

Proceeding Paper

Virtual Screening of Natural Compounds as Potential SARS-CoV-2 M^{pro} Inhibitors: A Molecular Docking and Molecular Dynamics Simulation Guided Approach [†]

Deepak K. Lokwani ^{1,*}, Sangita R. Chavan ¹, Aniket P. Sarkate ², Prabhu M. Natarajan ^{3,*}, Vidhya R. Umaphathy ⁴ and Shirish P. Jain ¹

¹ Rajarshi Shahu College of Pharmacy, Buldhana, Maharashtra 443001, India

² Department of Chemical Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra 431004, India

³ Department of Clinical Sciences, Center of Medical and Bio-Allied Health Sciences and Research, College of Dentistry, Ajman University, Ajman, United Arab Emirates

⁴ Department of Public Health Dentistry, Thai Moogambigai Dental College and Hospital, Chennai 600107, India

* Correspondence: dklokwani@gmail.com, prabhuperio@gmail.com

[†] Presented at the 27th International Electronic Conference on Synthetic Organic Chemistry (ECSOC-27), 15–30 November 2023; Available online: <https://ecsoc-27.sciforum.net/>.

Abstract: The coronavirus disease 2019 (COVID-19) pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has significantly impacted human lives, overburdened the healthcare system, and weakened global economies. The lack of specific drugs against SARS-CoV-2 is a significant hurdle towards the successful treatment of COVID-19. The SARS-CoV-2 Main protease (M^{pro}) is considered an appealing target because of its role in replication in host cells. Plant-derived natural compounds are being largely tested for their efficacy against COVID-19 targets to combat SARS-CoV-2 infection. To discover hit compounds that can be used alone or in combination with repositioned drugs, we curated a set of 224,205 natural product structures from the ZINC database and virtually screened against COVID-19 M^{pro}. The sequential docking protocols involving different level of exhaustiveness were performed to screen a library of natural compounds. Final 88 compounds were selected and post-processed using the MM-GBSA analysis for the generation of binding free energies. The top four compounds (ZINC000085626103, ZINC000085569275, ZINC000085625768 and ZINC000085488571) showing higher affinity against COVID-19 M^{pro} enzyme selected for MD simulation studies. The RMSD, RMSF and RoG analysis of the all four compound-protein complexes indicated the absolute stability during 100ns MD run. Further the post-MD simulation binding free energies were calculated for all four compounds and were found to be in range of –38.29 to –18.07 kcal/mol. The In-silico virtual screening results suggested that the selected natural compounds have the potential to be developed as a COVID-19 M^{pro} inhibitor and can be explored further for experimental research to evaluate the *in vitro* and *in vivo* efficacy of these compounds for the treatment of COVID-19.

Keywords: COVID-19 M^{pro} inhibitor; ZINC database; docking; MD simulations; MM-GBSA

Citation: Lokwani, D.K.; Chavan, S.R.; Sarkate, A.P.; Natarajan, P.M.; Umaphathy, V.R.; Jain, S.P. Virtual Screening of Natural Compounds as Potential SARS-CoV-2 M^{pro} Inhibitors: A Molecular Docking and Molecular Dynamics Simulation Guided Approach. *Chem. Proc.* **2023**, *14*, x. <https://doi.org/10.3390/xxxxx>

Academic Editor(s): Name

Published: 15 November 2023



Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A novel coronavirus SARS-CoV-2 has caused a worldwide pandemic and remained a severe threat to the entire human population due to the lack of specific therapeutic agents to control the sudden outbreak of the SARS-CoV-2 [1]. The SARS-CoV-2 main protease (M^{pro}) protein which is a vital target for drug discovery studies against the recent coronavirus pandemics [2]. In silico screening of phytochemical database has gained increasing interest in drug discovery research for the identification of new drugs leads or

drug molecules [3,4]. A virtual screening based on molecular docking emerges as an important tool for obtaining new antiviral molecules, where researchers can use this tool as a complementary approach so that the synthesis of new compounds or the repositioning of drugs can be assigned. The objective of the study is to perform Virtual Screening of Natural Compounds from Zinc Database to find out possible antiviral agents with protease inhibitory potential against SARS-CoV-2.

2. Result and Discussion

2.1. Docking Studies

In the pursuit of identifying the potential drug candidates targeting the COVID-19 Main Protease (Mpro) enzyme, we conducted a rigorous virtual screening process using docking studies employing the smina molecular docking software. The sequential docking protocols, involving varying levels of exhaustiveness were carried out to effectively screen the extensive library of approximately 224,205 natural product structures sourced from the ZINC database. Initially, all compounds were docked on Mpro enzyme with a default exhaustiveness setting of 8 and subsequently, the top 10% of compounds with the best docking scores from were selected for further screening. The obtained subset of approximately 10,000 compounds were further docked on Mpro enzyme with an exhaustiveness setting of 24. Once again, the top 10% of compounds with the highest docking scores were retained and in the final phase of screening, approximately 1000 compounds were subjected to rigorous docking simulations, employing an exhaustiveness setting of 48. Ultimately, we selected only the highest-scoring compound from this final set, ensuring the most stringent selection criteria. To calculate binding free energies, shedding light on the thermodynamic aspects of the ligand-receptor interactions and to gain deeper insights into the interactions between the selected 88 compounds and the Mpro enzyme, the post-docking MM-GBSA analysis were performed. Based on all above results, final four compounds namely ZINC000085626103, ZINC000085625768, ZINC000085488571 and ZINC000085569275 were selected with highest docking and MM-GBSA scores for the further ligand-enzyme interaction analysis and MD Simulation studies (Table 1, Figure 1).

Table 1. Docking results of selected compounds against Mpro enzyme.

Sr No	Compound Id	Docking Score	Ligand Efficacy	Free Binding Energy after Docking (Kcal/mol)
1	ZINC000085626103	-12.682	-0.278	-94.8
2	ZINC000085569275	-12.026	-0.463	-50.11
3	ZINC000085625768	-11.945	-0.291	-58.97
4	ZINC000085488571	-11.876	-0.276	-55.34

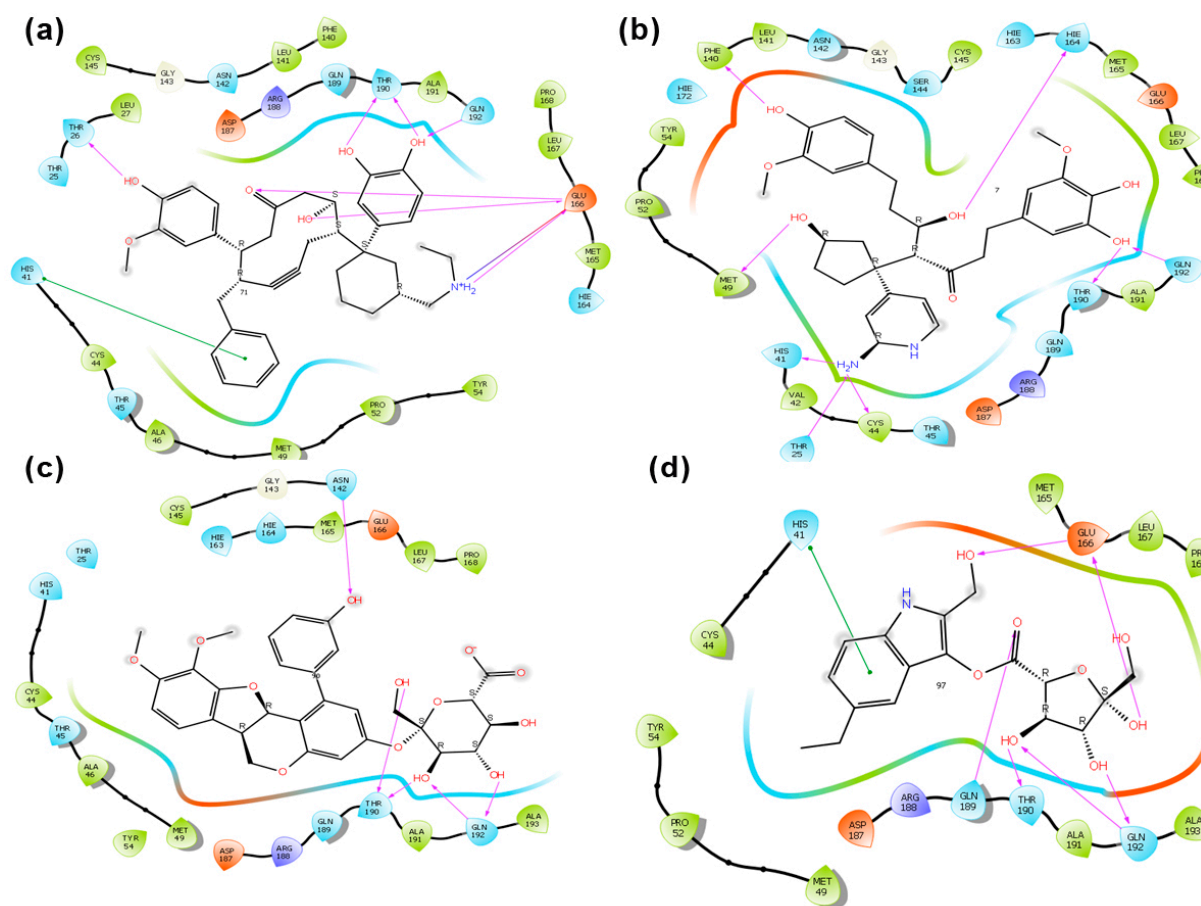


Figure 1. 2D images of (a) ZINC000085626103, (b) ZINC000085625768 (c) ZINC000085488571 and (d) ZINC000085569275 at active site of SARS-CoV-2 main protease (M^{pro}).

2.2. MD Simulation Studies

The comprehensive molecular dynamics (MD) simulations for a duration of 100 ns were carried out to validate the stability and dynamic behavior of the M^{pro} enzyme upon binding to all four selected compounds. The MD simulation study was also performed for apo protein structure to further support the analysis. The key parameters, such as the Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), Radius of Gyration (RoG), and ligand-protein interactions, were used to evaluate the molecular stability of complex systems, which helps to provide insight into the conformational changes that occur during compound-M^{pro} enzyme interaction (Figures 2 and 3).

The analysis of RMSD values for the M^{pro} enzyme-compounds complexes reveals that all systems exhibit remarkable stability over a timeframe of 100 ns (Figure 2). Like apo structure of M^{pro} enzyme, all four M^{pro} enzyme-compounds complexes were shown to stabilize below 0.30–0.35 nm of RMSD values. While RMSD values within the M^{pro} enzyme - ZINC000085625768 system and M^{pro} enzyme-ZINC000085569275 exhibits some degree of fluctuation, but these fluctuations tend to stabilize within the narrow range of 0.225–0.325 and 0.3 to 0.375 nm. Similarly, M^{pro} enzyme-ZINC000085626103 showed the minimal fluctuation as denoted by its RMSD values during entire simulation run, while M^{pro} enzyme-ZINC000085488571 showed highly stable compounds throughout MD simulation cycle with RMSD value tend to stabilise with the range of 0.15 to 0.275 nm. The analysis of RMSD indicated that the incorporation of all compounds into the active site of the M^{pro} enzyme leads to a consistent and steady behavioural pattern across these systems. The stability observed throughout MDs simulations emphasizes the potential therapeutic importance of these compounds in modulating the activity of the M^{pro} enzyme.

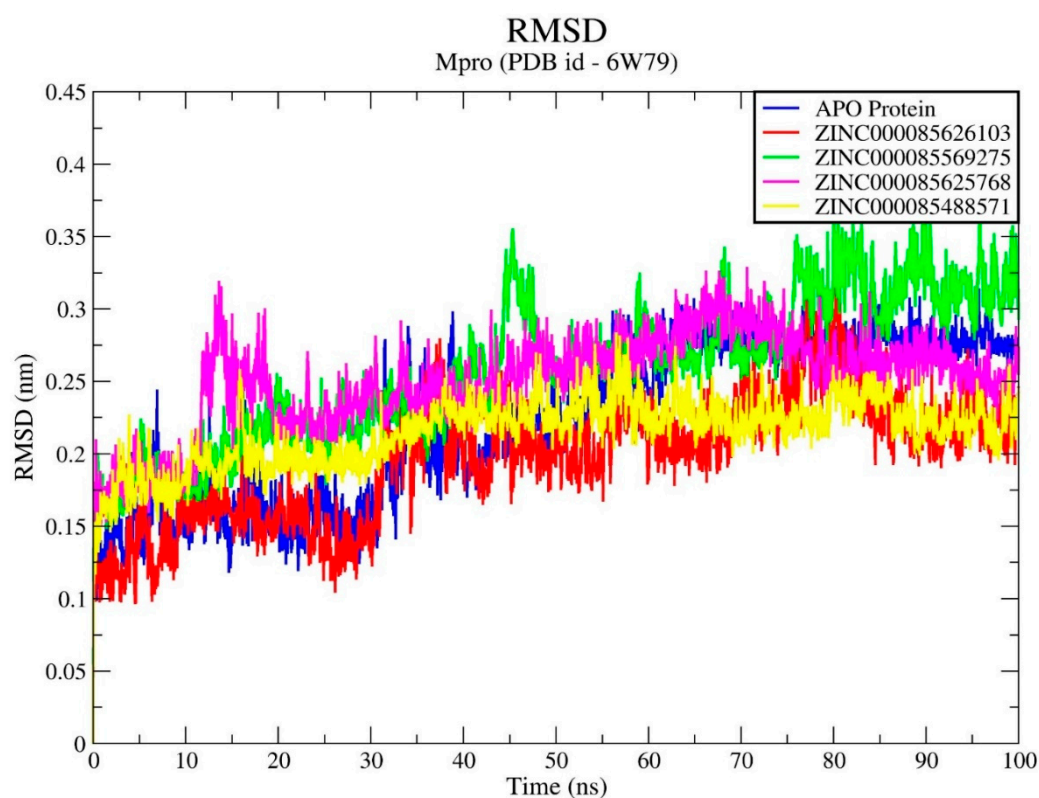


Figure 2. Time dependent review of RMSD for Mpro enzyme upon binding all compounds.

Furthermore, an in-depth analysis of RMSF of the $C\alpha$ atoms of amino acid residues was performed in all systems, as illustrated in Figure 3. The investigation unveiled that the $C\alpha$ atoms of amino acids located in the loop region of the enzyme noticed the most significant atomic fluctuations. Significantly, the most noteworthy fluctuations were primarily observed in the region encompassing amino acid residues 45 to 50 and 150 to 200, which corresponds to a domain associated with the loop located away from the active site. It is worth mentioning that similar levels of fluctuation were observed in all M^{pro} enzyme-compounds systems. The analysis of RMSF presented here the additional evidence of supporting the overall stability in the M^{pro} enzyme in complex with all four compounds. During the simulations, we continuously monitored the interactions between the M^{pro} enzyme and ligands, encompassing hydrogen bonds, hydrophobic interactions, ionic interactions, and water bridges. Figure 4, which presented stacked bar charts illustrating the protein-ligand contacts, revealed that both compounds engaged in a greater number of interactions aimed at stabilizing the complex with the M^{pro} enzyme throughout the simulations.

These overall observations further served as confirmation of the prevailing belief that both drugs form stable complexes within the active site of the M^{pro} enzyme with minimal structural changes.

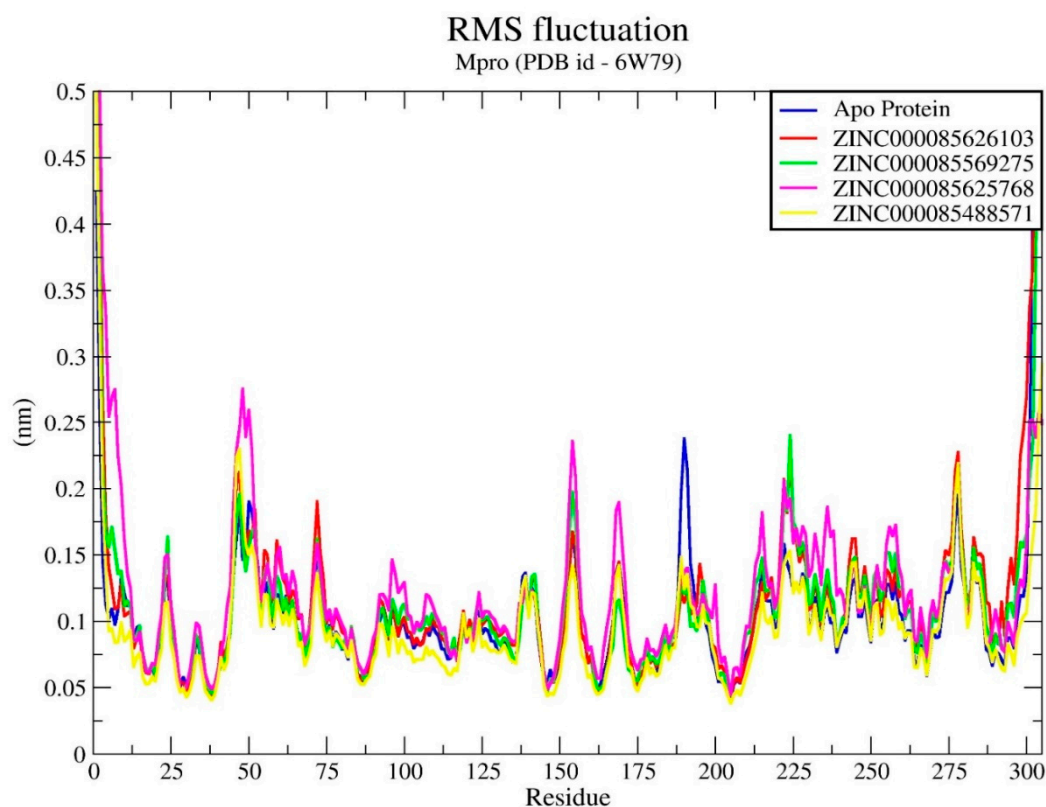


Figure 3. Time dependent evaluation of RMSF Rg for Mpro enzyme upon binding to all compounds.

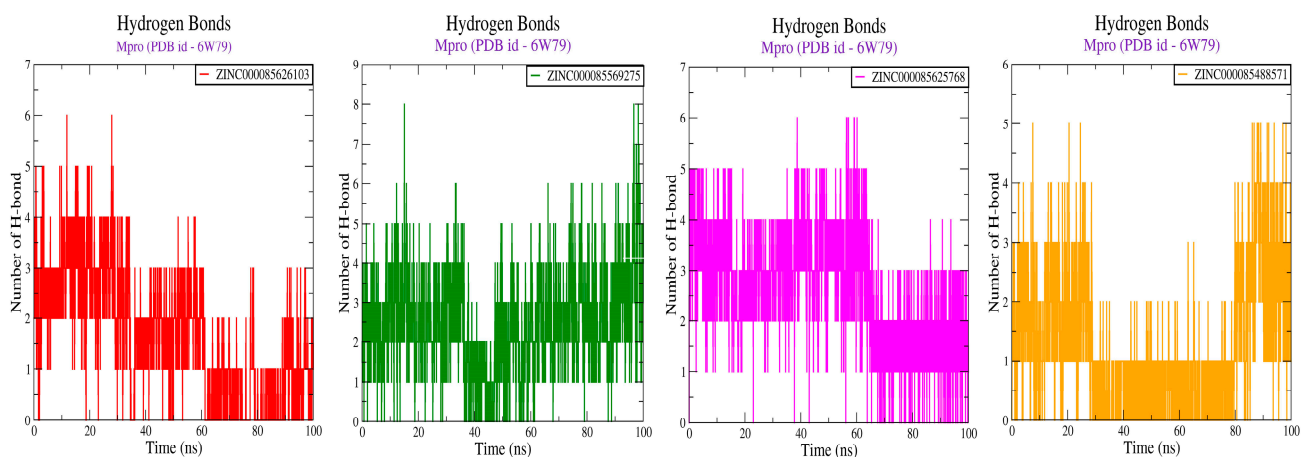


Figure 4. Bar Diagram showing Protein-ligand H-bonds obtained after 100ns MD simulation studies for compounds-Mpro complexes.

2.3. Binding Free Energy Calculations of the Complexes Using MM-GBSA Analysis

To assess the reliability of the binding affinity of all compounds with the Mpro enzyme, we conducted post-MD simulation MM-GBSA calculations. The MM-PBSA ΔG bind values were determined by assessing the energy difference between the bound and unbound states of the complexes. The average ΔG binding free energy values for ZINC000085626103, ZINC000085625768, ZINC000085488571 and ZINC000085569275 against the Mpro enzyme were found to be -19.17 ± 17.54 , -38.29 ± 5.84 , -25.84 ± 5.74 and -20.56 ± 5.53 kcal/mol, respectively (Table 2). These results indicate that all compounds

have a significant affinity for binding to the enzyme. Notably, ZINC000085625768 exhibited a higher affinity for the Mpro enzyme, suggesting that complexes formed with it may be more stable.

Table 2. MM-GBSA ΔG Binding free energy of docked compounds in complex with MAO-A enzyme.

Compounds	Delta G Gas	Delta G Solv	Delta G Total
ZINC000085626103	-108.43 ± 10.05	86.26 ± 14.38	-19.17 ± 17.54
ZINC000085625768	-82.90 ± 8.99	44.61 ± 5.67	-38.29 ± 5.84
ZINC000085488571	-15.96 ± 15.38	-9.87 ± 11.47	-25.84 ± 5.74
ZINC000085569275	-44.35 ± 11.83	23.79 ± 8.02	-20.56 ± 5.53

3. Conclusions

The four Compound namely ZINC000085626103, ZINC000085625768, ZINC000085488571 and ZINC000085569275 were found to exhibit remarkable binding affinities for the Mpro of SARS-CoV-2, after the screening of natural compound from ZINC database against SARS-CoV-2 Mpro enzyme. All four compounds were found to be highly effective against Mpro enzyme based on docking score, MMGBSA free binding energy and post processing ΔG binding Energy. These compounds could, therefore, serve as a starting point for the development as potent and successful antiviral drugs against the deadly COVID-19.

4. Methodology

4.1. Docking Methodology

The molecular docking investigation was conducted using the Smina molecular docking tool [5], following a procedure in alignment with the methodologies previously outlined by our research group [6–8]. Initially, the 3D structures of both compounds were obtained from PubChem database and were minimized utilizing the steepest descent method through the Open Babel chemical toolbox. For all enzymes, X-ray crystal structures were acquired from the Protein Data Bank, and their preparation for docking was performed via Dock Prep, an integrated tool within UCSF Chimera Software. The binding site was chosen by employing the coordinates of the co-crystal ligand of enzymes, with an additional 4 Å extension in each dimension. The lower energy conformers of both ligands were then subsequently docked within the selected active site of the enzyme, using the default scoring function of Smina. The academic version of Maestro software (Schrodinger, New York, NY, USA) was used to visualize score, poses, and save images of the docking results.

4.2. Molecular Dynamics (MD) Simulation

MD simulation studies of the compound-Mpro enzyme complexes were carried out using GROMACS 2021, following a methodology detailed in our prior publication [6]. In summary, we employed the CHARMM36 force field to establish the protein's topological structure. The topology and parameters for the ligands were generated using the AnteChamber Python Parser interface (ACPYPE) [9]. Subsequently, the system was placed within a cubic box with periodic boundary conditions (PBC), with the TIP3P water model utilized for solvation. Counter ions were introduced to achieve a neutralized system. To minimize the system's energy, we applied the steepest descent algorithm with a tolerance value set at $1000 \text{ kJ mol}^{-1} \text{ nm}^{-1}$. Following energy minimization, the system underwent equilibration under both the NVT and NPT ensembles, each lasting 1000 ps. The Berendsen algorithm was employed to control the thermostat and barostat during the equilibration process. Subsequently, the system was subjected to a production MD simulation spanning 100 nanoseconds (ns), with trajectory snapshots saved at 50 ps intervals,

resulting in approximately 2000 frames for subsequent analysis. Throughout the MD simulations, the temperature and pressure were held constant at 300 K and 1.01325 bar, respectively. We employed standard analysis techniques to compute parameters such as RMSD, RMSF, Rg and the formation of hydrogen bonds over the simulation duration. For post-simulation molecular mechanics with generalized Born and surface area (MM-GBSA) analysis, we utilized the gmx_MMPBSA tool [10].

References

1. Kumari, M.; Lu, R.-M.; Li, M.-C.; Huang, J.-L.; Hsu, F.-F.; Ko, S.-H.; Ke, F.-Y.; Su, S.-C.; Liang, K.-H.; Yuan, J.P.-Y.; et al. A Critical Overview of Current Progress for COVID-19: Development of Vaccines, Antiviral Drugs, and Therapeutic Antibodies. *J. Biomed. Sci.* **2022**, *29*, 68. <https://doi.org/10.1186/s12929-022-00852-9>.
2. Kronenberger, T.; Laufer, S.A.; Pillaiyar, T. COVID-19 Therapeutics: Small-Molecule Drug Development Targeting SARS-CoV-2 Main Protease. *Drug Discov. Today* **2023**, *28*, 103579. <https://doi.org/10.1016/j.drudis.2023.103579>.
3. Rani, I.; Kalsi, A.; Kaur, G.; Sharma, P.; Gupta, S.; Gautam, R.K.; Chopra, H.; Bibi, S.; Ahmad, S.U.; Singh, I.; et al. Modern Drug Discovery Applications for the Identification of Novel Candidates for COVID-19 Infections. *Ann. Med. Surg.* **2022**, *80*, 104125. <https://doi.org/10.1016/j.amsu.2022.104125>.
4. Das, A.P.; Agarwal, S.M. Recent Advances in the Area of Plant-Based Anti-Cancer Drug Discovery Using Computational Approaches. *Mol. Divers.* **2023**, 1–25. <https://doi.org/10.1007/s11030-022-10590-7>.
5. Koes, D.R.; Baumgartner, M.P.; Camacho, C.J. Lessons Learned in Empirical Scoring with Smina from the CSAR 2011 Benchmarking Exercise. *J. Chem. Inf. Model.* **2013**, *53*, 1893–1904. <https://doi.org/10.1021/ci300604z>.
6. Kardile, R.A.; Sarkate, A.P.; Lokwani, D.K.; Tiwari, S.V.; Azad, R.; Thopate, S.R. Design, Synthesis, and Biological Evaluation of Novel Quinoline Derivatives as Small Molecule Mutant EGFR Inhibitors Targeting Resistance in NSCLC: In Vitro Screening and ADME Predictions. *Eur. J. Med. Chem.* **2023**, *245*, 114889. <https://doi.org/10.1016/j.ejmech.2022.114889>.
7. Tiwari, S.V.; Sarkate, A.P.; Lokwani, D.K.; Pansare, D.N.; Gattani, S.G.; Sheaikh, S.S.; Jain, S.P.; Bhandari, S.V. Explorations of Novel Pyridine-Pyrimidine Hybrid Phosphonate Derivatives as Aurora Kinase Inhibitors. *Bioorganic Med. Chem. Lett.* **2022**, *67*, 128747. <https://doi.org/10.1016/j.bmcl.2022.128747>.
8. Thorat, N.M.; Khodade, V.S.; Ingale, A.P.; Lokwani, D.K.; Sarkate, A.P.; Thopate, S.R. Molecular Docking Studies and Application of 6-(1-Arylmethanamino)-2-Phenyl-4 H-Chromen-4-Ones as Potent Antibacterial Agents. *Polycycl. Aromat. Compd.* **2022**, 1–14. <https://doi.org/10.1080/10406638.2022.2150238>.
9. Sousa da Silva, A.W.; Vranken, W.F. ACPYPE—AnteChamber PYthon Parser interface. *BMC Res. Notes* **2012**, *5*, 367. <https://doi.org/10.1186/1756-0500-5-367>.
10. Valdés-Tresanco, M.S.; Valdés-Tresanco, M.E.; Valiente, P.A.; Moreno, E. gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS. *J. Chem. Theory Comput.* **2021**, *17*, 6281–6291. <https://doi.org/10.1021/acs.jctc.1c00645>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.