



OPEN ACCESS

EDITED BY

Selvam Chelliah,
Texas Southern University, United States

REVIEWED BY

Penke Vijaya Babu,
Tikvah Pharma Solutions Pvt Ltd., India
Narva Deshwar Kushwaha,
Wayne State University, United States

*CORRESPONDENCE

Safar M. Alqahtani,
✉ safar.alqahtani@psau.edu.sa

RECEIVED 04 July 2024

ACCEPTED 10 January 2025

PUBLISHED 05 February 2025

CITATION

Alqahtani SM (2025) Discovering broad-spectrum inhibitors for SARS-CoV-2 variants: a cheminformatics and biophysical approach targeting the main protease.
Front. Pharmacol. 16:1459581.
doi: 10.3389/fphar.2025.1459581

COPYRIGHT

© 2025 Alqahtani. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Discovering broad-spectrum inhibitors for SARS-CoV-2 variants: a cheminformatics and biophysical approach targeting the main protease

Safar M. Alqahtani*

Department of Pharmaceutical Chemistry, College of Pharmacy, Prince Sattam Bin Abdulaziz University, Al Kharj, Saudi Arabia

The COVID-19 pandemic caused by SARS-CoV-2 still lacks effective antiviral drugs. Therefore, a thorough receptor-based virtual screening study was conducted to screen different natural and synthetic drug libraries, such as the Asinex Antiviral, Seaweed Metabolite Database, Medicinal Fungi Secondary Metabolite and Therapeutics Library, and Comprehensive Marine Natural Products Database comprising 6,827, 1,191, 1,830, and 45,000 compounds, respectively, against the main protease enzyme of SARS-CoV-2. Accordingly, three drug molecules (BBB-26580140, BDE-32007849, and LAS-51378804) are highlighted as the best binding molecules to the main protease S1 pocket. The docking binding energy scores of BBB-26580140, BDE-32007849, and LAS-51378804 were -13.02 , -13.0 , and -12.56 kcal/mol, respectively. Compared to the control Z1741970824 molecule with a binding energy score of -11.59 kcal/mol, the lead structures identified herein showed robust hydrophilic and van der Waals interactions with the enzyme active site residues, such as His41 and Cys145, and achieved highly stable binding modes. The simulations showed a stable structure of the main protease enzyme with the shortlisted leads in the pocket, and the network of binding interactions remained intact during the simulations. The overall molecular mechanics with generalized Born and surface area solvation binding energies of the BBB-26580140, BDE-32007849, LAS-51378804, and control molecules are -53.02 , -56.85 , -55.44 , and -48.91 kcal/mol, respectively. Similarly, the net molecular mechanics Poisson–Boltzmann surface area binding energies of BBB-26580140, BDE-32007849, LAS-51378804, and control are -53.6 , -57.61 , -54.41 , and -50.09 kcal/mol, respectively. The binding entropy energies of these systems showed lower free energies, indicating their stable nature. Furthermore, the binding energies were revalidated using the water swap method that considers the role of the water molecules in bridging the ligands to the enzyme active site residues. The compounds also revealed good ADMET

properties and followed all major rules of drug-likeness. Thus, these compounds are predicted as promising leads and can be subjected to further experimental studies for evaluation of their biological activities.

KEYWORDS

COVID-19, SARS-CoV-2, BBB-26580140, BDE-32007849, LAS-51378804, molecular dynamics simulations, water swap method

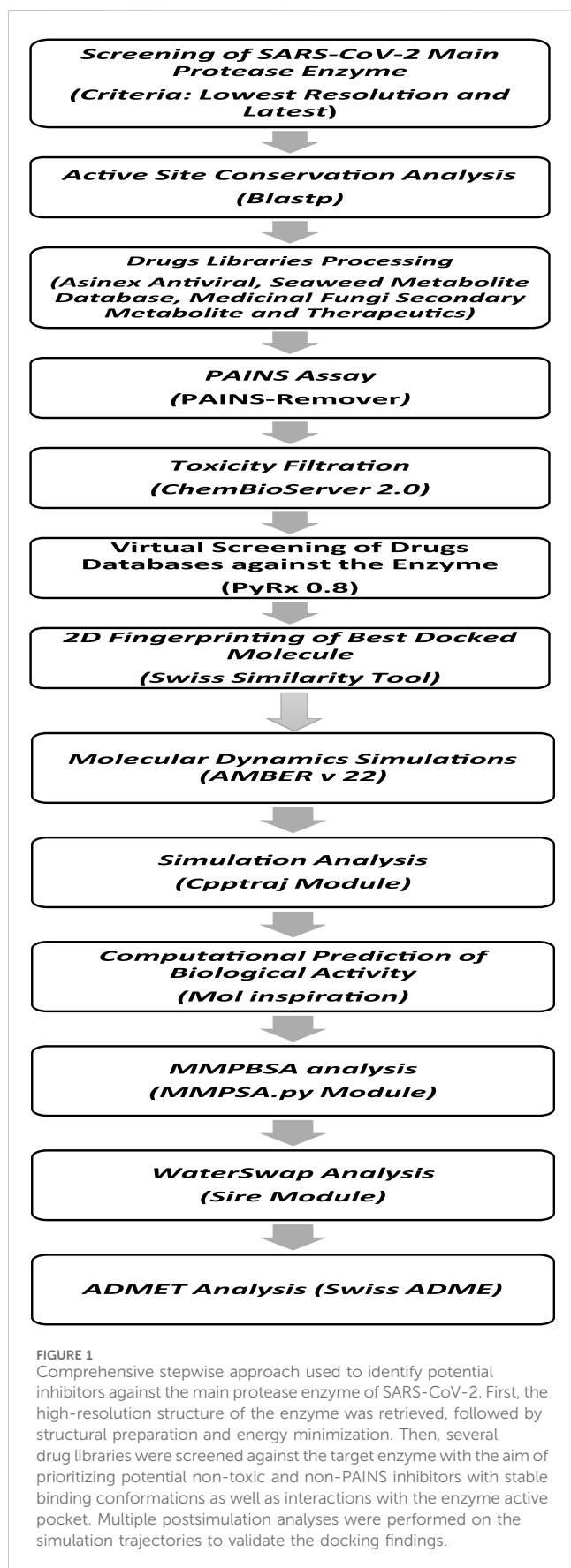
1 Introduction

Wuhan city in the Hubei province of China experienced an outbreak of pneumonia with unknown etiology around the end of December 2019. Up until 31 January 2020, the outbreak had grown significantly, infecting 9,720 people and resulting in the deaths of 213 people in China as well as 106 persons in 19 other countries (Carabelli et al., 2023). The coronavirus pandemic of 2019 first started in Wuhan city and was declared a pandemic by the World Health Organization (WHO) when the disease spread across the globe (Excler et al., 2021; Omolo et al., 2020). It was found to be caused by the highly contagious severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) microorganism (Viana et al., 2022). This virus is grouped under the Coronaviridae family and has an RNA genome with distinct crown-like spikes on the surface (Dai et al., 2020). As of 23 November 2024, this disease is responsible for 7,010,681 deaths resulting from 704,753,890 confirmed cases of infection in 231 countries (<https://www.worldometers.info/coronavirus/>). Although substantial efforts have been made in the last 2–3 years to better understand the biology of SARS-CoV-2, outbreaks of the disease are still being reported in many countries (Msemburi et al., 2023).

The COVID-19 pandemic was triggered by SARS-CoV-2. Since the coronavirus 3-chymotrypsin-like protease (3CLpro) regulates viral replication, it is considered as a key target and potential approach for the development of direct-acting rationally based antiviral drugs (Vandyck and Deval, 2021). Different mutants of the virus have been reported to be associated with the newer outbreaks (Khan et al., 2021). This high mutagenicity of the virus is attributable to its RNA genome, which acquires genetic adaptations to rapidly form new variants with different characteristics than the ancestral strain (Korber et al., 2020). The WHO has thus far reported five SARS-CoV-2 variants based on epidemiological data, namely the omicron (B.1.1.529), alpha (B.1.1.7), beta (B.1.351), gamma (P.1), and delta (B.1.617.2) types (Carabelli et al., 2023). The treatment options available for COVID-19 and recommended by WHO are as follows: baricitinib, sotrovimab, molnupirvir, remdesivir, sarilumab, and tocilizumab as IL-6 receptor blockers (Mousavi et al., 2022). Despite the availability of these drugs, COVID-19 remains an emergent disease for which development of new treatments is required. SARS-CoV-2 codes for at least four structural proteins that help with viral RNA synthesis, viral assembly process, and viral binding to the host cell receptors (Kim et al., 2020). The structural proteins include spike, membrane, envelope, nucleocapsid, and 16 non-structural proteins (Calligari et al., 2020). The main protein enzymes, also known as chymotrypsin-like and papain-like proteases, are

the two proteases vital for coronaviral proteolytic processing of polyproteins (Narayanan et al., 2022). The main protease is the key enzyme of SARS-CoV-2 biology as well as functionality and is significant as a drug target (Agost-Beltrán et al., 2022). The main protease enzyme was found to be the target in about 154 research studies, including both *in silico* and experimental works, and several compounds have been reported against it. Natural compounds are rich sources for drug development owing to their strong bioactivities and structural variety, which have long been recognized as important criteria. For instance, research has shown that plant-derived alkaloids and flavonoids can effectively block important enzymes linked to diseases because of their distinct chemical structures (Chopra and Dhingra, 2021). As proven by the production of sulfonamide derivatives with increased inhibitory activities against bacterial enzymes, synthetic compounds have the advantage of customization for particular biological targets integrating specific strategies; a new study has shown how natural scaffolds and synthetic shifts work collaboratively to improve the therapeutic design efficacy and specificity (Dzobo, 2022). Several inhibitors have been proposed against the main protease enzyme of SARS-CoV-2. For example, PF-07304814, PF-07304814, and PF-07304814 are phase 1 trial inhibitors proposed by Pfizer. Ebselen and masitinib proposed by Sound Pharmaceuticals and AB Science, respectively, are currently under phase 2 trials (Pang et al., 2023).

The development of new drug molecules is a complicated, expensive, and lengthy process that takes approximately 2–4 years of preclinical efforts and up to 6 years of clinical studies (Maltezou and Papa, 2011; Shaker et al., 2021), with the total procedural costs often being as much as 500 million dollars. Conventional drug discovery entails three vital steps, namely discovery and development of targets or leads, preclinical research, and clinical development (Yu and MacKerell, 2017). Drug discovery also includes hit screening, applications of medicinal chemistry, and structural/biological activity optimizations to limit the side effects of the drugs (Gediya and Njar, 2009); these steps often involve the use of either classical or reverse pharmacology. Both of these technologies have their pros and cons. Computational drug discovery has been proven significant in the recent past for its utility in identifying and optimizing lead structures (Van Drie, 2007). Computational and biophysics techniques could reduce time and deliver hits that can directly be used in experimental work. The present computational study aims to identify novel inhibitors against the main protease enzyme of SARS-CoV-2 using different techniques through computational tools like UCSF Chimera, Discovery Studio, AMBER package, and SwissADME.



2 Materials and methods

The methodological steps of this study, starting from enzyme retrieval to receptor-based virtual screening and biophysics techniques, are presented in Figure 1.

2.1 Main protease enzyme crystal structure searching, retrieval, and processing

An initial exhaustive search was performed for the best resolution and latest structure deposited in the protein data bank (PDB) from the start of the COVID-19 pandemic up to 5 September 2023. This allowed identification of the crystal structure of the protease enzyme (PDB ID: 5RF3) in complex with the compound Z1741970824 (Doungamath et al., 2020). The structure was deposited to the PDB on 15 March 2020 and released on the same day. The structure was also determined by X-ray diffraction and is available at the best resolution of 1.50 Å. This structure is expressed in *Escherichia coli*, which was used as the expression system. These findings suggest the determination of a high-quality main protease structure. Next, the structure was imported into UCSF Chimera v1.17 (Kaliappan and Bombay, 2018), where it was first visualized for missing residues and structural errors. The missing hydrogen atoms and charges were added using “AddH” and “Add Charge” features in Chimera v1.17 (Kaliappan and Bombay, 2018). Energy minimization was then carried out to relax the enzyme structure, and the process was completed using two algorithms. The first algorithm was the steepest descent method, which was implemented for 5,000 cycles with a step size of 0.02 Å. The second algorithm used was the conjugate gradient method, which was also implemented for 5,000 steps. Both these algorithms ensured the removal of steric clashes present in the enzyme structure.

2.2 Active site conservation analysis

The active site conservation analysis was performed using the main protease enzyme sequence of SARS-CoV-2 from the given PDB ID. The sequence was imported into the Blastp tool of NCBI and searched against the PDB (Mahram and Herbor dt, 2015). A multiple sequence alignment (MSA) was performed, and the conservation of two important catalytic residues (His41 and Cys145) at the S1 pocket of the enzyme was examined (Elsässer et al., 2017). A complete conservation was observed, illustrating the key mechanistic roles of these residues in enzyme catalysis.

2.3 Drug library selection, retrieval, and processing

A thorough literature search was conducted to determine the drug libraries that best meet the current research aim. The selection

of libraries was based on the following criteria: chemical scaffold diversity, easy availability for experimental assays, most extracted data from natural sources, and not much explored in past studies. Accordingly, the Asinex antiviral library (Yadav et al., 2021), Seaweed Metabolite Database (Davis and Vasanthi, 2011), Medicinal Fungi Secondary Metabolite and Therapeutics Library (MFSMTL) (Vivek-Ananth et al., 2021), and Comprehensive Marine Natural Products Database (CMNPD) (Sheokand et al., 2024) were found, which have 6,827, 1,191, 1,830, and 45,000 compounds, respectively. The Asinex antiviral library was designed from the perspectives of safety profile and profound antiviral activity. The Seaweed Metabolite Database comprises compounds extracted from marine algae. The MFSMTL is a curated database of natural products and contains secondary metabolites extracted from diverse medicinal fungi. The CMNPD is rich in marine compounds extracted from bacteria, fungi, green algae, sponges, cnidarians, echinoderms, etc. After retrieval, the libraries were subjected to pan-assay interference compounds (PAINS) filtration to remove false positive compounds (Whitty, 2011). Thereafter, toxicity filtration was performed to remove toxic chemical scaffolds from the filtered PAINS libraries; this was accomplished using the ChemBioServer 2.0 server (Athanasiadis et al., 2012). Then, the filtered compounds were imported into PyRx 0.8 for energy minimization (Dallakyan and Olson, 2015), where the libraries were energy minimized using the MM2 force field (Halgren, 1996) and converted to .pdbqt format.

2.4 Site-directed virtual screening

Site-directed virtual screening of the selected drug libraries against the S1 pocket of the main protease enzyme was conducted in PyRx 0.8 (Dallakyan and Olson, 2015). One of the essential enzymes that convert the viral polyprotein into useful components required for viral replication is the SARS-CoV-2 major protease (Mpro). This enzyme is an important target for treatment as it effectively stops the viral replication cycle (Pang et al., 2023). The grid boxes were set around His41:ND1 (x -axis: 12.676 Å, y -axis: -4.526 Å, z -axis: 21.049 Å) and Cys145: N (x -axis: 6.416 Å, y -axis: -5.110 Å, z -axis: 16.812 Å) (Doungamath et al., 2020). The box size along each dimension was set to 30 Å. The number of binding conformations generated for each compound at the active pocket of the enzyme was 100. Then, iterations with binding energy scores ≥ -5.00 kcal/mol were discarded. A control molecule in the form of Z1741970824 was used to test and validate the docking protocol. This control molecule was extracted from the crystal structure and docked with the enzyme at the same position using PyRx 0.8, and the root mean-squared deviation (RMSD) value was calculated using UCSF Chimera v1.17 (Kaliappan and Bombay, 2018). It was observed that both the crystal and PyRx 0.8 docked conformations were similar, with an RMSD value of 0.15 Å. The best-docked solutions from virtual screening were filtered and used in Discovery Studio Client v2021 to thoroughly examine the chemical interactions (Biovia, 2017). The binding conformations of the selected compounds with the enzyme were studied using UCSF Chimera v1.17 (Kaliappan and Bombay, 2018).

2.5 2D fingerprinting of the best-docked molecules

The 2D fingerprinting of the best-docked molecules was performed using the Swiss similarity online server (Zoete et al., 2016). The main objective here was to search for the derivatives of leads identified in the virtual screening process and to evaluate their binding affinities with the main protease enzyme.

2.6 Molecular dynamics simulations: systems preparation and production

Molecular dynamics simulation is a powerful biophysics tool that provides insights into complex dynamics over the simulation time (Ahmad et al., 2017; Rafi et al., 2020). The simulations were carried out using AMBER v22 (Case et al., 2022). The systems were preprocessed using the Antechamber program (Wang et al., 2001). The leap module was applied to add missing hydrogen atoms (Schafmeister et al., 1995). The systems were solvated into the OPC water model and counter ions were added subsequently to neutralize the charge on the complexes (Sengupta et al., 2021). To describe the receptors, the FF19SB force field was used. The compounds were treated using the general Amber force field 2 (GAFF2) (Tian et al., 2019; Vasetti et al., 2019), and the energy minimization was achieved using 500 iterations of the steepest descent and conjugate algorithms. During this process, the protein was fixed with a force constant of 250 kcal/mol/Å². The systems were heated to 300 K under a constant volume (NVT) ensemble for 20 ps. At the same time, a weak constraint of 10 kcal/mol/Å² was applied to the protein atoms. Then, ensemble equilibration of the systems was performed for approximately 1 ns under NPT. The production iterations comprised 100-ns runs with an average pressure of 1 atm using isotropic position scaling. Langevin dynamics was applied for temperature control in the presence of a collision frequency of 1 ps⁻¹ (Izaguirre et al., 2001). The long-term non-bonded interactions were handled using the particle mesh Ewald method (Petersen, 1995). The hydrogen atoms involved in bonding were constrained by the SHAKE method (Kräutler et al., 2001). The simulation trajectories were determined using the CPPTRAJ script of AMBER for structural stability investigations (Roe and Cheatham, 2013). Lastly, XMGRACE v5.1 was used to obtain the plots (Turner, 2005).

2.7 Estimating binding free energies

Estimation of the binding free energies is vital in the current drug discovery pipeline because of the modest computational requirements and prediction accuracy. The binding free energies are considered to be more accurate than the docking calculations as the former is based on simulation trajectory frames and considered as dynamic conformations of the intermolecular binding; the latter is based only on one docked intermolecular conformations (Wang et al., 2019). The binding free energies were estimated using the molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) approach in AMBER v22. The MM-PBSA was obtained using the AMBER MMPBSA.py module (Miller III et al., 2012). The initial

.prmtop files for the complexes, receptors, and ligands were generated using the ante-MMPBSA.py module. The MM-PBSA is based on the principle of subtraction of the receptor and ligand energies from that of the complex (Genheden and Ryde, 2015). Mathematically, the MM-PBSA is given by

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

Individually, the G terms above comprise the molecular mechanics energy, coupled temperature and entropic energy, and solvation free energy, as given by the following equation:

$$G = E_{\text{molecular mechanics}} - (TS) + (G_{\text{solvation}}).$$

The MM-PBSA energy equation was simulated on a total of 10,000 frames. The binding entropies for the complexes were obtained separately using normal mode analysis in AMBER (Genheden et al., 2012). Owing to the computationally intensive nature of the entropy energy calculations, only five frames were processed.

2.8 Water swap analysis

The water swap approach is considered to be a more robust and theoretically efficient method for predicting the absolute binding affinity of a guest ligand to a given macromolecule (Woods et al., 2014). Here, a protein ligand is swapped with an equal volume of water clusters present in the protein active pocket (Bergström and Larsson, 2018). This method entails constructing a λ -coordinate that allows the joining of a periodic box of water molecules with a periodic box of ligands. The energy change in the λ -coordinate is predicted by standard methods used in water swap (Ahmad et al., 2019) over a default of 1,000 iterations.

2.9 Prediction of ADMET properties

Predictions of the ADMET properties of the compounds are vital for selecting viable and effective molecules for *in vitro* and *in vivo* testing (Van De Waterbeemd and Gifford, 2003). Compound failure during clinical trials due to poor ADMET properties incur both financial and time costs to scientists (Jia et al., 2019). Therefore, it is preferable to select compounds that are predicted to possess favorable ADMET properties as these are more likely to clear clinical hurdles associated with their chemical attributes. The ADME properties of the compounds were determined using SwissADME (Daina et al., 2017), while their toxicities were predicted through the pkCSM server (Pires et al., 2015).

2.10 Computational prediction of the biological activity

The bioactivity scores of the selected compounds were evaluated using an online Mol inspiration server. This server predicts the bioactivity score of a given compound against the ion channels, kinases, proteases, G-protein-coupled receptors (GPCRs), and enzymes. The Mol inspiration server can be accessed using the following link: <https://www.molinspiration.com/cgi/properties>.

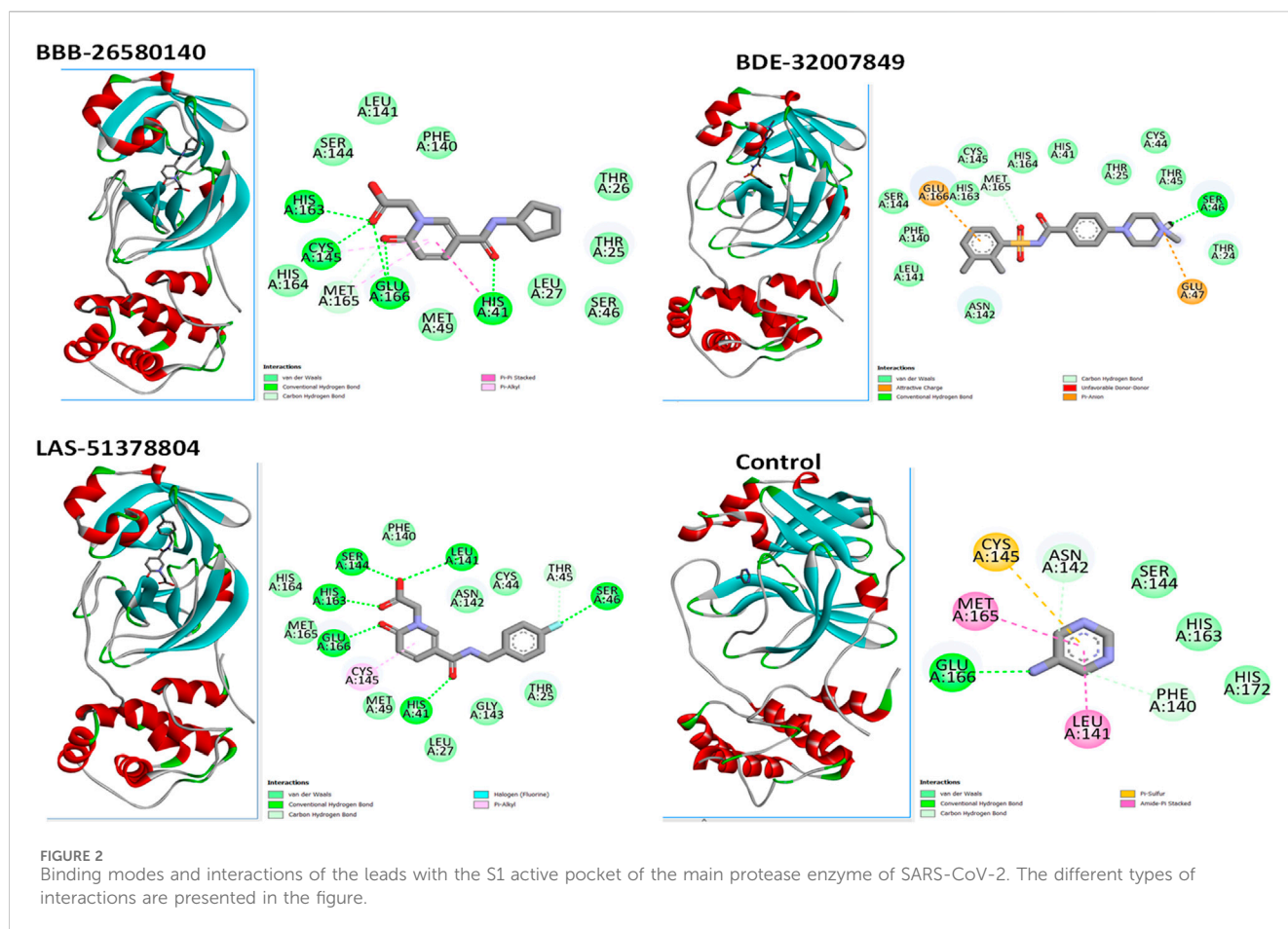
3 Results

3.1 Virtual screening for hit compounds

A structure-based virtual screening protocol was applied to screen compounds from the selected drug libraries. The non-PAINS and non-toxic compounds are only considered as discussed in the methods section. The screening identified three promising compounds from the libraries, namely BBB-26580140, BDE-32007849, and LAS-51378804, with binding energy scores of -13.02 , -13.0 , and -12.56 kcal/mol, respectively. These compounds showed better binding affinity scores than the control (Z1741970824; binding energy score: -11.59 kcal/mol). The lead structures also showed robust hydrophilic and hydrophobic interactions with the enzyme active site residues (His41 and Cys145) and achieved highly stable docked binding modes. The interaction patterns of the compounds were observed under broad and expanded conditions for the S1, S1', S2, and S3 active pocket regions. All three compounds showed deep bindings at the active pocket and were found to have adjusted well inside the pocket. The chemical names of BBB-26580140, BDE-32007849, and LAS-51378804 are 1-(carboxymethyl)-5-(cyclopentylcarbamoyl)pyridin-1-ium-2-olate, 4-(4-(((3-chloro-2-methylphenyl)sulfonyl)carbamoyl)phenyl)-1-methylpiperazin-1-ium, and 1-(carboxymethyl)-5-((4-fluorobenzyl)carbamoyl)pyridin-1-ium-2-olate, respectively. The BBB-26580140 compound formed hydrogen bonds with His41, Cys145, His163, and Glu166 with interaction lengths of 1.56 Å, 1.84 Å, 1.64 Å, and 2.0 Å, respectively; the compound also showed hydrophobic interactions with Leu141, Ser144, Phe140, Thr26, Thr25, Ser46, Leu27, Met49, Met165, and His164. The BDE-32007849 compound produced a hydrogen bond with Ser46 at a bond length of 1.87 Å; this hydrogen bond was formed through the compound terminal benzene ring. The remainder of the compound structure (N-((3-chloro-2-methylphenyl)sulfonyl)benzamide) formed hydrophobic contacts with His41, Thr25, Cys44, Thr45, Thr24, Ser144, Leu141, Phe140, Asn142, Glu47, and Glu166, among others. The LAS-51378804 compound interacted majorly through its 1-(carboxymethyl)-5-(methylcarbamoyl)pyridin-1-ium-2-olate part and formed hydrogen bonds with Ser144, His163, Glu166, and His41 at bond lengths of 1.56 Å, 2.01 Å, 1.56 Å, and 1.42 Å, respectively. Fluorobenzene also formed a hydrogen bond with Ser46 at a bond length of 2.14 Å. The docked conformations and chemical bondings of the compounds and control are given in Figure 2. These results illustrate that the compounds are deeply and strongly bound to the main protease enzyme S1 pocket, resulting in highly stable binding conformations. Furthermore, each of the lead molecules was used in a similarity-based 2D search, and the best results against each lead were tabulated.

3.2 2D fingerprinting of the best-docked molecules

Two-dimensional fingerprinting analysis was performed against the drug-like molecules from the PDB, which allowed shortlisting of 15 compounds for BBB-26580140, six compounds for BDE-32007849, and 16 compounds for LAS-51378804. The prediction scores against each of the derivatives are given in Table 1. The



compounds were redocked with the protease enzyme using the same protocols and parameters. None of the structurally similar compounds were found to have better binding than the parent structure.

3.3 Molecular dynamics simulations

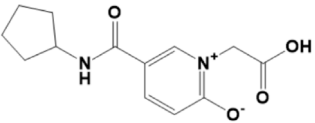
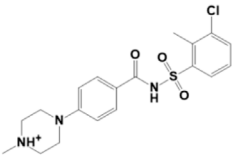
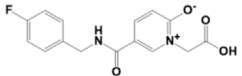
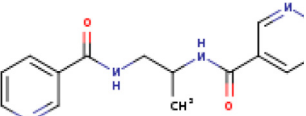
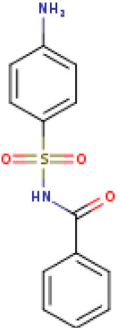
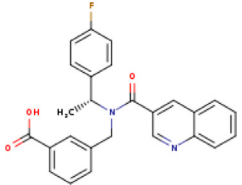
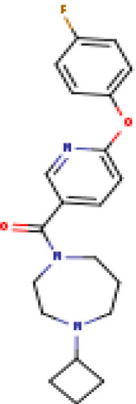

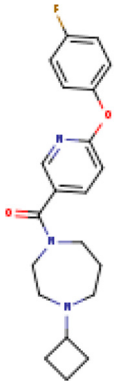
Study of the macromolecular structure is a key step in understanding the biology of a molecule. The macromolecules perform their functions by interacting with different ligands. Therefore, stable dynamic behaviors of the docked complexes were interpreted. In molecular dynamics simulations, stable conformations and interactions of the selected leads with the protease enzyme were disclosed. This was essential for the compounds to exert their biological actions by stopping enzyme activity. Statistical analyses were conducted based on different simulations, including the root mean-squared deviation (RMSD) (Ahmad et al., 2017; Maiorov and Crippen, 1994), root mean-squared fluctuation (RMSF) (Ahmad et al., 2017; Kuzmanic and Zagrovic, 2010), and radius of gyration (Rg) (Lobanov et al., 2008). These analyses were conducted by choosing the carbon alpha atoms of the docked systems. The RMSD was performed first; in contrast to the control used, the lead complexes have significantly stable behaviors with no major deviations in the dynamics. The mean RMSDs for BBB-26580140, BDE-32007849, LAS-51378804, and

control were 1.53 Å, 1.31 Å, 1.38 Å, and 1.89 Å, respectively (Figure 3A). These values indicate that the control system shows more structural variation in the presence of the control molecule. On the other hand, the lead systems were more tolerant of the presence of the compounds and showed strong intermolecular conformations, making the complexes overall stable in nature. The RMSF values of the complexes were calculated next; the mean RMSF values of BBB-26580140, BDE-32007849, LAS-51378804, and control were 1.35 Å, 1.20 Å, 1.25 Å, and 2.01 Å, respectively (Figure 3B). Here also, larger deviations in the control systems are depicted by the RMSF values; these changes may be attributed to the continuous movements of the control compound in the active pocket of the main protease enzyme, which results in pressure on the flexible regions. Next, the compact nature of the receptor enzyme in the presence of the inhibitors was evaluated. The mean Rg values of BBB-26580140, BDE-32007849, LAS-51378804, and control were 38.52 Å, 37.02 Å, 37.34 Å, and 40.25 Å, respectively (Figure 3C); the Rg values show the same trend as the RMSD and RMSF analyses.

3.4 Estimating the binding free energies

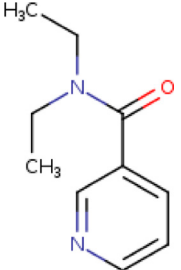
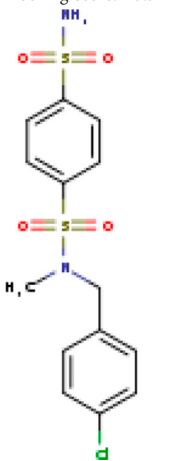
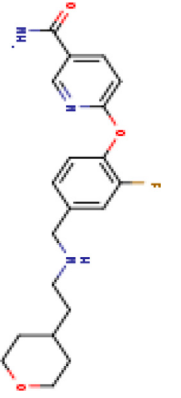
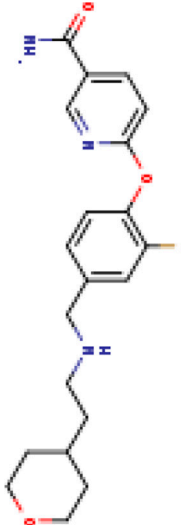
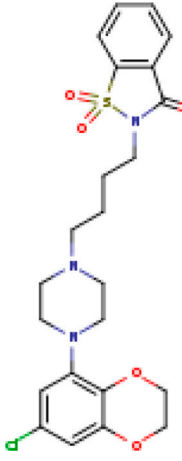
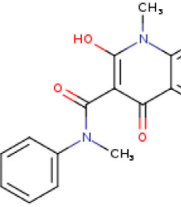
The overall objectives of the MM-PBSA and molecular mechanics with generalized Born and surface area (MM-GBSA) are to determine the differences in free energies between the bound

TABLE 1 Structure-based compounds similar to those of the leads.

Lead/Similarity-based scaffolds	BBB-26580140	BDE-32007849	LAS-51378804
Parent structure			
	<p>DB06397, Nicaraven Score: 0.540 Docking score: -8.21</p> 	<p>DB09355, Sulfabenzamide Score: 0.642 Docking score: -7.52</p> 	<p>DB15450, PF-05105679 Score: 0.542 Docking score: -6.34</p> 
	<p>DB12929, JNJ-39220675 Score: 0.536 Docking score: -9.62</p> 	<p>DB12418, Saccharin Score: 0.607 Docking score: -6.35</p> 	<p>DB12929, JNJ-39220675 Score: 0.539 Docking score: -7.54</p> 

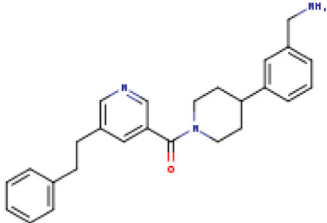
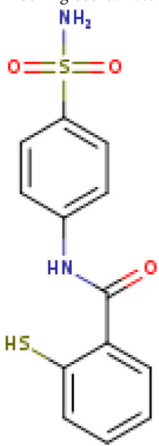
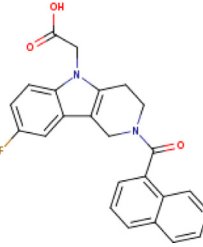
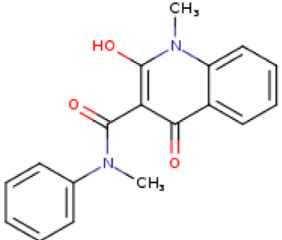
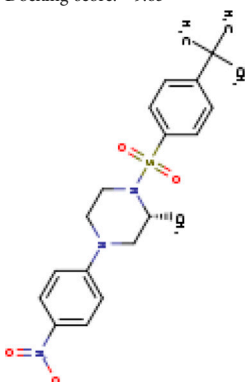
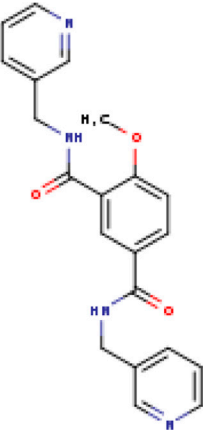
(Continued on following page)

TABLE 1 (Continued) Structure-based compounds similar to those of the leads.

Lead/Similarity-based scaffolds	BBB-26580140	BDE-32007849	LAS-51378804
	<p>DB13655, Nikethamide Score: 0.524 Docking score: -7.54</p> 	<p>DB07115, N-(4-chlorobenzyl)-N-methylbenzene-1 Score: 0.575 Docking score: -5.32</p> 	<p>DB12585, Ondelopran Score: 0.535 Docking score: -9.64</p> 
	<p>DB12585, Ondelopran Score: 0.523 Docking score: -10.25</p> 	<p>DB16040, DU125530 Score: 0.536 Docking score: -9.64</p> 	<p>DB11366, Roquinimex Score: 0.527 Docking score: -6.98</p> 

(Continued on following page)


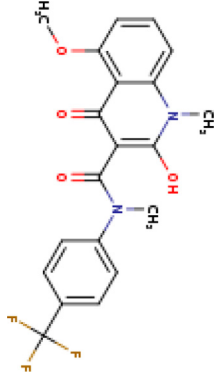
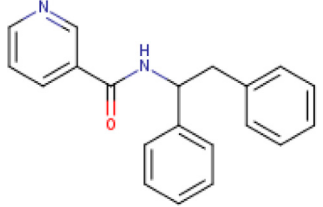
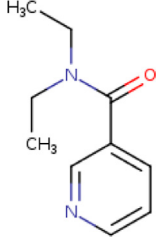
TABLE 1 (Continued) Structure-based compounds similar to those of the leads.

Lead/Similarity-based scaffolds	BBB-26580140	BDE-32007849	LAS-51378804
	<p>DB04764, [4-(3-aminomethyl-phenyl)-piperidin-1-yl]-(5-phenethyl- pyridin-3-yl)-methanone Score: 0.522 Docking score: -6.54</p> 	<p>DB07476, n-[4-(aminosulfonyl)phenyl]-2-mercaptobenzamide Score: 0.484 Docking score: -7.84</p> 	<p>DB12562, Setipiprant Score: 0.509 Docking score: -4.45</p> 
	<p>DB11366, Roquinimex Score: 0.503 Docking score: -10.065</p> 	<p>DB07049, (2R)-1-[(4-tert-butylphenyl)sulfonyl]-2-methyl-4-(4-nitrophenyl)piperazine Score: 0.482 Docking score: -9.65</p> 	<p>DB13327, Picotamide Score: 0.508 Docking score: -8.01</p> 

(Continued on following page)

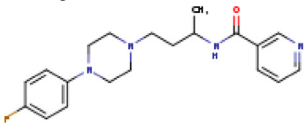
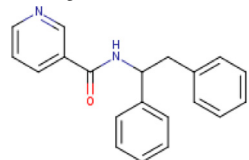
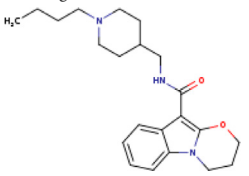
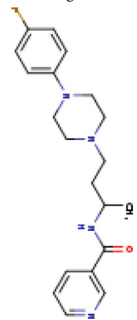
and

TABLE 1 (Continued) Structure-based compounds similar to those of the leads.

Lead/Similarity-based scaffolds	BBB-26580140	BDE-32007849	LAS-51378804
	<p>DB14805, DCFPyL F-18 Score: 0.500 Docking score: -8.65</p> 	Not predicted	<p>DB05861, Tasquinimod Score: 0.506 Docking score: -9.64</p> 
	<p>DB13531, Nicofetamide Score: 0.492 Docking score: -8.21</p> 	Not predicted	<p>DB13655, Nikethamide Score: 0.500 Docking score: -6.58</p> 

(Continued on following page)

TABLE 1 (Continued) Structure-based compounds similar to those of the leads.

Lead/Similarity-based scaffolds	BBB-26580140	BDE-32007849	LAS-51378804
	<p>DB13687, Niaprazine Score: 0.489 Docking score: -8.64</p> 	Not predicted	<p>DB13531, Nicofetamide Score: 0.496 Docking score: -9.64</p> 
	<p>DB04873, Piboserod Score: 0.488 Docking score: -8.54</p> 	Not predicted	<p>DB13687, Niaprazine Score: 0.493 Docking score: -8.64</p> 

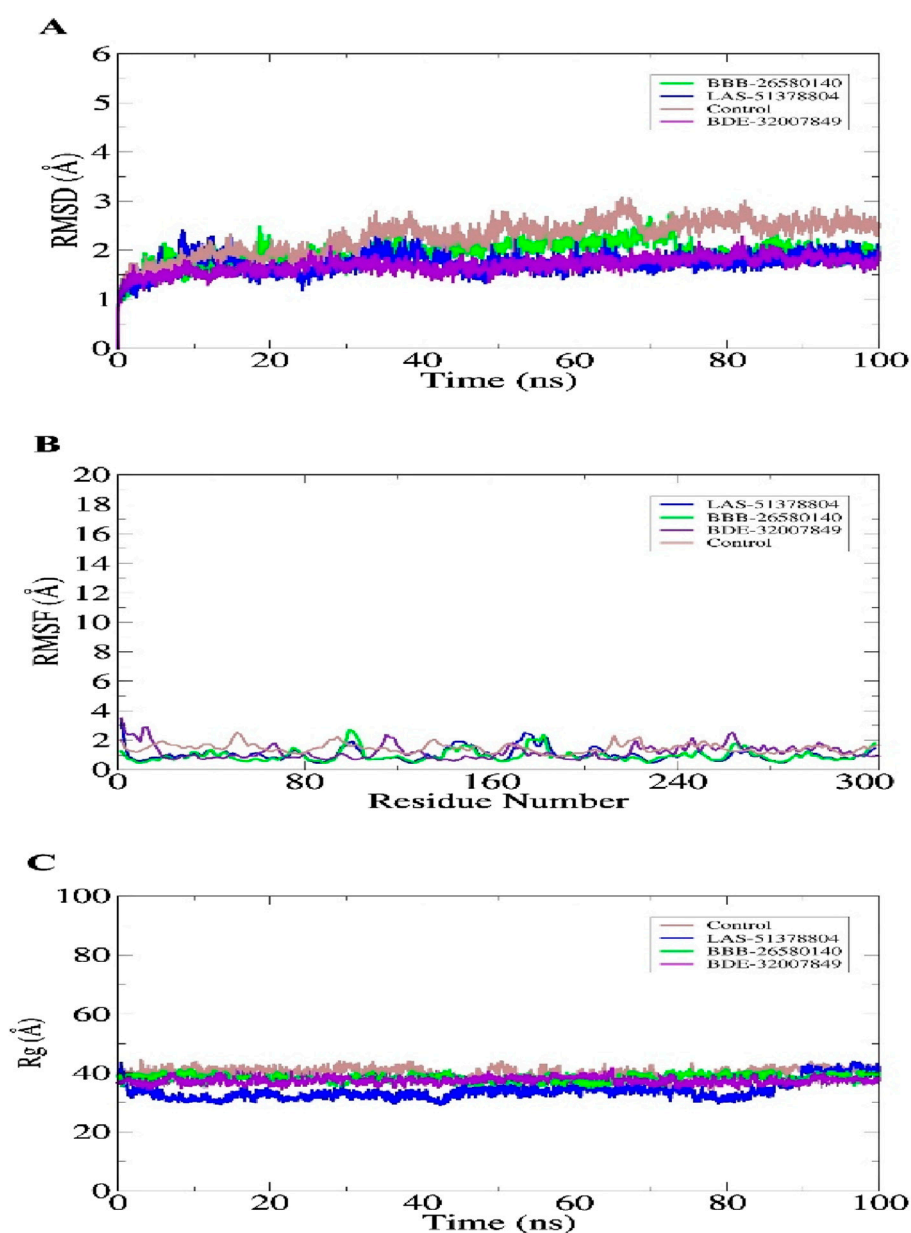


FIGURE 3
Statistical evaluations performed on the trajectories of the molecular dynamics simulations. The metrics included (A) RMSD, (B) RMSF, and (C) Rg. All units are in angstroms.

unbound states of the molecules in an aqueous solution and compare different conformations of the molecules (Genheden and Ryde, 2015; Wang et al., 2019). The MM-PBSA and MM-GBSA methods are now regularly applied in drug discovery pipelines as they offer more accurate information than the static intermolecular conformation docking studies (Tuccinardi, 2021). The MM-PBSA requires modest computational resources and is based on diverse molecular dynamics simulation frames (Zhang et al., 2017). The details regarding the MM-PBSA and MM-GBSA binding energies of the complexes are tabulated in Table 2. The values indicate that all three complexes and the control have robust binding energies and that the complexes are considerably stable in terms of the binding and docked

conformations. Between the MM-PBSA and MM-GBSA energies, the van der Waals energy demonstrated the most favorable contribution to the equilibrium of the complexes. The net van der Waals energy contributions of BBB-26580140, BDE-32007849, LAS-51378804, and control were found to be -50.36 , -52.01 , -50.61 , and -49.51 kcal/mol, respectively. The electrostatic energy showed the second most favorable contributions to the binding of ligands with the main protease enzyme. The net electrostatic energy contributions to the overall stabilization of the complexes are -16.34 , -17.38 , -15.00 , and -13.69 kcal/mol for BBB-26580140, BDE-32007849, LAS-51378804, and control, respectively. The overall solvation energies of the complexes

TABLE 2 Free energy estimates of the docked/control complexes. The values are statistically calculated in terms of kcal/mol.

Energy parameter	BBB-26580140	BDE-32007849	LAS-51378804	Control
MM-GBSA				
VDWAALS	-50.36	-52.01	-50.61	-49.51
EEL	-16.34	-17.38	-15.00	-13.69
DELTA G gas	-66.7	-69.39	-65.61	-63.2
DELTA G solv	13.68	12.54	10.17	14.29
DELTA TOTAL	-53.02	-56.85	-55.44	-48.91
MM-PBSA				
VDWAALS	-50.36	-52.01	-50.61	-49.51
EEL	-16.34	-17.38	-15.00	-13.69
DELTA G gas	-66.7	-69.39	-65.61	-63.2
DELTA G solv	13.10	11.78	11.20	13.11
DELTA TOTAL	-53.6	-57.61	-54.41	-50.09

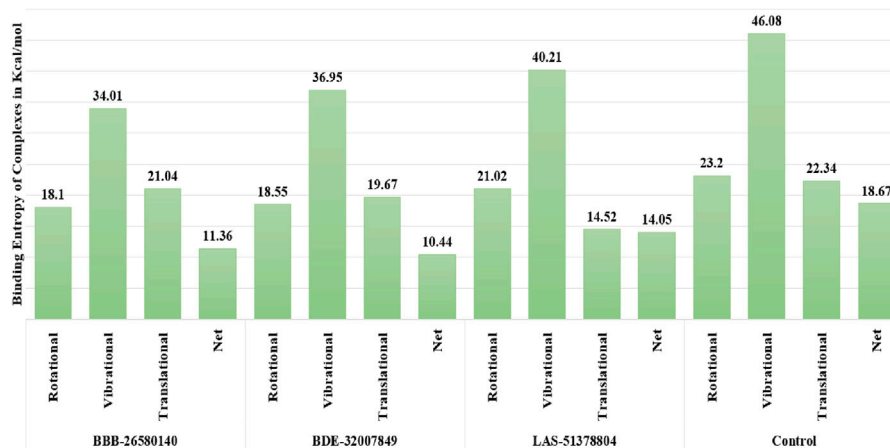


FIGURE 4 Binding entropy energy contribution of each complex. The total energy is split into the rotational, vibrational, and translational energy components.

were found to have non-favorable contributions, and most of these contributions were due to the polar solvation energies. Non-polar solvation energy seems to be vital for stabilization of the complexes. The net solvation energies of BBB-26580140, BDE-32007849, LAS-51378804, and control were found to be 13.68, 12.54, 10.17, and 14.29 kcal/mol for the MM-GBSA as well as 13.10, -57.61, -54.41, and -50.09 kcal/mol for the MM-PBSA, respectively. The overall MM-GBSA binding energies of BBB-26580140, BDE-32007849, LAS-51378804, and control were -53.02, -56.85, -55.44, and -48.91 kcal/mol, respectively. Similarly, the net MM-PBSA binding energies of BBB-26580140, BDE-32007849, LAS-51378804, and control were -53.6, -57.61, -54.41, and -50.09 kcal/mol, respectively.

3.5 Predictions of the entropy energies of the complexes

The random energy possessed by each complex was determined by entropy analysis. The random free energy enables the ligands to escape the enzyme active pocket and detach from it. It was found that the identified systems had lower free energies. The net entropy energies of BBB-26580140, BDE-32007849, LAS-51378804, and control were 11.36, 10.44, 14.05, and 18.67 kcal/mol, respectively. This additionally confirms that these systems have high intermolecular stabilities and good ligand binding strengths with the protease enzyme. The binding entropy energy findings of the complexes are reported in Figure 4.



FIGURE 5
Water-swap-based binding free energies of the complexes in units of kcal/mol.

3.6 Water swap energies of the complexes

Next, the water swap binding energies of the complexes were determined to validate and crosscheck the system stabilities. The water swap method is considered to be very robust in terms of confirming the binding affinities of the compounds to the main protease enzyme. Water molecules play significant roles in ligand binding with the receptor enzyme residues (Woods et al., 2014). The water swap approach is particularly useful for investigating the roles of the water molecules in bridging the ligands with the enzyme active site residues (Woods et al., 2011). The water swap method estimates the absolute binding free energy in terms of three algorithms, namely thermodynamic integration (TI), free energy perturbation (FEP), and Bennett's methods. Among the compounds, BBB-26580140 was found to have the most dominant and stable binding energy. The FEP, TI, and Bennett's energy values of the compounds were found to be -57.1 , -58.67 , and -58.6 kcal/mol, respectively. The details of the water swap energies of the complexes are given in Figure 5.

3.7 Predictions about drug-likeness and medicinal chemistry of the compounds

Before conducting experimental evaluations, predictions about the drug-likeness of the selected compounds were used to determine whether they could pass the prominent drug-like rules, such as the Lipinski rule of five (Lipinski, 2004), as well as the Veber (Veber et al., 2002), Egan (Egan et al., 2000), and Muegge (Ahmad et al., 2017) rules. Drug-likeness increases the chances that a compound would obtain clearance from the clinical trials for further marketing. Drug-likeness predictions allow us to compare and contrast the selected leads with known drug-like rules. The drug-like molecules also have improved oral bioavailabilities and can be transformed into successful drug molecules (Protti et al., 2021). All three leads and the control molecule were found to show favorable drug-like properties and achieve clearance based on the important drug-like rules, as summarized in Table 3. Similarly, the compounds were found to have good bioavailability scores,

meaning that high concentrations of the compounds can reach the target sites for biological actions. The compounds were also disclosed to show no alerts for PAINS, indicating that they may only bind to one specific biological macromolecule and not multiple ones (Whitty, 2011). Thus, there are fewer chances of obtaining false positive results. All the compounds were also found to have good synthetic accessibility scores, where lower scores imply that the compounds can be easily synthesized in the laboratory for biological activities.

3.8 Computation of the pharmacokinetic properties

To predict the bodily behaviors upon administration of the selected compounds, the pharmacokinetic properties of the inhibitors were determined (Lombardo et al., 2017). All the lead compounds and the control molecule showed very high gastrointestinal absorption, resulting in the maintenance of significant amounts at the target site for biological activity (Alex, 2003). All compounds were non-permeable to the blood-brain barrier, so the compounds are not expected to have any neurotoxic side effects (De Boer and Breimer, 1994). The compounds were non-substrates of P-glycoprotein, so they will not be expelled from the cells and will be available for biological functions (Wang et al., 2011). The compounds were also found to be non-inhibitors of the cytochrome isoforms (Cheng et al., 2011). Therefore, the compounds can be easily metabolized and excreted from the body after fulfilling their pharmacological actions. Table 4 lists the pharmacokinetic properties of the selected inhibitors and control.

3.9 Toxicity predictions

The toxicity features of compounds significantly impact their success as drugs. Hence, the toxicities of the selected compounds were evaluated. All compounds were found to be non-toxic, non-immunogenic, and non-mutagenic.

TABLE 3 Analysis of the medicinal chemistry properties and drug-likeness of the shortlisted inhibitors.

Drug-like rule	BBB-26580140	BDE-32007849	LAS-51378804	Control
Lipinski rule of five	Follow	Follow	Follow	Follow
Egan	Follow	Follow	Follow	Follow
Veber	Follow	Follow	Follow	Follow
Muegge	Follow	Follow	Follow	Follow
Bioavailability score	0.56	0.55	0.56	0.55
Pan-assay interference compounds	No alert	No alert	No alert	No alert
Synthetic accessibility	2.23	3.02	2.26	3.11

TABLE 4 Pharmacokinetic properties of the inhibitory compounds.

Property	BBB-26580140	BDE-32007849	LAS-51378804	Control
Drug absorption form the gastrointestinal tract	Maximum	Maximum	Maximum	Maximum
Penetration of blood–brain barrier	Not detected	Not detected	Not detected	Not detected
Compounds as P-gp substrates	Not detected	Not detected	Not detected	Not detected
Compounds as CYP1A2 receptor inhibitors	Not detected	Not detected	Not detected	Not detected
Compounds as CCYP2C19 receptor inhibitors	Not detected	Not detected	Not detected	Detected
Compounds as CYP2C9 receptor inhibitors	Not detected	Not detected	Not detected	Not detected
Compounds as CYP2D6 receptor inhibitors	Not detected	Not detected	Detected	Not detected
Compounds as CYP3A4 receptor inhibitors	Not detected	Not detected	Not detected	Not detected
Skin permeation (Log K_p)	-7.06 cm/s	-6.49 cm/s	-7.07 cm/s	-7.51 cm/s

TABLE 5 Predicted Mol inspiration bioactivities of the lead/control molecules.

Compound/control	Ligand for GPCR	Modulator of ion channel	Inhibitor of kinase	Ligand for nuclear receptor	Inhibitor of protease	Inhibitor of enzyme
BBB-26580140	0.04	0.09	-0.59	1.01	0.11	0.27
BDE-32007849	0.17	0.14	-0.15	0.15	0.16	0.23
LAS-51378804	0.00	0.09	-0.35	0.66	0.11	0.22
Control	0.13	0.06	-0.22	0.25	-0.005	0.15

3.10 Bioactivity predictions

The possible bioactivities of the leads and control were predicted (Table 5). The molecules with bioactivity scores >0.0 are considered to have considerable biological activities, while those with scores in the range of -0.50 to 0.00 are classified as modestly active and those with values <-0.50 are regarded as inactive (Khan et al., 2017). The findings presented in the table show that the compounds are biologically active and may exert physiological activity by binding to and interacting with the nuclear receptor ligands, protease enzyme, and GPCR ligands. Moreover, the bioactivities of the lead structures were predicted to be more pronounced than that of the control molecule. Among the leads, BDE-32007849 was found to have the most significant biological activities against the target enzymes.

4 Discussion

The COVID-19 pandemic that started in China and spread globally has resulted in the deaths of millions and left many with health disabilities (Cecconi et al., 2020). Although the disease is currently well-managed with extensive use of effective vaccines, there is an ongoing search for potent drugs (Chukwudozie et al., 2021). In this computational study, diverse drug library sources were screened in an attempt to identify compounds that show the best binding capabilities with the main protease enzyme of SARS-COV-2, the causative agent of the COVID-19 pandemic. As a result, three drug molecules (BBB-26580140, BDE-32007849, and LAS-51378804) were found, which showed highly stable conformations with the enzyme at the S1 pocket. The docking binding energy scores of BBB-26580140, BDE-32007849, and

LAS-51378804 were -13.02 , -13.0 , and -12.56 kcal/mol, respectively. These energy scores determine the compounds as lead structures with networks of strong intermolecular interactions. In contrast to the control molecule Z1741970824 (binding energy score: -11.59 kcal/mol), the three compounds showed robust hydrophilic and hydrophobic interactions. The interactions most important to the overall stabilities of the complexes involve the enzyme active site residues, such as His41 and Cys145. As the molecular-docking-based virtual screening often results in false positive findings, the predicted stabilities of the best molecular bindings with the main protease enzyme of SARS-CoV-2 were validated through molecular dynamics simulations. It was observed that the enzyme structures were stable in the presence of the shortlisted leads and that the networks of binding interactions were intact. The net MM-GBSA binding energies of BBB-26580140, BDE-32007849, LAS-51378804, and control were -53.02 , -56.85 , -55.44 , and -48.91 kcal/mol, respectively; the net MM-PBSA binding energies of BBB-26580140, BDE-32007849, LAS-51378804, and control were -53.6 , -57.61 , -54.41 , and -50.09 kcal/mol, respectively. These energies demonstrate the formation of strong intermolecular complexes. Moreover, the predicted binding entropy energies of the systems showed the presence of lower free energies, indicating that the systems had stable nature. From the drug-likeness perspective, the compounds were found to be non-toxic while fulfilling the important drug-like rules. The compounds also revealed good ADMET properties (Jia et al., 2019). Thus, the predicted compounds are promising lead structures that can be utilized in further *in vitro* and *in vivo* studies.

5 Conclusion

In this study, three drug molecules (BBB-26580140, BDE-32007849, and LAS-51378804) were identified as promising binders of the main protease enzyme of SARS-CoV-2 after extensive structure-based virtual screening of four libraries (Asinex Antiviral, Seaweed Metabolite Database, MFSMTL, and CMNPD with 6,827, 1,191, 1,830, and 45,000 compounds, respectively). All three compounds showed highly stable binding conformations at the S1 active pocket of the main protease enzyme and formed short-distance hydrophilic and hydrophobic contacts with catalytically active residues. The stable intermolecular conformations of the docked complexes were validated through molecular dynamics simulations, which showed minor local loop-mediated deviations and the absence of global changes. It was also noted that the enzyme was tolerant to the presence of the compounds when they occupied the active pocket. The compounds also revealed stable electrostatic and van der Waals energies as well as lower polar solvation energies. Additionally, the

compounds were proven to be drug-like and showed favorable medicinal chemistry properties along with non-toxic, non-mutagenic, and non-carcinogenic features. The compounds were found to show bioactivities against nuclear receptor ligands, protease enzymes, GPCRs, etc. The bottom line of this study is that the identified compounds are promising and may be used in further experimental tests to validate their real biological potentials.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material; further inquiries can be directed to the corresponding author.

Author contributions

SA: writing—original draft.

Funding

The author declares that financial support was received for the research, authorship, and/or publication of this article. This study is supported via funding from Prince sattam bin Abdulaziz University project number (PSAU/2024/R/1446).

Acknowledgments

This study is supported via funding from Prince sattam bin Abdulaziz University project number (PSAU/2024/R/1446).

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Agost-Beltrán, L., de la Hoz-Rodríguez, S., Bou-Iserte, L., Rodríguez, S., Fernández-de-la-Pradilla, A., and González, F. V. (2022). Advances in the development of SARS-CoV-2 mpro inhibitors. *Molecules* 27 (8), 2523. doi:10.3390/molecules27082523
- Ahmad, S., Raza, S., Abro, A., Liedl, K. R., and Azam, S. S. (2019). Toward novel inhibitors against KdsB: a highly specific and selective broad-spectrum bacterial enzyme. *J. Biomol. Struct. Dyn.* 37 (5), 1326–1345. doi:10.1080/07391102.2018.1459318
- Ahmad, S., Raza, S., Uddin, R., and Azam, S. S. (2017). Binding mode analysis, dynamic simulation, and binding free energy calculations of the MurF ligase from *Acinetobacter baumannii*. *J. Mol. Graph. Model.* 77, 72–85. doi:10.1016/j.jmglm.2017.07.024

- Alex, A. (2003). *Absorption and drug development*. Hoboken, New Jersey, USA: John Wiley and Sons, Inc.
- Athanasiadis, E., Cournia, Z., and Spyrou, G. (2012). ChemBioServer: a web-based pipeline for filtering, clustering and visualization of chemical compounds used in drug discovery. *Bioinformatics* 28 (22), 3002–3003. doi:10.1093/bioinformatics/bts551
- Bergström, C. A. S., and Larsson, P. (2018). Computational prediction of drug solubility in water-based systems: qualitative and quantitative approaches used in the current drug discovery and development setting. *Int. J. Pharm.* 540 (1–2), 185–193. doi:10.1016/j.ijpharm.2018.01.044
- Biovia, D. S. (2017). *Discovery studio visualizer*. San Diego, CA, USA.
- Calligari, P., Bobone, S., Ricci, G., and Bocedi, A. (2020). Molecular investigation of SARS-CoV-2 proteins and their interactions with antiviral drugs. *Viruses* 12 (4), 445. doi:10.3390/v12040445
- Carabelli, A. M., Peacock, T. P., Thorne, L. G., Harvey, W. T., Hughes, J., de Silva Thushan, I., et al. (2023). SARS-CoV-2 variant biology: immune escape, transmission and fitness. *Nat. Rev. Microbiol.* 21 (3), 162–177. doi:10.1038/s41579-022-00841-7
- Case, D. A., Duke, R. E., Walker, R. C., Skrynnikov, N. R., Cheatham III, T. E., Mikhailovskii, O., et al. (2022). AMBER 22 reference manual. Available at: <https://orbilu.uni.lu/handle/10993/52805>
- Cecconi, M., Formi, G., and Mantovani, A. (2020). Ten things we learned about COVID-19. *Intensive Care Med.* 46 (8), 1590–1593. doi:10.1007/s00134-020-06140-0
- Cheng, F., Yu, Y., Shen, J., Yang, L., Li, W., Liu, G., et al. (2011). Classification of cytochrome P450 inhibitors and noninhibitors using combined classifiers. *J. Chem. Inf. Model.* 51 (5), 996–1011. doi:10.1021/ci200028n
- Chopra, B., and Dhingra, A. K. (2021). Natural products: a lead for drug discovery and development. *Phytotherapy Res.* 35 (9), 4660–4702. doi:10.1002/ptr.7099
- Chukwudozie, O. S., Duru, V. C., Ndiribe, C. C., Aborode, A. T., Oyeibanji, V. O., and Emikpe, B. O. (2021). The relevance of bioinformatics applications in the discovery of vaccine candidates and potential drugs for COVID-19 treatment. *Bioinforma. Biol. Insights* 15, 11779322211002168. doi:10.1177/11779322211002168
- Dai, W., Zhang, B., Jiang, X.-M., Su, H., Li, J., Zhao, Y., et al. (2020). Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* 368 (6497), 1331–1335. doi:10.1126/science.abb4489
- Daina, A., Michielin, O., and Zoete, V. (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* 7 (January), 42717–42813. doi:10.1038/srep42717
- Dallakyan, S., and Olson, A. J. (2015). Small-molecule library screening by docking with PyRx. *Chem. Biol. Methods Protoc.* 1263, 243–250. doi:10.1007/978-1-4939-2269-7_19
- Davis, G. D. J., and Vasanthi, A. H. R. (2011). Seaweed metabolite database (SWMD): a database of natural compounds from marine algae. *Bioinformation* 5 (8), 361–364. doi:10.6026/97320630005361
- De Boer, A. G., and Breimer, D. D. (1994). The blood-brain barrier: clinical implications for drug delivery to the brain. *J. R. Coll. Physicians Lond.* 28 (6), 502–506. Available at: <https://pmc.ncbi.nlm.nih.gov/articles/PMC5401103/>
- Doungamath, A., Fearon, D., Gehrtz, P., Krojer, T., Lukacik, P., Owen, C. D., et al. (2020). Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. *Nat. Commun.* 11 (1), 5047. doi:10.1038/s41467-020-18709-w
- Dzobo, K. (2022). The role of natural products as sources of therapeutic agents for innovative drug discovery. *Compr. Pharmacol.* 408, 408–422. doi:10.1016/b978-0-12-820472-6.00041-4
- Egan, W. J., Merz, K. M., and Baldwin, J. J. (2000). Prediction of drug absorption using multivariate statistics. *J. Med. Chem.* 43 (21), 3867–3877. doi:10.1021/jm000292e
- Elsässer, B., Zauner, F. B., Messner, J., Soh, W. T., Dall, E., and Brandstetter, H. (2017). Distinct roles of catalytic cysteine and histidine in the protease and ligase mechanisms of human legumain as revealed by DFT-based QM/MM simulations. *ACS Catal.* 7 (9), 5585–5593. doi:10.1021/acscatal.7b01505
- Excler, J. L., Saville, M., Berkley, S., and Kim, J. H. (2021). Vaccine development for emerging infectious diseases. *Nat. Med.* 27 (4), 591–600. doi:10.1038/s41591-021-01301-0
- Gediya, L. K., and Njar, V. C. O. (2009). Promise and challenges in drug discovery and development of hybrid anticancer drugs. *Expert Opin. Drug Discov.* 4 (11), 1099–1111. doi:10.1517/17460440903341705
- Genheden, S., Kuhn, O., Mikulskis, P., Hoffmann, D., and Ryde, U. (2012). The normal-mode entropy in the MM/GBSA method: effect of system truncation, buffer region, and dielectric constant. *J. Chem. Inf. Model.* 52 (8), 2079–2088. doi:10.1021/ci3001919
- Genheden, S., and Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* 10 (5), 449–461. doi:10.1517/17460441.2015.1032936
- Halgren, T. A. (1996). Merck molecular force field. *J. Comput. Chem.* 17 (5–6), 490–519. doi:10.1002/(SICI)1096-987X(199604)17:5/6<520::AID-JCC2>3.0.CO;2-W
- Izguirre, J. A., Catarello, D. P., Wozniak, J. M., and Skeel, R. D. (2001). Langevin stabilization of molecular dynamics. *J. Chem. Phys.* 114 (5), 2090–2098. doi:10.1063/1.1332996
- Jia, C.-Y., Li, J.-Y., Hao, G.-F., and Yang, G.-F. (2019). A drug-likeness toolbox facilitates ADMET study in drug discovery. *Drug Discov. Today* 25, 248–258. doi:10.1016/j.drudis.2019.10.014
- Kaliappan, S., and Bombay, I. I. T. (2018). UCSF chimera-overview.
- Khan, A., Zia, T., Suleman, M., Khan, T., Ali, S. S., Abbasi, A. A., et al. (2021). Higher infectivity of the SARS-CoV-2 new variants is associated with K417N/T, E484K, and N501Y mutants: an insight from structural data. *J. Cell. Physiology* 236 (10), 7045–7057. doi:10.1002/jcp.30367
- Khan, T., Dixit, S., Ahmad, R., Raza, S., Azad, I., Joshi, S., et al. (2017). Molecular docking, PASS analysis, bioactivity score prediction, synthesis, characterization and biological activity evaluation of a functionalized 2-butanone thiosemicarbazone ligand and its complexes. *J. Chem. Biol.* 10, 91–104. doi:10.1007/s12154-017-0167-y
- Kim, D., Lee, J.-Y., Yang, J.-S., Kim, J. W., Kim, V. N., and Chang, H. (2020). The architecture of SARS-CoV-2 transcriptome. *Cell* 181, 914–921.e10. doi:10.1016/j.cell.2020.04.011
- Korber, B., Fischer, W. M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., et al. (2020). Spike mutation pipeline reveals the emergence of a more transmissible form of SARS-CoV-2. *bioRxiv*. doi:10.1101/2020.04.29.069054
- Kräutler, V., Van Gunsteren, W. F., and Hünenberger, P. H. (2001). A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J. Comput. Chem.* 22 (5), 501–508. doi:10.1002/1096-987x(20010415)22:5<501::aid-jcc1021>3.0.co;2-v
- Kuzmanic, A., and Zagrovic, B. (2010). Determination of ensemble-average pairwise root mean-square deviation from experimental B-factors. *Biophysical J.* 98 (5), 861–871. doi:10.1016/j.bpj.2009.11.011
- Lipinski, C. A. (2004). Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov. Today Technol.* 1 (4), 337–341. doi:10.1016/j.ddtec.2004.11.007
- Lobanov, M. Y., Bogatyreva, N. S., and Galzitskaya, O. V. (2008). Radius of gyration as an indicator of protein structure compactness. *Mol. Biol.* 42 (4), 623–628. doi:10.1134/s0026893308040195
- Lombardo, F., Desai, P. V., Arimoto, R., Desino, K. E., Fischer, H., Keefer, C. E., et al. (2017). *In silico* absorption, distribution, metabolism, excretion, and pharmacokinetics (ADME-PK): utility and best practices. An industry perspective from the international consortium for innovation through quality in pharmaceutical development. *J. Med. Chem.* 60 (22), 9097–9113. doi:10.1021/acs.jmedchem.7b00487
- Mahram, A., and Herboldt, M. C. (2015). NCBI BLASTP on high-performance reconfigurable computing systems. *ACM Trans. Reconfigurable Technol. Syst. (TRETS)* 7 (4), 1–20. doi:10.1145/2629691
- Maierov, V. N., and Crippen, G. M. (1994). Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins. *J. Mol. Biol.* 235, 625–634. doi:10.1006/jmbi.1994.1017
- Maltezos, H. C., and Papa, A. (2011). Crimean-Congo hemorrhagic fever: epidemiological trends and controversies in treatment. *BMC Med.* 9, 131. doi:10.1186/1741-7015-9-131
- Miller III, B. R., McGee Jr, T. D., Swails, J. M., Homeyer, N., Gohlke, H., and Roitberg, A. E. (2012). MMPBSA.py: an efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* 8 (9), 3314–3321. doi:10.1021/ct300418h
- Mousavi, S., Zare, S., Mirzaei, M., and Feizi, A. (2022). Novel drug design for treatment of COVID-19: a systematic review of preclinical studies. *Can. J. Infect. Dis. Med. Microbiol.* 2022, 2044282. doi:10.1155/2022/2044282
- Msemburi, W., Karlinsky, A., Knutson, V., Aleshin-Guendel, S., Chatterji, S., and Wakefield, J. (2023). The WHO estimates of excess mortality associated with the COVID-19 pandemic. *Nature* 613 (7942), 130–137. doi:10.1038/s41586-022-05522-2
- Narayanan, A., Narwal, M., Majowicz, S. A., Varricchio, C., Toner, S. A., Ballatore, C., et al. (2022). Identification of SARS-CoV-2 inhibitors targeting Mpro and PLpro using in-cell-protease assay. *Commun. Biol.* 5 (1), 169–217. doi:10.1038/s42003-022-03090-9
- Omolo, C. A., Soni, N., Fasiku, V. O., Mackraj, I., and Govender, T. (2020). Update on therapeutic approaches and emerging therapies for SARS-CoV-2 virus. *Eur. J. Pharmacol.* 883, 173348. doi:10.1016/j.ejphar.2020.173348
- Pang, X., Xu, W., Liu, Y., Li, H., and Chen, L. (2023). The research progress of SARS-CoV-2 main protease inhibitors from 2020 to 2022. *Eur. J. Med. Chem.* 257, 115491. doi:10.1016/j.ejmech.2023.115491
- Petersen, H. G. (1995). Accuracy and efficiency of the particle mesh Ewald method. *J. Chem. Phys.* 103 (9), 3668–3679. doi:10.1063/1.470043
- Pires, D. E. V., Blundell, T. L., and Ascher, D. B. (2015). pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J. Med. Chem.* 58 (9), 4066–4072. doi:10.1021/acs.jmedchem.5b00104
- Protti, Í. F., Rodrigues, D. R., Fonseca, S. K., Alves, R. J., de Oliveira, R. B., and Maltarollo, V. G. (2021). Do drug-likeness rules apply to oral prodrugs? *ChemMedChem* 16 (9), 1446–1456. doi:10.1002/cmdc.202000805
- Rafi, S., Yasmin, S., and Uddin, R. (2020). A molecular dynamic simulation approach: development of dengue virus vaccine by affinity improvement techniques. *J. Biomol. Struct. Dyn.* 40, 61–76. doi:10.1080/07391102.2020.1811771

- Roe, D. R., and Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* 9 (7), 3084–3095. doi:10.1021/ct400341p
- Schafmeister, C., Ross, W. S., and Romanovski, V. (1995). *The leap module of AMBER*. University of California.
- Sengupta, A., Li, Z., Song, L. F., Li, P., and Merz, Jr. K. M. (2021). Parameterization of monovalent ions for the OPC3, OPC, TIP3P-FB, and TIP4P-FB water models. *J. Chem. Inf. Model.* 61 (2), 869–880. doi:10.1021/acs.jcim.0c01390
- Shaker, B., Ahmad, S., Lee, J., Jung, C., and Na, D. (2021). *In silico* methods and tools for drug discovery. *Comput. Biol. Med.* 137, 104851. doi:10.1016/j.compbiomed.2021.104851
- Sheokand, D., Grewal, A., Kumar, V., Chauhan, R., Saini, V., and Kumar, A. (2024). *In silico* evaluation of marine phytochemicals as potential inhibitors of inhibitor-kappa B protein kinase (IkK) for alzheimer's disease therapeutics. *Biosci. Biotechnol. Res. Asia* 21 (3), 1063–1076. doi:10.13005/bbra/3286
- Tian, C., Kasavajhala, K., Belfon, K. A. A., Raguette, L., Huang, H., Miguez, A. N., et al. (2019). ff19SB: amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution. *J. Chem. Theory Comput.* 16 (1), 528–552. doi:10.1021/acs.jctc.9b00591
- Tuccinardi, T. (2021). What is the current value of MM/PBSA and MM/GBSA methods in drug discovery? *Expert. Opin. Drug. Discov.*, 1233–1237. Available at: <https://pubs.acs.org/doi/abs/10.1021/acs.jcim.6b00174>
- Turner, P. J. (2005). *Xmgrace*. Beaverton, OR: Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology. Version 5.1. 19.
- Van De Waterbeemd, H., and Gifford, E. (2003). ADMET *in silico* modelling: towards prediction paradise? *Nat. Rev. Drug Discov.* 2 (3), 192–204. doi:10.1038/nrd1032
- Van Drie, J. H. (2007). Computer-aided drug design: the next 20 years. *J. Computer-Aided Mol. Des.* 21 (10), 591–601. doi:10.1007/s10822-007-9142-y
- Vanduyck, K., and Deval, J. (2021). Considerations for the discovery and development of 3-chymotrypsin-like cysteine protease inhibitors targeting SARS-CoV-2 infection. *Curr. Opin. Virology* 49, 36–40. doi:10.1016/j.coviro.2021.04.006
- Vassetzki, D., Pagliari, M., and Procacci, P. (2019). Assessment of GAFF2 and OPLS-AA general force fields in combination with the water models TIP3P, SPCE, and OPC3 for the solvation free energy of druglike organic molecules. *J. Chem. Theory Comput.* 15 (3), 1983–1995. doi:10.1021/acs.jctc.8b01039
- Veber, D. F., Johnson, S. R., Cheng, H.-Y., Smith, B. R., Ward, K. W., and Kopple, K. D. (2002). Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* 45 (12), 2615–2623. doi:10.1021/jm020017n
- Viana, R., Moyo, S., Amoako, D. G., Tegally, H., Scheepers, C., Althaus, C. L., et al. (2022). Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature* 603 (7902), 679–686. doi:10.1038/s41586-022-04411-y
- Vivek-Ananth, R. P., Sahoo, A. K., Kumaravel, K., Mohanraj, K., and Samal, A. (2021). MeFSAT: a curated natural product database specific to secondary metabolites of medicinal fungi. *Rsc Adv.* 11 (5), 2596–2607. doi:10.1039/d0ra10322e
- Wang, E., Sun, H., Wang, J., Wang, Z., Liu, H., Zhang, J. Z. H., et al. (2019). End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. *Chem. Rev.* 119 (16), 9478–9508. doi:10.1021/acs.chemrev.9b00055
- Wang, J., Wang, W., Kollman, P. A., and Case, D. A. (2001). Antechamber: an accessory software package for molecular mechanical calculations. *J. Am. Chem. Soc.* 123, U403.
- Wang, Z., Chen, Y., Liang, H., Bender, A., Glen, R. C., and Yan, A. (2011). P-glycoprotein substrate models using support vector machines based on a comprehensive data set. *J. Chem. Inf. Model.* 51 (6), 1447–1456. doi:10.1021/ci2001583
- Whitty, A. (2011). Growing PAINS in academic drug discovery. *Future Med. Chem.* 3 (7), 797–801. doi:10.4155/fmc.11.44
- Woods, C. J., Malaisree, M., Hannongbua, S., and Mulholland, A. J. (2011). A water-swap reaction coordinate for the calculation of absolute protein-ligand binding free energies. *J. Chem. Phys.* 134 (5), 054114. doi:10.1063/1.3519057
- Woods, C. J., Malaisree, M., Michel, J., Long, B., McIntosh-Smith, S., and Mulholland, A. J. (2014). Rapid decomposition and visualisation of protein–ligand binding free energies by residue and by water. *Faraday Discuss.* 169, 477–499. doi:10.1039/c3fd00125c
- Yadav, R., Imran, M., Dhamija, P., Suchal, K., and Handu, S. (2021). Virtual screening and dynamics of potential inhibitors targeting RNA binding domain of nucleocapsid phosphoprotein from SARS-CoV-2. *J. Biomol. Struct. Dyn.* 39 (12), 4433–4448. doi:10.1080/07391102.2020.1778536
- Yu, W., and MacKerell, A. D. (2017). “Computer-aided drug design methods,” in *Antibiotics* (Springer), 85–106.
- Zhang, X., Perez-Sanchez, H., and C Lightstone, F. (2017). A comprehensive docking and MM/GBSA rescoring study of ligand recognition upon binding antithrombin. *Curr. Top. Med. Chem.* 17 (14), 1631–1639. doi:10.2174/1568026616666161117112604
- Zoete, V., Daina, A., Bovigny, C., and Michielin, O. (2016). *SwissSimilarity: a web tool for low to ultra high throughput ligand-based virtual screening*, 1399–1404. Available at: <https://www.tandfonline.com/doi/full/10.1080/17460441.2021.1942836>.