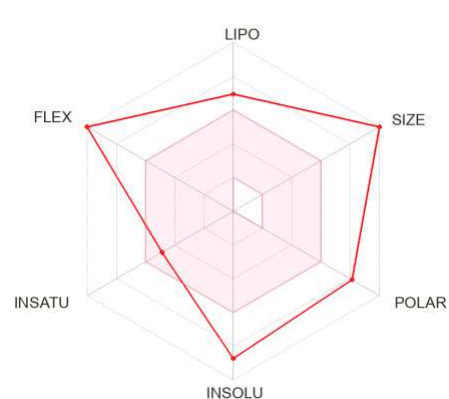
PubChem ID - 23658468



PubChem ID - 117637105



PubChem ID - 49836020



PubChem ID - 71661251

Figure 8. Bioavailability radar related to the physicochemical properties of the best compounds from the established and machine learning screening results. The pink area represents the optimal range for each of these properties LIPO, SIZE, POLAR, INSOLU, INSATU and FLEX.

3.5.3 Boiled Egg Plot

In the realm of drug development and discovery, the BOILED-Egg plot model offers a rapid, resilient, and statistically sound approach for determining the passive gastrointestinal absorption and brain penetration of small molecules. To generate the BOILED-Egg plot, both parameters were validated as eclipses on a cartesian plane, in conjunction with other notable parameters like MW, TPSA, MLOGP, GI, and BBB. Consequently, on the cartesian plane, if the compounds are present within the yolk region, symbolized by the yellow ellipse, there is a higher probability of traversing the blood-brain barrier (BBB). Conversely, if compounds are found in white regions, the likelihood of gastrointestinal absorption is increased. Moreover, if the compounds are found to be in gray regions beneath the ‘egg’ regions, they are considered

nonabsorptive and demonstrate mitigated brain penetration. Additionally, studies have indoctrinated that P-glycoproteins are crucial in the active efflux of molecules.

The egg plot in Figure 9, showcases the positioning for the given set (PubChem ID; 49836020 and 71661251) and ML compounds (PubChem ID; 23658468 and 117637105). The plot demonstrates that all four inhibitors reside within the white space, indicating a higher likelihood of GI absorption. Moreover, the analysis revealed that these molecules correspond with the P-glycoprotein substrate cadre (represented by blue dots). Furthermore, the top-hit compounds were evaluated based on their physiochemical and pharmacological properties, including the Water Partition Coefficient (WlogP) where values <5 indicate lower toxicity, oral administration, and specific binding levels. The topological polar surface area (TPSA), estimated to be less than 100 Å², further supports the probability of complete absorption. Therefore, it can be concluded that selected compounds possess therapeutic potential.

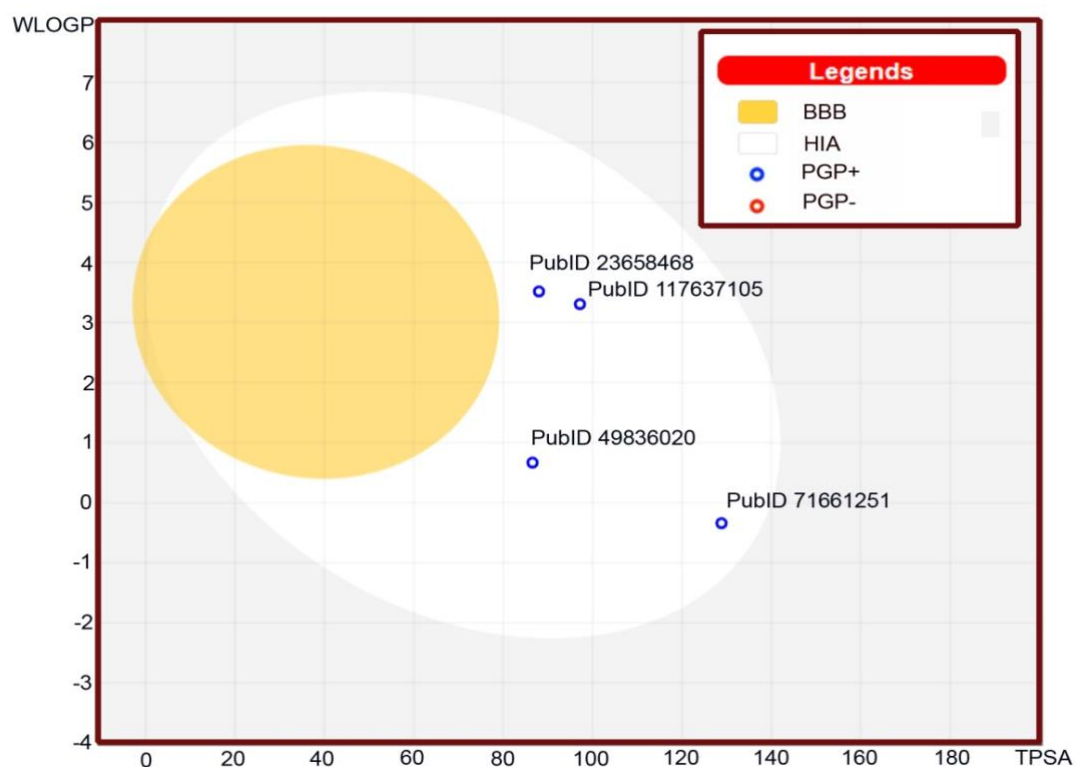


Figure 9. Predictive model brain or intestinal estimated permeation method (BOILED-Egg plot).

The AdmetSAR database renders a free and open-access web resource that provides an exclusive biological and chemical portfolio that is accessible via a query interface. Myriad ADMET factors such as excretion, metabolism, digestion, absorption, and toxicity level play

a paramount role in the development and discovery of novel drugs. The database is divided into 5 quantitative regression models, and 22 qualitative classifications, which mainly aim to allocate a comprehensive denouement with colossal precision based on estimation. Therefore, the database was employed to evaluate the properties of the top-hit established (PubChem ID;49836020, and 71661251) and ML (PubChem ID; 23658468, and 117637105) driven compounds. Comparative analysis was further governed by employing R-programming by generating HIA, BBB, Ames's toxicity, and LD50 values graph, as illustrated in Figure 10.

It is inferred that the best-hit ML compound (CID 23658468) displayed an augmented BBB and HIA index as compared to the best-hit established compound (CID 49836020). Consequently, the second-best ML compound (CID 117637105) exhibited mitigated BBB and HIA values as compared to the second-best established compound (CID 71661251). Higher HIA values correspond to the high absorption of the drug in the intestine after oral administration, and conversely, lower HIA values imply poorer absorption in the human intestines. Similarly, higher BBB permeability allows the compound or drug a better ability to cross the blood-brain barrier and enter the brain. It is highly advantageous when the therapeutic agent reaches the brain and exerts its effects, while lower BBB permeability implies that the compound has a limited ability to cross the blood-brain barrier. Therefore, the top-hit ML compound (CID 117637105) provides a better suite in the therapeutic cadre.

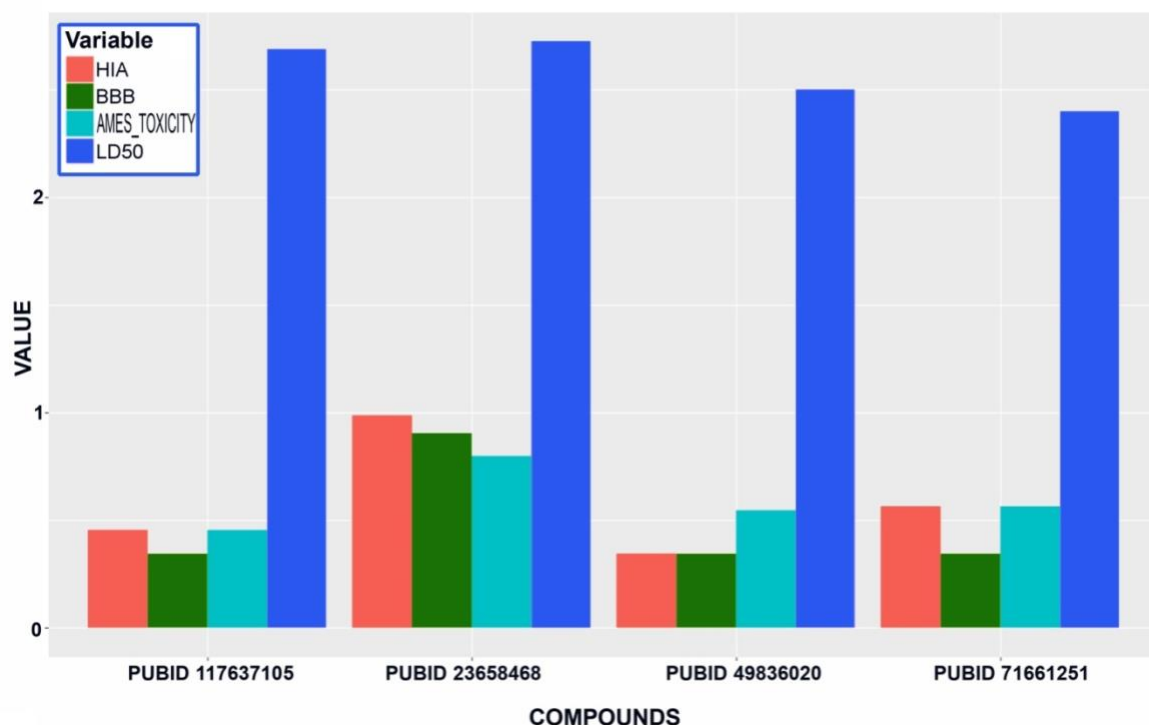


Figure 10. Comparative ADMET studies of BBB, HIA, AMES toxicity and LD50 of the established compound against machine learning-screened compounds

3.6 Molecular Dynamic Simulation

MD simulation was employed to evaluate the stability of the hit compounds binding to the target protein at the molecular level. The simulation ran for 100 nanoseconds, revealing effective dynamics and interplay between the protein and ligand following classical notion principles.

3.6.1 Protein Stability (RMSD)

The RMSD serves as a crucial aspect in simulation studies, offering insights into the monitoring of structural and conformational changes among a set of atoms in comparison to a reference structure. The minimum deviations in the RMSD are associated with higher stability and robust binding. The relationship is expressed as follows:

$$RMSD_x = \sqrt{\frac{1}{N} \sum_{i=1}^N \langle (r'_i(t_x) - r_i(t_{ref}))^2 \rangle}$$

where N indicates the number of atoms in the atom selection, t_{ref} represents the reference time where $t=0$, and r' denotes the position of the selected atoms in the x frame after alignment with the reference frame. The X frame is mentioned at the t^x time interval.

The stability of the protein-ligand complexes was evaluated by assessing the root-mean-square deviation (RMSD) of their backbones throughout a 100 ns simulation using the 'gmx rmsd' tool. The overall structural behavior appears to be comparatively rigid, and the complexes achieved a global minimum within an appropriate simulation timeframe. The permissible range of RMSD values was found to be below 0.34 nm. As illustrated in Figure 11, the RMSD backbone graph indicates that both compounds displayed minimal fluctuations in their dynamics, with average RMSD values of 0.24 nm for the ML compound (PubChem CID – 23658468): and 0.29 nm for the pre-established ligand (PubChem CID - 49836020). The results indicate that ML-based screened compound binds to protein effectively, resulting in the generation of a stable complex. The assessment of RMSD values offers numerical proof that reinforces the stability effectiveness, and success of the ML-based screening approach in pinpointing a promising hit with a favorable binding profile. This leads to the reliability of ML-based compounds for further exploration in drug development or related fields.

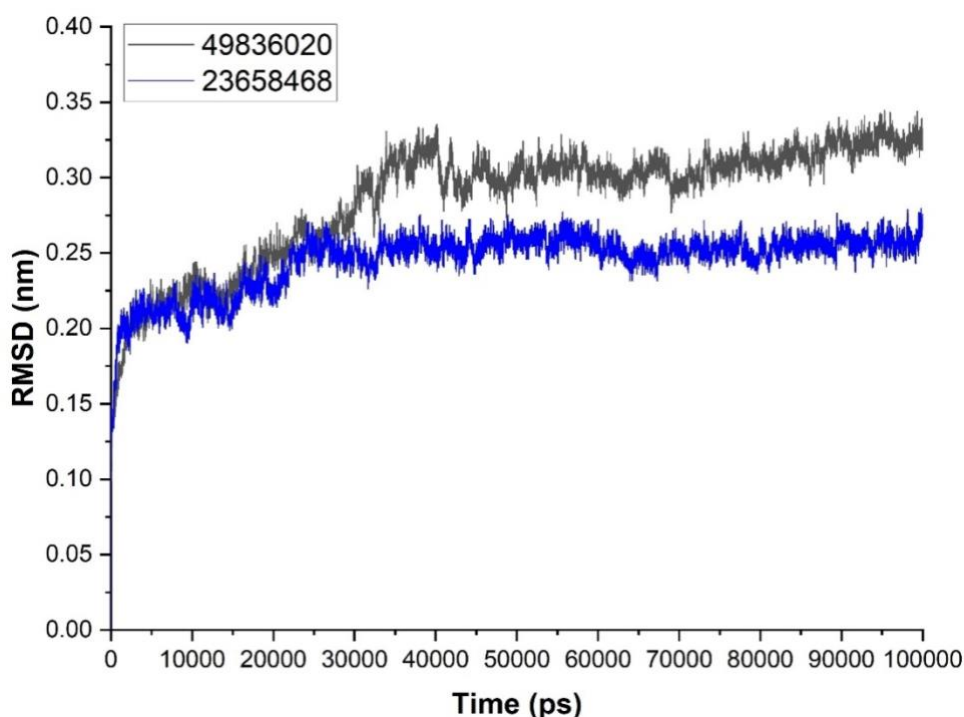


Figure 11. Graphical representation of RMSD values of the protein backbone from the simulated trajectories of both pre-established and ML strains screened throughout a 100 ns all-atom MD simulation.

3.6.2 Protein Compactness (Rg)

The radius of gyration (Rg) offers extensive understanding regarding the stability of folded or unfolded proteins around a common mass center. A lower radius of gyration signifies that the

atoms within the structure are positioned closer to their center of mass, indicating a higher level of compactness and stability. Conversely, a higher radius shows the distant location of atoms from the center of mass, thus reducing compactness and stability. The equations used for the computation of Rg are as follows:

$$r_{\text{RG}}^2 = \frac{\sum_{i=1}^N m_i (\mathbf{r}_i - \mathbf{r}_{\text{CM}})^2}{\sum_{i=1}^N m_i}$$

Where \mathbf{r}_i is the position of the atom at index i ; m_i is the mass of the atom at index i ; \mathbf{r}_{CM} is the center of mass; N is the number of atoms being counted; and r_{RG}^2 is the square of the radius of gyration. In the above equation, the radius of gyration is squared, which means it is the average distance of a system's atoms from their center of mass.

Gyration analysis was done in GROMACS using the *gmx gyrate* command to investigate the global stability of protein compactness during simulation. The obtained results, as depicted in Figure 12 demonstrate that the Rg value falls within the acceptable range of 2.04 and 2.14 nm indicating the consistent conformational changes in the structure throughout the simulation. Notably, the ML-screened ligands (PubChem CID-23658468) exhibited a reliable global stability, with an average Rg value of 2.08 nm, in comparison with the pre-established ligand (PubChem CID-49836020), possess Rg value of 2.10 nm. These findings suggest that the protein remained stably folded with the hits. Consequently, the Rg value is strongly correlated with the RMSD and is considered to be a more reliable indicator of stability in the presence of the screened hits.

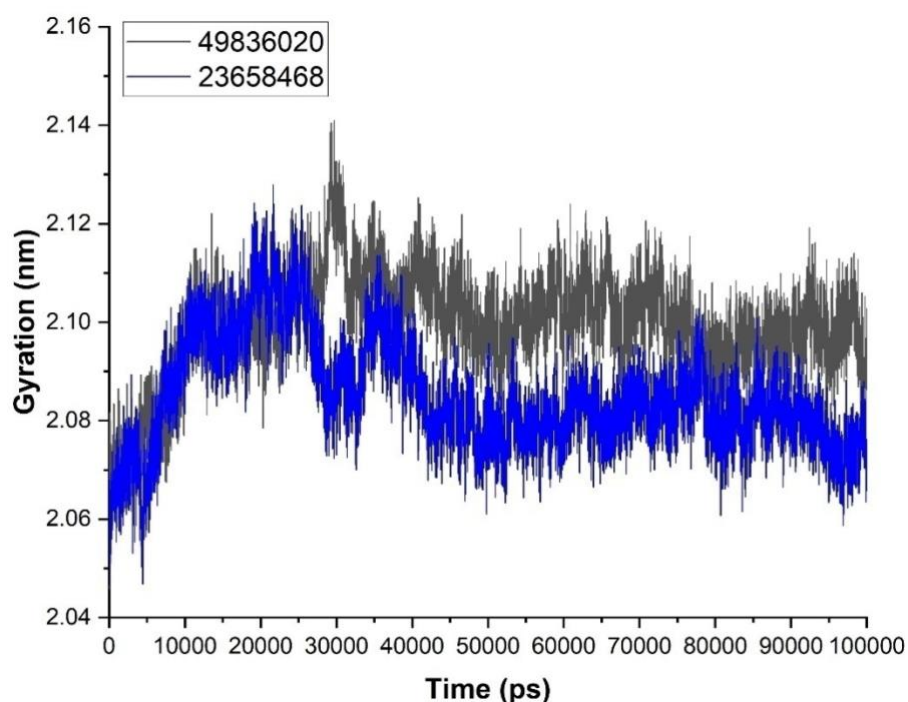


Figure 12. Trajectory of the radius of gyration (R_g) analysis to determine the global stability of ligands with targeted proteins.

3.6.3 H-Bond Interactions

Hydrogen bonds are of utmost importance in protein structural integrity and play a crucial role in protein-ligand interactions. These bonds serve as vital factors in the formation of protein-ligand complexes, offering specificity, stability, and affinity. Throughout the simulation period, the number of hydrogen bond donors and acceptors fluctuated between 0 and 15 (as shown in Figure 13). Notably, the ML-based hit (PubChem ID: 23658468) had a maximum of 15 hydrogen bond interactions, whereas the pre-established screened compound (PubChem ID: 49836020) exhibited 10 such interactions. This finding underscores the crucial role of hydrogen bonding in maintaining structural stability, optimizing fit and improving binding affinity.

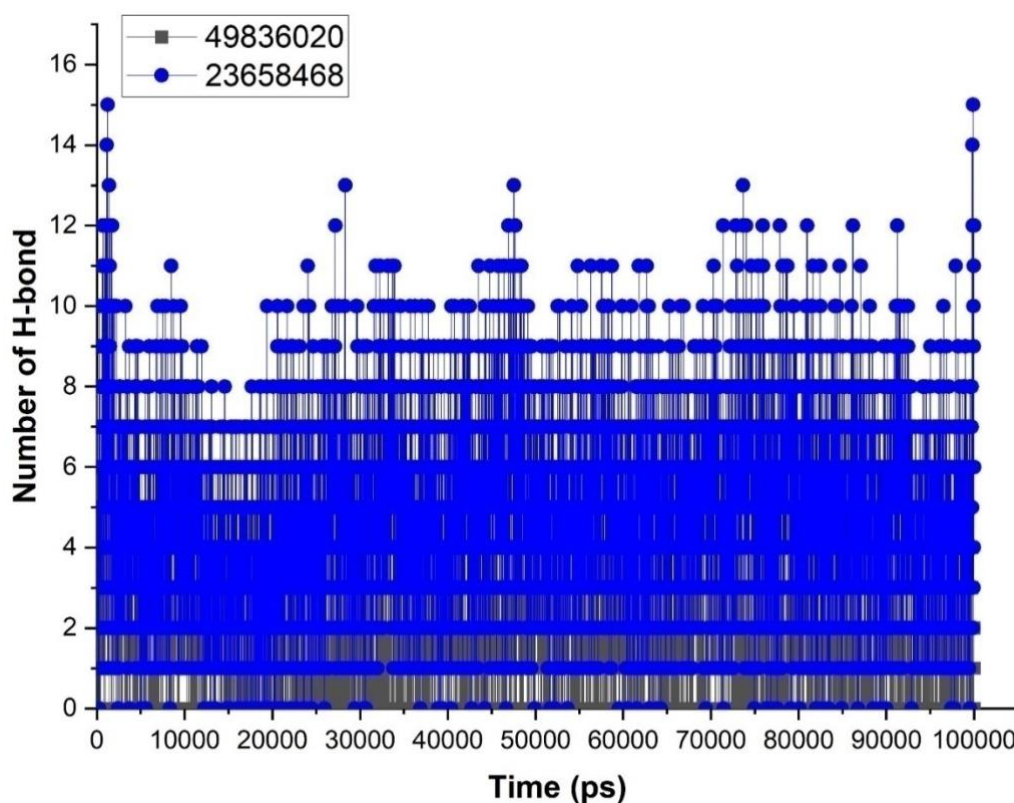


Figure 13. Trajectory analysis of the total number of hydrogen bonds between ligands (ML screened & Pre-established) and proteins.

Furthermore, the interaction signatures of the post-MD structure validated the robust binding interaction between the ML compound and the binding pocket residues corresponding to pre-established interaction, as depicted in Figure 14.

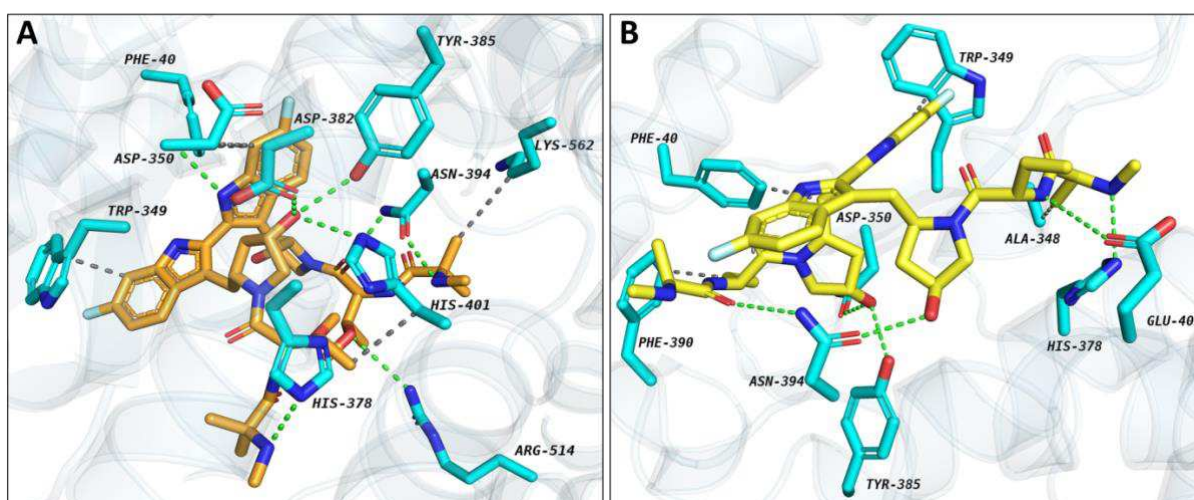


Figure 14. Interactive profiles of ML-screened compound (A) and Pre-established compound (B) after MD Simulation. [Green-Color Dotted Lines – H-Bonds, Gray-Color Dotted Lines – Hydrophobic; Cyan-Color – Residues]

3.7 Post MD Binding free energy calculation (MM/PBSA)

The MMPBSA results offer valuable insights into the thermodynamics of protein-ligand complexes, adding to the information acquired from MD simulations. A higher negative binding free energy highlights a stronger and more favorable interaction between the ligand and the protein. The ML ligand (PubChem ID: 23658468) exhibits a binding energy of approximately -132.90 Kcal/Mol significantly lower than the binding energy of the pre-established ligand (PubChem ID: 49836020) approximately -101.07 Kcal/Mol as shown in Figure 15. These results displayed that the ML compound is more energetically favorable in the complex. The aforementioned findings complement the MD simulation results, which showcased that ML-based screened hit displayed higher stability and retained a well-defined conformation through the simulation.

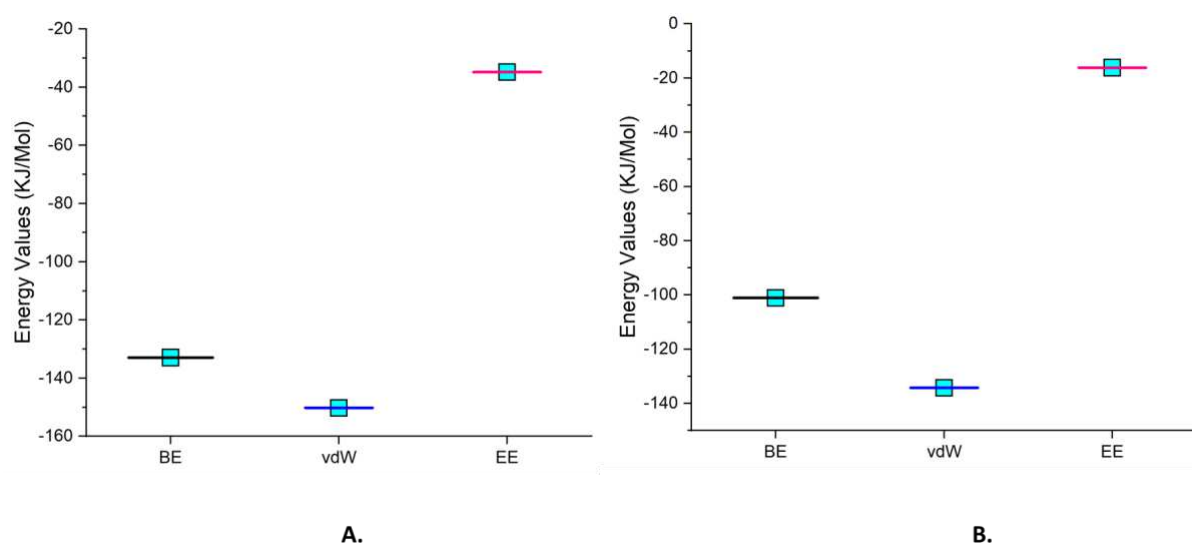


Figure 15. Predicted binding free energy using MM/PBSA: A. ML-screened ligand (PubChem CID – 23658468) and B. pre-established ligand, PubChem ID: 49836020 [BE – Binding Free Energy, vdW – Vander Walls Energy, EE – Electrostatic Energy]

In the end, the integration of MD simulation and MMPBSA results unveiled comprehensive insights into the behavior and stability of the protein-ligand complexes. The concordance of the results and the potential of ML-based screening molecules as a promising candidate for further research in drug development or other relevant applications are supported by the alignment between the two datasets. Furthermore, the ML-based screening method effectively identifies compounds with enhanced binding properties and provides insights for rational drug design endeavors.

Conclusion

SARS-CoV-2, the novel beta-coronavirus responsible for the COVID-19 pandemic, emerged as a lethal disease that infected several people worldwide. Studies have indoctrinated that the ACE2 receptor is a chief drug target for combating COVID-19. ACE2 regulates the signaling pathway of renin-angiotensin system (RAS), and converts angiotensin I and II into angiotensin 1-9 and angiotensin 1-7, respectively. The current study highlights the role of ACE2 as a potential drug target for identifying novel treatment strategies for impeding COVID-19. For this purpose, 201 established inhibitors belonging to two categories, namely, repurposed antiviral compounds, such as Birinapant, Remdesivir, Lopinavir, and Ivermectin, and ACE2 inhibitors, such as Quninapril, Moexipril, Benazepril, and Elanapril, were taken as test datasets. Further, Machine-learning algorithms were employed to govern the novel machine-learning drug compounds. The novel compound (PubChem ID: 23658468) was procured as a potential drug against the ACE2 receptor. This novel compound demonstrated an augmented affinity for the ACE2 receptor, which was confirmed via molecular docking and machine learning algorithms. The structural comparative stability analysis was conducted between the novel machine learning compounds and pre-established compounds via MD simulation approach. The pharmacophore designing of the selected compound (PubChem ID: 23658468) demonstrated effective interaction with the receptor protein, thus corroborating the results of the MD simulation. Moreover, analysis of the differences between the ADMET activity of the top two best pre-established and ML compounds revealed that the novel compound (PubChem ID: 23658468) exhibited the highest gastrointestinal absorption and BBB passive penetration. Thus, this novel compound is proposed to be a suitable therapeutic agent against the ACE2 receptor to prevent COVID-19. Thus, further therapeutic assessment can be conducted by employing the *In-vitro* studies.

AUTHOR CONTRIBUTIONS

A.N. was involved in conceptualization, investigation, methodology, project administration, supervision, writing, review, and editing. **AB, AK, KS, US,** and **UP** were involved in the molecular docking, molecular dynamics simulation, writing, review and editing. **VN, R.C, AK, TD, AP** and **LP** contributed to inhibitor collection, data curation, formal analysis, validation, and visualization. **FJBMJ** and **S.K.S.** contributed to the investigation, supervision, writing, review and editing.

Acknowledgments:

AN, AB, AK, KS, UP, and SKS thankfully acknowledge the MHRD-RUSA Phase 2.0 under [grant sanctioned wide letter no. F.24-51/ 2014-U, Policy (TN Multi-Gen), Department of Education, Government of India, dated 09.10.2018]; Tamil Nadu State Council for Higher Education (TANSCHHE) under [grant sanctioned wide letter no. AU: S.O. (P&D): TANSCHHE Projects: 117/2021, dated 31.03.2021, File No. RGP/2019-20/ALU/HECP-0048 dated 27.04.2021]; Department of Biotechnology, Ministry of Science and Technology, New Delhi, under Grant/Award [No. BT/PR40154/BTIS/137/ 34/2021, dated 07.03.2022]; DST-PURSE 2nd Phase Program [Order No. SR/PURSE Phase 2/38 (G dated 21.02.2017)& DST-FIST (SR/FST/LSI—667/2016)]; DBT-NNP Project, New Delhi, under Grant/Award [No. BT/PR40156/BTIS/54/2023 dated 06.02.2023] for providing financial assistance, infrastructure facilities in the lab. AB thanks DBT-NNP for the Fellowship.

Funding

We do not receive any fund for this study.

Availability of data and materials

Most of the data generated or analyzed during this study are included in this article. We have included additional data under the supplementary material. Moreover, queries can be directed to the corresponding author for any clarifications about the study if needed.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Wu, F., Zhao, S., Yu, B., Chen, Y.M., Wang, W., Song, Z.G., Hu, Y., Tao, Z.W., Tian, J.H., Pei, Y.Y. and Yuan, M.L., 2020. A new coronavirus associated with human respiratory disease in China. *Nature*, 579(7798), pp.265-269.
2. WHO. Coronavirus Disease (Covid-19) Pandemic. Available online: <https://www.who.int/emergencies/diseases/novelcoronavirus-2019> (accessed on 28 May 2021).
3. Cascella, M., Rajnik, M., Aleem, A., Dulebohn, S.C. and Di Napoli, R., 2022. Features, evaluation, and treatment of coronavirus (COVID-19). *Statpearls* [internet].

4. Bhatia, R., Ganti, S.S., Narang, R.K. and Rawal, R.K., 2020. Strategies and challenges to develop therapeutic candidates against the COVID-19 pandemic. *The Open Virology Journal*, 14(1).
5. Zhong, N.S., Zheng, B.J., Li, Y.M., Poon, L.L.M., Xie, Z.H., Chan, K.H., Li, P.H., Tan, S.Y., Chang, Q., Xie, J.P. and Liu, X.Q., 2003. Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. *The Lancet*, 362(9393), pp.1353-1358.
6. Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W., Si, H.R., Zhu, Y., Li, B., Huang, C.L. and Chen, H.D., 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *nature*, 579(7798), pp.270-273.
7. Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N. and Bi, Y., 2020. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The lancet*, 395(10224), pp.565-574.
8. Jiang, S., Hillyer, C. and Du, L., 2020. Neutralizing antibodies against SARS-CoV-2 and other human coronaviruses. *Trends in immunology*, 41(5), pp.355-359.
9. Jean, S., 2020. S, Lee P-I, Hsueh P-R. Treatment options for COVID-19: the reality and challenges. *J Microbiol Immunol Infect*, 53(3), pp.436-443.
10. McKee, D.L., Sternberg, A., Stange, U., Laufer, S. and Naujokat, C., 2020. Candidate drugs against SARS-CoV-2 and COVID-19. *Pharmacological research*, 157, p.104859.
11. Van Boheemen, S., De Graaf, M., Lauber, C., Bestebroer, T.M., Raj, V.S., Zaki, A.M., Osterhaus, A.D., Haagmans, B.L., Gorbalenya, A.E., Snijder, E.J. and Fouchier, R.A., 2012. Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans. *MBio*, 3(6), pp.00473-12.
12. Teralı, K., Baddal, B. and Gülcan, H.O., 2020. Prioritizing potential ACE2 inhibitors in the COVID-19 pandemic: Insights from a molecular mechanics-assisted structure-based virtual screening experiment. *Journal of Molecular Graphics and Modelling*, 100, p.107697.
13. Poduri, R., Joshi, G. and Jagadeesh, G., 2020. Drugs targeting various stages of the SARS-CoV-2 life cycle: Exploring promising drugs for the treatment of Covid-19. *Cellular signaling*, 74, p.109721.
14. Tai, W., He, L., Zhang, X., Pu, J., Voronin, D., Jiang, S., Zhou, Y., Du, L., 2020. Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus:

implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cellular & molecular immunology*, 17(6), pp.613-620.

15. Xu, X., Chen, P., Wang, J., Feng, J., Zhou, H., Li, X., Zhong, W, Hao, P., 2020. Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Science China Life Sciences*, 63(3), pp.457-460.
16. Shereen, M.A., Khan, S., Kazmi, A., Bashir, N. and Siddique, R., 2020. COVID-19 infection: Emergence, transmission, and characteristics of human coronaviruses. *Journal of advanced research*, 24, pp.91-98.
17. Wan, Y., Shang, J., Graham, R., Baric, R.S. and Li, F., 2020. Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS coronavirus. *Journal of virology*, 94(7), pp.e00127-20.
18. Ahmad, I., Pawara, R., Surana, S. and Patel, H., 2021. The Repurposed ACE2 Inhibitors: SARS-CoV-2 Entry Blockers of Covid-19. *Topics in Current Chemistry*, 379(6), pp.1-49.
19. Triggile, C.R., Bansal, D., Ding, H., Islam, M.M., Farag, E.A.B.A., Hadi, H.A. and Sultan, A.A., 2021. A comprehensive review of viral characteristics, transmission, pathophysiology, immune response, and management of SARS-CoV-2 and COVID-19 as a basis for controlling the pandemic. *Frontiers in immunology*, 12, p.338.
20. Khan, A., Benthin, C., Zeno, B., Albertson, T.E., Boyd, J., Christie, J.D., Hall, R., Poirier, G., Ronco, J.J., Tidswell, M. and HARDS, K., 2017. A pilot clinical trial of recombinant human angiotensin-converting enzyme 2 in acute respiratory distress syndrome. *Critical care*, 21(1), pp.1-9.
21. Kai, H. and Kai, M., 2020. Interactions of coronaviruses with ACE2, angiotensin II, and RAS inhibitors—lessons from available evidence and insights into COVID-19. *Hypertension Research*, 43(7), pp.648-654.
22. Senanayake, S.L., 2020. Drug repurposing strategies for COVID-19. *Future Drug Discovery*, 2(2).
23. Pushpakom, S., Iorio, F., Eyers, P. A., Escott, K. J., Hopper, S., Wells, A., et al.(2019). Drug Repurposing: Progress, Challenges and Recommendations. *Nat.Rev. Drug Discov.* 18, 41–58. doi:10.1038/nrd.2018.168.
24. Oprea, T.I., Bauman, J.E., Bologna, C.G., Buranda, T., Chigaev, A., Edwards, B.S., Jarvik, J.W., Gresham, H.D., Haynes, M.K., Hjelle, B. and Hromas, R., 2011. Drug repurposing from an academic perspective. *Drug Discovery Today: Therapeutic Strategies*, 8(3-4), pp.61-69.

25. Zhou, Y., Hou, Y., Shen, J., Huang, Y., Martin, W, Cheng, F., 2020. Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell discovery*, 6(1), pp.1-18.
26. Mule, S., Singh, A., Greish, K., Sahebkar, A., Kesharwani, P. and Shukla, R., 2022. Drug repurposing strategies and key challenges for COVID-19 management. *Journal of Drug Targeting*, 30(4), pp.413-429.
27. Li, G and De Clercq, E., 2020. Therapeutic options for the 2019 novel coronavirus (2019-nCoV), *Nat. Rev. Drug Discov.* <https://doi.org/10.1038/d41573-020-00016-0>.
28. Wang, Q., Zhang, Y., Wu, L., Niu, S., Song, C., Zhang, Z., Lu, G., Qiao, C., Hu, Y., Yuen, K.Y. and Wang, Q.(2020). Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell*, 181(4), 894-904.
29. Cantini, F., Niccoli, L., Matarrese, D., Nicastrì, E., Stobbione, P. and Goletti, D., 2020. Baricitinib therapy in COVID-19: A pilot study on safety and clinical impact. *Journal of Infection*, 81(2), pp.318-356.
30. Poduri, R., Joshi, G. and Jagadeesh, G., 2020. Drugs targeting various stages of the SARS-CoV-2 life cycle: Exploring promising drugs for the treatment of Covid-19. *Cellular signalling*, 74, p.109721.
31. Young, B., Tan, T.T. and Leo, Y.S., 2021. The place for remdesivir in COVID-19 treatment. *The Lancet Infectious Diseases*, 21(1), pp.20-21.
32. Cantini, F., Goletti, D., Petrone, L., Najafi Fard, S., Niccoli, L. and Foti, R., 2020. Immune therapy, or antiviral therapy, or both for COVID-19: a systematic review. *Drugs*, 80(18), pp.1929-1946.
33. Ledford, H., 2020. Coronavirus breakthrough: dexamethasone is first drug shown to save lives. *Nature*, 582(7813), pp.469-470.
34. Al-Khafaji, K., Al-Duhaidahawi, D. and Taskin Tok, T., 2021. Using integrated computational approaches to identify safe and rapid treatment for SARS-CoV-2. *Journal of Biomolecular Structure and Dynamics*, 39(9), pp.3387-3395.
35. Eldefrawi, A.T., Miller, E.R., Murphy, D.L. and Eldefrawi, M.E., 1982. [3H] Phencyclidine interactions with the nicotinic acetylcholine receptor channel and its inhibition by psychotropic, antipsychotic, opiate, antidepressant, antibiotic, antiviral, and antiarrhythmic drugs. *Molecular Pharmacology*, 22(1), pp.72-81.
36. Heidary, F., & Gharebaghi, R. (2020). Ivermectin: a systematic review from antiviral effects to COVID-19 complementary regimen. *The Journal of antibiotics*, 73(9), 593-602.
37. Chakraborty, C., Sharma, A.R., Bhattacharya, M., Agoramoorthy, G. and Lee, S.S., 2021. The drug repurposing for COVID-19 clinical trials provide very effective therapeutic combinations: lessons learned from major clinical studies. *Frontiers in Pharmacology*, 12.
38. Lou, Y., Liu, L., Yao, H., Hu, X., Su, J., Xu, K., Luo, R., Yang, X., He, L., Lu, X. and Zhao, Q., 2021. Clinical outcomes and plasma concentrations of baloxavir marboxil and favipiravir in COVID-19 patients: an exploratory randomized, controlled trial. *European Journal of Pharmaceutical Sciences*, 157, p.105631.
39. Touret, F. and de Lamballerie, X., 2020. Of chloroquine and COVID-19. *Antiviral research*, 177, p.104762.

40. Cortegiani, A., Ingoglia, G., Ippolito, M., Giarratano, A. and Einav, S., 2020. A systematic review on the efficacy and safety of chloroquine for the treatment of COVID-19. *Journal of critical care*, 57, pp.279-283.
41. Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.H., Nitsche, A. and Müller, M.A., 2020. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *cell*, 181(2), pp.271-280.
42. Mahmoud, A., Mostafa, A., Al-Karmalawy, A.A., Zidan, A., Abulkhair, H.S., Mahmoud, S.H., Shehata, M., Elhefnawi, M.M. and Ali, M.A., 2021. Telaprevir is a potential drug for repurposing against SARS-CoV-2: Computational and in vitro studies. *Heliyon*, 7(9), p.e07962.
43. Zeng, X., Song, X., Ma, T., Pan, X., Zhou, Y., Hou, Y., Zhang, Z., Li, K., Karypis, G. and Cheng, F., 2020. Repurpose open data to discover therapeutics for COVID-19 using deep learning. *Journal of proteome research*, 19(11), pp.4624-4636.
44. Wang, X. and Guan, Y., 2021. COVID-19 drug repurposing: a review of computational screening methods, clinical trials, and protein interaction assays. *Medicinal research reviews*, 41(1), pp.5-28.
45. Hung, I.F.N., Lung, K.C., Tso, E.Y.K., Liu, R., Chung, T.W.H., Chu, M.Y., Ng, Y.Y., Lo, J., Chan, J., Tam, A.R. and Shum, H.P., 2020. Triple combination of interferon beta-1b, lopinavir–ritonavir, and ribavirin in the treatment of patients admitted to hospital with COVID-19: an open-label, randomised, phase 2 trial. *The Lancet*, 395(10238), pp.1695-1704.
46. Center, S.K.C.C., ClinicalTrials. gov [Internet] Bethesda (MD): National Library of Medicine (US); 2000–. Vaccine Therapy With or Without Imiquimod in Treating Patients With Grade, 3.
47. Xue, X., Yu, H., Yang, H., Xue, F., Wu, Z., Shen, W., Li, J., Zhou, Z., Ding, Y., Zhao, Q. and Zhang, X.C., 2008. Structures of two coronavirus main proteases: implications for substrate binding and antiviral drug design. *Journal of virology*, 82(5), pp.2515-2527.
48. Blaising, J., Polyak, S.J. and Pécheur, E.I., 2014. Arbidol as a broad-spectrum antiviral: an update. *Antiviral research*, 107, pp.84-94.
49. Zhou, Y., Hou, Y., Shen, J., Huang, Y., Martin, W. and Cheng, F., 2020. Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell discovery*, 6(1), pp.1-18.
50. Wu, C., Liu, Y., Yang, Y., Zhang, P., Zhong, W., Wang, Y., Wang, Q., Xu, Y., Li, M., Li, X. and Zheng, M., 2020. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharmaceutica Sinica B*, 10(5), pp.766-788.
51. Al-Khafaji, K., Al-Duhaidahawi, D. and Taskin Tok, T., 2021. Using integrated computational approaches to identify safe and rapid treatment for SARS-CoV-2. *Journal of Biomolecular Structure and Dynamics*, 39(9), pp.3387-3395.
52. Shah, B., Modi, P. and Sagar, S.R., 2020. In silico studies on therapeutic agents for COVID-19: Drug repurposing approach. *Life sciences*, 252, p.117652.
53. Kandeel M, Al-Nazawi M. Virtual screening and repurposing of FDA approved drugs against COVID-19 main protease. *Life Sci.* 2020;251:117627
54. Mahanta, S., Chowdhury, P., Gogoi, N., Goswami, N., Borah, D., Kumar, R., Chetia, D., Borah, P., Buragohain, A.K. and Gogoi, B., 2021. Potential anti-viral activity of approved

- repurposed drug against main protease of SARS-CoV-2: an in silico based approach. *Journal of Biomolecular Structure and Dynamics*, 39(10), pp.3802-3811.
55. Pant S, Singh M, Ravichandiran V, Murty USN, Srivastava HK. Peptide-like and small-molecule inhibitors against Covid-19. *J Biomol Struct Dyn*. 2020;1-10
 56. Wang, J., 2020. Fast identification of possible drug treatment of coronavirus disease-19 (COVID-19) through computational drug repurposing study. *Journal of chemical information and modeling*, 60(6), pp.3277-3286.
 57. Odhar, H.A., Ahjel, S.W., Albeer, A.A.M.A., Hashim, A.F., Rayshan, A.M. and Humadi, S.S., 2020. Molecular docking and dynamics simulation of FDA approved drugs with the main protease from 2019 novel coronavirus. *Bioinformation*, 16(3), p.236.
 58. Uno, Y., 2020. Camostat mesilate therapy for COVID-19. *Internal and emergency medicine*, 15(8), pp.1577-1578.
 59. Ashour, N.A., Elmaaty, A.A., Sarhan, A.A., Elkaeed, E.B., Moussa, A.M., Erfan, I.A. and Al-Karmalawy, A.A., 2022. A Systematic Review of the Global Intervention for SARS-CoV-2 Combating: From Drugs Repurposing to Molnupiravir Approval. *Drug Design, Development and Therapy*, 16, p.685.
 60. Li, G. and De Clercq, E., 2020. Therapeutic options for the 2019 novel coronavirus (2019-nCoV). *Nature reviews Drug discovery*, 19(3), pp.149-150.
 61. Wei, L., Shang, J., Ma, Y., Xu, X., Huang, Y., Guan, Y., Duan, Z., Zhang, W., Gao, Z., Zhang, M. and Li, J., 2019. Efficacy and safety of 12-week interferon-based danoprevir regimen in patients with genotype 1 chronic hepatitis C. *Journal of Clinical and Translational Hepatology*, 7(3), p.221.
 62. Hall Jr, D.C. and Ji, H.F., 2020. A search for medications to treat COVID-19 via in silico molecular docking models of the SARS-CoV-2 spike glycoprotein and 3CL protease. *Travel medicine and infectious disease*, 35, p.101646.
 63. Ghanem, A., Emara, H.A., Muawia, S., Abd El Maksoud, A.I., Al-Karmalawy, A.A. and Elshal, M.F., 2020. Tanshinone IIA synergistically enhances the antitumor activity of doxorubicin by interfering with the PI3K/AKT/mTOR pathway and inhibition of topoisomerase II: in vitro and molecular docking studies. *New Journal of Chemistry*, 44(40), pp.17374-17381.
 64. Torres, J., Maheswari, U., Parthasarathy, K., Ng, L., Liu, D.X. and Gong, X., 2007. Conductance and amantadine binding of a pore formed by a lysine-flanked transmembrane domain of SARS coronavirus envelope protein. *Protein science*, 16(9), pp.2065-2071.
 65. Drożdżal, S., Rosik, J., Lechowicz, K., Machaj, F., Kotfis, K., Ghavami, S. and Łos, M.J., 2020. FDA approved drugs with pharmacotherapeutic potential for SARS-CoV-2 (COVID-19) therapy. *Drug resistance updates*, 53, p.100719.
 66. Pouya MA, Afshani SM, Maghsoudi AS, Hassani S, Mirnia K. Classification of the present pharmaceutical agents based on the possible effective mechanism on the COVID-19 infection. *DARU J Pharm Sci*. 2020;28:1–20
 67. Frediansyah, A., Tiwari, R., Sharun, K., Dhama, K. and Harapan, H., 2021. Antivirals for COVID-19: A critical review. *Clinical Epidemiology and global health*, 9, pp.90-98.
 68. Whitley, R., 2022. Molnupiravir—a step toward orally bioavailable therapies for Covid-19. *New England Journal of Medicine*, 386(6), pp.592-593..
 69. Vandyck, K. and Deval, J., 2021. Considerations for the discovery and development of 3-chymotrypsin-like cysteine protease inhibitors targeting SARS-CoV-2 infection. *Current Opinion in Virology*, 49, pp.36-40.

70. Sultana, J., Cutroneo, P. M., Crisafulli, S., Puglisi, G., Caramori, G., & Trifirò, G. (2020). Azithromycin in COVID-19 patients: pharmacological mechanism, clinical evidence and prescribing guidelines. *Drug safety*, 43, 691-698.
71. Cava C, Bertoli G, Castiglioni I. In silico discovery of candidate drugs against Covid-19. *Viruses*. 2020;12(4):404
72. Mittal, L., Kumari, A., Srivastava, M., Singh, M., & Asthana, S. (2021). Identification of potential molecules against COVID-19 main protease through structure-guided virtual screening approach. *Journal of Biomolecular Structure and Dynamics*, 39(10), 3662-3680.
73. Das, S., Sarmah, S., Lyndem, S., & Singha Roy, A. (2021). An investigation into the identification of potential inhibitors of SARS-CoV-2 main protease using molecular docking study. *Journal of Biomolecular Structure and Dynamics*, 39(9), 3347-3357.
74. Farag, A., Wang, P., Ahmed, M., & Sadek, H. (2020). Identification of FDA approved drugs targeting COVID-19 virus by structure-based drug repositioning.
75. Gimeno, A., Mestres-Truyol, J., Ojeda-Montes, M. J., Macip, G., Saldivar-Espinoza, B., Cereto-Massagué, A., ... & Garcia-Vallvé, S. (2020). Prediction of novel inhibitors of the main protease (M-pro) of SARS-CoV-2 through consensus docking and drug reposition. *International journal of molecular sciences*, 21(11), 3793.
76. Gupta, M. K., Vemula, S., Donde, R., Gouda, G., Behera, L., & Vadde, R. (2021). In-silico approaches to detect inhibitors of the human severe acute respiratory syndrome coronavirus envelope protein ion channel. *Journal of Biomolecular Structure and Dynamics*, 39(7), 2617-2627.
77. Beck, B. R., Shin, B., Choi, Y., Park, S., & Kang, K. (2020). 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*, 18, 784-790.
78. Elmezayen, A. D., Al-Obaidi, A., Şahin, A. T., & Yelekçi, K. (2021). Drug repurposing for coronavirus (COVID-19): in silico screening of known drugs against coronavirus 3CL hydrolase and protease enzymes. *Journal of Biomolecular Structure and Dynamics*, 39(8), 2980-2992.
79. Hall Jr, D. C., & Ji, H. F. (2020). A search for medications to treat COVID-19 via in silico molecular docking models of the SARS-CoV-2 spike glycoprotein and 3CL protease. *Travel medicine and infectious disease*, 35, 101646.
80. Batra R, Chan H, Kamath G, et al. Screening of therapeutic agents for COVID-19 using machine learning and ensemble docking simulations. *arXiv [q-bio.BM]*. 2020.
81. de Oliveira, O. V., Rocha, G. B., Paluch, A. S., & Costa, L. T. (2021). Repurposing approved drugs as inhibitors of SARS-CoV-2 S-protein from molecular modeling and virtual screening. *Journal of Biomolecular Structure and Dynamics*, 39(11), 3924-3933.
82. Torsello, A., Locatelli, V., Cella, S. G., Sanguini, A. M., & Berti, F. (2003). Moexipril and quinapril inhibition of tissue angiotensin-converting enzyme activity in the rat: evidence for direct effects in heart, lung and kidney and stimulation of prostacyclin generation. *Journal of endocrinological investigation*, 26(1), 79-83.
83. Maschio, G., Alberti, D., Janin, G., Locatelli, F., Mann, J. F., Motolese, M., ... & Angiotensin-Converting-Enzyme Inhibition in Progressive Renal Insufficiency Study Group. (1996). Effect of the angiotensin-converting-enzyme inhibitor benazepril on the

- progression of chronic renal insufficiency. *New England Journal of Medicine*, 334(15), 939-945.
84. Ravati, A., Junker, V., Kouklei, M., Ahlemeyer, B., Culmsee, C., & Kriegelstein, J. (1999). Enalapril and moexipril protect from free radical-induced neuronal damage in vitro and reduce ischemic brain injury in mice and rats. *European journal of pharmacology*, 373(1), 21-33.
85. Satish Sagar;Ashok Kumar Rathinavel;William E. Lutz;Lucas R. Struble;Surender Khurana;Andy T. Schnaubelt;Nitish Kumar Mishra;Chittibabu Guda;Nicholas Y. Palermo;Mara J. Broadhurst;Tobias Hoffmann;Kenneth W. Bayles;St. Patrick M. Reid;Gloria E. O. Borgstahl;Prakash Radhakrishnan; (2021). *Bromelain inhibits SARS-CoV-2 infection via targeting ACE-2, TMPRSS2, and spike protein* . *Clinical and Translational Medicine*, (), -. doi:10.1002/ctm2.281
86. Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N. H., Nitsche, A., Müller, M. A., Drosten, C., & Pöhlmann, S. (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*, 181(2), 271–280.e8.
87. Piepho, R. W. (2000). Overview of the angiotensin-converting-enzyme inhibitors. *American journal of health-system pharmacy*, 57(suppl_1), S3-S7.
88. Dales, N.A., Gould, A.E., Brown, J.A., Calderwood, E.F., Guan, B., Minor, C.A., Gavin, J.M., Hales, P., Kaushik, V.K., Stewart, M. and Tummino, P.J., 2002. Substrate-based design of the first class of angiotensin-converting enzyme-related carboxypeptidase (ACE2) inhibitors. *Journal of the American Chemical Society*, 124(40), pp.11852-11853.
89. Weber, M. A. (1992). Fosinopril: a new generation of angiotensin-converting enzyme inhibitors. *Journal of cardiovascular pharmacology*, 20, S7-12.
90. Malek Mahdavi, A. (2020). A brief review of interplay between vitamin D and angiotensin-converting enzyme 2: Implications for a potential treatment for COVID-19. *Reviews in medical virology*, 30(5), e2119.
91. Kumar, G., Kumar, D., & Singh, N. P. (2021). Therapeutic approach against 2019-nCoV by inhibition of ACE-2 receptor. *Drug research*, 71(04), 213-218.
92. Teralı, K., Baddal, B., & Gülcan, H. O. (2020). Prioritizing potential ACE2 inhibitors in the COVID-19 pandemic: Insights from a molecular mechanics-assisted structure-based virtual screening experiment. *Journal of Molecular Graphics and Modelling*, 100, 107697.
93. Schrodinger, LLC, NY, USA, 2009.
94. LigPrep, Schrodinger LLC, Ney York, NY.
95. Prime, Schrodinger, LLC, Ney York, NY.
96. Protein Preparation Wizard, Schrodinger, LLC, Ney York, NY.
97. Qikprop, Schrodinger, LLC, Ney York, NY.
98. Nayariseri, A., Khandelwal, R., Tanwar, P., Madhavi, M., Sharma, D., Thakur, G., Speck-Planche, A. and Singh, S.K. (2021). Artificial intelligence, big data and machine learning approaches in precision medicine & drug discovery. *Current drug targets*, 22(6), 631-655.
99. Nayariseri, A., Khandelwal, R., Madhavi, M., Selvaraj, C., Panwar, U., Sharma, K., Hussain, T. and Singh, S.K. (2020). Shape-based machine learning models for the

- potential novel COVID-19 protease inhibitors assisted by molecular dynamics simulation. *Current topics in medicinal chemistry*, 20(24), 2146-2167.
100. Jain, D., Udhvani, T., Sharma, S., Gandhe, A., Reddy, P.B., Nayarissari, A. and Singh, S.K. (2019). Design of novel JAK3 Inhibitors towards Rheumatoid Arthritis using molecular docking analysis. *Bioinformation*, 15(2), 68.
 101. Basak, S. C., Nayarissari, A., González-Díaz, H., & Bonchev, D. (2016). Editorial (Thematic Issue: chemoinformatics models for pharmaceutical design, part 1). *Current pharmaceutical design*, 22(33), 5041-5042.
 102. Basak, S. C., Nayarissari, A., González-Díaz, H., & Bonchev, D. (2016). Editorial (Thematic Issue: Chemoinformatics models for pharmaceutical design, part 2). *Current pharmaceutical design*, 22(34), 5177-5178.
 103. Mendonça-Junior, F. J., Scotti, M. T., Nayarissari, A., Zondegoumba, E. N., & Scotti, L. (2019). Natural bioactive products with antioxidant properties useful in neurodegenerative diseases. *Oxidative medicine and cellular longevity*, 2019.
 104. Shukla, P., Khandelwal, R., Sharma, D., Dhar, A., Nayarissari, A., & Singh, S. K. (2019). Virtual screening of IL-6 inhibitors for idiopathic arthritis. *Bioinformation*, 15(2), 121.
 105. Dunna, N. R., Bandaru, S., Raj Akare, U., Rajadhyax, S., Ravi Gutlapalli, V., Yadav, M., & Nayarissari, A. (2015). Multiclass comparative virtual screening to identify novel Hsp90 inhibitors: a therapeutic breast cancer drug target. *Current Topics in Medicinal Chemistry*, 15(1), 57-64.
 106. Vuree, S., Dunna, N.R., Khan, I.A., Alharbi, K.K., Vishnupriya, S., Soni, D., Shah, P., Chandok, H., Yadav, M. and Nayarissari, A. (2013). Pharmacogenomics of drug resistance in Breast Cancer Resistance Protein (BCRP) and its mutated variants. *journal of pharmacy research*, 6(7), 791-798.
 107. Nayarissari, A., Moghni, S. M., Yadav, M., Kharate, J., Sharma, P., Chandok, K. H., & Shah, K. P. (2013). In silico investigations on HSP90 and its inhibition for the therapeutic prevention of breast cancer. *journal of pharmacy research*, 7(2), 150-156.
 108. Grover, A., Katiyar, S. P., Singh, S. K., Dubey, V. K., & Sundar, D. (2012). A leishmaniasis study: structure-based screening and molecular dynamics mechanistic analysis for discovering potent inhibitors of spermidine synthase. *Biochim Biophys Acta*, 1824(12), 1476-1483. doi:10.1016/j.bbapap.2012.05.016
 109. Panwar, U., & Singh, S. K. (2018). Structure-based virtual screening toward the discovery of novel inhibitors for impeding the protein-protein interaction between HIV-1 integrase and human lens epithelium-derived growth factor (LEDGF/p75). *J Biomol Struct Dyn*, 36(12), 3199-3217. doi:10.1080/07391102.2017.1384400
 110. Patidar, K., Deshmukh, A., Bandaru, S., Lakkaraju, C., Girdhar, A., Vr, G., . . . Singh, S. K. (2016). Virtual Screening Approaches in Identification of Bioactive Compounds Akin to Delphinidin as Potential HER2 Inhibitors for the Treatment of Breast Cancer. *Asian Pac J Cancer Prev*, 17(4), 2291-2295. doi:10.7314/apjcp.2016.17.4.2291
 111. Praseetha, S., Bandaru, S., Nayarissari, A., & Sureshkumar, S. (2016). Pharmacological analysis of vorinostat analogues as potential anti-tumor agents targeting human histone deacetylases: an epigenetic treatment stratagem for cancers. *Asian Pacific Journal of Cancer Prevention*, 17(3), 1571-1576.

112. Khandekar, N., Singh, S., Shukla, R., Tirumalaraju, S., Bandaru, S., Banerjee, T., & Nayariseri, A. (2016). Structural basis for the in vitro known acyl-depsipeptide 2 (ADEP2) inhibition to Clp 2 protease from *Mycobacterium tuberculosis*. *Bioinformation*, 12(3), 92.
113. Gudala, S., Khan, U., Kanungo, N., Bandaru, S., Hussain, T., Parihar, M.S., Nayariseri, A. and Mundluru, H.P. (2016). Identification and pharmacological analysis of high efficacy small molecule inhibitors of EGF-EGFR interactions in clinical treatment of non-small cell lung carcinoma: A computational approach. *Asian Pacific Journal of Cancer Prevention*, 16(18), 8191-8196.
114. Gutlapalli, V. R., Sykam, A., Nayariseri, A., Suneetha, S., & Suneetha, L. M. (2015). Insights from the predicted epitope similarity between *Mycobacterium tuberculosis* virulent factors and its human homologs. *Bioinformation*, 11(12), 517.
115. Kelotra, S., Jain, M., Kelotra, A., Jain, I., Bandaru, S., Nayariseri, A., & Bidwai, A. (2015). An in silico appraisal to identify high affinity anti-apoptotic synthetic tetrapeptide inhibitors targeting the mammalian caspase 3 enzyme. *Asian Pacific Journal of Cancer Prevention*, 15(23), 10137-10142.
116. Bandaru, S., Tiwari, G., Akka, J., Kumar Marri, V., Alvala, M., Ravi Gutlapalli, V., Nayariseri, A. and Prasad Mundluru, H. (2015). Identification of high affinity bioactive Salbutamol conformer directed against mutated (Thr164Ile) beta 2 adrenergic receptor. *Current topics in medicinal chemistry*, 15(1), 50-56.
117. Tabassum, A., Rajeshwari, T., Soni, N., Raju, D.S.B., Yadav, M., Nayariseri, A. and Jahan, P. (2014). Structural characterization and mutational assessment of podocin—a novel drug target to nephrotic syndrome—an in silico approach. *Interdisciplinary Sciences: Computational Life Sciences*, 6(1), 32-39.
118. Nayariseri, A., Yadav, M., & Wishard, R. (2013). Computational evaluation of new homologous down regulators of Translationally Controlled Tumor Protein (TCTP) targeted for tumor reversion. *Interdisciplinary Sciences: Computational Life Sciences*, 5(4), 274-279.
119. Udhwani, T., Mukherjee, S., Sharma, K., Sweta, J., Khandekar, N., Nayariseri, A., & Singh, S. K. (2019). Design of PD-L1 inhibitors for lung cancer. *Bioinformation*, 15(2), 139.
120. Gokhale, P., Chauhan, A. P. S., Arora, A., Khandekar, N., Nayariseri, A., & Singh, S. K. (2019). FLT3 inhibitor design using molecular docking based virtual screening for acute myeloid leukemia. *Bioinformation*, 15(2), 104.
121. Sinha, K., Majhi, M., Thakur, G., Patidar, K., Sweta, J., Hussain, T., Nayariseri, A. and Singh, S.K. (2018). Computer-aided drug designing for the identification of high-affinity small molecule targeting cd20 for the clinical treatment of chronic lymphocytic leukemia (CLL). *Current topics in medicinal chemistry*, 18(29), 2527-2542.
122. Mysinger, M. M., Carchia, M., Irwin, J. J., & Shoichet, B. K. (2012). Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *Journal of medicinal chemistry*, 55(14), 6582-6594.
123. Imrie, F., Bradley, A. R., & Deane, C. M. (2021). Generating property-matched decoy molecules using deep learning. *Bioinformatics*, 37(15), 2134-2141.

124. Cleves, A. E., & Jain, A. N. (2020). Structure-and ligand-based virtual screening on DUD-E+: performance dependence on approximations to the binding pocket. *Journal of chemical information and modeling*, 60(9), 4296-4310.
125. Miranda-Quintana, R. A., Bajusz, D., Rácz, A., & Héberger, K. (2021). Differential consistency analysis: which similarity measures can be applied in drug discovery?. *Molecular Informatics*, 40(7), 2060017.
126. Nayariseri, A., & Hood, E. A. (2018). Advancement in microbial cheminformatics. *Current topics in medicinal chemistry*, 18(29), 2459-2461.
127. Wójcikowski, M., Zielenkiewicz, P., & Siedlecki, P. (2015). Open Drug Discovery Toolkit (ODDT): a new open-source player in the drug discovery field. *Journal of cheminformatics*, 7(1), 1-6.
128. Dhanda, S. K., Singla, D., Mondal, A. K., & Raghava, G. P. (2013). DrugMint: a webserver for predicting and designing of drug-like molecules. *Biology Direct*, 8(1), 1-12.
129. Chandrakar, B., Jain, A., Roy, S., Gutlapalli, V.R., Saraf, S., Suppahia, A., Verma, A., Tiwari, A., Yadav, M. and Nayariseri, A. (2013). Molecular modeling of Acetyl-CoA carboxylase (ACC) from *Jatropha curcas* and virtual screening for identification of inhibitors. *journal of pharmacy research*, 6(9), 913-918.
130. O' Hagan, S., Swainston, N., Handl, J., & Kell, D. B. (2015). A 'rule of 0.5' for the metabolite-likeness of approved pharmaceutical drugs. *Metabolomics*, 11, 323-339.
131. Nayariseri, A., & Singh, S. K. (2019). Functional inhibition of VEGF and EGFR suppressors in cancer treatment. *Current topics in medicinal chemistry*, 19(3), 178-179.
132. Yang, Z. Y., Yang, Z. J., Dong, J., Wang, L. L., Zhang, L. X., Ding, J. J., ... & Cao, D. S. (2019). Structural analysis and identification of colloidal aggregators in drug discovery. *Journal of chemical information and modeling*, 59(9), 3714-3726.
133. Wu, G., Liu, J., & Yue, X. (2019). Prediction of drug-disease associations based on ensemble meta paths and singular value decomposition. *BMC bioinformatics*, 20(3), 1-13.
134. Teodoro, M., Phillips, G., Kavraki, L., Satoru, M., Shamir, R., & Tagaki, T. (2000). Singular value decomposition of protein conformational motions: Application to HIV-1 protease. *Currents in Computational Molecular Biology*, 198-199.
135. Maruyama, K., Sheng, Y., Watanabe, H., Fukuzawa, K., & Tanaka, S. (2018). Application of singular value decomposition to the inter-fragment interaction energy analysis for ligand screening. *Computational and Theoretical Chemistry*, 1132, 23-34.
136. Patidar, K., Panwar, U., Vuree, S., Sweta, J., Sandhu, M. K., Nayariseri, A., & Singh, S. K. (2019). An in silico approach to identify high affinity small molecule targeting m-TOR inhibitors for the clinical treatment of breast cancer. *Asian Pacific journal of cancer prevention: APJCP*, 20(4), 1229.
137. Rao, D. M., Nayariseri, A., Yadav, M., & Patel, D. (2010). Comparative modeling of methylentetrahydrofolate reductase (MTHFR) enzyme and its mutational assessment: in silico approach. *Int J Bioinformatics Res*, 2, 5-9.
138. Ali, M. A., Vuree, S., Goud, H., Hussain, T., Nayariseri, A., & Singh, S. K. (2019). Identification of high-affinity small molecules targeting gamma secretase for the treatment of Alzheimer's disease. *Current topics in medicinal chemistry*, 19(13), 1173-1187.

139. Mulpuru, V., & Mishra, N. (2021). In silico prediction of fraction unbound in human plasma from chemical fingerprint using automated machine learning. *Acs Omega*, 6(10), 6791-6797.
140. Xie, Y., Liu, X., Ma, X., Duan, Y., Yao, Y., & Cai, Q. (2018). Small titanium-based MOFs prepared with the introduction of tetraethyl orthosilicate and their potential for use in drug delivery. *ACS applied materials & interfaces*, 10(16), 13325-13332.
141. Sweta, J., Khandelwal, R., Srinitha, S., Pancholi, R., Adhikary, R., Ali, M. A., ... & Singh, S. K. (2019). Identification of high-affinity small molecule targeting IDH2 for the clinical treatment of acute myeloid leukemia. *Asian Pacific journal of cancer prevention: APJCP*, 20(8), 2287.
142. Limaye, A., Sweta, J., Madhavi, M., Mudgal, U., Mukherjee, S., Sharma, S., Hussain, T., Nayariseri, A. and Singh, S.K. (2019). In silico insights on gd2: a potential target for pediatric neuroblastoma. *Current topics in medicinal chemistry*, 19(30), 2766-2781.
143. Nayariseri, A. (2019). Prospects of utilizing computational techniques for the treatment of human diseases. *Curr Top Med Chem*, 19(13), 1071-1074.
144. Nayariseri, A. (2020). Experimental and computational approaches to improve binding affinity in chemical biology and drug discovery. *Current Topics in Medicinal Chemistry*, 20(19), 1651-1660.
145. Sarmah, P., Konwar, P., Saikia, J., Borah, T., Verma, J. S., & Banik, D. (2023). Screening of potent inhibitor from *Aquilaria malaccensis* Lam. against arachidonic inflammatory enzymes: an insight from molecular docking, ADMET, molecular dynamics simulation and MM-PBSA approaches. *Journal of Biomolecular Structure and Dynamics*, 1-15.
146. Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific reports*, 7(1), 42717.
147. Ranjith, D., & Ravikumar, C. (2019). SwissADME predictions of pharmacokinetics and drug-likeness properties of small molecules present in *Ipomoea mauritiana* Jacq. *Journal of Pharmacognosy and Phytochemistry*, 8(5), 2063-2073.
148. Mukherjee, S., Abdalla, M., Yadav, M., Madhavi, M., Bhirdwaj, A., Khandelwal, R., Prajapati, L., Panicker, A., Chaudhary, A., Albrakati, A. and Hussain, T. (2022). Structure-Based Virtual Screening, Molecular Docking, and Molecular Dynamics Simulation of VEGF inhibitors for the clinical treatment of Ovarian Cancer. *Journal of Molecular Modeling*, 28(4), 1-21.
149. Yadav, M., Abdalla, M., Madhavi, M., Chopra, I., Bhirdwaj, A., Soni, L., Shaheen, U., Prajapati, L., Sharma, M., Sikarwar, M.S. and Albogami, S. (2022). Structure-Based Virtual Screening, Molecular Docking, Molecular Dynamics Simulation and Pharmacokinetic modelling of Cyclooxygenase-2 (COX-2) inhibitor for the clinical treatment of Colorectal Cancer. *Molecular Simulation*, 1-21.
150. Lorin, S., Rajaraman, D., Sonadevi, S., Jaganathan, R., Kumaradhas, P., Anthony, L. A., ... & Raja, K. (2023). Synthesis, quantum chemical studies, molecular docking, molecular dynamics simulation and ADMET studies on 2-(2, 3-dihydrobenzo [b][1, 4] dioxin-6-yl)-1, 4, 5-triphenyl-1 H-imidazole derivatives. *Molecular Physics*, e2295427.
151. Gadnayak, A., Dehury, B., Nayak, A., Jena, S., Sahoo, A., Panda, P. C., ... & Nayak, S. (2022). 'Mechanistic insights into 5-lipoxygenase inhibition by active principles

- derived from essential oils of Curcuma species: Molecular docking, ADMET analysis and molecular dynamic simulation study. *PLoS One*, 17(7), e0271956.
152. Mansour, M. A., AboulMagd, A. M., & Abdel-Rahman, H. M. (2020). Quinazoline-Schiff base conjugates: in silico study and ADMET predictions as multi-target inhibitors of coronavirus (SARS-CoV-2) proteins. *RSC advances*, 10(56), 34033-34045.
 153. Taghour, M. S., Elkady, H., Eldehna, W. M., El-Deeb, N., Kenawy, A. M., Abd El-Wahab, A. E., ... & Eissa, I. H. (2023). Discovery of new quinoline and isatine derivatives as potential VEGFR-2 inhibitors: Design, synthesis, antiproliferative, docking and MD simulation studies. *Journal of Biomolecular Structure and Dynamics*, 41(21), 11535-11550.
 154. Oussama, C., Abbdellah, E. L., Youssef, E. O., Bouachrine, M., & Ouammou, A. (2022). In Silico Prediction of Novel (TRIM24) Bromodomain Inhibitors: A Combination of 3D-QSAR, Molecular Docking, ADMET Prediction, and Molecular Dynamics Simulation. *Physical Chemistry Research*, 10(4), 519-535.
 155. Kawsar, S., Munia, N. S., Saha, S., & Ozeki, Y. (2024). In Silico Pharmacokinetics, Molecular Docking and Molecular Dynamics Simulation Studies of Nucleoside Analogs for Drug Discovery-A Mini Review. *Mini Reviews in Medicinal Chemistry*, 24(11), 1070-1088.
 156. Sharda, S., Khandelwal, R., Adhikary, R., Sharma, D., Majhi, M., Hussain, T., Nayariseri, A. and Singh, S.K. (2019). A computer-aided drug designing for pharmacological inhibition of mutant ALK for the treatment of non-small cell lung cancer. *Current topics in medicinal chemistry*, 19(13), 1129-1144.
 157. Yadav, M., Khandelwal, R., Mudgal, U., Srinitha, S., Khandekar, N., Nayariseri, A., Vuree, S. and Singh, S.K., (2019). Identification of potent VEGF inhibitors for the clinical treatment of glioblastoma, a virtual screening approach. *Asian Pacific journal of cancer prevention: APJCP*, 20(9), 2681.
 158. Nayariseri, A. (2020). Most promising compounds for treating COVID-19 and recent trends in antimicrobial & antifungal agents. *Current topics in medicinal chemistry*, 20(24), 2119-2125.
 159. Adhikary, R., Khandelwal, R., Hussain, T., Nayariseri, A., & Singh, S. K. (2021). Structural insights into the molecular design of ROS1 inhibitor for the treatment of non-small cell lung cancer (NSCLC). *Current Computer-Aided Drug Design*, 17(3), 387-401.
 160. Qureshi, S., Khandelwal, R., Madhavi, M., Khurana, N., Gupta, N., Choudhary, S.K., Suresh, R.A., Hazarika, L., Srija, C.D., Sharma, K. and Hindala, M.R. (2021). A multi-target drug designing for BTK, MMP9, proteasome and TAK1 for the clinical treatment of mantle cell lymphoma. *Current Topics in Medicinal Chemistry*, 21(9), 790-818.
 161. Natchimuthu, V., Abdalla, M., Yadav, M., Chopra, I., Bhirdwaj, A., Sharma, K., Ravi, S., Ravikumar, K., Alzahrani, K.J., Hussain, T. and Nayariseri, A. (2022). Synthesis, crystal structure, hirshfeld surface analysis, molecular docking and molecular dynamics studies of novel olanzapinium 2,5-d2, 5-dihydroxybenzoate as potential and active antipsychotic compound. *Journal of Experimental Nanoscience*, 17(1), 247-273.
 162. Mendonça-Junior, F. J. B., Scotti, M. T., Muratov, E. N., Scotti, L., & Nayariseri, A. (2021). Natural Bioactive Products with Antioxidant Properties Useful in Neurodegenerative Diseases 2020. *Oxidative Medicine and Cellular Longevity*, 2021.
 163. Accelrys, I. (2011). Discovery studio visualizer 3.0. *San Diego, CA*, 92121.

164. Nayarisseri, A., Abdalla, M., Joshi, I., Yadav, M., Bhirdwaj, A., Chopra, I., Khan, A., Saxena, A., Sharma, K., Panicker, A. and Panwar, U. & Singh, S. K. (2024). Potential inhibitors of VEGFR1, VEGFR2, and VEGFR3 developed through Deep Learning for the treatment of Cervical Cancer. *Scientific Reports*, 14(1), 13251.
165. Nayarisseri, A., Bandaru, S., Khan, A., Sharma, K., Bhirdwaj, A., Kaur, M., Ghosh, D., Chopra, I., Panicker, A., Kumar, A. and Saravanan, P. & Singh, S. K. (2024). Epigenetic dysregulation in cancers by isocitrate dehydrogenase 2 (IDH2). *Advances in protein chemistry and structural biology*, 141, 223-253.
166. Sharma, K., Panwar, U., Madhavi, M., Joshi, I., Chopra, I., Soni, L., ... & Singh, S. K. (2024). Unveiling the ESR1 conformational stability and screening potent inhibitors for breast cancer treatment. *Medicinal Chemistry*, 20(3), 352-368.
167. Maia, M. D. S., Mendonça-Junior, F. J. B., Rodrigues, G. C. S., Silva, A. S. D., Oliveira, N. I. P. D., Silva, P. R. D., ... & Scotti, L. (2023). Virtual screening of different subclasses of lignans with anticancer potential and based on genetic profile. *Molecules*, 28(16), 6011.
168. Bhirdwaj, A., Abdalla, M., Pande, A., Madhavi, M., Chopra, I., Soni, L., Vijayakumar, N., Panwar, U., Khan, M.A., Prajapati, L. and Gujrati, D. (2023). Structure-based virtual screening, molecular docking, molecular dynamics simulation of EGFR for the clinical treatment of glioblastoma. *Applied Biochemistry and Biotechnology*, 195(8), 5094-5119.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTable1.docx](#)