

INTERACTION OF BIOMOLECULES AND BIOACTIVE COMPOUNDS WITH THE SARS-COV-2 PROTEINS: MOLECULAR SIMULATIONS FOR THE FIGHT AGAINST COVID-19

EDITED BY: Mattia Falconi, Arvind Ramanathan and James Leland Olds
PUBLISHED IN: Frontiers in Molecular Biosciences



frontiers

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence.

The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714

ISBN 978-2-88976-575-1

DOI 10.3389/978-2-88976-575-1

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

INTERACTION OF BIOMOLECULES AND BIOACTIVE COMPOUNDS WITH THE SARS-COV-2 PROTEINS: MOLECULAR SIMULATIONS FOR THE FIGHT AGAINST COVID-19

Topic Editors:

Mattia Falconi, University of Rome Tor Vergata, Italy

Arvind Ramanathan, Argonne National Laboratory (DOE), United States

James Leland Olds, George Mason University, United States

Citation: Falconi, M., Ramanathan, A., Olds, J. L., eds. (2022). Interaction of Biomolecules and Bioactive Compounds With the SARS-CoV-2 Proteins: Molecular Simulations for the Fight Against Covid-19. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-575-1

Table of Contents

- 04 Editorial: Interaction of Biomolecules and Bioactive Compounds With the SARS-CoV-2 Proteins: Molecular Simulations for the Fight Against Covid-19**
Mattia Falconi, James Leland Olds and Arvind Ramanathan
- 06 Computational Insights Into the Effects of the R190K and N121Q Mutations on the SARS-CoV-2 Spike Complex With Biliverdin**
Zhiyuan Qu, Kaihang Li, Xiaoju Geng, Bo Huang and Jian Gao
- 15 A Computer-Aided Approach for the Discovery of D-Peptides as Inhibitors of SARS-CoV-2 Main Protease**
Jorge E. Hernández González, Raphael J. Eberle, Dieter Willbold and Mônica A. Coronado
- 32 Molecular Docking and Dynamics Studies to Explore Effective Inhibitory Peptides Against the Spike Receptor Binding Domain of SARS-CoV-2**
Suvro Biswas, Shafi Mahmud, Mohasana Akter Mita, Shamima Afrose, Md. Robiul Hasan, Mst. Sharmin Sultana Shimu, Md. Abu Saleh, Gomaa Mostafa-Hedeab, Mohammed Alqarni, Ahmad J. Obaidullah and Gaber El-Saber Batiha
- 42 Phytochemicals of Euphorbia hirta L. and Their Inhibitory Potential Against SARS-CoV-2 Main Protease**
Ruel Cayona and Evelyn Creencia
- 57 SARS-CoV-2 Spike Protein Unlikely to Bind to Integrins via the Arg-Gly-Asp (RGD) Motif of the Receptor Binding Domain: Evidence From Structural Analysis and Microscale Accelerated Molecular Dynamics**
Houcemeddine Othman, Haifa Ben Messaoud, Oussema Khamessi, Hazem Ben-Mabrouk, Kais Ghedira, Avani Bharuthram, Florette Treurnicht, Ikechukwu Achilonu, Yasien Sayed and Najet Srairi-Abid
- 68 Computational Repurposing of Drugs and Natural Products Against SARS-CoV-2 Main Protease (M^{pro}) as Potential COVID-19 Therapies**
Sakshi Piplani, Puneet Singh, Nikolai Petrovsky and David A. Winkler
- 83 Antigen–Antibody Complex-Guided Exploration of the Hotspots Conferring the Immune-Escaping Ability of the SARS-CoV-2 RBD**
Kit-Man Fung, Shu-Jung Lai, Tzu-Lu Lin and Tien-Sheng Tseng
- 102 Potential of Natural Alkaloids From Jadwar (*Delphinium denudatum*) as Inhibitors Against Main Protease of COVID-19: A Molecular Modeling Approach**
Anuj Kumar, Mansi Sharma, Christopher D. Richardson and David J. Kelvin
- 116 Evaluation of Clove Phytochemicals as Potential Antiviral Drug Candidates Targeting SARS-CoV-2 Main Protease: Computational Docking, Molecular Dynamics Simulation, and Pharmacokinetic Profiling**
Arun Chandra Manivannan, Arunkumar Malaisamy, Murugesw Eswaran, Arun Meyyazhagan, Vijaya Anand Arumugam, Kannan R. R. Rengasamy, Balamuralikrishnan Balasubramanian and Wen-Chao Liu



Editorial: Interaction of Biomolecules and Bioactive Compounds With the SARS-CoV-2 Proteins: Molecular Simulations for the Fight Against Covid-19

Mattia Falconi^{1*}, James Leland Olds^{2*} and Arvind Ramanathan^{3*}

¹Department of Biology, University of Rome "Tor Vergata", Rome, Italy, ²Schar School of Policy and Government, George Mason University, Fairfax, VA, United States, ³Argonne National Laboratory (DOE), Lemont, IL, United States

Keywords: SARS-CoV-2, mpro protease, spike protein, computational methods, molecular simulation and docking

Editorial on the Research Topic

Interaction of Biomolecules and Bioactive Compounds with the SARS-CoV-2 Proteins: Molecular Simulations for the Fight Against Covid-19

The Covid-19 pandemic, which we are still experiencing, has fostered the interest and study of all researchers worldwide who have made their expertise available to contribute to solving this global problem (Sharma et al., 2021). Computational scientists were able to continue their research when the experimentalists often had to leave their laboratories to prevent the spread of this dangerous virus. This pandemic has shown how computational approaches (Śledź and Caflisch, 2018; Patel et al., 2020; Romeo et al., 2020) can help the understanding of the structural basis underlying possible coronavirus inhibition mechanisms and how they may contribute to accelerating the discovery of novel treatment methods (Gurung et al., 2021). Although the release of very effective vaccines has largely helped people control this disease and reduce its burden on the worldwide population, there is still a lack of effective, safe, and broad-spectrum antiviral drugs to treat infected patients and stem future generations epidemics.

Thanks to the solution of the molecular structures composing the SARS-CoV-2 virion (Arya et al., 2021), many are the targets offered to molecular simulators who have designed various classes of molecules, peptides, or have selected proteins and antibodies to stem the spread of this threatening coronavirus.

This Special Issue collects computational research having two main coronavirus proteins as a target that allow and promote the SARS-CoV-2 infection, i.e., the main protease of the virus (Mpro) and the Spike glycoprotein.

Some of the collected papers deal with virtual screening applications associated with *in silico* or/and experimental validation of natural compounds or peptides that target the Mpro protease.

This target has been chosen to block the viral proteins processing. During host cell infection, the viral genome acts as messenger RNA. It directs the synthesis of two large polyproteins (pp1a and pp1ab), containing small proteins necessary to produce new viral particles inside infected cells. This set of proteins includes a replication/transcription complex, several structural proteins needed to build virions and two proteases (Wu et al., 2020; Zhou et al., 2020). The proteases play an essential role in cutting the two large polyproteins into smaller functional proteins. The SARS-CoV-2 main protease Mpro weighs 33.8 kDa and makes the most cuts. Mpro, essential for viral replication and absent in human cells, represents an optimal target for developing new antiviral drugs: blocking its functions would be lethal for the virus but safe for humans.

OPEN ACCESS

Edited by:

Liliang Zhang,
Shandong First Medical University,
China

Reviewed by:

Peiqi Yin,
Albert Einstein College of Medicine,
United States

*Correspondence:

Mattia Falconi
falconi@uniroma2.it
James Leland Olds
jolds@gmu.edu
Arvind Ramanathan
ramanathana@anl.gov

Specialty section:

This article was submitted to
Biological Modeling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 23 May 2022

Accepted: 07 June 2022

Published: 22 June 2022

Citation:

Falconi M, Olds JL and Ramanathan A
(2022) Editorial: Interaction of
Biomolecules and Bioactive
Compounds With the SARS-CoV-2
Proteins: Molecular Simulations for the
Fight Against Covid-19.
Front. Mol. Biosci. 9:950891.
doi: 10.3389/fmolb.2022.950891

In this regard, to disable the SARS-CoV-2 Mpro activity, Piplani et al. carried out a computational repurposing of a series of drugs and natural products to be used as potential novel COVID-19 therapies; Cayona and Creencia tested the inhibitory potential of phytochemicals from the plant *Euphorbia hirta* L.; Kumar et al. evaluate the potential as inhibitors of natural alkaloids from Jadwar (*Delphinium denudatum*), and finally, Manivannan et al. evaluate the clove phytochemicals, a traditional natural therapeutic that comprises important bioactive compounds, as possible antiviral drug candidate targeting Mpro. The study has been carried out using molecular docking, molecular dynamics simulation and pharmacokinetic profiling. Interesting results have been obtained by Hernández González et al. designing and suggesting D-peptides as Mpro inhibitors. These authors use MM-GBSA free energy calculations, molecular dynamics (MD) simulations, and *in vitro* enzymatic assays of the four top-scoring D-tetrapeptides, all of which caused 55–85% inhibition of Mpro activity, thus highlighting the suitability of the devised approach.

The other section of this Special Issue turns to the other main research target i.e., the Spike glycoprotein (S) of the virus.

According to a key-lock model, viruses continuously evolve the proteins on their surface to enhance the interaction with the receptors on the cells and enter them more efficiently. This is also the case of SARS-CoV-2 Spike glycoprotein (the key) and the human Angiotensin Converting Enzyme 2 receptor (hACE2, the lock).

Spike protein is one of the most interesting and studied proteins that contribute to the binding with the host receptor and viral pathogenesis. Spike decorates the virus surface and is responsible for the viral surface corona appearance (Zhou et al., 2020). The virus uses this protein as a key to enter host cells (Tortorici et al., 2019). It acts by binding the receptor on target cells, inducing endocytosis of virions, catalyzing the fusion between cell and viral membranes and ensuring the entry of

viral genomic RNA into the cytoplasm of cells. Spike protein is also the main target of the immune system, activating it and inducing the production of antibodies. For this reason, it is considered the primary target of antiviral drugs and vaccines, constituting a rich source of helpful information for the design of molecules able to inhibit its function and, therefore, potentially usable as therapeutic treatments.

In this regard, to explore effective inhibitory peptides against the Spike RBD of SARS-CoV-2, Biswas et al. have applied molecular docking and MD techniques on 23 antimicrobial peptides selected from literature. The obtained computational insights helped to understand the decrease in binding affinity of biliverdin with Spike caused by the R190K and N121Q mutations. Fung et al. have systematically investigated the RBD variants that markedly destabilize the binding to six neutralizing antibodies through in-depth mutational scanning; and finally Othman et al., using structural analysis and microscale accelerated MD, have explored the possibility of Spike Protein to bind integrins, proposed as host receptors for SARS-CoV-2, through the Arg-Gly-Asp (RGD) motif of the RBD.

In conclusion, deciphering the structure of critical virus constituents (Arya et al., 2021) has permitted computational researchers to apply their techniques and work to hypothesize a molecular solution to the pandemic problem. The experimental world has also greatly benefited from computational suggestions, discovering the usefulness of the simulation techniques.

We expect that this Special Issue will be helpful to experimentalists and clinicians and that it will further stimulate the use of novel and low-cost molecules to counteract the SARS-CoV-2 virus.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial and intellectual contribution to the work, and approved it for publication.

REFERENCES

- Arya, R., Kumari, S., Pandey, B., Mistry, H., Bihani, S. C., Das, A., et al. (2021). Structural Insights into SARS-CoV-2 Proteins. *J. Mol. Biol.* 433 (2), 166725. doi:10.1016/j.jmb.2020.11.024
- Gurung, A. B., Ali, M. A., Lee, J., Farah, M. A., and Al-Anazi, K. M. (2021). An Updated Review of Computer-Aided Drug Design and its Application to COVID-19. *Biomed. Res. Int.* 2021, 8853056. doi:10.1155/2021/8853056
- Romeo, A., Iacovelli, F., and Falconi, M. (2020). Targeting the SARS-CoV-2 Spike Glycoprotein Prefusion Conformation: Virtual Screening and Molecular Dynamics Simulations Applied to the Identification of Potential Fusion Inhibitors. *Virus Res.* 286, 198068. doi:10.1016/j.virusres.2020.198068
- Patel, L., Shukla, T., Huang, X., Ussery, D. W., and Wang, S. (2020). Machine Learning Methods in Drug Discovery. *Molecules* 25 (22), 5277. doi:10.3390/molecules25225277
- Sharma, A., Ahmad Farouk, I., and Lal, S. K. (2021). COVID-19: A Review on the Novel Coronavirus Disease Evolution, Transmission, Detection, Control and Prevention. *Viruses* 13 (2), 202. doi:10.3390/v13020202
- Śledź, P., and Caffisch, A. (2018). Protein Structure-Based Drug Design: from Docking to Molecular Dynamics. *Curr. Opin. Struct. Biol.* 48, 93–102. doi:10.1016/j.sbi.2017.10.010
- Tortorici, M. A., Walls, A. C., Lang, Y., Wang, C., Li, Z., Koerhuis, D., et al. (2019). Structural Basis for Human Coronavirus Attachment to Sialic Acid Receptors. *Nat. Struct. Mol. Biol.* 26 (6), 481–489. doi:10.1038/s41594-019-0233-y

- Wu, F., Zhao, S., Yu, B., Chen, Y. M., Wang, W., Song, Z. G., et al. (2020). A New Coronavirus Associated with Human Respiratory Disease in China. *Nature* 579 (7798), 265–269. doi:10.1038/s41586-020-2008-3
- Zhou, P., Yang, X. L., Wang, X. G., Hu, B., Zhang, L., Zhang, W., et al. (2020). A Pneumonia Outbreak Associated with a New Coronavirus of Probable Bat Origin. *Nature* 579 (7798), 270–273. doi:10.1038/s41586-020-1212-7

Conflict of Interest: The authors declare 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.

Copyright © 2022 Falconi, Olds and Ramanathan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.



Computational Insights Into the Effects of the R190K and N121Q Mutations on the SARS-CoV-2 Spike Complex With Biliverdin

Zhiyuan Qu¹, Kaihang Li¹, Xiaoju Geng¹, Bo Huang¹ and Jian Gao^{1,2*}

¹Jiangsu Key Laboratory of New Drug Research and Clinical Pharmacy, Xuzhou Medical University, Xuzhou, China, ²Xuzhou Medical University Technology Transfer Center Co., Ltd., Xuzhou Medical University, Xuzhou, China

OPEN ACCESS

Edited by:

James Leland Olds,
George Mason University,
United States

Reviewed by:

Lance Hellman,
Nevada State College, United States
Qiang Shao,
Shanghai Institute of Materia Medica
(CAS), China

*Correspondence:

Jian Gao
gaojian@xzhmu.edu.cn

*ORCID:

Jian Gao
orcid.org/0000-0001-6723-0490

Specialty section:

This article was submitted to
Biological Modeling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 09 October 2021

Accepted: 15 November 2021

Published: 13 December 2021

Citation:

Qu Z, Li K, Geng X, Huang B and Gao J
(2021) Computational Insights Into the
Effects of the R190K and N121Q
Mutations on the SARS-CoV-2 Spike
Complex With Biliverdin.
Front. Mol. Biosci. 8:791885.
doi: 10.3389/fmolb.2021.791885

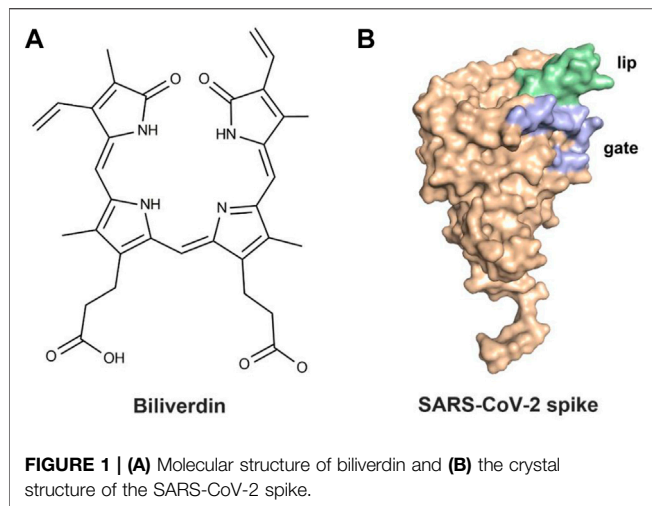
The SARS-CoV-2 spike has been regarded as the main target of antibody design against COVID-19. Two single-site mutations, R190K and N121Q, were deemed to weaken the binding affinity of biliverdin although the underlying molecular mechanism is still unknown. Meanwhile, the effect of the two mutations on the conformational changes of “lip” and “gate” loops was also elusive. Thus, molecular dynamics simulation and molecular mechanics/generalized Born surface area (MM/GBSA) free energy calculation were conducted on the wild-type and two other SARS-CoV-2 spike mutants. Our simulations indicated that the R190K mutation causes Lys190 to form six hydrogen bonds, guided by Asn99 and Ile101, which brings Lys190 closer to Arg102 and Asn121, thereby weakening the interaction energy between biliverdin and Ile101 as well as Lys190. For the N121Q mutation, Gln121 still maintained a hydrogen bond with biliverdin; nevertheless, the overall binding mode deviated significantly under the reversal of the side chain of Phe175. Moreover, the two mutants would stabilize the lip loop, which would restrain the meaningful upward movement of the lip. In addition, N121Q significantly promoted the gate loop deviating to the biliverdin binding site and compressed the site. This work would be useful in understanding the dynamics binding biliverdin to the SARS-CoV-2 spike.

Keywords: SARS-CoV-2 spike, biliverdin, molecular dynamics simulation, MM/GBSA calculation, mutation

INTRODUCTION

Since the advent of novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), significant threats have been posed to the human population worldwide (Wang et al., 2020). Millions of infections and deaths have been caused by this severe epidemic (Hu et al., 2021). The spike protein trimers, a protruded structure that exists on the SARS-CoV-2 virions, are able to bind to a surface receptor on the cell and accommodate fusion of the viral and cellular membranes when they are glycosylated. For these reasons, with perfect conformational flexibility (Ke et al., 2020), the SARS-CoV-2 spike is a crucial viral antigen and the target in designing antibodies, which lead to its support for the current critical SARS-CoV-2 vaccine development efforts (Tregoning et al., 2020).

The previous study (Rosa et al., 2021) revealed that the SARS-CoV-2 spike could bind biliverdin (Figure 1A), the tetrapyrrole product of heme metabolism. The tetrapyrrole interaction pocket to the spike N-terminal domain (NTD)’s deep cleft was mapped with the aid of cryo-electron microscopy and X-ray crystallography (Chi et al., 2020; Liu et al., 2020; Zost et al., 2020). The relevance between mutations in the SARS-CoV-2 spike NTD and viral escape from antibody immunity (Graham et al.,



2021; Kemp et al., 2021) has been proven *via* observations in circulating viral strains (Tegally et al., 2021). The reactivity of SARS-CoV-2 spike with immune sera was intensively decreased by biliverdin. Meanwhile, a subset of neutralizing antibodies was also inhibited by biliverdin. Characterized neutralizing antibodies primarily bind the spike C-terminal domain (referred to as the receptor-binding domain, RBD) (Ju et al., 2020; Walls et al., 2020; Wrapp et al., 2020). The access to the dominant epitope of SARS-CoV-2 spike NTD can be controlled by an allosteric mechanism regulated through the recruitment of a metabolite.

After binding with biliverdin, a solvent-exposed loop “gate” in the wild type (WT) swings out of the way, allowing for Fab binding, which is complemented by an upward movement of a β -hairpin “lip” (Figure 1B) (Rosa et al., 2021). Moreover, the binding affinity of biliverdin bound to SARS-CoV-2 S1 is profoundly affected by the R190K and N121Q amino acid substitutions, with the corresponding K_d values significantly weakening (Rosa et al., 2021). To our knowledge, the dynamic effects of the two mutations on the lip and the gate loops have not been illustrated yet. Furthermore, the sharp decreased molecular mechanism of biliverdin binding affinity, caused by the two mutations, is still unknown. Accordingly, a combined strategy of molecular dynamics (MD) simulation and molecular mechanics/generalized Born (GB) surface area (MM/GBSA) free energy calculation was conducted on the complexes of wild-type and two mutated SARS-CoV-2 S1 proteins in the present study.

MATERIALS AND METHODS

Construction of Simulation Systems

To figure out the impact of single-site mutations on the interactions of SARS-CoV-2 spike binding with biliverdin, three systems were constructed, including Spike^{WT}/Biliverdin, Spike^{N121Q}/Biliverdin, and Spike^{R190K}/Biliverdin. In detail, the Spike^{WT}/Biliverdin system was acquired from its corresponding X-ray crystal structure (PDB ID: 7b62), while the other two single-site mutations were constructed based on the system of Spike^{WT}/Biliverdin in Sybyl-X2.1.

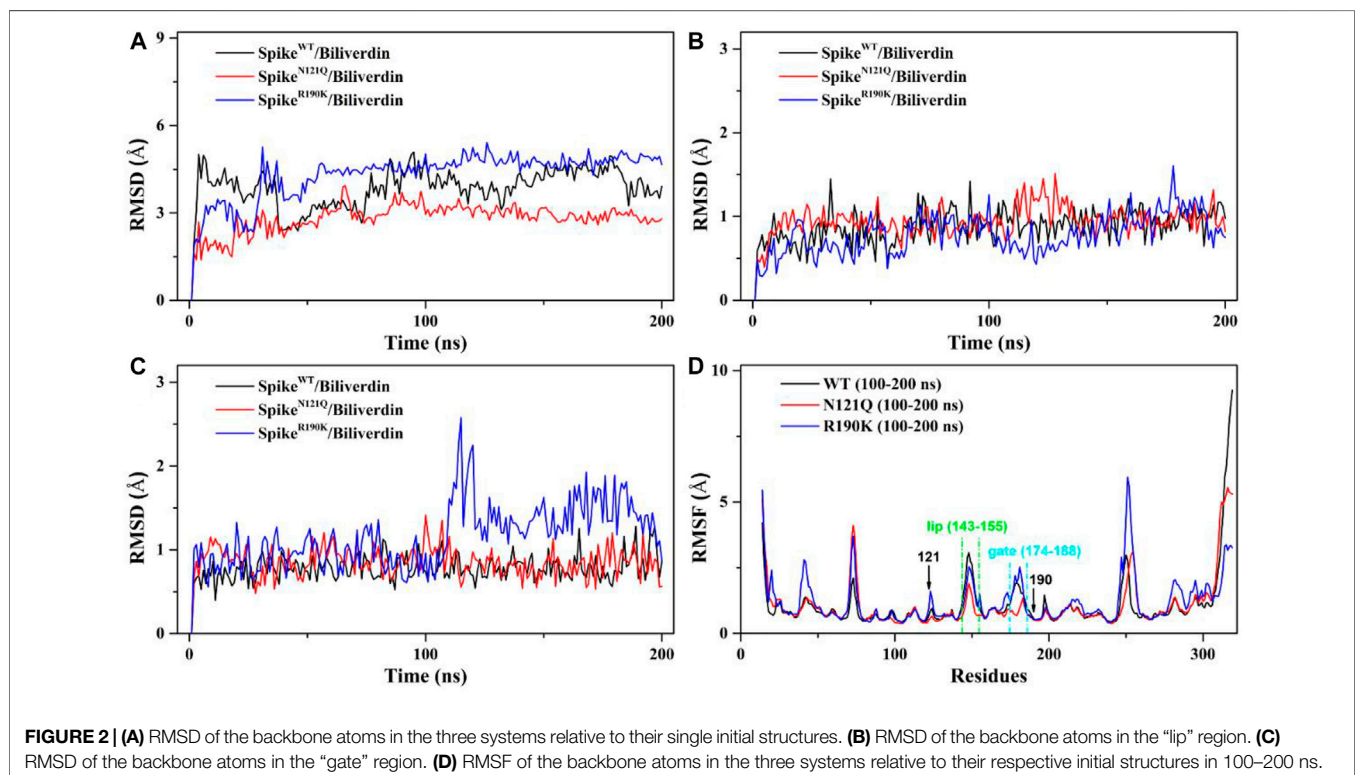
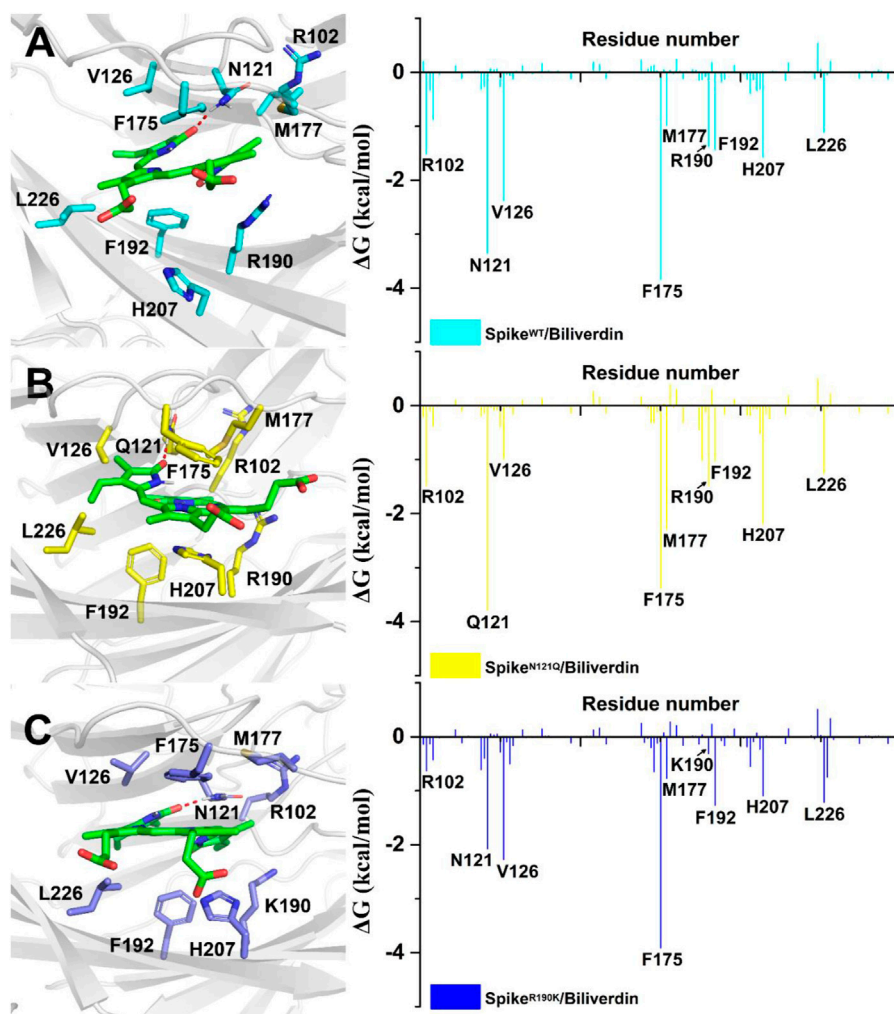


TABLE 1 | Binding free energies and individual energy terms of biliverdin in three systems calculated in MM/GBSA (kcal/mol).

System	Spike ^{WT} /Biliverdin	Spike ^{N121Q} /Biliverdin	Spike ^{R190K} /Biliverdin
ΔE_{vdw}	-49.74 ± 2.26	-51.13 ± 2.69	-47.41 ± 2.71
ΔE_{ele}	-93.82 ± 13.25	-116.72 ± 13.73	-79.64 ± 13.34
ΔG_{GB}	117.03 ± 11.08	139.95 ± 11.81	103.19 ± 11.71
ΔG_{SA}	-6.43 ± 0.45	-6.27 ± 0.68	-6.31 ± 0.54
$T\Delta S$	-14.35 ± 3.14	-15.72 ± 3.25	-14.78 ± 3.16
ΔG_{bind}	-18.60 ± 3.52	-18.45 ± 3.25	-15.39 ± 3.76
K_d (nM) (Rosa et al., 2021)	9.8	16,800	1,500

**FIGURE 3** | Residues around biliverdin and MM/GBSA decomposition results of the total binding free energies per residue for (A) Spike^{WT}/Biliverdin, (B) Spike^{N121Q}/Biliverdin, and (C) Spike^{R190K}/Biliverdin.

MD Simulation and MM/GBSA Free Energy Calculation

With the aid of AMBER 12 (Case et al., 2005; Han et al., 2021) software, molecule energy minimization and MD simulation of the three systems were performed to reach the most stable conformation and decrease atomic energy. Some missing

hydrogen atoms of the protein complex and ligand were added with the assistance of the tleap module in AMBER 12 software. The ligand biliverdin was minimized using the HF/6-31* optimization in the Gaussian 09 program (Frisch et al., 2016). In AMBER 12 software, the electrostatic potentials, derived from the restrained electrostatic potential (RESP) fitting

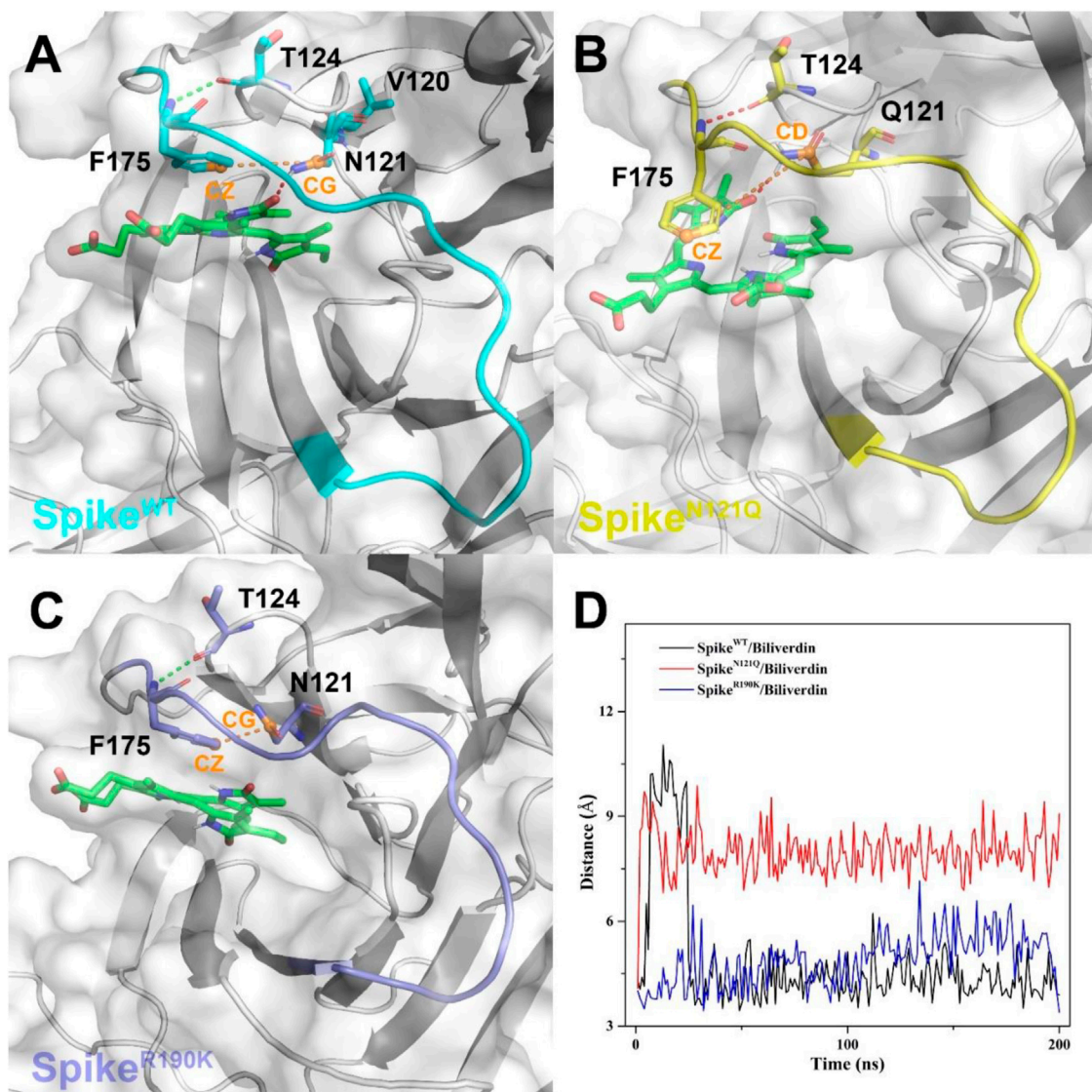


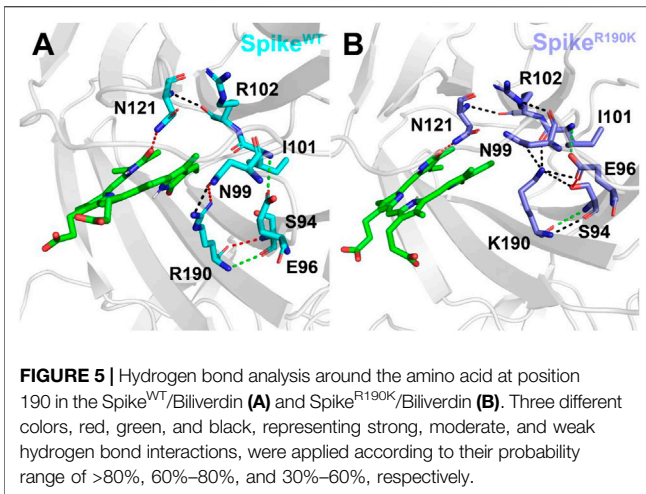
FIGURE 4 | The binding interaction between the SARS-CoV-2 spike and biliverdin in (A) Spike^{WT}/Biliverdin, (B) in Spike^{N121Q}/Biliverdin, and (C) in Spike^{R190K}/Biliverdin. (D) The distance between the CZ atom of Phe175 and the CG atom of Asn121 (or CD atom for the case of Gln121) was measured to reveal the deflection process of Phe175.

technique in the Gaussian program, created the partial charges. Generated by the antechamber module in AMBER 12 (Wang et al., 2006), the field parameters and the partial charges for ligand biliverdin were established. The parameter for biliverdin was set by employing the general AMBER force field (GAFF) (Wang et al., 2004), and the standard AMBER force field (ff03) (Duan et al., 2003) was applied to define protein parameters in the following MD simulation.

Forming a rectangular box of TIP3P, the water molecules which extended 12 Å away from any solute atoms wrapped up the three systems. Appropriate numbers of K⁺ were added to neutralize those systems. Every single system was initially energy minimized *via* three steps with the help of a sander module, as described in a previous study (Chen H. et al.,

2019). Then each system was heated gradually in the NVT ensemble from 0 to 300 K in 100 ps. Under a constant temperature of 300 K, a 200 ns MD simulation with a 2.0 fs time step was carried out for each system. Also, the SHAKE procedure was used to restrict all bonds involving at least one hydrogen atom. The particle mesh Ewald (PME) was used to deal with long-range electrostatic interactions. Concerning the following binding free energy calculation, the coordinates were saved every 100 ps in the sampling process.

With the aim of calculating the binding free energy, the calculation of MM/GBSA was carried out by using MM/GBSA in AMBER 12 software *via* the following equation (Kollman et al., 2000; Chen Q. et al., 2019; An et al., 2020; Shi et al., 2021).



$$\begin{aligned}\Delta G_{\text{bind}} &= G_{\text{complex}} - G_{\text{protein}} - G_{\text{ligand}} \\ &= \Delta E_{\text{MM}} + \Delta G_{\text{GB}} + \Delta G_{\text{SA}} - T\Delta S \\ &= \Delta E_{\text{vdw}} + \Delta E_{\text{ele}} + \Delta G_{\text{GB}} + \Delta G_{\text{SA}} - T\Delta S\end{aligned}$$

ΔE_{MM} , the gas-phase interaction energy, is composed of two components: ΔE_{vdw} (van der Waals energy) and ΔE_{ele} (electrostatic energy) (Genheden and Ryde, 2015). Both ΔG_{GB} and ΔG_{SA} mean the components of the desolvation free energy. While the former is polar, the latter one is nonpolar. The polar desolvation free energy, with dielectric constants of the solvent and the solute set to 80 and 1, respectively, was calculated by the GB models developed by Onufriev et al. (2004). Normal-mode analysis was applied in evaluating the entropy contribution to the binding free energy.

RESULTS AND DISCUSSION

Overall Structure and Dynamics

The root mean square deviation (RMSD) values of the whole protein backbone atoms were calculated to explore the conformation stability of three systems during the 200 ns MD simulation (Figure 2). The plot shows that all three systems reached equilibrium after 100 ns. The RMSD values of the Spike^{WT}/Biliverdin and Spike^{N121Q}/Biliverdin are around 3 Å, while that of the Spike^{R190K}/Biliverdin mutation is around 4 Å (Figure 2A). Two single-site mutations do not induce significant conformation change to the protein since the RMSD values of two mutated systems stabilize in a very short time. The previous study [5] revealed that the SARS-CoV-2 spike contains the motion of loop regions including “gate” (residues 174–188) and “lip” (residues 143–155). To further assess the dynamic changes of two regions, the same RMSD calculation was conducted. Compared with the lip loop (Figure 2B), residues in the gate take on intensively small volatility and maintain a lower RMSD value, mostly below 1.0 Å (Figure 2C). It is noticeable that Spike^{R190K}/Biliverdin fluctuates in both regions while Spike^{WT}/Biliverdin and Spike^{N121Q}/Biliverdin are quite stable. Three plots together proved that 200 ns was enough for three systems to relax.

We also employed the root mean fluctuation (RMSF) calculation of every amino acid residue based on a 100–200 ns MD trajectory to study the fluctuation of individual residues on the SARS-CoV-2 spike (Figure 2D). It can be clearly seen that the residues of Spike^{N121Q}/Biliverdin and Spike^{R190K}/Biliverdin have changed after comparative analysis. Especially at loop regions of the gate and lip, the amino acid residues of Spike^{N121Q}/Biliverdin fluctuated less obviously than did those of Spike^{WT}/Biliverdin and Spike^{R190K}/Biliverdin in two regions. In line with the above RMSD analyses, Spike^{R190K}/Biliverdin is the most unstable system as its RMSF value is much higher than the other two systems in most cases. However, for the lip loop as well as residues at positions 121 and 190, there is only a little difference in the fluctuation of the RMSF value in all three systems. Also, since the possible interaction with surrounding residues, the RMSF value of the gate loop in the Spike^{N121Q}/Biliverdin is lower than those of the other two systems. In total, it is reasonable to carry out the following binding free energy calculation and free energy decomposition analysis based on the last 100 ns trajectories.

Binding Free Energy Calculated by MM/GBSA

To analyze how the binding affinity between biliverdin and the SARS-CoV-2 spike is affected by two single-site mutations, MM/GBSA free energy calculation was conducted. According to Table 1, the binding free energy (ΔG_{bind}) of Spike^{WT}/Biliverdin (−18.60 kcal/mol) is stronger than that of Spike^{R190K}/Biliverdin (−15.39 kcal/mol), which is in line with the fact that the R190K mutant would increase the Km value from 9.8 to 1,500 nM (Rosa et al., 2021). Unexpectedly, the calculated ΔG_{bind} of Spike^{N121Q}/Biliverdin is −18.45 kcal/mol, comparable to the case of WT, which is different from the previous findings that the N121Q mutant is more sensitive to the binding interaction of biliverdin to SARS-CoV-2 spike (Rosa et al., 2021). With regard to the puzzling question, it will be discussed in the following binding mode analysis. Nevertheless, as for all three systems, the van der Waals interaction (ΔE_{vdw}) plays a crucial role in the total binding free energy. Since the electrostatic interaction (ΔE_{ele}) is completely counteracted by the polar desolvation energy (ΔG_{GB}), the net of electrostatic interactions ($\Delta E_{\text{ele}} + \Delta G_{\text{GB}}$) is even unfavorable to the binding affinities. Additionally, the differences in entropy contribution ($T\Delta S$) of the three systems are not obvious.

To further investigate detailed binding modes between biliverdin and the SARS-CoV-2 spike, MM/GBSA free energy decomposition analysis that decomposes the total binding free energies into ligand–residue pairs is carried out accordingly. Eight residues, Asn121, Val126, Phe175, Met177, Arg190, Phe192, His207, and Leu226, play a dominant role in biliverdin binding, among which the side chain of Asn121 forms a hydrogen bond with biliverdin (Figure 3A). However, although the amino acid in position 121 mutated from asparagine to glutamine in Spike^{N121Q}/Biliverdin, the side chain of Gln121 can still form a hydrogen bond with biliverdin although its contribution to the binding free energy changes slightly (Figure 3B). The binding pose of

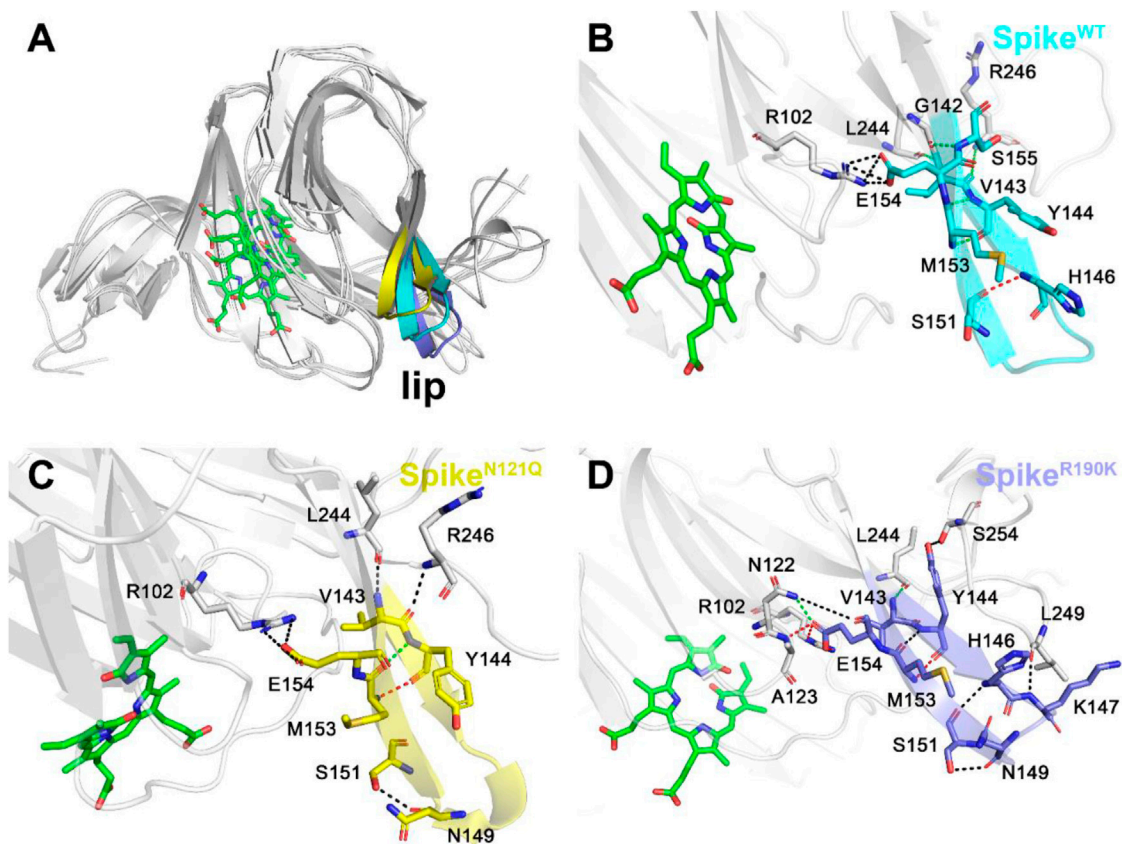


FIGURE 6 | Three different colors, red, green, and black, representing strong, moderate, and weak hydrogen bond interactions, were applied according to their probability range of >80%, 60%–80%, and 30%–60%, respectively. **(A)** Comparison of the lip loops in three systems. **(B)** Illustration of the hydrogen bond interaction of the lip in Spike^{WT}/Biliverdin, **(C)** in Spike^{N121Q}/Biliverdin, and **(D)** in Spike^{R190K}/Biliverdin.

biliverdin in the Spike^{N121Q}/Biliverdin changes greatly compared with the one in the Spike^{WT}/Biliverdin. To be specific, biliverdin tends to approach His207 and away from Val126 since the length of the side chain at position 121 would become longer when the asparagine mutates to glutamine. In order to form the hydrogen bond, the biliverdin extends outside and deflects accordingly, whose dynamic procedure will be introduced in the following hydrogen bond analysis on the “gate” loop. For the Spike^{R190K}/Biliverdin (Figure 3C), its binding mode is similar to the Spike^{WT}/Biliverdin as the same amino acids contribute to the binding free energy. However, the contributions of Lys190 and Arg102 drop from -1.37 to -0.31 kcal/mol and from -1.58 to -0.70 kcal/mol, respectively.

Uncover the Possible Molecular Mechanisms of Two Single-Site Mutations on the Binding Interaction Between the SARS-CoV-2 Spike and Biliverdin

To further investigate how the N121Q mutant affects biliverdin binding affinity, we draw attention to hydrogen bonds around residue 121 in all three systems. In Spike^{WT}/Biliverdin, one

hydrogen bond is formed between the residue Asn121 and biliverdin. The side chain benzene ring of Phe175 can have a relative harmony π - π stack interaction with biliverdin (Figure 4A). After residue Asn121 mutates into Gln121, however, as the residue Gln121 side-chain elongates while the hydrogen bond between Gln121 and the biliverdin remains, the steric hindrance is generated between the side chains of Gln121 and Phe175, which makes the Phe175 side chain benzene ring deflect outward (Figure 4B). With two carboxyl groups exposed to the solvent, the biliverdin tends to move outward as the entire gate loop moves downward, narrowing the biliverdin binding pocket. Similar to that in Spike^{WT}/Biliverdin, the gate loop does not change since the side chain benzene ring of residue Phe175 does not deflect in Spike^{R190K}/Biliverdin (Figure 4C). To further verify the deflection process, we measure the distances between the CZ atom of Phe175 and the CG atom of Asn121 (or CD atom for the case of Gln121) (Figure 4D). It is not hard to find that for Spike^{N121Q}/Biliverdin, the distance increased to 7.5 Å in a short time while the distances of the other two systems remained at 4.5 Å, which means that the deflection of Phe175 occurred very soon after the beginning of the dynamics simulation. The N121Q mutant likely leads the biliverdin binding as a differential pose

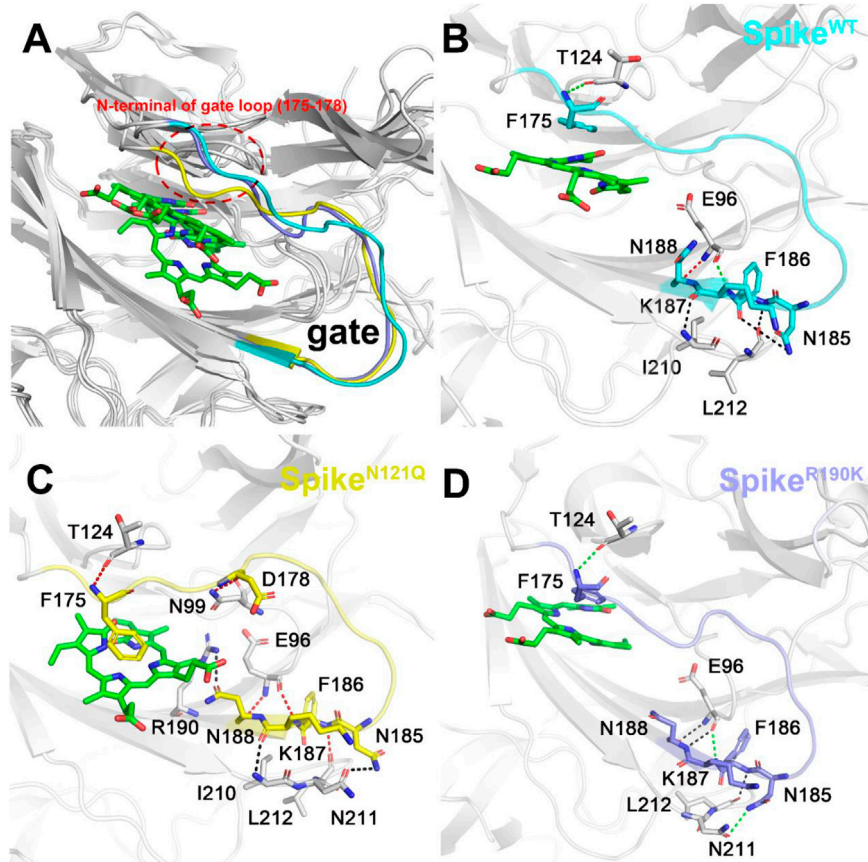


FIGURE 7 | Three different colors, red, green, and black, representing strong, moderate, and weak hydrogen bond interactions, were applied according to their probability range of >80%, 60%–80%, and 30%–60%, respectively. **(A)** Comparison of the gate loops in three systems. Illustration of hydrogen bond interaction of the lip **(B)** in Spike^{WT}/Biliverdin, **(C)** in Spike^{N121Q}/Biliverdin, and **(D)** in Spike^{R190K}/Biliverdin.

when concerning the wild-type state. In other words, to counteract the unfavorable factor from the deflection of Phe175, compound biliverdin tends to be in an inactive state by induced fit. In turn, the induced-fit effect would result in the irreducible binding affinity of biliverdin to the SpikeN121Q protein.

As mentioned above, the mutation of R190K reduced the affinities of the biliverdin to Arg102 and Lys190. To explore the molecular mechanism, we also conducted a detailed analysis of hydrogen bonds near the amino acid at position 190, and the results are shown in **Figure 5**. It can be seen that in Spike^{WT}/Biliverdin, Arg190 can form four hydrogen bonds with Asn99 and Ser94. Meanwhile, Asn121 forms one hydrogen bond with Arg102 in addition to one hydrogen bond with the compound biliverdin. According to **Figure 5A**, these two hydrogen bond networks are independent and unrelated, and the larger cavity enables the biliverdin to be deeply embedded and bound stably. In Spike^{R190K}/Biliverdin (**Figure 5B**), however, Lys190 could be simultaneously integrated with Ser94, Glu96, Asn99, and Ile101, forming six hydrogen bonds in total. While the hydrogen bond between Asn121 and Arg102 continues to

exist, Arg102 also forms hydrogen bonds with Asn99. Guided by Asn99 and Ile101, Lys190 is brought closer to Arg102 and Asn121, which leads to the narrow cavity formed by these amino acids and is not conducive to the compound biliverdin insertion, widening the distances between biliverdin and Ile101 as well as Lys190 and reducing their respective energy contributions.

Effect of Two Single-Site Mutations on the Conformations of “Lip” and “Gate” by Hydrogen Bond Analysis

Previous research (Rosa et al., 2021) has already validated that the antibody binding to the SARS-CoV-2 spike NTD is inhibited by biliverdin *via* an allosteric mechanism which is associated with two loop regions of “lip” and “gate”. It will be meaningful to study the possible conformation of both the two loop regions in Spike^{WT}/Biliverdin and the two mutants. Thus, hydrogen bond analyses were conducted on the two loops based on the last 100 ns trajectories (**Supplementary Table S1**). The cutoff length value of forming hydrogen bonds is set at 3 Å. For a clearer observation, three different colors,

red, green, and black, representing strong, moderate, and weak hydrogen bond interactions, respectively, were applied according to their probability range of >80%, 60%–80%, and 30%–60%, respectively.

After observing three systems overlapping together (**Figure 6A**), it can be concluded that two single-site mutations have posed significant changes to the lip region. As for the lip of Spike^{WT}/Biliverdin (**Figure 6B** and **Supplementary Figure S1A**), one strong hydrogen bond is formed between His146 and Ser151, and five moderate hydrogen bonds are formed between Met153 and Tyr144 (two hydrogen bonds are formed), Gly142 and Ser155, Val143 and Arg246, and Leu244 and Val143 while Arg102 and Glu154 form four weak hydrogen bonds. Concerning Spike^{N121Q}/Biliverdin (**Figure 6C**), the advent of residues Leu244 and Arg246 resulted in a new hydrogen bond formed each with Val143 and Tyr144, leading to the loop₂₄₄₋₂₆₁ approaching the lip (**Figure 6C** and **Supplementary Figure S1B**). Moreover, compared with Spike^{N121Q}/Biliverdin, new hydrogen bonds form between Asn122 and Glu154, Ala123 and Glu154, Ser254 and Tyr144, and Leu249 and Lys147 in the Spike^{R190K}/Biliverdin (**Figure 6D**); it reveals that not only loop₂₄₄₋₂₆₁ but also loop₁₂₂₋₁₂₃ interacts with the lip (**Supplementary Figure S1C**). The stabilized lip is likely to exist in two mutated systems, which would restrain the meaningful upward movement of the lip as discussed in previous literature (Rosa et al., 2021), ultimately interfering with the antibody fixing on the SARS-CoV-2 spike.

The possible conformation changes of the gate loop derived from two mutants were also studied (**Figure 7**). It is not hard to find that the N121Q mutation had the N-terminal residues of the gate (residues 175 to 178) dramatically changed in contrast to the other two systems (**Figure 7A**). The hydrogen bond network of the Spike^{WT}/Biliverdin system is characterized by one strong hydrogen bond formed between Asn188 and Glu96, two moderate ones formed between Glu96 and Lys187 and between Phe175 and Thr124, and another three weak hydrogen bonds (**Figure 7B**). Intriguingly, a strong Asn99–Asp178 hydrogen bond is formed in Spike^{N121Q}/Biliverdin (**Figure 7C**), which triggers the gate loop approaching and narrowing the binding pocket of the biliverdin. As for Spike^{R190K}/Biliverdin (**Figure 7D**), an additional moderate hydrogen bond is formed by Asn185 and Asn211 to stabilize the similar gate conformation in Spike^{WT}/Biliverdin.

CONCLUSION

The global spread of the novel coronavirus disease 2019 (COVID-19), one of the deadliest pandemics in modern history, has been unprecedented since it first emerged. The SARS-CoV-2 spike is currently the main target of antibody design. Previous studies have reported the crystal structure of the SARS-CoV-2 spike and biliverdin and have come up with the idea that two single-site mutations, R190K and N121Q, will weaken the binding affinity of the biliverdin although the potential molecular mechanism is still unknown. As a result, this project studied the WT and two other mutants of the SARS-CoV-2 spike, employing MD simulation and MM/

GBSA free energy calculation. Our simulations confirmed that the R190K mutation causes amino acid 190 to form six hydrogen bonds with surrounding residues Ser94, Glu96, Asn99, and Ile101, which, guided by Asn99 and Ile101, brings Lys190 closer to Arg102 and Asn121, thereby weakening the interaction energy between biliverdin and Ile101 as well as Lys190. However, in the case of N121Q mutation, although Q121 still maintained hydrogen bond interaction with biliverdin, the overall binding mode deviated significantly under the reversal of the benzene ring of the Phe175 side chain compared with WT. Moreover, we found that R190K and N121Q mutants would stabilize the lip loop. The stabilized lip was likely to exist in two mutated systems, which would restrain the meaningful upward movement of the lip, ultimately interfering with the antibody fixing on the SARS-CoV-2 spike. In addition, N121Q significantly promoted the gate loop deviating to the biliverdin binding site and compressed the site so that biliverdin could not maintain the binding mode of WT. However, the R190K mutation had little influence on the gate loop and could also make the structural change of the gate loop similar to WT. Studying the interaction between the SARS-CoV-2 spike and the biliverdin and two single-site mutations using simulation methods, we hope, will support future antibody research targeting the SARS-CoV-2 spike.

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

Conceptualization, JG and ZQ; methodology, ZQ, KL and XG; data curation, ZQ and XG; writing—original draft preparation, ZQ and BH; writing—review and editing, JG and ZQ; visualization, ZQ; project administration, JG; funding acquisition, JG. All authors have read and agreed to the published version of the manuscript.

FUNDING

This research was funded by the Natural Science Foundation of Jiangsu Province (grant number BK20201157), the Six Talent Peaks Project in Jiangsu Province (grant number YY-046), and the Qinglan Project of Jiangsu Province of China to JG.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2021.791885/full#supplementary-material>

REFERENCES

- An, Y., Meng, C., Chen, Q., and Gao, J. (2020). Discovery of Small Molecule Sirt1 Activator Using High-Throughput Virtual Screening, Molecular Dynamics Simulation, Molecular Mechanics Generalized Born/surface Area (MM/GBSA) Calculation, and Biological Evaluation. *Med. Chem. Res.* 29, 255–261. doi:10.1007/s00044-019-02479-2
- Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., et al. (2005). The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* 26, 1668–1688. doi:10.1002/jcc.20290
- Chen, H., Wang, Y., Gao, Z., Yang, W., and Gao, J. (2019a). Assessing the Performance of Three Resveratrol in Binding with SIRT1 by Molecular Dynamics Simulation and MM/GBSA Methods: the Weakest Binding of Resveratrol 3 to SIRT1 Triggers a Possibility of Dissociation from its Binding Site. *J. Comput. Aided Mol. Des.* 33, 437–446. doi:10.1007/s10822-019-00193-0
- Chen, Q., Wang, Y., Shi, S., Li, K., Zhang, L., and Gao, J. (2019b). Insights into the Interaction Mechanisms of the Proviral Integration Site of Moloney Murine Leukemia Virus (Pim) Kinases with Pan-Pim Inhibitors PIM447 and AZD1208: A Molecular Dynamics Simulation and MM/GBSA Calculation Study. *Int. J. Mol. Sci.* 20, 5410. doi:10.3390/ijms20215410
- Chi, X., Yan, R., Zhang, J., Zhang, G., Zhang, Y., Hao, M., et al. (2020). A Neutralizing Human Antibody Binds to the N-Terminal Domain of the Spike Protein of SARS-CoV-2. *Science* 369, 650–655. doi:10.1126/science.abc6952
- Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., et al. (2003). A point-charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-phase Quantum Mechanical Calculations. *J. Comput. Chem.* 24, 1999–2012. doi:10.1002/jcc.10349
- Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2016). *Expanding the Limits of Computational Chemistry*. Wallingford Ct: Gaussian, Inc.
- Genheden, S., and Ryde, U. (2015). The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities. *Expert Opin. Drug Discov.* 10, 449–461. doi:10.1517/17460441.2015.1032936
- Graham, C., Seow, J., Huettner, I., Khan, H., Kouphou, N., Acors, S., et al. (2021). Neutralization Potency of Monoclonal Antibodies Recognizing Dominant and Subdominant Epitopes on SARS-CoV-2 Spike Is Impacted by the B.1.1.7 Variant. *Immunity* 54, 1276–1289. e1276. doi:10.1016/j.immuni.2021.03.023
- Han, D., Wang, H., Wujieti, B., Zhang, B., Cui, W., and Chen, B.-Z. (2021). Insight into the Drug Resistance Mechanisms of GS-9669 Caused by Mutations of HCV NS5B Polymerase via Molecular Simulation. *Comput. Struct. Biotechnol. J.* 19, 2761–2774. doi:10.1016/j.csbj.2021.04.026
- Hu, B., Guo, H., Zhou, P., and Shi, Z.-L. (2021). Characteristics of SARS-CoV-2 and COVID-19. *Nat. Rev. Microbiol.* 19, 141–154. doi:10.1038/s41579-020-00459-7
- Ju, B., Zhang, Q., Ge, J., Wang, R., Sun, J., Ge, X., et al. (2020). Human Neutralizing Antibodies Elicited by SARS-CoV-2 Infection. *Nature* 584, 115–119. doi:10.1038/s41586-020-2380-z
- Ke, Z., Oton, J., Qu, K., Cortese, M., Zila, V., Mckean, L., et al. (2020). Structures and Distributions of SARS-CoV-2 Spike Proteins on Intact Virions. *Nature* 588, 498–502. doi:10.1038/s41586-020-2665-2
- Kemp, S. A., Collier, D. A., Datir, R. P., Ferreira, I. A. T. M., Gayed, S., Jahun, A., et al. (2021). SARS-CoV-2 Evolution during Treatment of Chronic Infection. *Nature* 592, 277–282. doi:10.1038/s41586-021-03291-y
- Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., et al. (2000). Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* 33, 889–897. doi:10.1021/ar000033j
- Liu, L., Wang, P., Nair, M. S., Yu, J., Rapp, M., Wang, Q., et al. (2020). Potent Neutralizing Antibodies against Multiple Epitopes on SARS-CoV-2 Spike. *Nature* 584, 450–456. doi:10.1038/s41586-020-2571-7
- Onufriev, A., Bashford, D., and Case, D. A. (2004). Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model. *Proteins* 55, 383–394. doi:10.1002/prot.20033
- Rosa, A., Pye, V. E., Graham, C., Muir, L., Seow, J., Ng, K. W., et al. (2021). SARS-CoV-2 Can Recruit a Heme Metabolite to Evade Antibody Immunity. *Sci. Adv.* 7, eabg7607. doi:10.1126/sciadv.abg7607
- Shi, S., Wang, Q., Liu, S., Qu, Z., Li, K., Geng, X., et al. (2021). Characterization the Performances of Twofold Resveratrol Integrated Compounds in Binding with SIRT1 by Molecular Dynamics Simulation and Molecular Mechanics/generalized Born Surface Area (MM/GBSA) Calculation. *Chem. Phys.* 544, 111108. doi:10.1016/j.chemphys.2021.111108
- Tegally, H., Wilkinson, E., Lessells, R. J., Giandhari, J., Pillay, S., Msomi, N., et al. (2021). Sixteen Novel Lineages of SARS-CoV-2 in South Africa. *Nat. Med.* 27, 440–446. doi:10.1038/s41591-021-01255-3
- Tregoning, J. S., Brown, E. S., Cheeseman, H. M., Flight, K. E., Higham, S. L., Lemm, N. M., et al. (2020). Vaccines for COVID-19. *Clin. Exp. Immunol.* 202, 162–192. doi:10.1111/cei.13517
- Walls, A. C., Park, Y.-J., Tortorici, M. A., Wall, A., McGuire, A. T., and Veesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* 181, 281–292. doi:10.1016/j.cell.2020.02.058
- Wang, J., Wang, W., Kollman, P. A., and Case, D. A. (2006). Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graphics Model.* 25, 247–260. doi:10.1016/j.jmkgm.2005.12.005
- Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004). Development and Testing of a General Amber Force Field. *J. Comput. Chem.* 25, 1157–1174. doi:10.1002/jcc.20035
- Wang, M.-Y., Zhao, R., Gao, L.-J., Gao, X.-F., Wang, D.-P., and Cao, J.-M. (2020). SARS-CoV-2: Structure, Biology, and Structure-Based Therapeutics Development. *Front. Cel. Infect. Microbiol.* 10, 587269. doi:10.3389/fcimb.2020.587269
- Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C.-L., Abiona, O., et al. (2020). Cryo-EM Structure of the 2019-nCoV Spike in the Prefusion Conformation. *Science* 367, 1260–1263. doi:10.1126/science.abb2507
- Zost, S. J., Gilchuk, P., Case, J. B., Binshtein, E., Chen, R. E., Nkolola, J. P., et al. (2020). Potently Neutralizing and Protective Human Antibodies against SARS-CoV-2. *Nature* 584, 443–449. doi:10.1038/s41586-020-2548-6

Conflict of Interest: JG was employed by the company Xuzhou Medical University Technology Transfer Center Co., Ltd.

The remaining authors declare 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.

Copyright © 2021 Qu, Li, Geng, Huang and Gao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.



A Computer-Aided Approach for the Discovery of D-Peptides as Inhibitors of SARS-CoV-2 Main Protease

Jorge E. Hernández González^{1,2}, Raphael J. Eberle^{3,4}, Dieter Willbold^{3,4,5} and Mônica A. Coronado^{3*}

¹Multisuser Center for Biomolecular Innovation, IBILCE, Universidade Estadual Paulista (UNESP), São Jose do Rio Preto, Brazil, ²Laboratory for Molecular Modeling and Dynamics, Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, Cidade Universitária Ilha do Fundão, Rio de Janeiro, Brazil, ³Institute of Biological Information Processing (IBI-7, Structural Biochemistry), Forschungszentrum Jülich, Jülich, Germany, ⁴Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, Universitätsstraße, Düsseldorf, Germany, ⁵JuStruct: Jülich Centre for Structural Biology, Forschungszentrum Jülich, Jülich, Germany

OPEN ACCESS

Edited by:

Mattia Falconi,
University of Rome Tor Vergata, Italy

Reviewed by:

Aditya K. Padhi,
RIKEN Yokohama, Japan
K. Venkateswara Swamy,
MIT School of Bioengineering
Sciences and Research, India

*Correspondence:

Mônica A. Coronado
m.coronado@fz-juelich.de

Specialty section:

This article was submitted to
Biological Modeling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 16 November 2021

Accepted: 30 December 2021

Published: 24 January 2022

Citation:

Hernández González JE, Eberle RJ,
Willbold D and Coronado MA (2022) A
Computer-Aided Approach for the
Discovery of D-Peptides as Inhibitors
of SARS-CoV-2 Main Protease.
Front. Mol. Biosci. 8:816166.
doi: 10.3389/fmolb.2021.816166

The SARS-CoV-2 main protease, also known as 3-chymotrypsin-like protease (3CL^{Pro}), is a cysteine protease responsible for the cleavage of viral polyproteins pp1a and pp1ab, at least, at eleven conserved sites, which leads to the formation of mature nonstructural proteins essential for the replication of the virus. Due to its essential role, numerous studies have been conducted so far, which have confirmed 3CL^{Pro} as an attractive drug target to combat Covid-19 and have reported a vast number of inhibitors and their co-crystal structures. Despite all the ongoing efforts, D-peptides, which possess key advantages over L-peptides as therapeutic agents, have not been explored as potential drug candidates against 3CL^{Pro}. The current work fills this gap by reporting an *in silico* approach for the discovery of D-peptides capable of inhibiting 3CL^{Pro} that involves structure-based virtual screening (SBVS) of an *in-house* library of D-tripeptides and D-tetrapeptides into the protease active site and subsequent rescoring steps, including Molecular Mechanics Generalized-Born Surface Area (MM-GBSA) free energy calculations and molecular dynamics (MD) simulations. *In vitro* enzymatic assays conducted for the four top-scoring D-tetrapeptides at 20 μ M showed that all of them caused 55–85% inhibition of 3CL^{Pro} activity, thus highlighting the suitability of the devised approach. Overall, our results present a promising computational strategy to identify D-peptides capable of inhibiting 3CL^{Pro}, with broader application in problems involving protein inhibition.

Keywords: SARS-CoV-2, 3CL^{Pro}, D-peptide, virtual screening, molecular dynamics simulation

INTRODUCTION

Covid-19 is a pandemic disease caused by the novel acute respiratory syndrome coronavirus 2 (SARS-CoV-2). As of December 12th, 2021, over 269 million confirmed Covid-19 cases and 5.3 million related deaths had been reported since the start of the pandemic (World Health Organization, 2021). SARS-CoV-2, together with SARS-CoV and Middle East Respiratory Syndrome (MERS) coronaviruses responsible for two significant outbreaks during the current century, are enveloped and single-stranded RNA viruses (Payne, 2017; Wu F. et al., 2020; Wang et al., 2020). During their replication, coronaviruses encode several accessory proteins and two replicase polyproteins (pp1a and pp1ab) (Marra et al., 2003; Rota et al., 2003; Ziebuhr, 2005; Wu F. et al., 2020; Yan and Wu, 2021), which are proteolytically processed by two cysteine proteases, i.e., the papain-

like protease (PL^{Pro}) and the main protease, also called 3-chymotrypsin-like protease (3CL^{Pro}). The latter cleaves the pp1a and pp1b at 11 conserved sites by recognizing the XXXLQAXXX and XXXLQSXXX sequence motifs, thus generating nonstructural proteins (NSPs) essential for the viral replication (Gorbalenya et al., 1989; Hegyi and Ziebuhr, 2002; Kiemer et al., 2004; Yan and Wu, 2021).

The essential role played by SARS-CoV-2 3CL^{Pro} during the viral replication has encouraged the search for anti-Covid drugs targeting this protease. Numerous potent orthosteric inhibitors of 3CL^{Pro}, most of them of peptide-based or peptidomimetic nature, have been reported so far (Amin et al., 2021; Chia et al., 2021; Sabbah et al., 2021; Yan and Gao, 2021). These compounds have shown significant inhibitory activity, not only against the protease but also against the viral replication in cell cultures. Meanwhile, the crystal structures of 3CL^{Pro} in complex with a myriad of inhibitors and compound fragments have been deposited in the Protein Data Bank (PDB) and provide useful structural information for the rational design of new drugs (Mengist et al., 2021). Some of these structures have revealed the existence of allosteric binding sites in the surface of 3CL^{Pro}, which can also be exploited to search for noncompetitive inhibitors (Douangamath et al., 2020; Gunther et al., 2021). More recently, the 3CL^{Pro} peptidomimetic inhibitor PF-07321332 has shown promising results in phase I clinical trials, thus paving the way toward the discovery of an effective antiviral (Owen et al., 2021). All these results underscore the importance of 3CL^{Pro} as an attractive drug target to combat Covid-19.

Even though diverse scaffolds of 3CL^{Pro} inhibitors have been identified, D-peptides remain unexplored. These molecules are made up of D-amino acids, i.e., amino acids whose chiral Ca atoms have the opposite stereochemical configuration to that observed in the amino acids that commonly form the natural proteins (the L-amino acids). This structural feature endows D-peptides with key advantages over the L-peptides, such as higher stability to proteolysis, improved intestinal absorption upon oral administration, and low or missing immunogenicity. These properties, along with others shared with L-peptides, e.g., lower manufacturing costs and higher binding affinity and specificity for the target receptors in comparison with small molecules, make D-peptides attractive therapeutic agents (Wiesehan and Willbold, 2003; Funke and Willbold, 2009; Sun et al., 2012; Liu et al., 2016; Garton et al., 2018). Remarkably, α -helical D-peptides designed *in silico* were reported to block the binding of the SARS-CoV-2 spike protein receptor-binding domain (RBD) to the human angiotensin-converting enzyme 2 (ACE2), the molecule that mediates the virus internalization into human cells, thus leading to the inhibition of viral infection *in vitro* (Valiente et al., 2021). The previous results provide an excellent example of the use of D-peptides as promising anti-Covid drug candidates.

Structure-based virtual screening (SBVS) of diverse ligand databases, many of them containing drug repurposing candidates and natural products, has been extensively applied to identify potential 3CL^{Pro} inhibitors (Wu C. et al., 2020; Chowdhury et al., 2020; Jukic et al., 2020; Meyer-Almes, 2020; Olubiyi et al., 2020; Selvaraj et al., 2020; Federico et al., 2021; Gogoi et al., 2021; Guedes et al., 2021; Kumar et al., 2021; Lokhande et al., 2021; Naik et al., 2021; Rajpoot et al., 2021;

Rehman et al., 2021; Sisakht et al., 2021). In several cases, this approach has led to the successful identification of compounds displaying *in vitro* inhibitory activity against 3CL^{Pro} (Ghahremanpour et al., 2020; Gupta et al., 2020; Jin et al., 2020; Li et al., 2020; Alves et al., 2021; Banerjee et al., 2021; Gunther et al., 2021; Guo et al., 2021; Gupta et al., 2021; Hamdy et al., 2021; Pathak et al., 2021; Yang et al., 2021). On the other hand, protein-peptide docking remains far more challenging compared to other small molecules due to the higher flexibility of peptides (Rentzsch and Renard, 2015; Ciemny et al., 2018; Hashemi et al., 2021). Nonetheless, at least one work has reported two L-pentapeptides as potential 3CL^{Pro} inhibitors by screening a 70,000-peptide library (Porto, 2021), using AutoDock Vina for the docking simulations (Trott and Olson, 2010). Remarkably, AutoDock Vina outperformed other freely-available docking algorithms, such as AutoDock and ZDOCK (Chen and Weng, 2002; Morris et al., 2009), in a benchmark study that presented a pipeline for peptide SBVS (Ansar and Vetrivel, 2019).

Encouraged by the previous findings, this study presents D-peptides as 3CL^{Pro} inhibitors. The computational workflow employed for D-peptide identification, which will be fully described in Materials and Methods, selects the best binders to the protease active site through SBVS and a series of rescoring steps combining Molecular Mechanics Generalized-Born Surface Area (MM-GBSA) free energy calculations (Gohlke et al., 2003; Gohlke and Case, 2004; Miller et al., 2012) and molecular dynamics (MD) simulations (Hou et al., 2011; Hernandez Gonzalez et al., 2021). The four top-ranked D-peptides were purchased and tested *in vitro* to evaluate their inhibitory activity against 3CL^{Pro}. Remarkably, all the tested D-peptides caused 3CL^{Pro} inhibition at 20 μ M during primary assays, resulting in up to 85% loss of proteolytic activity in certain cases. Therefore, the devised workflow led to promising results potentially extensible to broader applications related to protein inhibition.

MATERIALS AND METHODS

Preparation of the Protein Structure for Virtual Screening and MD Simulations

The crystal structure of free 3CL^{Pro} (PDB: 6Y2E, resolution 1.75 Å) (Zhang et al., 2020b) was chosen to conduct SBVSs and MD simulations with identified D-peptides. Protonation at pH 7.2 was performed using the PDB2PQR Web Server (<https://server.poissonboltzmann.org/pdb2pqr>) (Dolinsky et al., 2007). The protonated structure was then converted into the pdbqt file required for SBVS with the program *prepare_receptor4.py* of AutoDockTools 4 (Morris et al., 2009). MD simulations of the 3CL^{Pro}/peptide complexes were also performed using the predicted protonation states of the ionizable protein residues.

Building an In-House Library of D-Tripeptides and D-Tetrapeptides

An in-house library of capped D-tripeptides and D-tetrapeptides (for brevity's sake the term "capped" will be omitted hereinafter when referring to the D-peptides) was built using *tleap* of

Amber20 (Case et al., 2020). Briefly, *tleap* was called inside three or four nested loops, depending on the peptide length, each iterating over all the different amino acids. For HIS, its two different neutral tautomers were considered, thus totalizing 21 residue types. The *sequence* command of *tleap* was employed to create each peptide, with acetyl (ACE) and N-methyl amide (NME) capping groups being added at the N- and C-termini, respectively. The default L configuration of the C α atoms was then inverted to D configuration using the *flip* command of *tleap* (Case et al., 2020). Of note, the chiral centres of ILE and THR side-chains were not inverted, thus being modeled as D-allo-isoleucine and D-allo-threonine diastereomers, which will be referred to as ILE and THR hereinafter. The D-peptides were embedded into TIP3P octahedral solvation boxes, with edges spanning at least 10 Å from the solute surface, and counter-ions (Na⁺) were added to neutralize the system net charge. Topology and coordinate files for every solvated D-peptide were finally generated and saved for subsequent steps.

The systems were subjected to two rounds of energy minimization (EM) using *pmemd.MPI* of Amber20 in order to obtain a suitable conformation of each D-peptide in solution (Case et al., 2020). The first EM step consisted of 500 cycles of steepest descents (SD) followed by 500 cycles of conjugate gradient (CG) minimization, and both were carried out in the presence of harmonic restraints applied to the D-peptide heavy atoms ($k = 10 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^2$). The second EM was performed with no harmonic restraints and, as before, involved 500 cycles of SD followed by 500 cycles of CG minimization. The energy-minimized D-peptides were then stripped off the solvent and counter-ions, saved as *pdb* files with *cpptraj* of Amber20 (Case et al., 2020), and converted into *pdbqt* files with *prepare_ligand4.py* of AutoDock Tools 4 (Morris et al., 2009). This step completed the preparation of the peptide library containing 9,261 D-tripeptides and 194,481 D-tetrapeptides. Libraries of larger peptides were not prepared, as docking algorithms tend to produce less accurate results for molecules bearing many freely-rotatable bonds (Rentzsch and Renard, 2015; Ciemny et al., 2018; Hashemi et al., 2021).

Structure-based Virtual Screening

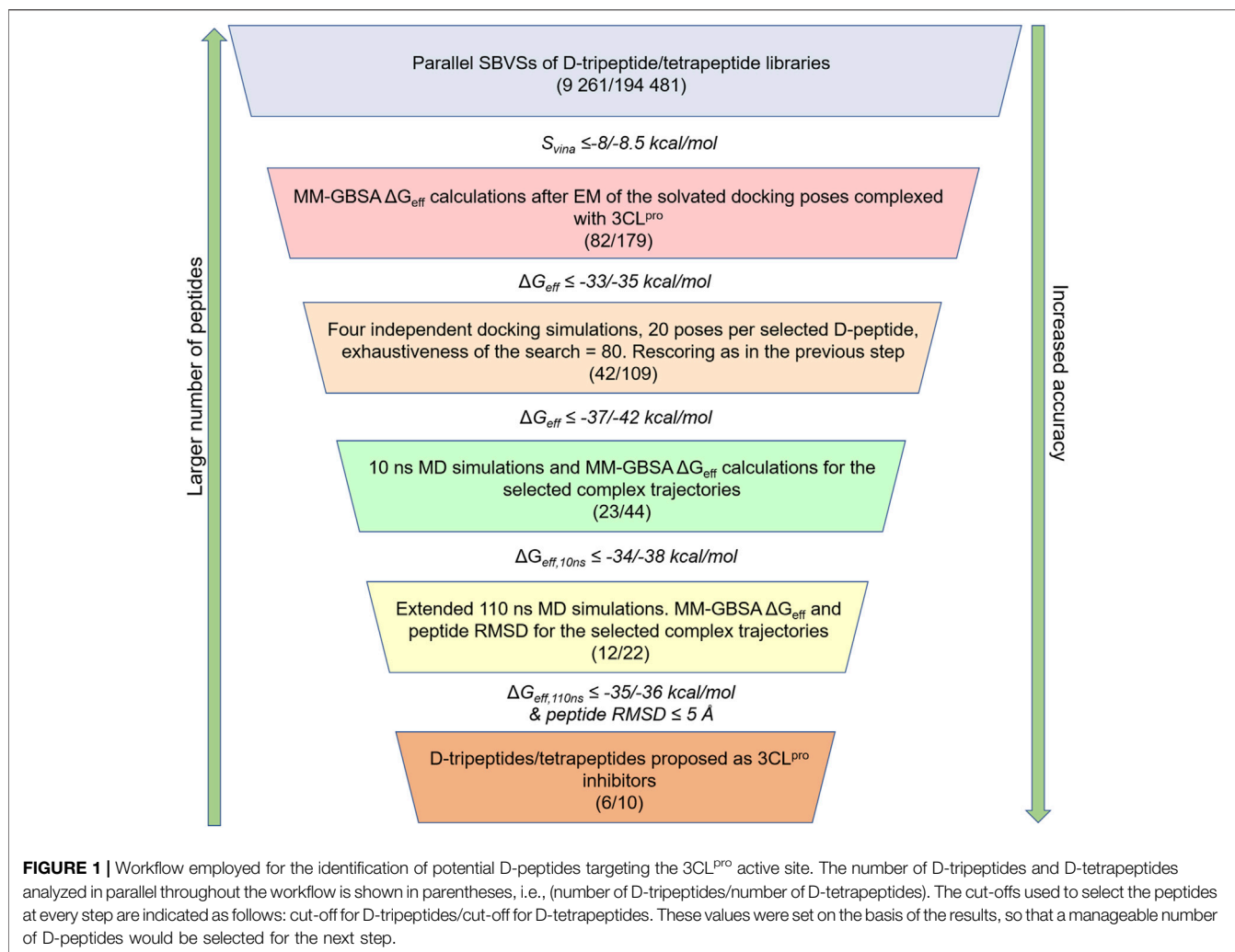
Despite 3CL^{pro} being a homodimer in solution, a monomer was chosen for SBVS and subsequent post-docking rescoring steps, as each active site in the functional homodimer is formed by residues belonging to an individual chain. A $19.5 \text{ \AA} \times 18.0 \text{ \AA} \times 22.0 \text{ \AA}$ box spanning the whole active site of 3CL^{pro} (PDB: 6Y2E) was then built using the Autodock/Vina plugin of Pymol (Supplementary Figure S1), and docking of D-peptides was performed with AutoDock Vina v1.12 (DeLano, 2002; Seeliger and de Groot, 2010; Trott and Olson, 2010). Default parameters, i.e., 9 poses per ligand, the exhaustiveness of the search equal to 8, and an energy difference of 3 kcal/mol between the best and worst poses were set for the docking simulations during SBVSs. D-tripeptide and D-tetrapeptide libraries were screened and ranked separately to reduce the impact of ligand-size bias (Chang et al., 2010). Based on the obtained AutoDock Vina scores (S_{vina}), the 82 and 179 top-ranked D-tripeptides and D-tetrapeptides, respectively, were selected for subsequent rescoring steps (Figure 1).

Post-docking Rescoring Steps

The selected D-tripeptides and D-tetrapeptides underwent in parallel a series of rescoring steps involving MM-GBSA free energy calculations (Gohlke et al., 2003; Gohlke and Case, 2004; Miller et al., 2012; Case et al., 2020), docking with increased exhaustiveness of the search and MD simulations (Figure 1). First, each of the nine docking poses per selected D-peptide was rescored based on their MM-GBSA effective free energies (ΔG_{eff}), calculated for the energy-minimized solvated complexes. The D-peptides were then re-ranked according to the ΔG_{eff} values of their respective lowest-energy poses. The best D-peptides were then docked again into the 3CL^{pro} active site with AutoDock Vina. For this step, the exhaustiveness of the search was increased to 80, and four independent and randomly-initialized docking simulations, each generating 20 different poses (80 poses in total), were run for every D-peptide. As before, lowest-energy poses were identified by rescoring the 80 docking poses per ligand using their corresponding MM-GBSA ΔG_{eff} values after EM of the 3CL^{pro}/peptide complexes. Finally, 10 ns of MD simulations, subsequently extended to 110 ns for the best candidates, were conducted, and average MM-GBSA ΔG_{eff} values were calculated for the generated trajectories ($\Delta G_{\text{eff},10\text{ns}}$ and $\Delta G_{\text{eff},110\text{ns}}$, respectively). The D-peptides with average $\Delta G_{\text{eff},110\text{ns}}$ and Root Mean Squared Deviation (RMSD) values (see *Trajectory Analyses* section) lower than the indicated cut-offs were proposed as potential 3CL^{pro} inhibitors (Figure 1).

MD Simulations Setup

The 3CL^{pro}/peptide complexes were parametrized with Amber ff14SB force-field (Maier et al., 2015) and embedded in octahedral simulation boxes with edges spanning, at least, 10 Å away from the solute surface, filled with TIP3P waters (Price and Brooks, 2004) and sufficient Na⁺ counter-ions to neutralize the systems (Supplementary Table S1) (Li et al., 2015). All previous steps were carried out with *tleap* of Amber20 (Case et al., 2020). Again, as individual chains form the two active sites of the 3CL^{pro} homodimer, MD simulations were carried out with monomeric 3CL^{pro} in complex with the selected D-peptides. This approach was deemed sufficiently accurate to assess the complex stability and significantly reduced the computational demand by decreasing the size of the simulated systems. The solvated complexes underwent two consecutive EMs identical to those conducted during the preparation of the D-peptide libraries. Each energy-minimized system was equilibrated prior to the productive run through a 1 ns NVT heating using a linear temperature gradient from 10 to 298.15 K, followed by a 1 ns NPT equilibration at $p = 1 \text{ bar}$ and $T = 298.15 \text{ K}$. Both equilibration steps were carried out in the presence of harmonic restraints ($k = 10 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^2$) applied to the complex heavy atoms. Then, four consecutive 1 ns MD simulations in the NPT ensemble, in which the harmonic constant was lowered from 8 to $2 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^2$ in $2 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^2$ strides, were carried out. Subsequently, 10 ns NVT MD simulations, extended to 110 ns for several systems showing favorable $\Delta G_{\text{eff},10\text{ns}}$ values, were conducted. Finally, the MD simulations corresponding to the four experimentally tested D-peptides in complex with 3CL^{pro} were extended to 1 μs . Replicate 1 μs MD simulations for these systems were also performed by subjecting the last frame of the respective 110 ns



trajectories to EM, heating using randomly-initialized atomic velocities drawn from a Maxwell-Boltzmann distribution, and NPT equilibration steps, as described hereinbefore.

The program *pmemd.cuda* of Amber20 was used to run all MD simulations (Salomon-Ferrer et al., 2013; Case et al., 2020). Periodic boundary conditions were set during both EM and MD simulations, and long-range electrostatic interactions (for distances >9 Å) were handled with the Particle Mesh Ewald (PME) algorithm (Darden et al., 1993). Temperature control was carried out using the Berendsen weak coupling algorithm (Berendsen et al., 1984) during heating and the Langevin thermostat (Schneider and Stoll, 1978), with a collision frequency of 2 ps^{-1} , during both NPT equilibration and NVT production runs. The Berendsen barostat (Berendsen et al., 1984), with a relaxation time of 2 ps, was employed to control the pressure during NPT equilibrations. Covalent bonds involving hydrogen atoms of the solute and water molecules were constrained with the SHAKE (Ryckaert et al., 1977) and SETTLE (Miyamoto and Kollman, 1992) algorithms. The equation of motion of the simulated systems was integrated

using the *leap-frog* algorithm with a timestep of 2 fs (Case et al., 2020).

MM-GBSA Free Energy Calculations

MM-GBSA free energy calculations were conducted with the *MMPBSA.py* program of Amber20 (Gohlke et al., 2003; Gohlke and Case, 2004; Miller et al., 2012; Case et al., 2020). The single trajectory approach, in which the free ligand and the free receptor trajectories are extracted from that of the complex, was adopted in all cases after stripping off the solvent and ions (Miller et al., 2012). The GB-neck2 implicit solvation model was employed to determine the polar solvation free energy component (Nguyen et al., 2013; Case et al., 2020), as it yielded a good correlation with experimental results in a set of protein-peptide complexes (Hernandez Gonzalez et al., 2017). Moreover, the calculations were performed using the *mbondi3* set of atomic radii, a salt concentration of 0.1 M and external and internal dielectric constants of 80 and 1, respectively. The surface tension and the offset values were set to $0.0072 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{Å}^{-2}$ and zero, respectively, in order to estimate the nonpolar free-energy

component from the variation of the solvent accessible surface areas (SASAs) of the interacting molecules, i.e., 3CL^{pro} and D-peptides (Case et al., 2020). In turn, SASA values were obtained with the Linear Combination of Pair-wise Overlaps (LCPO) algorithm included in Amber20 suite (Weiser et al., 1999; Case et al., 2020) using a probe radius of 1.4 Å (Connolly, 1983). As mentioned before, the $\Delta G_{\text{eff},10\text{ns}}$ and $\Delta G_{\text{eff},110\text{ns}}$ values obtained from the 10 and 110 ns MD simulations after discarding the first 2 and 40 ns, respectively, allowed us to select the most promising D-peptides at the two last steps of the workflow shown in Figure 1.

Principal Component Analysis

Principal component analysis (PCA) (Amadei et al., 1993) was carried out for the 1 μ s trajectories of 3CL^{pro} in complex with the experimentally-tested D-peptides. Using this technique, we sought to reduce the phase space dimensionality by projecting the system's motion along the two eigenvectors, known as principal components (PCs), of the highest variance (=largest eigenvalues), as our main interest here was to identify the different conformations of the D-peptides in the 3CL^{pro} active site during the long MD simulations. PCA was performed only for the Ca atoms of the former molecules. The replicate 1 μ s MD simulations of each system were concatenated and fitted using *cpptraj* of Amber 20 (Case et al., 2020). Trajectory fitting was carried out concerning the 3CL^{pro} backbone atoms belonging to the chymotrypsin-like (ChT-like) domains, i.e., domains I and II, residues 8 to 183 (Tahir Ul Qamar et al., 2020), which contain the active site. This step ensured that the displacements of the D-peptide Ca atoms during the long MD simulations were measured relative to the enzyme's active site and eliminated the influence of domain III motions during fitting. The program *gmx covar* of Gromacs v5.1.4 (Abraham et al., 2015) was employed to calculate the covariance matrices of the D-peptide Ca positions along the fitted trajectories and their corresponding sets of eigenvalues and eigenvectors. Two-dimensional (2D) projections of the trajectories onto the first two eigenvectors (PC1 and PC2) were obtained with *gmx anaeig* of Gromacs v5.1.4.

PCA was combined with free energy landscape (FEL) visualization and clustering to determine central structures for the different conformations sampled during the MD simulations (Papaleo et al., 2009). FELs were obtained from the 2D projections and depicted as heatmaps by discretizing the 2D phase space into 700 square bins of equal size and counting the number of points within each. The free energy value corresponding to bin i (ΔG_i) was then calculated from the probability of finding the system into that bin (p_i) using the equation:

$$\Delta G_i = -RT \ln(p_i) \quad (1)$$

where R is the gas constant and T , the temperature (298.15 K).

Finally, the trajectories were split using the K-means algorithm with random initial seeds, implemented as an option of the *cluster* command of *cpptraj* (Case et al., 2020), by using the PC1 and PC2 values as a metric. The number of clusters in each case was set to the main FEL basins observed in the corresponding heatmap.

Each newly generated trajectory was subsequently clustered to determine its central structure through RMSD clustering (see next section for details).

Trajectory Analyses

The 110 ns MD simulations were clustered in order to select the central structure of each analyzed 3CL^{pro}/peptide complex. This step was performed with the *cluster* command of *cpptraj* using the average linkage algorithm (Shao et al., 2007; Case et al., 2020). The RMSD for the heavy atoms of the peptide and 3CL^{pro} residues lying within a 4 Å cut-off was chosen as a metric for clustering the trajectories. This procedure was also applied to determine the central structures corresponding to the main energy minima observed in the PC1 vs PC2 projections of the complexes subjected to PCA. The main central structures were selected for structural representation using Pymol 2.1.0 (DeLano, 2002). RMSD values for the D-peptide heavy atoms along the 110 ns trajectories were calculated with *rms* of *cpptraj* after fitting all frames with respect to the 3CL^{pro} backbone atoms belonging to the ChT-like domains in the corresponding starting structures ($t = 0$). These RMSD values were averaged during the last 20 ns of the trajectories to assess whether the binding modes of the D-peptides sampled at the end of the MD simulations deviated significantly from those of the starting structures. Root mean square fluctuations (RMSFs) were calculated with the *rmsf* command of *cpptraj* (Case et al., 2020). Finally, hydrogen bonds (H-bonds) formed at the complex interfaces during the MD simulations were determined with *hbond* command of *cpptraj* (Case et al., 2020), using the following geometric criteria: a donor-acceptor distance ≤ 3.5 Å and a donor-H-acceptor angle $\geq 120^\circ$.

D-Peptide Synthesis

Synthetic D-enantiomeric peptides used in this study (4P1, 4P2, 4P3, and 4P4) were synthesized by Genscript (Leiden, NL), with a purity of $\geq 90\%$. The D-peptides were acetylated at the N-terminus and methylated at the C-terminus.

Cloning, Expression, and Purification of SARS-CoV-2 3CL^{pro}

The codon-optimized cDNA encoding SARS-CoV-2 3CL^{pro} (Uniprot entry: P0DTD1, virus strain: hCoV-19/Wuhan/WIV04/2019) was synthesized and implemented in the ampicillin-resistant vector pGEX-6P-3 (BioCat GmbH, Heidelberg, Germany). The construct contains an N-terminal GST-tag and a PreScission protease cleavage site (LEFLFQGP). Expression and purification were performed as described before (Eberle et al., 2021).

Primary 3CL^{pro} Enzymatic Inhibition Assay

All measurements were performed in triplicate in 20 mM Tris pH 7.2, 200 mM NaCl, 1 mM EDTA, and 1 mM TCEP as described previously (Zhang et al., 2020a; Zhang et al., 2020b; Ma et al., 2020; Eberle et al., 2021). 20 μ M of the peptides (4P1, 4P2, 4P3, and 4P4) were pipetted into a Corning 96-Well plate (Sigma Aldrich), 3CL^{pro} was added to a final concentration of 500 nM,

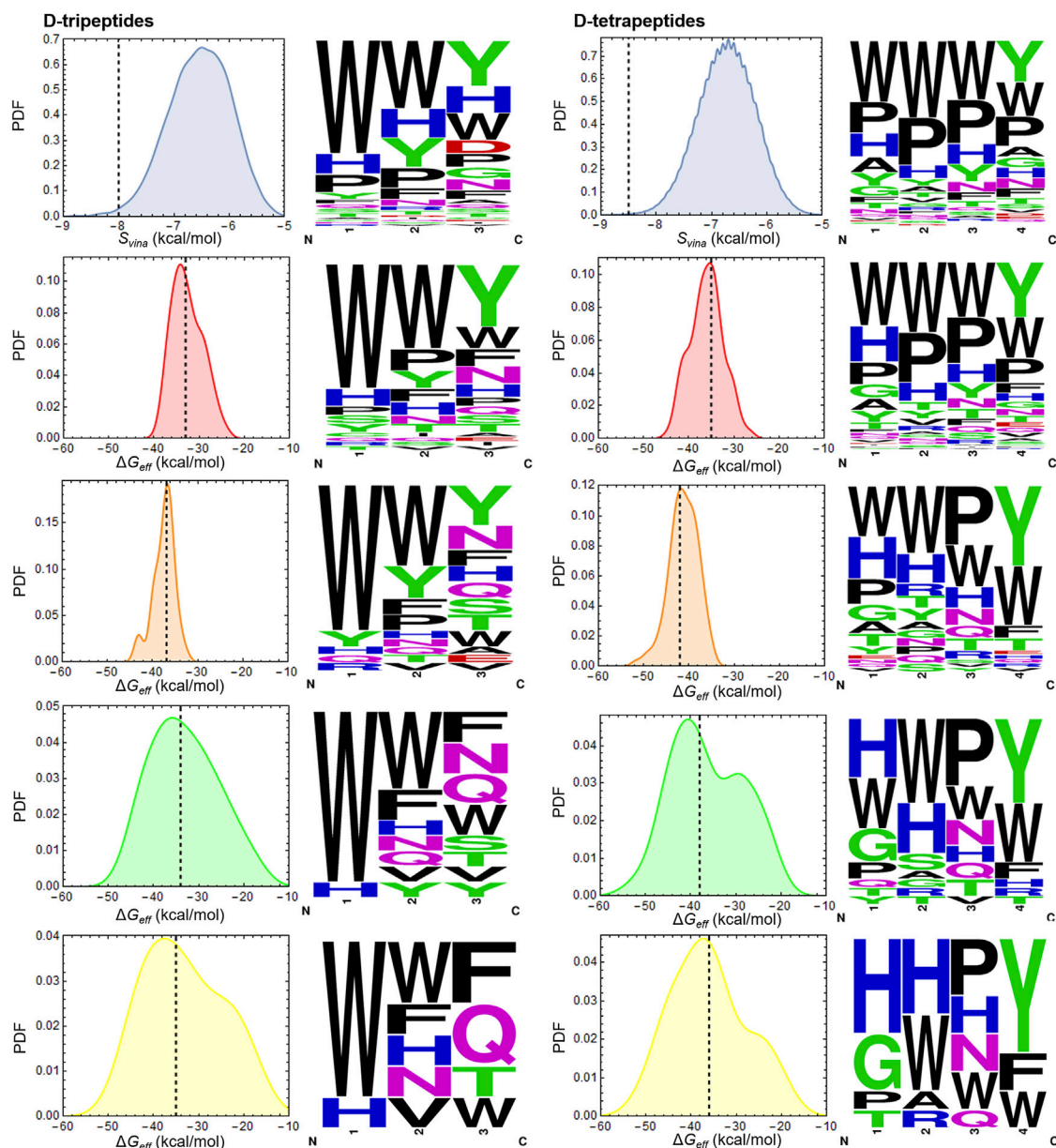


FIGURE 2 | Energy distributions and D-peptide sequences at every step of the *in silico* workflow. Each histogram is colored as the corresponding step of the workflow shown in **Figure 1**. Dashed lines indicate the cut-off values employed to select the best candidates at every step. Peptides with $S_{vina}/\Delta G_{eff}$ values to the left of the dashed lines were selected. PDF is the abbreviation for probability density function. Sequence logos of the D-peptides identified at every step are shown to the right of the corresponding histograms. Letter sizes are proportional to the frequency of occurrence of the indicated residues in each position. N and C below the graphs stand for the N- and C- termini, respectively. Sequence logos were generated at <https://weblogo.berkeley.edu/logo.cgi>. A D-tetrapeptide was excluded as potential 3CL^{pro} inhibitor after filtering according to the RMSD value in the last rescoring step (**Figure 1**, not shown). The last sequence logo of D-tetrapeptides does not contain the excluded peptide.

and the mixture was incubated for 30 min. Subsequently, the enzymatic reaction was initiated by adding the fluorogenic substrate DABCYL-KTSAVLQ↓SGFRKME-EDANS (Bachem, Switzerland) to a final concentration of 50 μ M. The gradual release of fluorescent 5-((2-Aminoethyl)amino)naphthalene-1-sulfonic acid (EDANS) was monitored for 30 min with 60 s intervals. The excitation and emission wavelengths were 360 and 460 nm, respectively, using an Infinite 200 PRO plate reader (Tecan, Männedorf, Switzerland). The

temperature was set to 37°C. The results are shown as mean value \pm standard deviation (STD).

Statistical Analyses

Block averaging was conducted using *gmx analyze* of Gromacs v.5.1.4 (Hess, 2002; Abraham et al., 2015) to estimate standard errors of the mean (SEMs) from time-dependent values collected from MD simulations, such as the reported $\Delta G_{eff,110ms}$ and RMSD

TABLE 1 | D-peptides identified *in silico* as potential 3CL^{pro} inhibitors.

Peptide ID	Peptide sequence ^a	$\Delta G_{eff, 110ns}$ (kcal/mol) ^b	RMSD (Å) ^c
D-tripeptides			
3P1	ACE-TRP-TRP-THR-NME	-44.2 ± 0.5	2.58 ± 0.03
3P2	ACE-TRP-ASN-PHE-NME	-43.8 ± 1.0	2.6 ± 0.3
3P3	ACE-TRP-PHE-GLN-NME	-40.4 ± 2.0	4.22 ± 0.05
3P4	ACE-TRP-VAL-PHE-NME	-40.3 ± 1.3	2.3 ± 0.3
3P5	ACE-TRP-TRP-GLN-NME	-36.5 ± 0.6	4.43 ± 0.04
3P6	ACE-HIE-HID-TRP-NME	-35.6 ± 1.3	1.22 ± 0.06
D-tetrapeptides			
4P1	ACE-GLY-TRP-ASN-TYR-NME	-50.2 ± 0.9	3.60 ± 0.04
4P2	ACE-GLY-TRP-HIE-TRP-NME	-45.6 ± 0.7	3.0 ± 0.9
4P3	ACE-HIE-ALA-PRO-TRP-NME	-44.9 ± 0.9	1.17 ± 0.05
4P4	ACE-HIE-HIE-PRO-TYR-NME	-44.9 ± 1.5	2.70 ± 0.09
4P5	ACE-THR-HIE-TRP-TYR-NME	-44.5 ± 2.0	3.574 ± 0.08
4P6	ACE-HIE-HIE-ASN-TYR-NME	-42.8 ± 0.9	3.20 ± 0.04
4P7	ACE-HIE-TRP-PRO-PHE-NME	-39.2 ± 0.7	3.33 ± 0.06
4P8	ACE-HIE-HIE-HID-TYR-NME	-38.8 ± 0.7	3.5 ± 0.2
4P9	ACE-PRO-TRP-GLN-PHE-NME	-38.7 ± 0.5	2.20 ± 0.02
4P10	ACE-GLY-ARG-TRP-TYR-NME	-37.5 ± 2.0	3.36 ± 0.09

^aResidues are shown in three-letter code and separated by hyphens. ACE and NME are the N- and C-terminal caps added to the D-peptides. HIE and HID are HIS tautomers.

^bMM-GBSA average effective free energies calculated over the last 70 ns of each 110 ns MD trajectory ± SEMs estimated through block averaging. See instantaneous ΔG_{eff} values vs time plots for every system along their respective 110 ns MD simulations in **Supplementary Figure S2**.

^cMean RMSD values for the peptide heavy atoms with respect to the starting structure ($t = 0$) calculated over the last 20 ns of each 110 ns MD trajectory ± standard errors of the mean estimated through block averaging. See RMSD values vs time plots for all systems along their respective 110 ns MD simulation in **Supplementary Figure S3**.

mean values. On the other hand, the statistical significance of the residual activity mean values' differences was performed with GraphPad Prism software version 8 (GraphPad, 2018) and was assessed with one-way analyses of variance (ANOVA), followed by Tukeys' multiple comparison test. Significant differences were considered at $p < 0.01$ (**) and $p < 0.001$ (***).

RESULTS

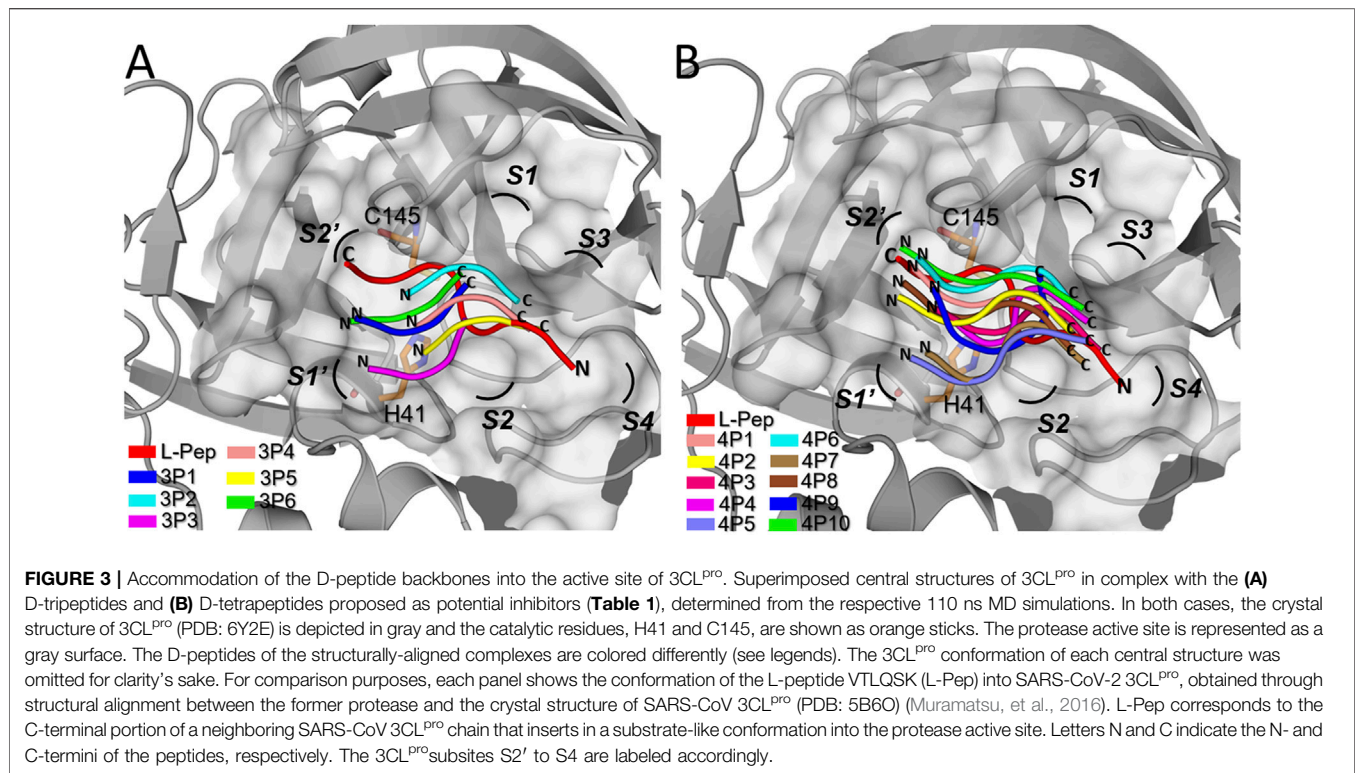
D-Peptides Predicted as 3CL^{pro} Inhibitors

The D-peptides proposed as potential 3CL^{pro} inhibitors were selected through a workflow involving SBVS and several rescoring steps (Figures 1, 2). The top-ranked D-peptides, according to the S_{vina} values, were made up mainly of aromatic and hydrophobic residues, with TRP being the residue most frequently found in all positions except the C-terminus, in which TYR was the most abundant (Figure 2). Certain regions of 3CL^{pro} active site, such as the S2 pocket, are hydrophobic (Jin et al., 2020), which can favor the binding of peptides containing the aforementioned residues. However, the already known bias of AutoDock Vina and other docking algorithms toward larger compounds can also be at play here (Pan et al., 2003; Chang et al., 2010). The subsequent rescoring steps aimed to correct this bias and enrich selected D-peptides' lists with accurate hits. In fact, it can be observed from the sequence logos, shown in Figure 2, that even though TRP was still prevalent at different positions of the D-peptides, other residues became progressively more abundant throughout the workflow steps, which was particularly apparent for D-tetrapeptides, in which HIS and PRO were found to be predominant in positions 1 to 3 after completing the workflow (Figure 2).

Apart from HIS, other residues with polar side-chains, such as ASN, GLN, THR, and ARG, occurred more frequently among the D-peptides selected after the last rescoring step than among those selected from the SBVS. Consequently, the D-peptides prioritized after the workflow completion were, on average, smaller in size than those ranked in the top positions by the docking algorithm. Overall, the rescoring steps tended to upweight the occurrence of intermolecular H-bonds mediated by the side-chains of D-peptide residues at the expense of ligand size. These results, in turn, suggest a reduced impact of ligand-size bias on the final set of chosen D-peptides. The D-tripeptides and D-tetrapeptides identified as potential 3CL^{pro} inhibitors are shown in Table 1. All these D-peptides fulfilled the conditions set throughout the workflow steps. They showed a good affinity for the enzyme and stability of the initial docking pose during the MD simulations, measured in terms of $\Delta G_{eff, 110ns}$ and peptide RMSD mean values, respectively (Table 1 and Supplementary Figures S2, S3).

Structural Features of the Predicted 3CL^{pro}/D-peptide Interfaces

The central structures of the selected D-peptides (Table 1) in complex with 3CL^{pro} obtained after clustering the respective 110 ns MD simulations are shown in cartoon representation in Figure 3. Moreover, for comparison purposes, the L-peptide VTLQSK (L-Pep) is depicted at the 3CL^{pro} active site (Figure 3). L-Pep corresponds to the C-terminus of the homologue SARS-CoV 3CL^{pro} and inserts into the active site of a neighboring protease chain in the PDB structure 5B6O (Muramatsu et al., 2016), thus allowing the template-based modeling of the SARS-CoV-2 3CL^{pro}/L-Pep complex. Of note,



the backbones of all the analyzed D-peptides adopt an orientation opposite (=retro-binding) to that of L-Pep (Figures 3A,B). As we will show below in more detail, the retro-binding enables the formation of key interactions between the active site residues and the backbone atoms of the D-peptide present at the 3CL^{pro}/L-Pep interface.

It can be seen that there is a partial overlap between the L-Pep and D-peptide backbones, especially involving the N- and C-terminal portions of L-pep and D-tetrapeptides, respectively (Figure 3). Therefore, several key interactions mediated by the peptide backbones and neighboring residues on the S side of the 3CL^{pro} active site are expected to be preserved in the predicted complexes. The D-tetrapeptide backbones span a larger stretch of the 3CL^{pro} active site when compared to the D-tripeptides and resemble, to a greater extent, the overall accommodation of L-Pep, which could explain, in turn, the generally more favorable free energy values obtained for the D-tetrapeptides in complex with 3CL^{pro} (Table 1). On the other hand, the backbone accommodation was more divergent in the N terminal regions of the D-peptides, not only relative to L-pep but among themselves (Figure 3). This might arise from the plasticity of the 3CL^{pro} active site (Kneller et al., 2020) and the inability of the small D-peptides to satisfy interactions equivalent to those observed for L-peptides.

The number of intermolecular H-bonds formed along the MD simulations can be used to indicate complex stability (Lokhande et al., 2021). Therefore, we decided to calculate the time profiles of such interactions in the predicted 3CL^{pro}/D-peptide complexes (Supplementary Figure S4). The graphs show that all the D-peptides in Table 1 formed several H-bonds with 3CL^{pro}

active site residues during the 110 ns MD simulations of the complexes (Supplementary Figure S4). The average number of intermolecular H-bonds ranged from 3 to 8, depending on the complex. This result underscores the good complementarity of the identified D-peptides to the 3CL^{pro} active site.

The analysis of the interactions occurring at the 3CL^{pro}/L-Pep complex provides valuable information to study the 3CL^{pro}/D-peptide interfaces. Therefore, a detailed structural representation of this complex was included in Figure 4. L-Pep extends along with the S4 to S2' subsites of the enzyme, displaying the N-terminal VAL residue at position P4 and the C-terminal LYS at P2'. THR at P3 is exposed mainly to the solvent, whereas LEU and GLN at P2 and P1, respectively, insert into well-defined pockets. As can be observed, several amide nitrogen (N) and carbonyl oxygen (O) atoms of L-Pep backbone engage in H-bond formation with 3CL^{pro} residues, e.g., VAL(N)-T190(O), THR(N)-E166(O), THR(O)-E166(N), GLN(N)-H164(O), GLN(O)-C145(N), GLN(O)-G143(N), and LYS(N)-T26(O) (Figure 4). Of note, the amide oxygen (OE1) of L-Pep GLN side-chain forms a key H-bond with the protonated N atom of H163 imidazole ring H163(NE2), which helps explain the preference of 3CL^{pro} for the former residue at P1 (Singh et al., 2020). Other H-bonds mediated by the side-chains of L-Pep residues are GLN(NE2)-F140(O) and LYS(NZ)-G143(O) at the S1 and S2' subsites, respectively (Figure 4).

Structural representations of the interfaces of 3CL^{pro} in complex with the top-ranked D-tripeptides (Table 1) are shown in Figure 4. For brevity's sake, only the complexes having $\Delta G_{eff,110ns}$ values < -40 kcal/mol were included in the figure (Table 1). The backbone polar atoms of the D-tripeptides

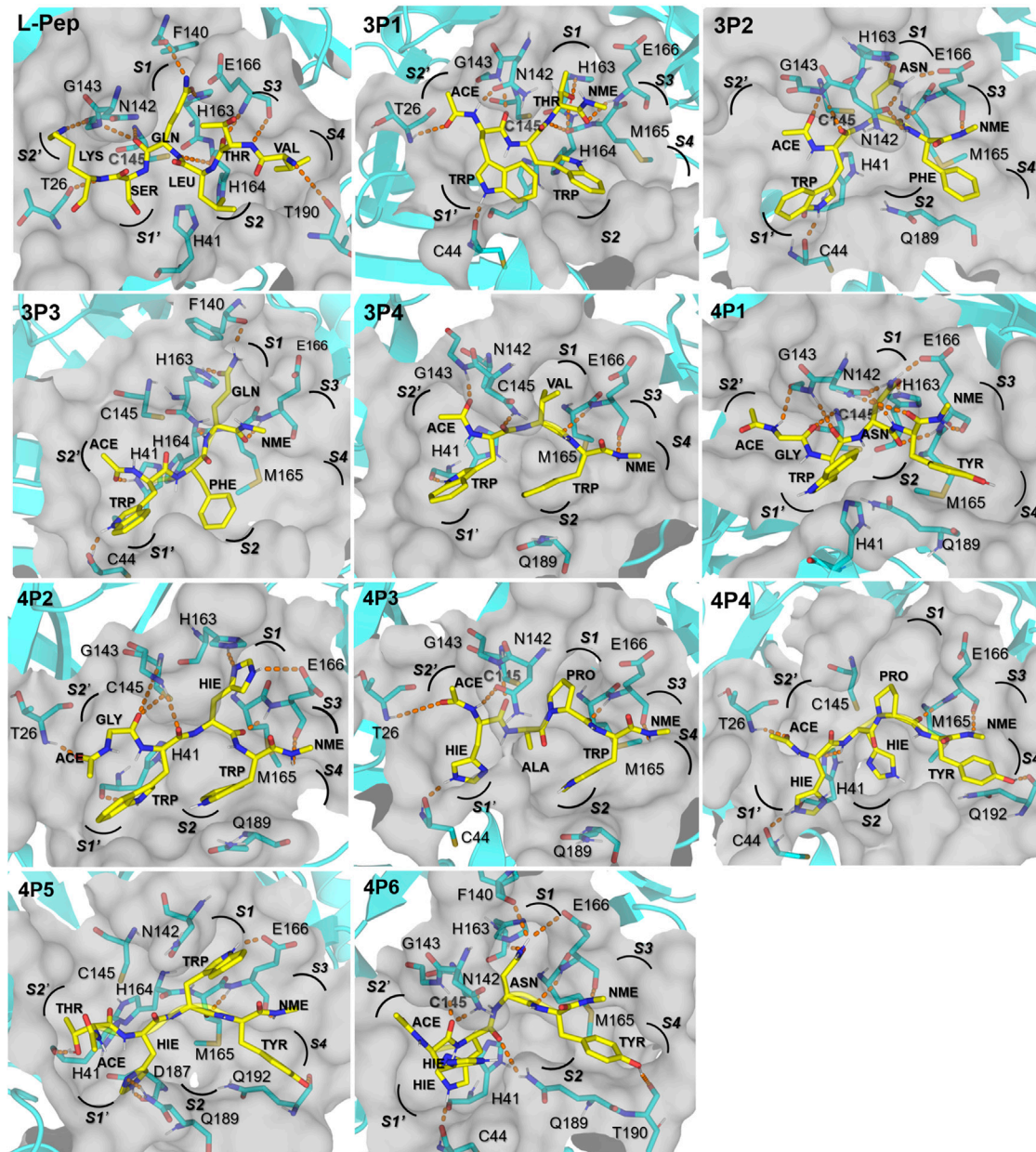


FIGURE 4 | Structural representation of top-ranked D-tripeptides and D-tetrapeptides in complex with 3CL^{PRO}. All D-peptides are shown as yellow sticks and their residues are labeled in bold and in the three-letter code. 3CL^{PRO} residues forming H-bonds with the peptides plus the catalytic residues H41 and C145 are labeled and represented as cyan sticks. The 3CL^{PRO} active site cavity is depicted as a transparent gray surface. H-bonds between the D-peptides and the 3CL^{PRO} residues with occupancies >25% during the respective 110 ns MD simulations are displayed as orange dashed lines. Subsites S4 to S2' are labeled in bold and italic. Structural representations of the remaining D-peptides included in **Table 1** can be found in **Supplementary Figure S5**.

form a network of H-bonds with 3CL^{PRO} residues, some of them equivalent to those observed at the interface of the L-Pep complex. For example, the H-bond involving an O atom of the D-tripeptides and E166(N) occurs in all the analyzed interfaces. An additional H-bond between a D-tripeptide N atom and E166(O) was found in the 3P2 and 3P4 complexes (**Figure 4**). H-bonds mediated by G143(N) and C145(N) and a D-tripeptide backbone O atom also occur in most depicted

complexes. Differently from the L-Pep complex, the amide group of N142 can form H-bonds with polar atoms of 3P1, 3P2, and 3P4 backbone, thus providing additional stabilizing interactions. This result underscores that the flexible side-chain of N142 can adopt conformations in solution that facilitate transient H-bonding to the ligands.

Three top-ranked D-tripeptides (3P1, 3P3, and 3P4) possess aromatic residues TRP or PHE inserted into the S2 pocket of

3CL^{pro}, which prefers hydrophobic residues (Chuck et al., 2010; Rut et al., 2021), whereas PHE of 3P2 occupies the S4 subsite (Figure 4). The S1 subsite accommodates the D-tripeptide residues THR, ASN, GLN, and VAL. Of note, ASN(OD1) and GLN(OE1) of 3P2 and 3P3, respectively, form the key H-bond with H163(NE2) observed for L-Pep GLN (Figure 4). The ASN side-chain can mimic the interactions established by the L-Pep GLN amide group because the backbone of 3P2 leans toward the entrance of the S1 pocket, thus shortening the distance to reach the bottom of this subsite (Figure 4). Like L-Pep GLN, the amide group of 3P2 ASN and 3P3 GLN can form additional H-bonds at S1, such as ASN(ND2)/GLN(NE2)-F140(O) or ASN(ND2)/GLN(NE2)-E166(OE1,2). Interestingly, the hydroxyl oxygen of 3P1 THR, THR (OG1), is also able to interact with H163(NE2), which could explain the favorable $\Delta G_{eff,110ns}$ value obtained for this D-peptide (Table 1) despite not bearing GLN or ASN at P1. Finally, it is worth noting that all the D-tripeptides proposed as 3CL^{pro} inhibitors possess TRP at the N-terminus, except for 3P6 that contains HIE (Table 1). Our results indicate that TRP accommodates favorably at the S1' pocket, and the nitrogen of the indole group atom (NE1) can form H-bonds with C44(O)/H41(O) (Figure 4).

As done for the D-tripeptides, the interfaces of the top-ranked D-tetrapeptides ($\Delta G_{eff,110ns} < -40$ kcal/mol, Table 1) in complex with 3CL^{pro} are depicted in Figure 4. The backbone of the selected D-tetrapeptides span the 3CL^{pro} active site from the S4 to the S2' subsites and establish several polar interactions with the neighboring residues. Key H-bonds with E166(N) and E166(O), observed in the 3CL^{pro}/L-Pep complex, occur in all the analyzed D-tetrapeptide complexes. Other H-bonds, such as those involving C145(N), G143(N), the N142 side-chain amide group, T26(N) and Q192(NE2), can be found at various interfaces (Figure 4).

Four out of the six top-ranked D-tetrapeptides, i.e., 4P1, 4P4, 4P5, and 4P6, contain TYR at the C-terminus, the most abundant residue occurring at this position in the whole set of identified D-tetrapeptides (Figure 2 and Table 1). The predicted structures suggest that TYR accommodates at the S4 subsite, where the aromatic ring sits on the pocket base, mainly formed by M165 and Q192, and the side-chain hydroxyl oxygen, TYR (OH), can form H-bonds with T190(O)/Q192(O) (Figure 4). The other two D-tetrapeptides, 4P2 and 4P3, insert their C-terminal TRP residues into the S2 pocket of 3CL^{pro}. The latter D-peptide also accommodates its second residue, ALA, on the opposite side of the same pocket. Likewise, 4P4 can accommodate its second residue, HIE2, at the S2 pocket (Figure 4). A closer look at their sequences and the predicted structures of their complexes with 3CL^{pro} reveals that PRO at P1 of 4P3 and 4P4 bends the D-peptide backbones in a way that makes it feasible for upstream residues to interact with S2. Like PRO, ASN was found at the S1 pocket of 3CL^{pro} in complex with two other D-tetrapeptides, i.e., 4P1 and 4P6 (Figure 4). However, unlike the former residue, which leaves the S1 pocket largely unoccupied, ASN can form H-bonds equivalent to L-Pep GLN. At the same position, 4P2 HIE interacts with H163(NE2) and E166(OE1,2). Of note, HIS is the second most favorable residue at P1 according to substrate specificity profiling conducted for SARS-CoV 3CL^{pro} (Chuck et al., 2010). However, to reach the bottom of the S1 pocket,

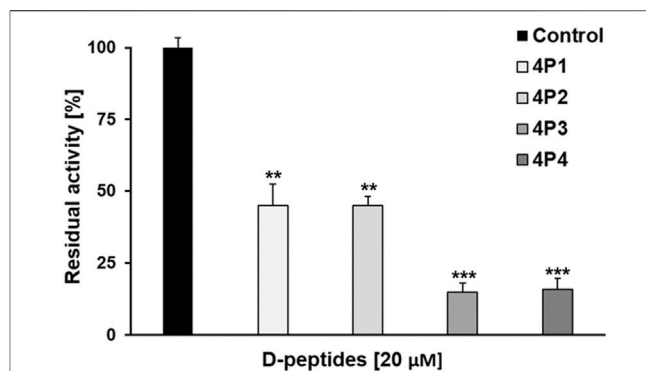


FIGURE 5 | Preliminary inhibition tests of 4P1-4 against 3CL^{pro}. 4P1 and 4P2 inhibit the virus protease activity by more than 60%. 4P3 and 4P4 inhibit the virus protease activity more than 80%. Data shown are the mean \pm STD from three independent measurements ($n = 3$). Asterisks mean that the data differs from the control (0 μ M inhibitor) significantly at $p < 0.01$ (**) and $p < 0.001$ (***), level according to ANOVA and Tukey's test.

HIS, like ASN, requires a backbone accommodation closer to the pocket entrance, which is accessible to the D-peptides according to our predictions. The specificity for HIS at P1 in L-peptides has been explained by proposing that this residue interacts with the N142 side-chain (Chuck et al., 2010). TRP was also found at the S1 subsite of the 3CL^{pro}/4P5 complex (Figure 4). In this case, the residue does not penetrate deeply into the subsite but forms the H-bond TRP(NE1)-E166(OE1,2).

HIE and TRP, in that order, are the most abundant residues of the top-ranked D-tetrapeptides placed at the S1' subsite (Figure 4). HIE (NE2) is capable of forming H-bonds with C44(O) at the interfaces of 3CL^{pro} in complex with 4P3, 4P4, and 4P6 or with Q189(NE2) and D187(O) at the interface of the 3CL^{pro}/4P5 complex. Interestingly, the latter D-peptide is the only one bearing an N-terminal THR residue, which forms H-bonds with H41(O). In this case, the ACE cap is sticking out to the solvent instead of lying at the S2' subsite, as in the remaining complexes of the identified D-tetrapeptides (Figure 4).

In vitro Inhibitory Activity of the Four Top-Ranked D-Tetrapeptides Against 3CL^{pro}

The energetic and structural analyses presented in the previous sections demonstrated that the selected D-tetrapeptides displayed, in general, more favorable free energy values and better complementarity with the 3CL^{pro} active site than D-tripeptides (Table 1 and Figures 3, 4). Therefore, to assess the validity of the computational workflow for D-peptide identification, we decided to test *in vitro* the inhibitory activity of the four top-ranked D-tetrapeptides, 4P1 to 4P4 (Table 1).

The results of the primary screening are shown in Figure 5. All the D-tetrapeptides, tested at a final concentration of 20 μ M, significantly reduced the activity of 3CL^{pro} relative to the control assay. In fact, the percentage of residual activity of the protease dropped below 20% upon incubation with 4P3 and 4P4, which are the most potent D-peptides. On the other hand, both 4P1 and

4P2 inhibited roughly 55% of the enzymatic activity under such conditions. Our results indicate that the devised computational workflow successfully identified promising D-peptides displaying inhibitory potency against 3CL^{pro} in the micromolar concentration range.

Identification of the Most Stable Binding Modes of 4P1, 4P2, 4P3, and 4P4 to 3CL^{pro} Combining Long MD Simulations, PCA, FEL and Clustering

Replicate 1 μ s MD simulations were conducted for 3CL^{pro} in complex with the experimentally-tested D-tetrapeptides to assess the time stability of binding modes previously proposed from shorter MD simulations (Figure 4) and the possible occurrence of alternate bound conformations. First, we noticed large deviations in the RMSD values for all 3CL^{pro} backbone atoms concerning the starting structures along several 1 μ s trajectories (Supplementary Figure S6). However, it became apparent through visual inspection that such deviations are caused by large motions of domain III relative to the ChT-like domains. This was corroborated after calculating the RMSD for the backbone atoms of ChT-like domains, as very stable time profiles were obtained in this case (Supplementary Figure S6). Moreover, RMSF values for all 3CL^{pro} plus D-peptide backbone atoms calculated after fitting the trajectories in respect of the backbone atoms of ChT-like domains show the large relative fluctuations of domain III (Supplementary Figure S6), which are likely to arise from the fact that 3CL^{pro} was simulated in the monomeric state to reduce the computational demand. The relatively loose interactions between domain III and the ChT-like domains suggest that the latter might suffice to simulate complexes with active site ligands, thus increasing the MD simulations performance. On the other hand, the RMSF profiles sharply drop for residues beyond 306, which belong to the D-tetrapeptides (see region 3 in Supplementary Figure S6). Along with the intermolecular H-bond time profiles (Supplementary Figure S7), this result indicates that the D-tetrapeptides keep forming favorable interactions with 3CL^{pro} active site residues during all the simulation time.

The RMSD time profiles for 4P1 backbone atoms calculated after fitting the trajectories with respect to the backbone atoms of ChT-like domains indicate that the D-peptide remains bound in conformations similar to that of the starting structure throughout the microsecond-long MD simulations (Supplementary Figure S8). Nonetheless, we observed several transitions between slightly different conformations in both backward and forward directions during the replicate MD simulations of this complex. The previous result was confirmed by the FEL obtained by projecting the concatenated 1 μ s trajectories onto PC1 and PC2 (Figure 6). Indeed, two main local minima or basins (termed 4P1-0 and 4P1-1, the former being more populated) are observed in the FEL heatmap (Figure 6). The 3CL^{pro}/4P1 central structures corresponding to the main basins display slight differences in the conformation adopted by 4P1 ASN side-chain and ACE-GLY in the S1 and S2' subsites, respectively (Supplementary Figure S9). Interestingly, the 4P1-1

conformation suggests that ASN(OD1) and ASN(ND2) at P1 can form alternate and less prevalent H-bonds with C145(N) and S144(OG), respectively (Supplementary Figure S9).

The long MD simulations for 3CL^{pro}/4P2 and 3CL^{pro}/4P3 complexes show stable peptide RMSD patterns (Supplementary Figure S8). FEL heatmaps in Figure 6 also indicate that 4P2 and 4P3 sampled conformations around a single energy basin during most of the simulation time. The central structures of both complexes corresponding to their respective minima (4P2-0 and 4P3-0) (Supplementary Figure S9) are very similar to those previously calculated from the 110 ns MD simulations (Figure 4). However, we found minor differences in the accommodation of 4P2-0 HIE side-chain at P1, which forms H-bonds with S144(O) instead of the H-bond with H163(NE2), more prevalent in the shorter MD simulation (Figure 4 and Supplementary Figure S9). Further calculations indicated that both H-bonds occur during the 2 μ s concatenated MD simulations with 67 and 20% occupancies, respectively. Moreover, the 4P2-0 GLY residue at S2' interacts preferentially with T26(OG1) rather than with G143(N) and C145(N) (Figure 4 and Supplementary Figure S9). On the other hand, no appreciable differences between 4P3-0 and the central structure of 4P3 corresponding to the shorter MD simulation are observed (Figure 4 and Supplementary Figure S9).

Contrary to what was observed for the previously-analyzed complexes, significant variations in the peptide RMSD patterns occur along the two replicate 1 μ s MD simulations of the 3CL^{pro}/4P4 complex (Supplementary Figure S8). Because of such instabilities, a third 1 μ s MD simulation was run for this complex, which also shows wide peptide RMSD variations (Supplementary Figure S8). In agreement with the depicted RMSD patterns, the FEL for 3CL^{pro}/4P4 indicates the existence of four main energy minima, termed 4P4-0 to 4P4-3, involving relatively large motions along PC1 and PC2 (Figure 6). The central structures corresponding to those minima also display appreciable divergence, especially 4P4-1 and 4P4-3 with respect to 4P4-0 and 4P4-2 (Supplementary Figure S9).

In principle, the stability of the sampled minima can be estimated from their relative abundances, i.e., the fraction of trajectory frames belonging to each minimum. However, under-sampling can still occur even in microsecond-long simulations, especially if transitions between different states do not occur several times in both directions during the simulation time, as in the present case (Supplementary Figure S8). This issue precludes the accurate calculation of conformational population sizes at equilibrium. Therefore, we decided to calculate the binding free energies (ΔG_{bind}) for the four main conformations of the 3CL^{pro}/4P4 complex sampled during the long MD simulations (see Supplementary Text S1 and references cited therein). Interestingly, the results show that conformations 4P4-0 and 4P4-2 have similar ΔG_{bind} values, considering the uncertainties of the calculated mean values (-5.2 ± 0.4 and -5.8 ± 0.4 kcal/mol, respectively, Supplementary Table S2) and the errors of ~ 1 kcal/mol associated with the employed technique (Aldeghi et al., 2016). Of note, inhibition constants (K_i) ranging from 50 to 100 μ M are expected from the previous ΔG_{bind} values, in agreement with the micromolar inhibition potency displayed

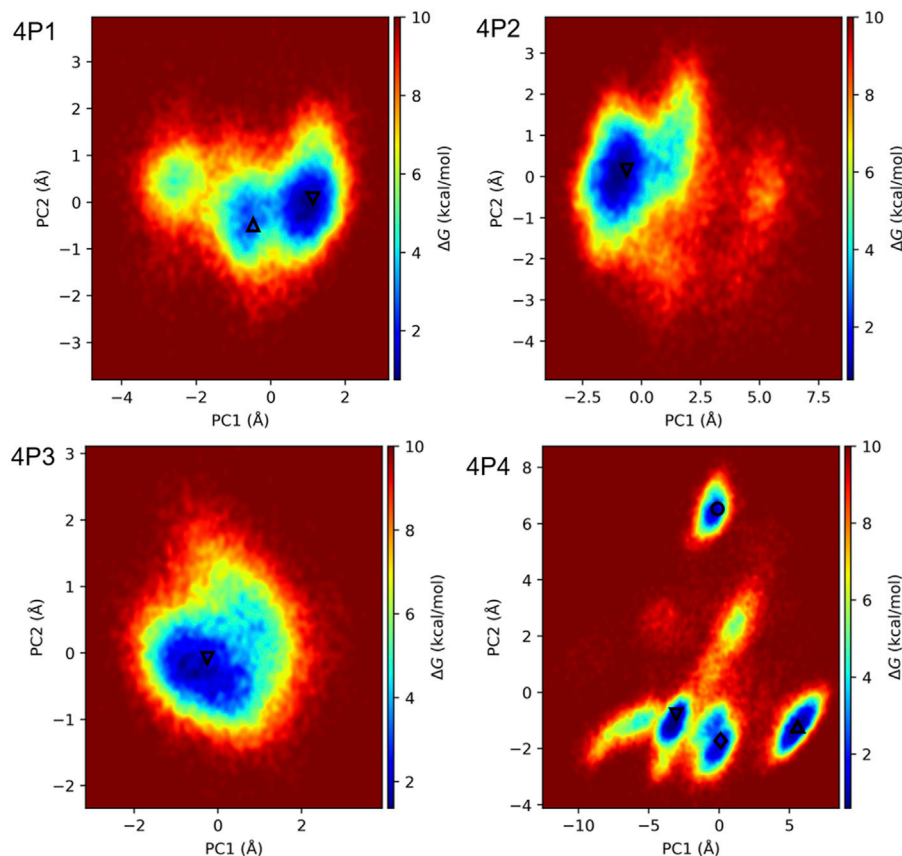


FIGURE 6 | FEL heatmaps for 3CL^{pro} in complex with 4P1, 4P2, 4P3, and 4P4. FELs were obtained by projecting the concatenated replicate 1 μ s trajectories of the complexes onto PC1 and PC2, associated with the motions of the D-peptide C α atoms. All trajectories were fitted to the respective initial structures with respect to the backbone atoms of 3CL^{pro} in the ChT-like domains. The local energy minima observed in the heatmaps are indicated with the symbols Δ , ∇ , \diamond , and \circ , which rank the minima according to their relative sizes in decreasing order ($\Delta > \nabla > \diamond > \circ$). PC1 and PC2 were chosen to generate FELs because they account for more than 68% of the motions of the D-peptide C α atoms in all cases (**Supplementary Figure S10**).

by 4P4 (**Figure 5**). The other two conformations, i.e., 4P1-1 and 4P4-3, are significantly less stable (-0.8 ± 0.4 and -1.9 ± 0.3 kcal/mol, respectively, **Supplementary Table S2**). Overall, the free energy calculations show that the relative stabilities of the different sampled conformations do not match the results expected from the relative sizes of the four main energy minima observed in the 3CL^{pro}/4P4 FEL.

The two lowest-energy and nearly-isoenergetic conformations of 4P4 (4P-0 and 4P4-2, **Supplementary Figure S9**) differ mainly in the accommodation of HIE1 and ACE in S' side of the active site. In fact, 4P4-0 HIE1 forms an H-bond with C44(O), whereas 4P4-2 HIE1 forms an H-bond with H41(O) (**Supplementary Figure S9**). The remaining residues in both conformations occupy the same subsites, although some small differences in their positions are observed (**Supplementary Figure S9**). On the other hand, the 3CL^{pro}/4P4 central structure determined from the 110 ns MD simulation (**Figure 4**) adopts roughly the same conformation of 4P4-0 HIE1 at S1', whereas the accommodation of the remaining residues in the former structure resembles to a larger extent that of 4P4-2 (**Figure 4** and **Supplementary Figure S9**).

DISCUSSION

This work reports the *in silico* identification and *in vitro* validation of promising D-peptide inhibitors of SARS-CoV-2 3CL^{pro}. An in-house D-peptide library was built from scratch to search for potential 3CL^{pro} inhibitors through a computational workflow comprising SBVS with AutoDock Vina and several rescoring steps (**Figure 1**). This workflow was already employed elsewhere to predict nonpeptidic allosteric inhibitors against the malarial protease falcipain-2 (Hernandez Gonzalez et al., 2021) and was adapted here for D-peptide SBVS. To our knowledge, our study is the first to employ a docking-based approach to screen D-peptide libraries against protein targets.

Protein/peptide docking remains a challenging task due to the high flexibility of peptides, which undermines the prediction of accurate bound conformations (Rentzsch and Renard, 2015; Ciemny et al., 2018; Hashemi et al., 2021). To address this issue, existing protein/peptide docking methodologies have relied on different strategies. Peptide poses can be generated on the fly, and the most favorable ones can be selected according to their binding energy scores (Morris et al., 1998; Ewing et al.,

2001; Staneva and Wallin, 2009; Hashemi et al., 2021). However, this approach, coined as *de novo* docking, is not suitable for larger peptides, as exhaustive conformational sampling becomes prohibitive due to the presence of many freely-rotatable covalent bonds (Yan et al., 2017; Ansar and Vetrivel, 2019). This hurdle can be potentially overcome by conducting ensemble docking, which involves the generation of peptide conformations that can be subsequently docked into the protein binding site through rigid docking (Yan et al., 2017; Ciemny et al., 2018; Zhou et al., 2018; Ansar and Vetrivel, 2019). Moreover, MD simulation-based refinement steps can improve the accuracy of docked peptide poses (Trellet et al., 2013; Schindler et al., 2015; Yan et al., 2016; Ciemny et al., 2018; Hashemi et al., 2021).

Despite the availability of multiple peptide/protein docking tools (Ciemny et al., 2018; Hashemi et al., 2021), their efficient integration into SBVS campaigns is not straightforward (Ansar and Vetrivel, 2019). Recently, Ansar and Vetrivel presented PepVis, a pipeline for peptide SBVS following an ensemble docking approach (Ansar and Vetrivel, 2019). Their pipeline involves the generation of multiple peptide conformations that are rigidly docked into the targets and subsequently rescored and refined. As in PepVis, the D-peptide library screened in our work was created from the sequence and the starting structures were solvated and energy-minimized. However, to keep things simpler, we intentionally built very small D-peptide (three and four residues) libraries, thus avoiding issues like predicting the peptide secondary structure (Ansar and Vetrivel, 2019). This was possible in our case, as we concluded by visually inspecting several crystal structures of 3CL^{Pro} in complex with peptidomimetic inhibitors that D-tripeptides and especially D-tetrapeptides are large enough to occupy the main pockets of the enzyme's active site. Given the small size of the screened D-peptides, the ensemble approach was deemed unnecessary as it would have required additional computational steps. Instead, bound conformations during SBVS were generated and scored by the docking algorithm, as customary for small molecules. Subsequent rescoring steps in our workflow, such as generating multiple poses with AutoDock Vina using increased exhaustiveness of the search and different random seeds and MD simulations, aimed to enhance the conformational sampling of the analyzed D-peptides.

To prepare and screen libraries of larger D-peptides starting from the sequence, additional steps related to the peptide structure prediction and conformational search will be required. Existing methods and pipelines devised for protein/peptide docking (Yan et al., 2017; Ciemny et al., 2018; Ansar and Vetrivel, 2019; Hashemi et al., 2021) can be readily used if mirror images of the L-peptide structures (Garton et al., 2018) are created prior to the search for bound poses. However, the most straightforward strategy, inspired by mirror-image phage display experiments to identify D-peptide ligands (Schumacher et al., 1996), is to invert the configuration of the target protein C α atoms and use an L-peptide library for SBVS.

The D-peptides proposed here as 3CL^{Pro} inhibitors are mostly made up of aromatic residues plus HIS. Initially, we were intrigued by the fact that LEU was not found in any position of the selected D-peptides, although both SARS-CoV and SARS-CoV-2 3CL^{Pro}s prefer this residue at P2 (Chuck et al., 2010; Rut et al., 2021). However, at least in multiple heterochiral peptides

reported by Rut et al., the preference for LEU at P2 is restricted to the L-enantiomer (Rut et al., 2021), which might explain why LEU is absent from the identified D-peptides. Furthermore, the predicted variable accommodation of D-peptide backbones along the 3CL^{Pro} active site, divergent from the canonical conformation of the L-peptides (Figure 3), implies that specificity profiles obtained for L-peptide substrates cannot be straightforwardly extrapolated to small D-peptides. Finally, it is worth noting that the S2 subsite of 3CL^{Pro} can accommodate bulky aromatic moieties, e.g., 3-fluoro-L-PHE (PDB: 6M0K) (Dai et al., 2020), 4-nitro-L-phenylalanine (PHE(4-NO₂)) and 2,3-dihydro-L-tryptophan (Dht) (Rut et al., 2021), thus demonstrating that there is room in this subsite for large residues like those observed in the predicted D-peptides.

Apart from the previous factors, the absence of aliphatic residues in the identified D-peptides in favor of aromatic residues may arise from a persistent bias in our workflow toward larger ligands. In fact, it is known that Autodock Vina and other docking algorithms tend to overestimate the affinity of large ligands (Chang et al., 2010). Nonetheless, as mentioned before, the impact of such bias on the final results was reduced by the subsequent rescoring steps combining MM-GBSA free energy calculations and MD simulations. Interestingly, in a previous work, two L-peptides, HHYWH and HYWWT, identified as potential 3CL^{Pro} inhibitors using AutoDock Vina (Porto, 2021), showed a high content of HIS, TYR, and TRP, in resemblance to our results. This coincidence seems to reinforce the occurrence of a bias toward the former residues in the docking algorithm.

Despite the inaccuracies of the employed computational techniques pointed out earlier, the inhibition assays carried out for the four top-ranked D-peptides (4P1, 4P2, 4P3, and 4P4) validated our predictions. The tested D-tetrapeptides displayed significant inhibition of 3CL^{Pro} activity at 20 μ M, causing 55–85% loss of activity in all cases. Moreover, the FELs obtained from microsecond-long MD simulations conducted for 3CL^{Pro} bound to 4P1, 4P2, and 4P3 showed the stability of such complexes during the simulation time and that they sampled conformations around one or two similar main energy minima. Conversely, 4P4 sampled several well-separated energy minima and relatively-large peptide RMSD variations along the replicate 1 μ s MD trajectories. However, further energetic analyses indicated that this D-peptide coexists as two nearly-isoenergetic conformations with binding free energies consistent with its experimental inhibitory potency. In general, the main conformations of the tested D-peptides obtained from the long MD simulations were similar to those of the 110 ns MD simulations that were conducted as part of the presented *in silico* workflow. Therefore, we believe that short MD simulations are sufficient to identify promising ligands. On the other hand, longer MD simulations can be conducted after experimental validation to predict more accurate complex conformations that can be used as starting points for structure-based optimization of the hits.

The predicted structures of 3CL^{Pro} in complex with the selected D-peptides indicate the occurrence of significant intermolecular H-bonds present in the available crystal structures of this protease and its close homologue SARS-CoV

3CL^{Pro}. In this sense, we observed the formation of H-bonds between the amide oxygen of ASN/GLN of the D-peptides and H163(NE2), which explains the strong preference for GLN at P1 (Singh et al., 2020), in several predicted complexes, including that of the tested D-peptide 4P1. Other D-peptides, such as 3P1 and 4P2, were found to accommodate THR and HIE at the S1 subsite and to either mimic the H-bonds formed between GLN in L-peptides and H163(NE2) or to form alternate stabilizing interactions within the subsite. Of note, D-amino acids with side-chains smaller than GLN could reach the bottom of the S1 subsite because the backbones of the analyzed D-peptides can lie closer to the pocket entrance than the L-peptide backbones. Overall, the previous results show that our workflow was able to capture the interactions underlying the fine-tuned specificity of the S1 subsite of 3CL^{Pro} (Rut et al., 2021).

D-peptides are considered attractive therapeutic agents (Liu et al., 2016). However, this type of molecule has not been explored as potential ligands of 3CL^{Pro} active site until now. Therefore, the tested D-tetrapeptides 4P1, 4P2, 4P3, and 4P4 expand the chemical repertoire of known 3CL^{Pro} inhibitors that can help combat Covid-19. Beyond this concrete example, the computational workflow presented here can contribute to the fast discovery of small D-peptide ligands targeting different 3CL^{Pro} variants that can arise under viral adaptation to drug pressure (Padhi and Tripathi, 2021), as well as other proteins of interest.

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.

REFERENCES

- Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., et al. (2015). GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* 1-2, 19–25. doi:10.1016/j.softx.2015.06.001
- Aldeghi, M., Heifetz, A., Bodkin, M. J., Knapp, S., and Biggin, P. C. (2016). Accurate Calculation of the Absolute Free Energy of Binding for Drug Molecules. *Chem. Sci.* 7, 207–218. doi:10.1039/c5sc02678d
- Alves, V. M., Bobrowski, T., Melo-Filho, C. C., Korn, D., Auerbach, S., Schmitt, C., et al. (2021). QSAR Modeling of SARS-CoV MproInhibitors Identifies Sufugolix, Cenicriviroc, Proglumetacin, and Other Drugs as Candidates for Repurposing against SARS-CoV-2. *Mol. Inf.* 40, 2000113. doi:10.1002/minf.202000113
- Amadei, A., Linssen, A. B. M., and Berendsen, H. J. C. (1993). Essential Dynamics of Proteins. *Proteins* 17, 412–425. doi:10.1002/prot.340170408
- Amin, S. A., Banerjee, S., Ghosh, K., Gayen, S., and Jha, T. (2021). Protease Targeted COVID-19 Drug Discovery and its Challenges: Insight into Viral Main Protease (Mpro) and Papain-like Protease (PLpro) Inhibitors. *Bioorg. Med. Chem.* 29, 115860. doi:10.1016/j.bmc.2020.115860
- Ansar, S., and Vetrivel, U. (2019). PepVis: An Integrated Peptide Virtual Screening Pipeline for Ensemble and Flexible Docking Protocols. *Chem. Biol. Drug Des.* 94, 2041–2050. doi:10.1111/cbdd.13607
- Banerjee, S., Yadav, S., Banerjee, S., Fakayode, S. O., Parvathareddy, J., Reichard, W., et al. (2021). Drug Repurposing to Identify Nilotinib as a Potential SARS-CoV-2 Main Protease Inhibitor: Insights from a Computational and *In Vitro* Study. *J. Chem. Inf. Model.* 61, 5469–5483. doi:10.1021/acs.jcim.1c00524
- Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984). Molecular Dynamics with Coupling to an External bath. *J. Chem. Phys.* 81, 3684–3690. doi:10.1063/1.448118
- Case, D. A., Belfon, K., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., Cheatham, T. E., III, et al. (2020). *Amber 2020*. San Francisco: University of California.
- Chang, M. W., Ayeni, C., Breuer, S., and Torbett, B. E. (2010). Virtual Screening for HIV Protease Inhibitors: a Comparison of AutoDock 4 and Vina. *PLoS One* 5, e11955. doi:10.1371/journal.pone.0011955
- Chen, R., and Weng, Z. (2002). Docking Unbound Proteins Using Shape Complementarity, Desolvation, and Electrostatics. *Proteins* 47, 281–294. doi:10.1002/prot.10092
- Chia, C. S. B., Xu, W., and Shuyi Ng, P. (2021). A Patent Review on SARS Coronavirus Main Protease (3CL^{Pro}) Inhibitors. *Chem. Med. Chem.* doi:10.1002/cmdc.202100576
- Chowdhury, K. H., Chowdhury, M. R., Mahmud, S., Tareq, A. M., Hanif, N. B., Banu, N., et al. (2020). Drug Repurposing Approach against Novel Coronavirus Disease (COVID-19) through Virtual Screening Targeting SARS-CoV-2 Main Protease. *Biology* 10, 2. doi:10.3390/biology10010002
- Chuck, C.-P., Chong, L.-T., Chen, C., Chow, H.-F., Wan, D. C.-C., and Wong, K.-B. (2010). Profiling of Substrate Specificity of SARS-CoV 3CL^{Pro}. *PLoS One* 5, e13197. doi:10.1371/journal.pone.0013197
- Ciemny, M., Kurcinski, M., Kamel, K., Kolinski, A., Alam, N., Schueler-Furman, O., et al. (2018). Protein-peptide Docking: Opportunities and Challenges. *Drug*

AUTHOR CONTRIBUTIONS

JEHG, RJE, MAC Methodology, Investigation, Validation, Formal analysis. MAC Conceptualisation. JEHG Writing. DW Resources and Editing. RJE, MAC Reviewing and Editing.

FUNDING

This work was supported in part by São Paulo Research Foundation (FAPESP), Grant 2020/10214-1, and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Edital 09/2020 Pandemias, Grant 88881.503808/2020-01.

ACKNOWLEDGMENTS

We would like to thank the support of the Institute of Biological Information Processing (IBI-7) Forschungszentrum Jülich, Germany and of the National Laboratory for Scientific Computing (LNCC/MCTI, Brazil) for providing HPC resources of the SDumont supercomputer, which have contributed to the research results reported within this work (<http://sdumont.lncc.br>). We are grateful to our Brazilian partner, whom the federal government is constantly neglecting.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2021.816166/full#supplementary-material>

- Discov. Today* 23, 1530–1537. doi:10.1016/j.drudis.2018.05.006
- Connolly, M. L. (1983). Analytical Molecular Surface Calculation. *J. Appl. Cryst.* 16, 548–558. doi:10.1107/S0021889883010985
- 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, 1331–1335. doi:10.1126/science.abb4489
- Darden, T., York, D., and Pedersen, L. (1993). Particle Mesh Ewald: AnN-Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* 98, 10089–10092. doi:10.1063/1.464397
- DeLano, W. L. (2002). *PyMOL*. San Carlos, CA: PyMOL, 700.
- Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G., et al. (2007). PDB2PQR: Expanding and Upgrading Automated Preparation of Biomolecular Structures for Molecular Simulations. *Nucleic Acids Res.* 35, W522–W525. doi:10.1093/nar/gkm276
- Douangamath, 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, 5047. doi:10.1038/s41467-020-18709-w
- Eberle, R. J., Olivier, D. S., Amaral, M. S., Gering, I., Willbold, D., Arni, R. K., et al. (2021). The Repurposed Drugs Suramin and Quinacrine Cooperatively Inhibit SARS-CoV-2 3CL^{pro} *In Vitro*. *Viruses* 13, 873. doi:10.3390/v13050873
- Ewing, T. J. A., Makino, S., Skillman, A. G., and Kuntz, I. D. (2001). DOCK 4.0: Search Strategies for Automated Molecular Docking of Flexible Molecule Databases. *J. Comput. Aided Mol. Des.* 15, 411–428. doi:10.1023/a:101115820450
- Federico, L. B., Silva, G. M., da Silva Hage-Melim, L. I., Gomes, S. Q., Barcelos, M. P., Galindo Francischini, I. A., et al. (2021). Identification of Known Drugs as Potential SARS-CoV-2 Mpro Inhibitors Using Ligand- and Structure-Based Virtual Screening. *Future Med. Chem.* 13, 1353–1366. doi:10.4155/fmc-2021-0025
- Funke, S. A., and Willbold, D. (2009). Mirror Image Phage Display-A Method to Generate D-Peptide Ligands for Use in Diagnostic or Therapeutical Applications. *Mol. Biosyst.* 5, 783–786. doi:10.1039/b904138a
- Garton, M., Nim, S., Stone, T. A., Wang, K. E., Deber, C. M., and Kim, P. M. (2018). Method to Generate Highly Stable D-Amino Acid Analogs of Bioactive Helical Peptides Using a Mirror Image of the Entire PDB. *Proc. Natl. Acad. Sci. USA* 115, 1505–1510. doi:10.1073/pnas.1711837115
- Ghahremanpour, M. M., Tirado-Rives, J., Deshmukh, M., Ippolito, J. A., Zhang, C.-H., Cabeza de Vaca, I., et al. (2020). Identification of 14 Known Drugs as Inhibitors of the Main Protease of SARS-CoV-2. *ACS Med. Chem. Lett.* 11, 2526–2533. doi:10.1021/acsmmedchemlett.0c00521
- Gogoi, M., Borkotoky, M., Borchetia, S., Chowdhury, P., Mahanta, S., and Barooah, A. K. (2021). Black tea Bioactives as Inhibitors of Multiple Targets of SARS-CoV-2 (3CL^{pro}, PL^{pro} and RdRp): a Virtual Screening and Molecular Dynamic Simulation Study. *J. Biomol. Struct. Dyn.* 1, 24. doi:10.1080/07391102.2021.1897679
- Gohlke, H., and Case, D. A. (2004). Converging Free Energy Estimates: MM-PB(GB)SA Studies on the Protein-Protein Complex Ras-Raf. *J. Comput. Chem.* 25, 238–250. doi:10.1002/jcc.10379
- Gohlke, H., Kiel, C., and Case, D. A. (2003). Insights into Protein-Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras-Raf and Ras-RalGDS Complexes. *J. Mol. Biol.* 330, 891–913. doi:10.1016/s0022-2836(03)00610-7
- Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1989). Coronavirus Genome: Prediction of Putative Functional Domains in the Non-structural Polyprotein by Comparative Amino Acid Sequence Analysis. *Nucl. Acids Res.* 17, 4847–4861. doi:10.1093/nar/17.12.4847
- GraphPad (2018). *One-way ANOVA Followed by Dunnett's Multiple Comparisons Test Was Performed Using GraphPad Prism Version 8.0.0 for Windows*. San Diego, California, USA: GraphPad. Available at: www.graphpad.com.
- Guedes, I. A., Costa, L. S. C., Dos Santos, K. B., Karl, A. L. M., Rocha, G. K., Teixeira, I. M., et al. (2021). Drug Design and Repurposing with DockThor-VS Web Server Focusing on SARS-CoV-2 Therapeutic Targets and Their Non-synonym Variants. *Sci. Rep.* 11, 5543. doi:10.1038/s41598-021-84700-0
- Günther, S., Reinke, P. Y. A., Fernández-García, Y., Lieske, J., Lane, T. J., Ginn, H. M., et al. (2021). X-ray Screening Identifies Active Site and Allosteric Inhibitors of SARS-CoV-2 Main Protease. *Science* 372, 642–646. doi:10.1126/science.abb7945
- Guo, S., Xie, H., Lei, Y., Liu, B., Zhang, L., Xu, Y., et al. (2021). Discovery of Novel Inhibitors against Main Protease (Mpro) of SARS-CoV-2 via Virtual Screening and Biochemical Evaluation. *Bioorg. Chem.* 110, 104767. doi:10.1016/j.bioorg.2021.104767
- Gupta, A., Rani, C., Pant, P., Vijayan, V., Vikram, N., Kaur, P., et al. (2020). Structure-Based Virtual Screening and Biochemical Validation to Discover a Potential Inhibitor of the SARS-CoV-2 Main Protease. *ACS Omega* 5, 33151–33161. doi:10.1021/acsomega.0c04808
- Gupta, Y., Maciorowski, D., Zak, S. E., Jones, K. A., Kathayat, R. S., Azizi, S.-A., et al. (2021). Bisindolylmaleimide IX: A Novel Anti-SARS-CoV2 Agent Targeting Viral Main Protease 3CL^{pro} Demonstrated by Virtual Screening Pipeline and *In-Vitro* Validation Assays. *Methods* 195, 57–71. doi:10.1016/j.ymeth.2021.01.003
- Hamdy, R., Fayed, A., Mostafa, A., Shama, N. M. A., Mahmoud, S. H., Mehta, C. H., et al. (2021). Iterated Virtual Screening-Assisted Antiviral and Enzyme Inhibition Assays Reveal the Discovery of Novel Promising Anti-SARS-CoV-2 with Dual Activity. *Int. J. Mol. Sci.* 22, 9057. doi:10.3390/ijms22169057
- Hashemi, Z. S., Zarei, M., Fath, M. K., Ganji, M., Farahani, M. S., Afsharnouri, F., et al. (2021). In Silico Approaches for the Design and Optimization of Interfering Peptides against Protein-Protein Interactions. *Front. Mol. Biosci.* 8, 669431. doi:10.3389/fmolb.2021.669431
- Hegyí, A., and Ziebuhr, J. (2002). Conservation of Substrate Specificities Among Coronavirus Main Proteases. *J. Gen. Virol.* 83, 595–599. doi:10.1099/0022-1317-83-3-595
- Hernández González, J. E., Hernández Alvarez, L., Pascutti, P. G., and Valiente, P. A. (2017). Predicting Binding Modes of Reversible Peptide-Based Inhibitors of Falcipain-2 Consistent with Structure-Activity Relationships. *Proteins* 85, 1666–1683. doi:10.1002/prot.25322
- Hernández González, J. E., Salas-Sarduy, E., Hernández Alvarez, L., Barreto Gomes, D. E., Pascutti, P. G., Oostenbrink, C., et al. (2021). In Silico identification of Noncompetitive Inhibitors Targeting an Uncharacterized Allosteric Site of Falcipain-2. *J. Comput. Aided Mol. Des.* 35, 1067–1079. doi:10.1007/s10822-021-00420-7
- Hess, B. (2002). Determining the Shear Viscosity of Model Liquids from Molecular Dynamics Simulations. *J. Chem. Phys.* 116, 209–217. doi:10.1063/1.1421362
- Hou, T., Wang, J., Li, Y., and Wang, W. (2011). Assessing the Performance of the Molecular mechanics/Poisson Boltzmann Surface Area and Molecular Mechanics/generalized Born Surface Area Methods. II. The Accuracy of Ranking Poses Generated from Docking. *J. Comput. Chem.* 32, 866–877. doi:10.1002/jcc.21666
- Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., et al. (2020). Structure of Mpro from SARS-CoV-2 and Discovery of its Inhibitors. *Nature* 582, 289–293. doi:10.1038/s41586-020-2223-y
- Jukić, M., Janežič, D., and Bren, U. (2020). Ensemble Docking Coupled to Linear Interaction Energy Calculations for Identification of Coronavirus Main Protease (3CL^{pro}) Non-covalent Small-Molecule Inhibitors. *Molecules* 25, 5808. doi:10.3390/molecules25245808
- Kiemer, L., Lund, O., Brunak, S., and Blom, N. (2004). Coronavirus 3CL^{pro} Proteinase Cleavage Sites: Possible Relevance to SARS Virus Pathology. *BMC Bioinformatics* 5, 72. doi:10.1186/1471-2105-5-72
- Kneller, D. W., Phillips, G., O'Neill, H. M., Jedrzejczak, R., Stols, L., Langan, P., et al. (2020). Structural Plasticity of SARS-CoV-2 3CL^{pro} Active Site Cavity Revealed by Room Temperature X-ray Crystallography. *Nat. Commun.* 11, 3202. doi:10.1038/s41467-020-16954-7
- Kumar, A., Mishra, D. C., Angadi, U. B., Yadav, R., Rai, A., and Kumar, D. (2021). Inhibition Potencies of Phytochemicals Derived from Sesame against SARS-CoV-2 Main Protease: A Molecular Docking and Simulation Study. *Front. Chem.* 9, 744376. doi:10.3389/fchem.2021.744376
- Li, P., Song, L. F., and Merz, K. M., Jr. (2015). Systematic Parameterization of Monovalent Ions Employing the Nonbonded Model. *J. Chem. Theor. Comput.* 11, 1645–1657. doi:10.1021/ct500918t
- Li, Z., Li, X., Huang, Y.-Y., Wu, Y., Liu, R., Zhou, L., et al. (2020). Identify Potent SARS-CoV-2 Main Protease Inhibitors via Accelerated Free Energy Perturbation-Based Virtual Screening of Existing Drugs. *Proc. Natl. Acad. Sci. USA* 117, 27381–27387. doi:10.1073/pnas.2010470117
- Liu, M., Li, X., Xie, Z., Xie, C., Zhan, C., Hu, X., et al. (2016). D-peptides as Recognition Molecules and Therapeutic Agents. *Chem. Rec.* 16, 1772–1786. doi:10.1002/tcr.201600005

- Lokhande, K. B., Doiphode, S., Vyas, R., and Swamy, K. V. (2021). Molecular Docking and Simulation Studies on SARS-CoV-2 Mpro Reveals Mitoxantrone, Leucovorin, Birinapant, and Dynasore as Potent Drugs against COVID-19. *J. Biomol. Struct. Dyn.* 39, 7294–7305. doi:10.1080/07391102.2020.1805019
- Ma, C., Sacco, M. D., Hurst, B., Townsend, J. A., Hu, Y., Szeto, T., et al. (2020). Boceprevir, GC-376, and Calpain Inhibitors II, XII Inhibit SARS-CoV-2 Viral Replication by Targeting the Viral Main Protease. *Cell Res* 30, 678–692. doi:10.1038/s41422-020-0356-z
- Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theor. Comput.* 11, 3696–3713. doi:10.1021/acs.jctc.5b00255
- Marra, M. A., Jones, S. J. M., Astell, C. R., Holt, R. A., Brooks-Wilson, A., Butterfield, Y. S. N., et al. (2003). The Genome Sequence of the SARS-Associated Coronavirus. *Science* 300, 1399–1404. doi:10.1126/science.1085953
- Mengist, H. M., Dilnessa, T., and Jin, T. (2021). Structural Basis of Potential Inhibitors Targeting SARS-CoV-2 Main Protease. *Front. Chem.* 9, 622898. doi:10.3389/fchem.2021.622898
- Meyer-Almes, F.-J. (2020). Repurposing Approved Drugs as Potential Inhibitors of 3CL-Protease of SARS-CoV-2: Virtual Screening and Structure Based Drug Design. *Comput. Biol. Chem.* 88, 107351. doi:10.1016/j.compbiolchem.2020.107351
- Miller, B. R., 3rd, McGee, T. D., Jr., 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. Theor. Comput.* 8, 3314–3321. doi:10.1021/ct300418h
- Miyamoto, S., and Kollman, P. A. (1992). Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* 13, 952–962. doi:10.1002/jcc.540130805
- Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., et al. (1998). Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem.* 19, 1639–1662. doi:10.1002/(SICI)1096-987X(19981115)19:14<1639:AID-JCC10>3.0.CO;2-B
- Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., et al. (2009). AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* 30, 2785–2791. doi:10.1002/jcc.21256
- Muramatsu, T., Takemoto, C., Kim, Y.-T., Wang, H., Nishii, W., Terada, T., et al. (2016). SARS-CoV 3CL Protease Cleaves its C-Terminal Autoprocessing Site by Novel Subsite Cooperativity. *Proc. Natl. Acad. Sci. USA* 113, 12997–13002. doi:10.1073/pnas.1601327113
- Naik, S. R., Bharadwaj, P., Dingelstad, N., Kalyanamoorthy, S., Mandal, S. C., Ganesan, A., et al. (2021). Structure-based Virtual Screening, Molecular Dynamics and Binding Affinity Calculations of Some Potential Phytocompounds against SARS-CoV-2. *J. Biomol. Struct. Dyn.* 1, 18. doi:10.1080/07391102.2021.1891969
- Nguyen, H., Roe, D. R., and Simmerling, C. (2013). Improved Generalized Born Solvent Model Parameters for Protein Simulations. *J. Chem. Theor. Comput.* 9, 2020–2034. doi:10.1021/ct3010485
- Olubiyi, O. O., Olagunju, M., Keutmann, M., Loschwitz, J., and Strodel, B. (2020). High Throughput Virtual Screening to Discover Inhibitors of the Main Protease of the Coronavirus SARS-CoV-2. *Molecules* 25, 3193. doi:10.3390/molecules25143193
- Owen, D. R., Allerton, C. M. N., Anderson, A. S., Aschenbrenner, L., Avery, M., Berritt, S., et al. (2021). An Oral SARS-CoV-2 M Pro Inhibitor Clinical Candidate for the Treatment of COVID-19. *Science* 374, 1586–1593. doi:10.1126/science.abl4784
- Padhi, A. K., and Tripathi, T. (2021). Targeted Design of Drug Binding Sites in the Main Protease of SARS-CoV-2 Reveals Potential Signatures of Adaptation. *Biochem. Biophysical Res. Commun.* 555, 147–153. doi:10.1016/j.bbrc.2021.03.118
- Pan, Y., Huang, N., Cho, S., and MacKerell, A. D., Jr. (2003). Consideration of Molecular Weight during Compound Selection in Virtual Target-Based Database Screening. *J. Chem. Inf. Comput. Sci.* 43, 267–272. doi:10.1021/ci020055f
- Papaleo, E., Mereghetti, P., Fantucci, P., Grandori, R., and De Gioia, L. (2009). Free-energy Landscape, Principal Component Analysis, and Structural Clustering to Identify Representative Conformations from Molecular Dynamics Simulations: the Myoglobin Case. *J. Mol. Graphics Model.* 27, 889–899. doi:10.1016/j.jmgm.2009.01.006
- Pathak, N., Chen, Y.-T., Hsu, Y.-C., Hsu, N.-Y., Kuo, C.-J., Tsai, H. P., et al. (2021). Uncovering Flexible Active Site Conformations of SARS-CoV-2 3CL Proteases through Protease Pharmacophore Clusters and COVID-19 Drug Repurposing. *ACS Nano* 15, 857–872. doi:10.1021/acsnano.0c07383
- Payne, S. (2017). Family Coronaviridae. *Viruses* 2017, 149–158. doi:10.1016/B978-0-12-803109-4.00017-9
- Porto, W. F. (2021). Virtual Screening of Peptides with High Affinity for SARS-CoV-2 Main Protease. *Comput. Biol. Med.* 133, 104363. doi:10.1016/j.compbimed.2021.104363
- Price, D. J., and Brooks, C. L., 3rd (2004). A Modified TIP3P Water Potential for Simulation with Ewald Summation. *J. Chem. Phys.* 121, 10096–10103. doi:10.1063/1.1808117
- Rajpoot, S., Alagumuthu, M., and Baig, M. S. (2021). Dual Targeting of 3CLpro and PLpro of SARS-CoV-2: A Novel Structure-Based Design Approach to Treat COVID-19. *Curr. Res. Struct. Biol.* 3, 9–18. doi:10.1016/j.crstbi.2020.12.001
- Rehman, M. T., AlAjmi, M. F., and Hussain, A. (2021). Natural Compounds as Inhibitors of SARS-CoV-2 Main Protease (3CLpro): A Molecular Docking and Simulation Approach to Combat COVID-19. *Cpd* 27, 3577–3589. doi:10.2174/1381612826999201116195851
- Rentsch, R., and Renard, B. Y. (2015). Docking Small Peptides Remains a Great challenge: an Assessment Using AutoDock Vina. *Brief. Bioinform.* 16, 1045–1056. doi:10.1093/bib/bbv008
- Rota, P. A., Oberste, M. S., Monroe, S. S., Nix, W. A., Campagnoli, R., Icenogle, J. P., et al. (2003). Characterization of a Novel Coronavirus Associated with Severe Acute Respiratory Syndrome. *Science* 300, 1394–1399. doi:10.1126/science.1085952
- Rut, W., Grobörz, K., Zhang, L., Sun, X., Zmudzinski, M., Pawlik, B., et al. (2021). SARS-CoV-2 Mpro Inhibitors and Activity-Based Probes for Patient-Sample Imaging. *Nat. Chem. Biol.* 17, 222–228. doi:10.1038/s41589-020-00689-z
- Ryckaert, J.-P., Ciccoliti, G., and Berendsen, H. J. C. (1977). Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* 23, 327–341. doi:10.1016/0021-9991(77)90098-5
- Sabbah, D. A., Hajjo, R., Bardaweel, S. K., and Zhong, H. A. (2021). An Updated Review on SARS-CoV-2 Main Proteinase (MPro): Protein Structure and Small-Molecule Inhibitors. *Curr. Top. Med. Chem.* 21, 442–460. doi:10.2174/1568026620666201207095117
- Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S., and Walker, R. C. (2013). Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theor. Comput.* 9, 3878–3888. doi:10.1021/ct400314y
- Schindler, C. E. M., de Vries, S. J., and Zacharias, M. (2015). Fully Blind Peptide-Protein Docking with pepATTRACT. *Structure* 23, 1507–1515. doi:10.1016/j.str.2015.05.021
- Schneider, T., and Stoll, E. (1978). Molecular-dynamics Study of a Three-Dimensional One-Component Model for Distortive Phase Transitions. *Phys. Rev. B* 17, 1302–1322. doi:10.1103/PhysRevB.17.1302
- Schumacher, T. N. M., Mayr, L. M., Minor, D. L., Jr., Milhollen, M. A., Burgess, M. W., and Kim, P. S. (1996). Identification of D-Peptide Ligands through Mirror-Image Phage Display. *Science* 271, 1854–1857. doi:10.1126/science.271.5257.1854
- Seeliger, D., and de Groot, B. L. (2010). Ligand Docking and Binding Site Analysis with PyMOL and Autodock/Vina. *J. Comput. Aided Mol. Des.* 24, 417–422. doi:10.1007/s10822-010-9352-6
- Selvaraj, C., Panwar, U., Dinesh, D. C., Boura, E., Singh, P., Dubey, V. K., et al. (2020). Microsecond MD Simulation and Multiple-Conformation Virtual Screening to Identify Potential Anti-COVID-19 Inhibitors against SARS-CoV-2 Main Protease. *Front. Chem.* 8, 595273. doi:10.3389/fchem.2020.595273
- Shao, J., Tanner, S. W., Thompson, N., and Cheatham, T. E. (2007). Clustering Molecular Dynamics Trajectories: 1. Characterizing the Performance of Different Clustering Algorithms. *J. Chem. Theor. Comput.* 3, 2312–2334. doi:10.1021/ct700119m
- Singh, E., Khan, R. J., Jha, R. K., Amera, G. M., Jain, M., Singh, R. P., et al. (2020). A Comprehensive Review on Promising Anti-viral Therapeutic Candidates Identified against Main Protease from SARS-CoV-2 through Various

- Computational Methods. *J. Genet. Eng. Biotechnol.* 18, 69. doi:10.1186/s43141-020-00085-z
- Sisakht, M., Mahmoodzadeh, A., and Darabian, M. (2021). Plant-derived Chemicals as Potential Inhibitors of SARS-CoV -2 Main Protease (6LU7), a Virtual Screening Study. *Phytotherapy Res.* 35, 3262–3274. doi:10.1002/ptr.7041
- Staneva, I., and Wallin, S. (2009). All-atom Monte Carlo Approach to Protein-Peptide Binding. *J. Mol. Biol.* 393, 1118–1128. doi:10.1016/j.jmb.2009.08.063
- Sun, N., Funke, S. A., and Willbold, D. (2012). Mirror Image Phage Display - Generating Stable Therapeutically and Diagnostically Active Peptides with Biotechnological Means. *J. Biotechnol.* 161, 121–125. doi:10.1016/j.jbiotec.2012.05.019
- Tahir Ul Qamar, M., Alqahtani, S. M., Alamri, M. A., and Chen, L.-L. (2020). Structural Basis of SARS-CoV-2 3CL^{pro} and Anti-COVID-19 Drug Discovery from Medicinal Plants. *J. Pharm. Anal.* 10, 313–319. doi:10.1016/j.jpha.2020.03.009
- Trellet, M., Melquiond, A. S. J., and Bonvin, A. M. J. J. (2013). A Unified Conformational Selection and Induced Fit Approach to Protein-Peptide Docking. *PLoS One* 8, e58769. doi:10.1371/journal.pone.0058769
- Trott, O., and Olson, A. J. (2009). AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* 31, 21334. doi:10.1002/jcc.21334
- Valiente, P. A., Wen, H., Nim, S., Lee, J., Kim, H. J., Kim, J., et al. (2021). Computational Design of Potent D-Peptide Inhibitors of SARS-CoV-2. *J. Med. Chem.* 64, 14955–14967. doi:10.1021/acs.jmedchem.1c00655
- Wang, C., Horby, P. W., Hayden, F. G., and Gao, G. F. (2020). A Novel Coronavirus Outbreak of Global Health Concern. *The Lancet* 395, 470–473. doi:10.1016/S0140-6736(20)30185-9
- Weiser, J. r., Shenkin, P. S., and Still, W. C. (1999). Approximate Atomic Surfaces from Linear Combinations of Pairwise Overlaps (LCPO). *J. Comput. Chem.* 20, 217–230. doi:10.1002/(SICI)1096-987X(19990130)20:2<217::AID-JCC4>3.0.CO;2-A
- Wiesehan, K., and Willbold, D. (2003). Mirror-image Phage Display: Aiming at the Mirror. *Chembiochem* 4, 811–815. doi:10.1002/cbic.200300570
- World Health Organization (2021). Coronavirus Disease (COVID-19) Weekly Epidemiological Update and Weekly Operational Update. Available at: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports> (Accessed December 18th, 2021).
- Wu, C., Liu, Y., Yang, Y., Zhang, P., Zhong, W., Wang, Y., et al. (2020a). Analysis of Therapeutic Targets for SARS-CoV-2 and Discovery of Potential Drugs by Computational Methods. *Acta Pharmaceutica Sinica B* 10, 766–788. doi:10.1016/j.apsb.2020.02.008
- Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Song, Z.-G., et al. (2020b). A New Coronavirus Associated with Human Respiratory Disease in China. *Nature* 579, 265–269. doi:10.1038/s41586-020-2008-3
- Yan, C., Xu, X., and Zou, X. (2016). Fully Blind Docking at the Atomic Level for Protein-Peptide Complex Structure Prediction. *Structure* 24, 1842–1853. doi:10.1016/j.str.2016.07.021
- Yan, F., and Gao, F. (2021). An Overview of Potential Inhibitors Targeting Non-structural Proteins 3 (PL^{pro} and Mac1) and 5 (3CL^{pro}/M^{pro}) of SARS-CoV-2. *Comput. Struct. Biotechnol. J.* 19, 4868–4883. doi:10.1016/j.csbj.2021.08.036
- Yan, S., and Wu, G. (2021). Potential 3-chymotrypsin-like Cysteine Protease Cleavage Sites in the Coronavirus Polypeptides Pp1a and Pp1ab and Their Possible Relevance to COVID-19 Vaccine and Drug Development. *FASEB j.* 35, e21573. doi:10.1096/fj.202100280RR
- Yan, Y., Zhang, D., and Huang, S.-Y. (2017). Efficient Conformational Ensemble Generation of Protein-Bound Peptides. *J. Cheminform* 9, 59. doi:10.1186/s13321-017-0246-7
- Yang, J., Lin, X., Xing, N., Zhang, Z., Zhang, H., Wu, H., et al. (2021). Structure-Based Discovery of Novel Nonpeptide Inhibitors Targeting SARS-CoV-2 M^{pro}. *J. Chem. Inf. Model.* 61, 3917–3926. doi:10.1021/acs.jcim.1c00355
- Zhang, L., Lin, D., Kusov, Y., Nian, Y., Ma, Q., Wang, J., et al. (2020a). α -Ketoamides as Broad-Spectrum Inhibitors of Coronavirus and Enterovirus Replication: Structure-Based Design, Synthesis, and Activity Assessment. *J. Med. Chem.* 63, 4562–4578. doi:10.1021/acs.jmedchem.9b01828
- Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., et al. (2020b). Crystal Structure of SARS-CoV-2 Main Protease Provides a Basis for Design of Improved α -ketoamide Inhibitors. *Science* 368, 409–412. doi:10.1126/science.abb3405
- Zhou, P., Jin, B., Li, H., and Huang, S.-Y. (2018). HPEPDOCK: a Web Server for Blind Peptide-Protein Docking Based on a Hierarchical Algorithm. *Nucleic Acids Res.* 46, W443–W450. doi:10.1093/nar/gky357
- Ziebuhr, J. (2005). The Coronavirus Replicase. *Curr. Top. Microbiol. Immunol.* 287, 57–94. doi:10.1007/3-540-26765-4_3

Conflict of Interest: The authors declare 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.

Copyright © 2022 Hernández González, Eberle, Willbold and Coronado. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.



Molecular Docking and Dynamics Studies to Explore Effective Inhibitory Peptides Against the Spike Receptor Binding Domain of SARS-CoV-2

Suvro Biswas^{1†}, Shafi Mahmud^{1*†}, Mohasana Akter Mita¹, Shamima Afrose¹, Md. Robiul Hasan¹, Mst. Sharmin Sultana Shimu¹, Md. Abu Saleh^{1*}, Gomaa Mostafa-Hedeab^{2,3}, Mohammed Alqarni⁴, Ahmad J. Obaidullah⁵ and Gaber El-Saber Batiha⁶

OPEN ACCESS

Edited by:

Arvind Ramanathan,
Argonne National Laboratory (DOE),
United States

Reviewed by:

Soumya Lipsa Rath,
National Institute of Technology
Warangal, India
Weiwei Xue,
Chongqing University, China

*Correspondence:

Shafi Mahmud
shafimahmudfz@gmail.com
Md. Abu Saleh
saleh@ru.ac.bd

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Biological Modeling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 08 October 2021

Accepted: 24 November 2021

Published: 27 January 2022

Citation:

Biswas S, Mahmud S, Mita MA,
Afrose S, Hasan MR,
Sultana Shimu MS, Saleh MA,
Mostafa-Hedeab G, Alqarni M,
Obaidullah AJ and
Batiha GE-S (2022) Molecular Docking
and Dynamics Studies to Explore
Effective Inhibitory Peptides Against
the Spike Receptor Binding Domain
of SARS-CoV-2.
Front. Mol. Biosci. 8:791642.
doi: 10.3389/fmolb.2021.791642

¹Department of Genetic Engineering and Biotechnology, University of Rajshahi, Rajshahi, Bangladesh, ²Pharmacology Department and Health Research Unit-medical College, Jouf University, Jouf, Saudi Arabia, ³Pharmacology Department, Faculty of Medicine, Beni-Suef University, Beni Suef, Egypt, ⁴Department of Pharmaceutical Chemistry, College of Pharmacy, Taif University, Taif, Saudi Arabia, ⁵Drug Exploration and Development Chair (DEDC), Department of Pharmaceutical Chemistry, College of Pharmacy, King Saud University, Riyadh, Saudi Arabia, ⁶Department of Pharmacology and Therapeutics, Faculty of Veterinary Medicine, Damanhour University, Damanhour, Egypt

The spread of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has become a pandemic due to the high transmission and mortality rate of this virus. The world health and economic sectors have been severely affected by this deadly virus, exacerbated by the lack of sufficient efficient vaccines. The design of effective drug candidates and their rapid development is necessary to combat this virus. In this study, we selected 23 antimicrobial peptides from the literature and predicted their structure using PEP-FOLD 3.5. In addition, we docked them to the SARS-CoV-2 spike protein receptor-binding domain (RBD) to study their capability to inhibit the RBD, which plays a significant role in virus binding, fusion and entry into the host cell. We used several docking programs including HDOCK, HPEPDOCK, ClusPro, and HawkDock to calculate the binding energy of the protein-peptide complexes. We identified four peptides with high binding free energy and docking scores. The docking results were further verified by molecular dynamics (MD) simulations to characterize the protein-peptide complexes in terms of their root-mean-square fluctuation (RMSF), root-mean-square deviation (RMSD), radius of gyration (Rg), solvent-accessible surface area (SASA), and hydrogen bond formation. Allergenicity and toxicity predictions suggested that the peptides we identified were non-allergenic and non-toxic. This study suggests that these four antimicrobial peptides could inhibit the RBD of SARS-CoV-2. Future *in vitro* and *in vivo* studies are necessary to confirm this.

Keywords: SARS-CoV-2, peptides, RBD, peptide-protein docking, molecular dynamics

INTRODUCTION

The whole world is currently experiencing a pandemic which originated in the Chinese city of Wuhan in Hubei province in late December 2019. This life-threatening agent was named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by WHO, which declared it as “the first pandemic of the 21st century” (De Wit et al., 2016; Dong et al., 2020; Gorbalenya et al., 2020; Li Q et al., 2020; Li X et al., 2020; Machhi et al., 2020; Zhu et al., 2020). SARS-CoV-2 is linear single-

stranded positive sense enveloped RNA virus which contains a crown-like spike on its surface. SARS-CoV-2 has a genome size ranging from 26 to 32 kilobases and a virion size of roughly 80–120 nm in diameter (Li et al., 2005; Cui et al., 2019; Chen, 2020; Gorbalenya et al., 2020; Lu et al., 2020; Machhi et al., 2020; Wrapp et al., 2020; Yin, 2020). At present, 224 countries and territories are affected by SARS-CoV-2 viral infection. As of November 16, 2021, there have been a total of 254,807,373 confirmed cases and 5,126,239 deaths (<https://www.worldometers.info/coronavirus/>). SARS-CoV-2 is considered the third most highly pathogenic coronavirus. Its genome encodes four structural proteins: a helical nucleocapsid protein (N), an envelope protein (E), membrane/matrix protein (M) which has a significant role in viral assembly, and the spike surface glycoprotein (S), which facilitates viral entry into the host cell (Ashour et al., 2020; Dehelean et al., 2020; Khan et al., 2020). Several studies suggest that SARS-CoV-2 is zoonotic in origin, with 79.9% nucleotide sequence identity with SARS-CoV, 51.8% identity with MERS-CoV, and 87.6–89% identity with the bat-origin SARS-like coronavirus (bat-SL-CoVZC45) (Dehelean et al., 2020; Machhi et al., 2020; Ren et al., 2020; Wu et al., 2020; Zhang and Holmes, 2020).

The S proteins of coronavirus consist of spike monomers with two subunits, S1 and S2 (Gui et al., 2017; Yuan et al., 2017; Kirchdoerfer et al., 2018; Song et al., 2018; Lan et al., 2020). The S1 subunit contains the receptor-binding domain (RBD) and N-terminal domain (NTD) which are responsible for virus binding and entry. The RBD is located in the middle part of the S1 subunit and is used as an antigen to raise antibodies that interrupt virus-host binding (Xiao et al., 2003; Babcock et al., 2004; He et al., 2004; Wong et al., 2004; Lan et al., 2020; Dejnirattisai et al., 2021). The S2 domain has a proposed fusion peptide and two heptad repeats (HR1 and HR2) that facilitate cell membrane fusion between viral and target cells following proteolytic activation (Wild et al., 1994; He et al., 2004; Shang et al., 2020; Kim et al., 2021). The RBD-containing S1 domain also contains the SD1 and SD2 subdomains at the C-terminus. Although both the NTD and the RBD are immunogenic, the RBD contains the interaction surface for ACE2. Due to the fact that the receptor-binding site (RBS) is incompletely driven into the down state, the RBD solely engages with the up state of ACE2 (Lan et al., 2020; Premkumar et al., 2020; Dejnirattisai et al., 2021; Yuan et al., 2021).

Usually, many years of research are required before vaccines enter clinical trial. However, in a record period, scientists and researchers have made great efforts to develop secure, efficient, and active SARS-CoV-2 vaccines. Currently, 13 vaccines have been approved for early or limited use, and 8 vaccines have been approved for complete use (<https://www.nytimes.com/interactive/2020/science/coronavirus-vaccine-tracker>). On December 31, 2020, the WHO prepared an emergency use listing (EUL) for a vaccine named 'BNT162b2/COMIRNATY Tozinameran (INN)' manufactured by Pfizer. SK Bio and the State Institute of India generated "AZD1222" and "Covishield" vaccines that received an EUL on February 16, 2021. The "Ad26.COV 2. S" developed by Janssen (a subsidiary of Johnson & Johnson) was displayed on March 12, 2021. Moderna developed

the "mRNA-1273" vaccine. The Sinopharm vaccine and the Sinovac-CoronaVac have also been granted EUL by the WHO [[https://www.who.int/news-room/q-a-detail/coronavirus-disease-\(covid-19\)-vaccines](https://www.who.int/news-room/q-a-detail/coronavirus-disease-(covid-19)-vaccines)]. To date, only 2.2% of people in low-income countries have had at least one SARS-CoV-2 vaccine dose (<https://ourworldindata.org/covid-vaccinations>).

In this study, the selected peptides were docked to the RBD of SARS-CoV-2, leading to the identification of four peptides with high binding free energy. These peptide-RBD complexes were subsequently subjected to molecular dynamics study. Structural attributes and conformations of the docked complexes were obtained from the MD simulations, and suggested stiff and inflexible interactions between the RBD active site and the hit peptides. In comparison to earlier studies, we utilized multiple docking programs in combination to identify four peptides with high binding affinity to the active site of the RBD. Several previous studies have suggested antiviral effects of small molecules and peptides against SARS-CoV-2 through binding to the RBD. These studies identified molecules that, although they were predicted to bind to the RBD, did not interact directly with the RBD active site (Rathod et al., 2020; Padhi et al., 2021; Priya et al., 2021). In contrast, all of the peptides we identified were predicted to bind directly to the RBD active site. Although a few previous studies identified peptides that formed a single non-bonded interaction with the RBD active site, the predicted binding energies were lower than ours using the same docking software (Chowdhury et al., 2020; Hossain et al., 2021) and the complexes were less stable in MD simulations.

MATERIALS AND METHODS

Peptide Screening and Preparation

In this study, we started from 27 peptide molecules that were previously identified in the venom of the wild bee *Hylaeus signatus* and which were screened for antimicrobial activity (Nešuta et al., 2016). Three peptides were excluded as they contained D-amino acids. Additionally, one further peptide was excluded as its amino acid sequence was incompletely characterized. The PEP-FOLD 3.5 webserver was used to predict the peptide structures from the amino acid sequences of the peptides (Lamiabile et al., 2016). This webserver uses a Hidden Markov Model suboptimal sampling algorithm to predict the peptide structures. The resulting peptide structures were used as the starting point for 20 ns molecular dynamics simulations, and the root mean square deviations of the alpha carbon atoms were calculated. The final frames of these molecular dynamics simulations were used for further studies.

Protein Preparation

The three-dimensional structure of the spike receptor-binding domain (RBD) of SARS-CoV-2 at 2.43 Å resolution (PDB ID: 6M0J) was retrieved from the Protein Data Bank. The protein structure was prepared by removing heteroatoms and water molecules using Discovery Studio (Discovery Studio, 2009). Additionally, energy minimization of the protein structure was performed using the AMBER14 (Case et al., 2014) force field in

YASARA software (Krieger et al., 2013). Molecular docking and dynamics studies used this energy minimized protein structure.

Molecular Docking

The peptides and RBD protein were uploaded as ligand and receptor molecules respectively to the HDock, HPEPDOCK, and ClusPro web servers. HDock uses a combined template-based and template-free algorithm in an automatic manner (Yan et al., 2017), while HPEPDOCK uses a hierarchical algorithm (Zhou et al., 2018). After docking the peptides to the protein using HDock and HPEPDOCK, ClusPro was used to calculate binding energies (Comeau et al., 2004). The top ten peptides with the highest docking scores were selected for further evaluation. These ten peptides were docked to the RBD using the HawkDock web server. For further analysis, the four highest-scoring peptides were chosen based on their binding free energy and docking scores from HawkDock. On the HawkDock server, the HawkRank scoring system, the ATTRACT docking algorithm improved in groups, and MM/GBSA free energy decomposition analysis are implemented on a multipurpose platform (Weng et al., 2019). PyMOL and Discovery Studio (Discovery Studio, 2009) were used for structural analysis of the top four protein-peptide complexes.

Molecular Dynamics Simulation

Molecular dynamics simulations was performed in YASARA dynamics (Land and Humble, 2018) using the AMBER14 force field (Wang et al., 2004). The docked peptide-protein complexes were initially cleaned, optimized and the hydrogen bond network was oriented. A cubic simulation cell was created with periodic boundary conditions and the TIP3P water model was used (Harrach and Drossel, 2014). The simulation cell was extended by 20 Å in each direction beyond the protein-peptide complexes. The physical conditions of the simulation cells were set at 298 K, pH 7.4, and 0.9% NaCl (Krieger and Vriend, 2015). The initial energy minimization of the simulation cells were conducted by steepest gradient approaches with simulated annealing methods (5,000 cycles). The time step of the simulation was 2.0 fs. Long-range electrostatic interactions were calculated by the Particle Mesh Ewald (PME) method with a cut-off radius of 8.0 Å (Essmann et al., 1995; Krieger et al., 2006; Harvey and De Fabritiis, 2009). The simulation trajectories were saved every 100 ps. The simulations were run for 100 ns at constant pressure and temperature, using a Berendsen thermostat. The simulation trajectories were used to calculate root mean square deviations, root mean square fluctuations, solvent accessible surface areas, radii of gyration, and hydrogen bonds (Adji et al., 2021; Dutta et al., 2021; Obaidullah et al., 2021).

The per residue energy contribution of the peptide-protein structures was calculated using the pyDockEneRes webserver. The first, last, and average structure was extracted from simulations trajectories and utilized as input entry, and the average higher energy from the hotspot residues were tabulated. This tool can be utilized for the identification of the hotspot residues (Romero-Durana et al., 2020).

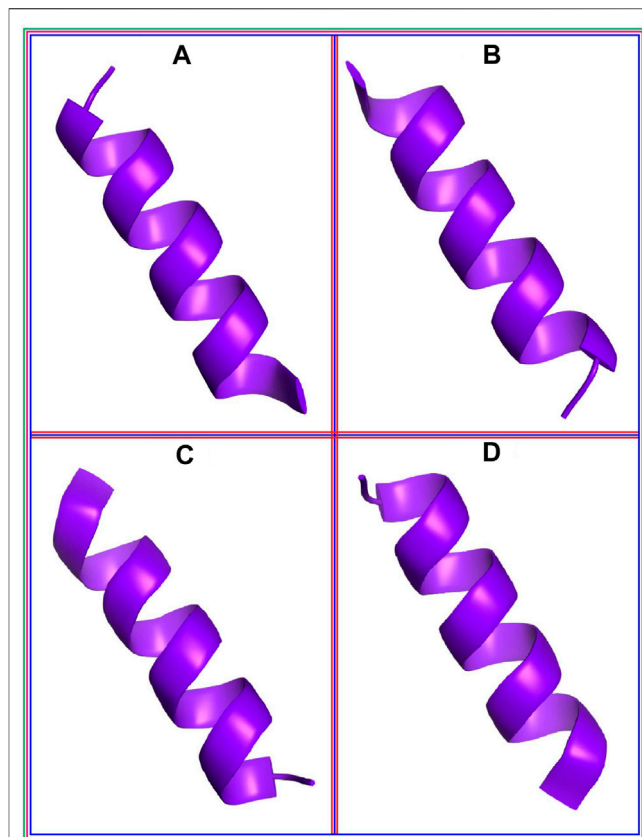


FIGURE 1 | The best four peptide molecules based on the binding free energy in the docking program; (A) P2, (B) P18, (C) P20, and (D) P21 peptide molecules.

$$\begin{aligned} E_{resipyDockSCEle_vdw} &= \sum sidechainiscEiscele + Eiscv dw \\ &= -\Delta\Delta G \text{ (Romero-Durana et al., 2020)} \end{aligned}$$

Allergenicity and Toxicity Prediction

AllerTOP (Dimitrov et al., 2014a) and AllergenFP (Dimitrov et al., 2014b) webserver were used to predict the allergenicity of the peptides. AllergenFP implements five E-descriptor-based fingerprinting, whilst AllerTOP uses both k-nearest neighbor (kNN) and amino acid E-descriptors to predict the allergenicity of peptides. The ToxinPred (Gupta et al., 2013) webserver was utilized to predict the toxicity of the peptides.

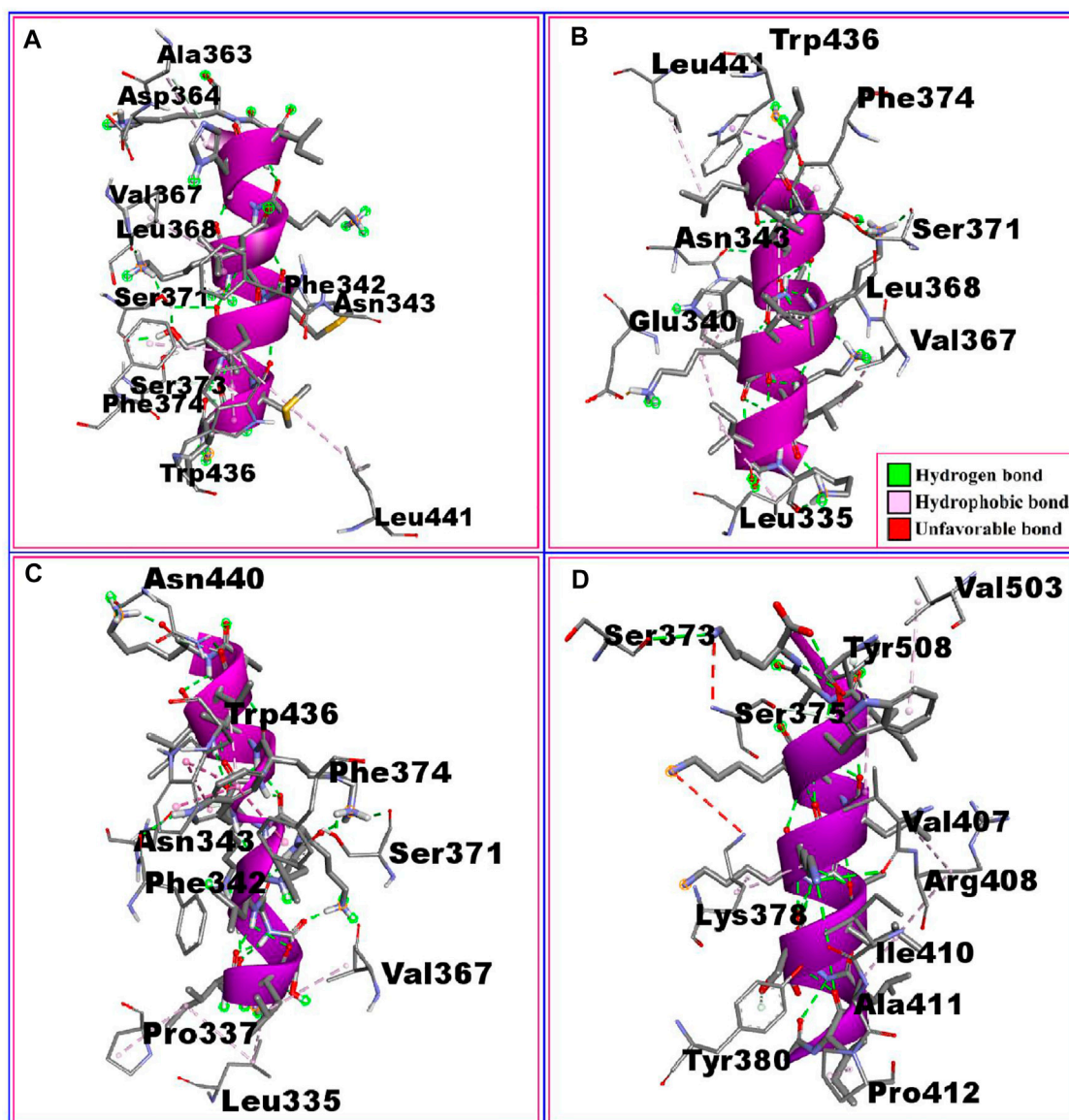
RESULTS AND DISCUSSION

Molecular Docking

Venom extracted from hymenopteran insects, including the solitary bee *Hylaeus signatus*, is a prominent source of antimicrobial peptides (AMPs). Several α -helical amphipathic AMPs, referred to as HYL, have been identified in the solitary bee, with antimicrobial activity against distinct strains of pathogenic bacteria and fungi as well as the ability to lyse cancer cells

TABLE 1 | The binding free energy and the dock score of the best four peptide molecules; P2, P18, P20, and P21

Peptide ID	Sequence	Dock score	Binding free energy of complex (kcal/mol)
P2	GIMSSLMKKLKAHIAK	-2,660.34	-28.15
P18	GILSSLWKKLKKIIAK	-2,815.13	-37.2
P20	GILSSLLKKWKKIIAK	-2,848.16	-40.38
P21	GILSSLLKKLKKWIAK	-2,735.49	-31.17

**FIGURE 2 |** The non-bonded interaction of the P2, P18, P20, P21 peptides and the RBD protein from SARS-CoV-2 at certain simulation times. Here, (A–D), represents the binding interactions between the P2, P18, P20, P21 peptides and the RBD protein after 0ns of simulation time respectively.

(Slaninová et al., 2011; Slaninová et al., 2012; Nešuta et al., 2016). Additionally, HYL had low hemolytic activity, suggesting that it may be safe for use in humans (Nešuta et al., 2016).

Notably, the earlier-discovered antimicrobial peptides, including cecropin-A from *Hyalophora cecropia*, exhibited

antiviral activity (Both, 1980). Cecropin-A was shown to have antiviral activity against HSV-1, HIV-1, and JV, whereas melittin from *A. mellifera* had antiviral activity against influenza A viruses, HSV-1, HIV-1, JV, TMV, RSV, VSV, enterovirus, and coxsackievirus (Albiol Matanic and Castilla, 2004). In addition,

TABLE 2 | The non-bonded interactions between the P2, P18, P20, P21 peptides and receptor-binding domain (RBD) protein of SARS-CoV-2 after 0ns of simulation times.

Peptide name	Protein	Bond distance (Å)	Interaction category
P2	ALA363	2.68	Hydrogen bond
	ASP364	3.36	Hydrophobic bond
	VAL367	2.77	Hydrogen bond
	LEU368	4.10	Hydrophobic bond
	SER371	2.80	Hydrogen bond
	SER373	2.95	Hydrogen bond
	PHE374	3.31	Hydrophobic bond
	TRP436	4.06	Hydrophobic bond
	PHE342	3.49	Hydrophobic bond
	ASN343	2.73	Hydrogen bond
P18	LEU441	3.53	Hydrophobic bond
	TRP436	3.97	Hydrophobic bond
	PHE374	3.41	Hydrophobic bond
	LEU441	3.39	Hydrophobic bond
	SER371	2.84	Hydrogen bond
	ASN343	3.11	Hydrogen bond
	LEU368	3.79	Hydrophobic bond
	GLU340	2.70	Hydrogen bond
	VAL367	3.77	Hydrophobic bond
	LEU335	3.96	Hydrophobic bond
P20	ASN440	2.94	Hydrogen bond
	TRP436	3.60	Hydrophobic bond
	PHE374	3.44	Hydrophobic bond
	ASN343	2.87	Hydrogen bond
	PHE342	2.77	Hydrogen bond
	SER371	3.69	Hydrogen bond
	PRO337	3.47	Hydrophobic bond
	VAL367	4.02	Hydrophobic bond
	LEU335	3.31	Hydrophobic bond
	VAL503	2.72	Hydrophobic bond
P21	SER373	3.01	Unfavorable bond
	TYR508	3.45	Hydrophobic bond
	SER375	2.49	Hydrogen bond
	VAL407	2.92	Hydrogen bond
	LYS378	3.83	Hydrophobic bond
	ARG408	3.33	Hydrophobic bond
	ILE410	3.11	Hydrogen bond
	ALA411	3.42	Hydrophobic bond
	TYR380	3.56	Hydrophobic bond
	PRO412	3.55	Hydrophobic bond

alloferon 1 and alloferon 2 from *C. vicina* were effective against influenza viruses (Chernysh et al., 2002), TnGlv1 and TnGlv2 from *Trichoplusia ni* were effective against AcMNPV (Moreno-Habel et al., 2012), attC and dptB from *Drosophila melanogaster* were effective against SINV (Huang et al., 2013), and a myristoylated peptide from *H. virescens* was effective against HIV-1 and HSV-1 (Ourth, 2004; Feng et al., 2020). Therefore, we hypothesized that the HYL antimicrobial peptides from the solitary bee could have antiviral activity and consequently devised an *in silico* study to test this hypothesis. We found that several HYL peptides bind the RBD of SARS-CoV-2 *in silico* with favorable binding energy and stable conformations through 100 ns of molecular dynamics simulation.

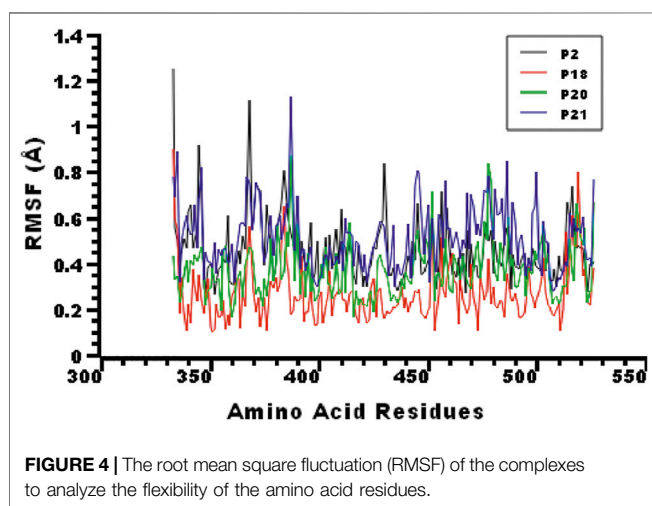
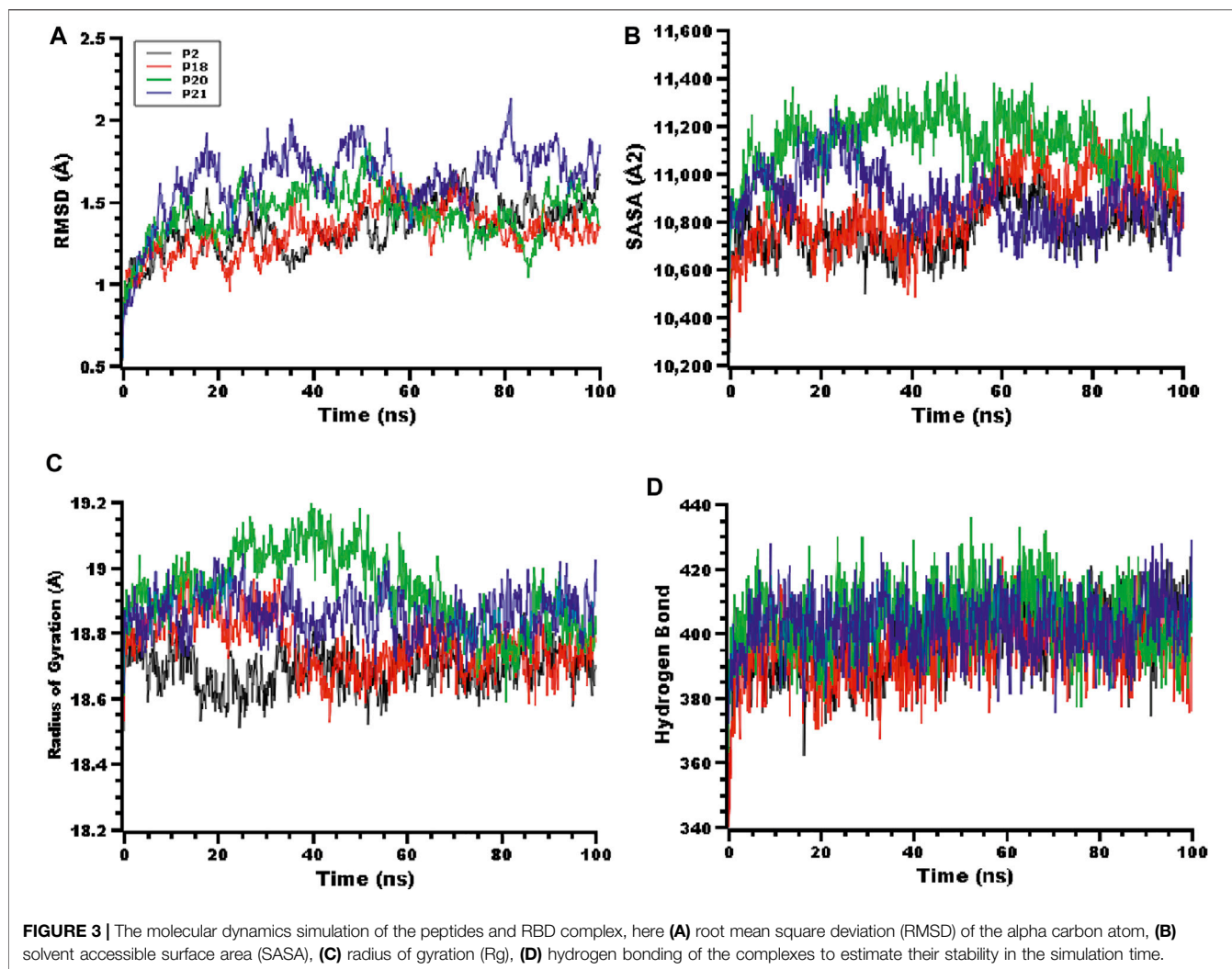
Molecular docking can be utilized to rapidly assess the binding affinities and modes between a target substrate, such as a protein, and diverse ligands, including small peptides, to assist antiviral drug design (Willett and Glen, 1995). Virtual screening can

identify the preferred binding orientation, optimal conformation, and binding sites of protein and peptide molecules (Alonso et al., 2006; Liu et al., 2018). Our computational approach confers improved structural precision and rapid prediction of both the existence and strength of binding through the use of multiple diverse docking algorithms aimed at achieving accuracy (Bartuzi et al., 2017).

The predicted peptide structures from PEP-FOLD 3.5 are shown in Supplementary Figure 1. The peptide structures were optimized by molecular dynamics simulations in which the root mean square deviations of the peptides were found to be small (Supplementary Figure 2). After predicting the peptide structures, all 23 peptides were docked to the RBD of SARS-CoV-2 by using the HDock, ClusPro, and HPEPDOCK webserver (Supplementary Table 1). Binding energies and/or docking scores were recorded for the ten highest-scoring peptides (Supplementary Table 2). Moreover, the binding free energies and docking scores from HawkDock (Supplementary Table 3) were used to select the four highest-scoring peptides (Figure 1). The highest docking score was $-2,848.16$ kcal/mol for peptide P20, followed by $-2,815.13$, $-2,735.49$, and $-2,660.34$ kcal/mol for P18, P21, and P2, respectively. The maximum binding free energy was -40.38 kcal/mol for the P20 peptide, followed by -37.2 , -31.17 , and -28.15 for P18, P21, and P2, respectively (Table 1).

The interactions between the RBD protein and the four highest-scoring peptides are shown in Figure 2. The P2:RBD complex had five hydrogen bonds at the RBD residues ALA363, VAL367, SER371, SER373, and ASN343, and six hydrophobic interactions at ASP364, LEU368, PHE374, TRP436, PHE342, and LEU441. The P18:RBD complex formed six hydrophobic interactions at TRP436, PHE374, LEU441, LEU368, VAL367, and LEU335, and three hydrogen bonds at SER371, ASN343, and GLU340. The P20:RBD complex had five hydrophobic interactions at TRP436, PHE374, PRO337, VAL367, and LEU335, and four hydrogen bonds at ASN440, ASN343, PHE342, and SER371. The P21:RBD complex had seven hydrophobic interactions at VAL503, TYR508, LYS378, ARG408, ALA411, TYR380, PRO412 position, one unfavorable bond at SER373 and three hydrogen bonds at SER375, VAL407, ILE410 position (Table 2; Figure 2).

The interactions between the RBD protein and the remaining six of the top ten peptides are shown in supplementary figure 3. The P3:RBD complex formed two hydrogen bonds at ALA344 and ASN343, five hydrophobic interactions at LEU441, LEU335, PHE374, VAL362, and VAL367, and one unfavorable bond at ARG509. The P6:RBD complex had five hydrophobic interactions at LEU441, PHE374, LEU368, VAL367, and LEU335, and six hydrogen bonds at ASN440, ARG509, TRP436, ASN343, PHE342, and GLY339 position. The P16:RBD complex had four hydrophobic interactions at CYS391, CYS525, ALA522, and VAL362, two hydrogen bonds at ASP389 and THR523, one salt bridge at ASP389 and one unfavorable bond at CYS361. The P17:RBD complex had five hydrophobic interactions at LEU335, VAL367, LEU368, PHE342, and TRP436, and three hydrogen bonds at SER371, ASN343, and ASN440. The P22:RBD complex formed five hydrogen bonds at ASN440, ASN343, SER371, GLU340, and ASP364, three hydrophobic interactions



at TRP436, VAL367, and LEU335, and one salt bridge at GLU340. The P23:RBD complex formed six hydrophobic interactions at ALA522, THR333, CYS391, CYS525, VAL362, and LEU390 and

two hydrogen bonds at GLY526 and ASP389 (Supplementary table 4, Supplementary figure 3).

Molecular Dynamics

Molecular dynamics simulations were conducted to explore the structural stability of the docked peptide-protein complexes. The root mean square deviations (RMSD) from the peptide-protein systems were calculated from the simulation trajectories. **Figure 3A** indicates that the peptide-protein complexes involving the P2, P18, P20, and P21 peptides had an initial upwards RMSD trend, suggesting flexibility of the complexes. The peptide-protein systems subsequently stabilized after 30 ns and maintained their integrity for the remaining 70 ns of the simulations. The P21 complex had a comparatively higher RMSD than the other three peptide complexes, which indicates the more flexible nature of this complex. However, all four peptide-protein complexes had RMSD lower than 2.5 Å, which indicates the stable nature of the complexes over the whole simulations.

The solvent-accessible surface area (SASA) of the complexes was also explored. **Figure 3B** indicates that the P20 and P21 complexes increased in SASA upon binding with the protein target, and after 30 ns the P21 complex decreased in SASA,

TABLE 3 | The per residue energy contribution from RBD of SARS-CoV-2 where energy contribution was considered <2 kcal/mol.

Complex	Residues	Energy
P2	Leu335	-4.27
	Glu340	-2.28
	Val367	-2.74
	Phe374	-3.37
P18	Asp364	-2.94
	Ser373	-2.21
	Ser375	-2.36
	Trp380	-2.21
P20	Leu335	-3.68
	Asp364	-3.71
	Val367	-3.10
	Trp436	-2.50
P22	Leu335	-5.70
	Phe342	-4.07
	Asn343	-2.60
	Asp364	-4.49
	Val367	-5.96
	Ser373	-2.35
	Phe374	-3.90
	Glu484	-2.21

maintaining a similar SASA trend to the rest of the complexes. The P20 complex stabilized after 50 ns and maintained its integrity. The radius of gyration (Rg) of the complexes was also explored to understand the compactness of the complexes. High Rg indicates an extended nature whereas a low Rg indicates a more compact structure. **Figure 3C** indicates that P20 had slightly higher Rg than other complexes which indicates that this complex is less compact. The other complexes had smaller deviations, which indicate the rigid nature of these complexes. Hydrogen bonding can play a crucial role in determining the stability of protein complexes. **Figure 3D** indicates that all complexes had a stable hydrogen bonding profile over the 100 ns of simulation.

The root mean square fluctuations (RMSF) of the complexes were explored to understand the flexibility on a residue-by-residue basis. **Figure 4** indicates that almost every residue of the complexes had an RMSF lower than 1.4 Å, indicating the stable nature of the protein-peptide systems.

After 100 ns of simulation time, the P2:RBD complex was stabilized by five hydrogen bonds at ASN343, SER371, ASN370, ASP364, and ALA363, as well as five hydrophobic interactions at LEU441, TRP436, LEU368, VAL367, and LEU335. The P18:RBD complex had two hydrogen bonds at SER371 and GLU340, as well as six hydrophobic interactions at TRP436, PHE374, PHE342, LEU368, VAL367 and LEU335. The P20:RBD complex formed two hydrogen bonds at ASN440 and SER371, in addition to three

hydrophobic interactions at TRP436, LEU368 and LEU335. The P21:RBD complex had four hydrogen bonds at ILE410, LYS378, TYR508 and PHE374, in addition to three hydrophobic interactions at PRO412, ALA411 and VAL407 (Supplementary table 5, Supplementary figure 4).

In the P2:RBD simulation, the interactions at LEU441, TRP436, ASN343, SER371, LEU368, VAL367, ASP364, and ALA363 were occupied throughout the whole simulation period. Similarly for TRP436, PHE374, SER371, GLU340, LEU368, VAL367, and LEU335 in the P18:RBD simulation, ASN440, TRP436, SER371, and LEU335 in the P20:RBD simulation, and PRO412, ALA411, ILE410, LYS378, VAL407, and TYR508 in the P21:RBD simulation (**Table 2; Figure 2**, Supplementary table 5, Supplementary figure 4).

The active sites of the RBD protein include amino acid residues 340, 374, 375, 378, 403, 420, 477, 478, 499, 543, and 546 (Lan et al., 2020). At the start of the simulation, the P2 peptide bound to the active site of RBD via a hydrophobic interaction with the conserved residue PHE374 (bond distance 3.31360), whereas the P18 peptide interacted with the RBD active site at both PHE374 (Bond distance 3.41730) and GLU340 (bond distance 2.70389) with a hydrophobic interaction and a hydrogen bond respectively. The P20 peptide bound to the active site of RBD at PHE374 (bond distance 3.44294) via a hydrophobic interaction, and the P21 peptide interacted with the RBD active site at SER375 (bond distance 2.49462) and LYS378 (bond distance 3.83247) position with a hydrogen bond and a hydrophobic interaction respectively (**Table 2; Figure 2**).

After 100 ns of simulation time, the P18 peptide was bound to the active site of RBD at two conserved residues, PHE374 (bond distance 3.44205) and GLU340 (bond distance 2.73681), via a hydrophobic interaction and a hydrogen bond respectively. The P21 peptide interacted with the active site of RBD via LYS378 (bond distance 2.81702) and PHE374 (bond distance 2.87497) via two hydrogen bonds (Supplementary table 5, Supplementary figure 4). These interactions of the peptides with the active site of the RBD indicate that these peptides may inhibit the RBD with high binding affinity.

Previously suggested multifunctional peptides, including ALPEEVIQHTFNLKSQ and DIENLIKSQ from *Bacillus*-fermented soybean, failed to bind to the active site of the RBD. However, our top four peptides bound to the active site. In addition, P18 and P20 had higher binding energy than DIENLIKSQ when employing the same HawkDock server (Padhi et al., 2021). The computationally designed peptide APASMFLGKGDHEILM made no interactions with the active site when docked using the same HPEPDOCK server that we used for our top four peptides (Priya et al., 2021). Another peptide

TABLE 4 | The allergenicity and toxicity profiling of the best four-peptide molecules.

Peptide ID	Sequence	Allergenicity prediction		Toxicity prediction through toxinPred
		AllergenFP v.1.0	AllerTOP v. 2.0	
P2	GIMSSLMKKLKAHIAK	Probable allergen	Probable allergen	Non-toxic
P18	GILSSLLWKLLKKIIAK	Probable non-allergen	Probable non-allergen	Non-toxic
P20	GILSSLLKKWKIIAK	Probable non-allergen	Probable non-allergen	Non-toxic
P21	GILSSLLKKLKKWIAK	Probable non-allergen	Probable allergen	Non-toxic

modeling and screening study suggested that AVP0671 can bind to the RBD, albeit not at the active site, and the HDOCK scores were not reported. Meanwhile, our top four suggested peptides not only exhibited active site binding but also had a more favorable HDOCK docking score (Rathod et al., 2020). Antiviral peptides S2P25 and S2P26 were predicted to form one bond at the active site, but had lower docking interaction energies than our P21 and P18 peptides when comparing the Cluspro scores (Chowdhury et al., 2020). Additionally, AVP1795, identified from the computational screening of 645 antiviral peptides, exhibited only one interaction (ARG403) at the active site, which was occupied for $\geq 90\%$ of the time in MD simulations, whereas our peptides were stably bound over the whole simulation period (Hossain et al., 2021).

Multiple methodology developed to understand the hotspot residues in protein-protein interactions which are based on snapshot from MD simulation sampling (Gohlke et al., 2003). In addition to pyDockEneRes, mm_pbsa.pl tools successfully implemented for hotspot residues identifications from cytokines and receptor interface (Du et al., 2020) as well as RBD from SARS-CoV-2 and antibodies (Yang et al., 2021). Moreover, the pyDockEneRes tools enables the hotspot residues identifications from the protein-protein interactions which is key to understand the biological process at molecular level (Romero-Durana et al., 2020). In our study, the P2 peptide and RBD complex, four important residues contribute to overall energy contributions: Leu335, Glu340, Val367, and Phe374; in the P18-RBD complex, hotspot residues include Asp164, Ser373, Ser375, and Trp380 (Table 3). Leu335, Asn343, Asp364, Val367, Ser373, Phe374, and Glu484 were major energy contributing residues in the P20-RBD complex, whereas Leu335, Phe342, Asn343, Asp364, Val367, Ser373, Phe374, and Glu484 were key energy contributing residues in the P22-RBD complex.

Allergenicity and Toxicity Prediction

An allergenic antigen can activate Th2 cells resulting in stimulation of B cells to generate immunoglobulin E (IgE) that binds to Fc ϵ RI and activates eosinophils leading to inflammation and tissue shrinkage (Dimitrov et al., 2013). The online AllerTOP tool, which assesses allergens using E-descriptors affined with amino acid attributes, was utilized to predict the allergenicity of our peptides (Dimitrov et al., 2014a). Three distinct web-based approaches were utilized to assess toxicity (Table 4). According to the AllergenFP v1.0 webserver, P2 is a probable allergen, whereas P18, P20, and P21 are probably non-allergenic. The same result was obtained using the AllerTOP v2.0 webserver. In addition, all four peptides (P2, P18, P20, and P21) were categorized as non-toxic by the ToxinPred web server.

REFERENCES

- Adj, A., Niode, N. J., Memah, V. V., Posangi, J., Wahongan, G. J. P., Ophinni, Y., et al. (2021). Designing an Epitope Vaccine against Dermatophagoides Pteronyssinus: An In Silico Study. *Acta Tropica* 222, 106028. doi:10.1016/j.actatropica.2021.106028
- Albiol Matanic, V. C., and Castilla, V. (2004). Antiviral Activity of Antimicrobial Cationic Peptides against Junin Virus and Herpes

CONCLUSION

To develop novel therapeutics against SARS-CoV-2, targeting the spike protein RBD by designing antiviral peptides could identify promising leads. In this study, 23 peptides were docked to the RBD of SARS-CoV-2, leading to the identification of four peptides with high binding affinity. Molecular dynamics studies demonstrated that docked peptides were not ductile in nature, but were instead rigidly bound. Moreover, allergenicity and toxicity profiling of the peptides suggest that they have no allergenic or toxic properties. Finally, our study may facilitate the development of efficient drugs against SARS-CoV-2 through further *in vitro* studies.

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 authors.

AUTHOR CONTRIBUTIONS

SM: Conceptualization, investigation, data curation, validation, software, formal analysis, writing, editing, visualization, and supervision; SB: investigation, data curation, validation, software, formal analysis, writing, editing, and visualization; MM: formal analysis, writing, and visualization; SA: formal analysis and writing; MH: formal analysis and writing; SS: formal analysis and writing, MS: editing and supervision; GM-H: editing and writing, MA: editing and writing; AO: editing and writing, GB: editing and writing. All authors have read and agreed to the current version of the manuscript.

FUNDING

The authors would like to extend their sincere appreciation to Taif University Researchers supporting Project number (TURSP 2020/309) Taif University, Taif, Saudi Arabia.

ACKNOWLEDGMENTS

The authors would like to extend their sincere appreciation to Taif University Researchers supporting Project number (TURSP 2020/309) Taif University, Taif, Saudi Arabia.

Simplex Virus. *Int. J. Antimicrob. Agents* 23, 382–389. doi:10.1016/j.ijantimicag.2003.07.022

Alonso, H., Bliznyuk, A. A., and Gready, J. E. (2006). Combining Docking and Molecular Dynamic Simulations in Drug Design. *Med. Res. Rev.* 26, 531–568. doi:10.1002/med.20067

Ashour, H. M., Elkhatib, W. F., Rahman, M. M., and Elshabrawy, H. A. (2020). Insights into the Recent 2019 Novel Coronavirus (Sars-coV-2) in Light of Past Human Coronavirus Outbreaks. *Pathogens* 9, 186–215. doi:10.3390/pathogens9030186

- Babcock, G. J., Eshaki, D. J., Thomas, W. D., and Ambrosino, D. M. (2004). Amino Acids 270 to 510 of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein Are Required for Interaction with Receptor. *J. Virol.* 78, 4552–4560. doi:10.1128/jvi.78.9.4552-4560.2004
- Bartuzi, D., Kaczor, A., Targowska-Duda, K., and Matysiuk, D. (2017). Recent Advances and Applications of Molecular Docking to G Protein-Coupled Receptors. *Molecules* 22, 340. doi:10.3390/molecules22020340
- Both, C. (1980). From Hemolymph of Immunized Pupae of *Hyalophora cecropia*. *Eur. J. Biochem.* 16, 7–16.
- Case, D. A., Babin, V., Berryman, J. T., Betz, R. M., Cai, Q., Cerutti, D. S., et al. (2014). Amber14, Version AMBER14. Available at: <http://ambermd.org/> (Accessed October, 2015).
- Chen, J. (2020). Pathogenicity and Transmissibility of 2019-nCoV-A Quick Overview and Comparison with Other Emerging Viruses. *Microbes Infect.* 22, 69–71. doi:10.1016/j.micinf.2020.01.004
- Chernysh, S., Kim, S. I., Bekker, G., Pleskach, V. A., Filatova, N. A., Anikin, V. B., et al. (2002). Antiviral and Antitumor Peptides from Insects. *Proc. Natl. Acad. Sci.* 99, 12628–12632. doi:10.1073/pnas.192301899
- Chowdhury, S. M., Talukder, S. A., Khan, A. M., Afrin, N., Ali, M. A., Islam, R., et al. (2020). Antiviral Peptides as Promising Therapeutics against SARS-CoV-2. *J. Phys. Chem. B* 124, 9785–9792. doi:10.1021/acs.jpcc.0c05621
- Comeau, S. R., Gatchell, D. W., Vajda, S., and Camacho, C. J. (2004). ClusPro: A Fully Automated Algorithm for Protein-Protein Docking. *Nucleic Acids Res.* 32, W96–W99. doi:10.1093/nar/gkh354
- Cui, J., Li, F., and Shi, Z.-L. (2019). Origin and Evolution of Pathogenic Coronaviruses. *Nat. Rev. Microbiol.* 17, 181–192. doi:10.1038/s41579-018-0118-9
- De Wit, E., Van Doremalen, N., Falzarano, D., and Munster, V. J. (2016). SARS and MERS: Recent Insights into Emerging Coronaviruses. *Nat. Rev. Microbiol.* 14, 523–534. doi:10.1038/nrmicro.2016.81
- Dehelean, C. A., Lazureanu, V., Coricovac, D., Mioc, M., Oancea, R., Marcovici, I., et al. (2020). SARS-CoV-2: Repurposed Drugs and Novel Therapeutic Approaches-Insights into Chemical Structure-Biological Activity and Toxicological Screening. *J. Clin. Med.* 9, 2084. doi:10.3390/jcm9072084
- Dejnirattisai, W., Zhou, D., Ginn, H. M., Duyvesteyn, H. M. E., Supasa, P., Case, J. B., et al. (2021). The Antigenic Anatomy of SARS-CoV-2 Receptor Binding Domain. *Cell* 184, 2183–2200. doi:10.1016/j.cell.2021.02.032
- Dimitrov, I., Flower, D. R., and Doytchinova, I. (2013). AllerTOP - a Server for In Silico Prediction of Allergens. *BMC Bioinformatics* 14, S4. doi:10.1186/1471-2105-14-S6-S4
- Dimitrov, I., Bangov, I., Flower, D. R., and Doytchinova, I. (2014a). AllerTOP v.2-a Server for In Silico Prediction of Allergens. *J. Mol. Model.* 20, 2278. doi:10.1007/s00894-014-2278-5
- Dimitrov, I., Naneva, L., Doytchinova, I., and Bangov, I. (2014b). AllergenFP: Allergenicity Prediction by Descriptor Fingerprints. *Bioinformatics* 30, 846–851. doi:10.1093/bioinformatics/btt619
- Discovery Studio (2009). *Discovery Studio, Version 2.5*. San Diego, CA, USA: Accelrys Inc.
- Dong, E., Du, H., and Gardner, L. (2020). An Interactive Web-Based Dashboard to Track COVID-19 in Real Time. *Lancet Infect. Dis.* 20, 533–534. doi:10.1016/S1473-3099(20)30120-1
- Du, Q., Qian, Y., and Xue, W. (2020). Molecular Simulation of Oncostatin M and Receptor (OSM-OSMR) Interaction as a Potential Therapeutic Target for Inflammatory Bowel Disease. *Front. Mol. Biosci.* 7, 29. doi:10.3389/fmolb.2020.00029
- Dutta, M., Tareq, A. M., Rakib, A., Mahmud, S., Sami, S. A., Mallick, J., et al. (2021). Phytochemicals from *Leucas zeylanica* Targeting Main Protease of SARS-CoV-2: Chemical Profiles, Molecular Docking, and Molecular Dynamics Simulations. *Biology* 10, 789. doi:10.3390/biology10080789
- Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995). A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* 103, 8577–8593. doi:10.1063/1.470117
- Feng, M., Fei, S., Xia, J., Labropoulou, V., Swevers, L., and Sun, J. (2020). Antimicrobial Peptides as Potential Antiviral Factors in Insect Antiviral Immune Response. *Front. Immunol.* 11, 1–12. doi:10.3389/fimmu.2020.02030
- Gohlke, H., Kiel, C., and Case, D. A. (2003). Insights into Protein-Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras-Raf and Ras-RalGDS Complexes. *J. Mol. Biol.* 330, 891. doi:10.1016/S0022-2836(03)00610-7
- Gorbalenya, A. E., Krupovic, M., Mushegian, A., Kropinski, A. M., Siddell, S. G., Varsani, A., et al. (2020). The New Scope of Virus Taxonomy: Partitioning the Virosphere into 15 Hierarchical Ranks. *Nat. Microbiol.* 5, 668–674. doi:10.1038/s41564-020-0709-x
- Gui, M., Song, W., Zhou, H., Xu, J., Chen, S., Xiang, Y., et al. (2017). Cryo-electron Microscopy Structures of the SARS-CoV Spike Glycoprotein Reveal a Prerequisite Conformational State for Receptor Binding. *Cell Res.* 27, 119–129. doi:10.1038/cr.2016.152
- Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., and Raghava, G. P. S. (2013). In Silico Approach for Predicting Toxicity of Peptides and Proteins. *PLoS One* 8, e73957. doi:10.1371/journal.pone.0073957
- Harrach, M. F., and Drossel, B. (2014). Structure and Dynamics of TIP3P, TIP4P, and TIP5P Water Near Smooth and Atomistic walls of Different Hydroaffinity. *J. Chem. Phys.* 140, 174501. doi:10.1063/1.4872239
- Harvey, M. J., and De Fabritiis, G. (2009). An Implementation of the Smooth Particle Mesh Ewald Method on GPU Hardware. *J. Chem. Theor. Comput.* 5, 2371–2377. doi:10.1021/ct900275y
- He, Y., Zhou, Y., Liu, S., Kou, Z., Li, W., Farzan, M., et al. (2004). Receptor-binding Domain of SARS-CoV Spike Protein Induces Highly Potent Neutralizing Antibodies: Implication for Developing Subunit Vaccine. *Biochem. Biophys. Res. Commun.* 324, 773–781. doi:10.1016/j.bbrc.2004.09.106
- Sakib, M. M. H., Nishat, A. A., Islam, M. T., Raihan Uddin, M. A., Iqbal, M. S., Bin Hossen, F. F., et al. (2021). Computational Screening of 645 Antiviral Peptides against the Receptor-Binding Domain of the Spike Protein in SARS-CoV-2. *Comput. Biol. Med.* 136, 104759. doi:10.1016/j.combiomed.2021.104759
- Huang, Z., Kingsolver, M. B., Avadhanula, V., and Hardy, R. W. (2013). An Antiviral Role for Antimicrobial Peptides during the Arthropod Response to Alphavirus Replication. *J. Virol.* 87, 4272–4280. doi:10.1128/jvi.03360-12
- Khan, S., Siddique, R., Shereen, M. A., Ali, A., Liu, J., Bai, Q., et al. (2020). Emergence of a Novel Coronavirus, Severe Acute Respiratory Syndrome Coronavirus 2: Biology and Therapeutic Options. *J. Clin. Microbiol.* 58, e00187. doi:10.1128/jcm.00187-20
- Kim, C., Ryu, D.-K., Lee, J., Kim, Y.-I., Seo, J.-M., Kim, Y.-G., et al. (2021). A Therapeutic Neutralizing Antibody Targeting Receptor Binding Domain of SARS-CoV-2 Spike Protein. *Nat. Commun.* 12, 1–10. doi:10.1038/s41467-020-20602-5
- Kirchdoerfer, R. N., Wang, N., Pallesen, J., Wrapp, D., Turner, H. L., Cottrell, C. A., et al. (2018). Stabilized Coronavirus Spikes Are Resistant to Conformational Changes Induced by Receptor Recognition or Proteolysis. *Sci. Rep.* 8, 1–11. doi:10.1038/s41598-018-34171-7
- Krieger, E., and Vriend, G. (2015). New Ways to Boost Molecular Dynamics Simulations. *J. Comput. Chem.* 36, 996–1007. doi:10.1002/jcc.23899
- Krieger, E., Nielsen, J. E., Spronk, C. A. E. M., and Vriend, G. (2006). Fast Empirical pKa Prediction by Ewald Summation. *J. Mol. Graph. Model.* 25, 481–486. doi:10.1016/j.jmgm.2006.02.009
- Krieger, E., Koraimann, G., and Vriend, G. (2013). YASARA—Yet Another Scientific Artificial Reality Application. YASARA.org.
- Lamiable, A., Thévenet, P., Rey, J., Vavrusa, M., Derreumaux, P., and Tufféry, P. (2016). PEP-FOLD3: Faster de Novo Structure Prediction for Linear Peptides in Solution and in Complex. *Nucleic Acids Res.* 44, W449–W454. doi:10.1093/nar/gkw329
- Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., et al. (2020). Structure of the SARS-CoV-2 Spike Receptor-Binding Domain Bound to the ACE2 Receptor. *Nature* 581, 215–220. doi:10.1038/s41586-020-2180-5
- Land, H., and Humble, M. S. (2018). YASARA: A Tool to Obtain Structural Guidance in Biocatalytic Investigations. *Methods Mol. Biol.* 1685, 43–67. doi:10.1007/978-1-4939-7366-8_4
- Li, T., Zhang, Y., Fu, L., Yu, C., Li, X., Li, Y., et al. (2005). siRNA Targeting the Leader Sequence of SARS-CoV Inhibits Virus Replication. *Gene Ther.* 12, 751–761. doi:10.1038/sj.gt.3302479
- Li, Q. Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., et al. (2020). Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N. Engl. J. Med.* 382, 1199–1207. doi:10.1056/nejmoa2001316
- Li, X., Zai, J., Zhao, Q., Nie, Q., Li, Y., Foley, B. T., et al. (2020). Evolutionary History, Potential Intermediate Animal Host, and Cross-species Analyses of SARS-CoV-2. *J. Med. Virol.* 92, 602–611. doi:10.1002/jmv.25731
- Liu, Z., Liu, Y., Zeng, G., Shao, B., Chen, M., Li, Z., et al. (2018). Application of Molecular Docking for the Degradation of Organic Pollutants in the Environmental Remediation: A Review. *Chemosphere* 203, 139–150. doi:10.1016/j.chemosphere.2018.03.179

- Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., et al. (2020). Genomic Characterisation and Epidemiology of 2019 Novel Coronavirus: Implications for Virus Origins and Receptor Binding. *Lancet* 395, 565–574. doi:10.1016/S0140-6736(20)30251-8
- Machhi, J., Herskovitz, J., Senan, A. M., Dutta, D., Nath, B., Oleynikov, M. D., et al. (2020). The Natural History, Pathobiology, and Clinical Manifestations of SARS-CoV-2 Infections. *J. Neuroimmune Pharmacol.* 15, 359–386. doi:10.1007/s11481-020-09944-5
- Moreno-Habel, D. A., Biglang-awa, I. M., Dulce, A., Luu, D. D., Garcia, P., Weers, P. M. M., et al. (2012). Inactivation of the Budded Virus of Autographa Californica M Nucleopolyhedrovirus by Gloverin. *J. Invertebr. Pathol.* 110, 92–101. doi:10.1016/j.jip.2012.02.007
- Nešuta, O., Hexnerová, R., Buděšínský, M., Slaninová, J., Bednářová, L., Hadravová, R., et al. (2016). Antimicrobial Peptide from the Wild Bee *Hylaeus Signatus* Venom and its Analogues: Structure-Activity Study and Synergistic Effect with Antibiotics. *J. Nat. Prod.* 79, 1073–1083. doi:10.1021/acs.jnatprod.5b01129
- Obaidullah, A. J., Alanazi, M. M., Alsaif, N. A., Albassam, H., Almezizia, A. A., Alqahtani, A. M., et al. (2021). Immunoinformatics-guided Design of a Multi-Epitope Vaccine Based on the Structural Proteins of Severe Acute Respiratory Syndrome Coronavirus 2. *RSC Adv.* 11, 18103–18121. doi:10.1039/d1ra02885e
- Ourth, D. D. (2004). Antiviral Activity against Human Immunodeficiency Virus-1 *In Vitro* by Myristoylated-Peptide from *Heliothis virescens*. *Biochem. Biophys. Res. Commun.* 320, 190–196. doi:10.1016/j.bbrc.2004.05.137
- Padhi, S., Sanjukt, S., Chourasia, R., Labala, R. K., Singh, S. P., and Rai, A. K. (2021). A Multifunctional Peptide from Bacillus Fermented Soybean for Effective Inhibition of SARS-CoV-2 S1 Receptor Binding Domain and Modulation of Toll like Receptor 4: A Molecular Docking Study. *Front. Mol. Biosci.* 8, 1–14. doi:10.3389/fmolb.2021.636647
- Premkumar, L., Segovia-Chumbez, B., Jadi, R., Martinez, D. R., Raut, R., Markmann, A. J., et al. (2020). The Receptor-Binding Domain of the Viral Spike Protein Is an Immunodominant and Highly Specific Target of Antibodies in SARS-CoV-2 Patients. *Sci. Immunol.* 5, 1–10. doi:10.1126/SCIIMMUNOL.ABC8413
- Rathod, S. B., Prajapati, P. B., Punjabi, L. B., Prajapati, K. N., Chauhan, N., and Mansuri, M. F. (2020). Peptide Modelling and Screening against Human ACE2 and Spike Glycoprotein RBD of SARS-CoV-2. *Silico Pharmacol.* 8, 1–9. doi:10.1007/s40203-020-00055-w
- Ren, L.-L., Wang, Y.-M., Wu, Z.-Q., Xiang, Z.-C., Guo, L., Xu, T., et al. (2020). Identification of a Novel Coronavirus Causing Severe Pneumonia in Human: a Descriptive Study. *Chin. Med. J. (Engl)* 133, 1015–1024. doi:10.1097/CM9.0000000000000722
- Romero-Durana, M., Jiménez-García, B., and Fernández-Recio, J. (2020). PyDockEneRes: Per-Residue Decomposition of Protein-Protein Docking Energy. *Bioinformatics* 36, 2284–2285. doi:10.1093/bioinformatics/btz884
- Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., et al. (2020). Structural Basis of Receptor Recognition by SARS-CoV-2. *Nature* 581, 221–224. doi:10.1038/s41586-020-2179-y
- Slaninová, J., Putnová, H., Borovičková, L., Šácha, P., Čerovský, V., Monincová, L., et al. (2011). The Antifungal Effect of Peptides from Hymenoptera Venom and Their Analogs. *Cent. Eur. J. Biol.* 6, 150–159. doi:10.2478/s11535-010-0111-4
- Slaninová, J., Mlsová, V., Kroupová, H., Alán, L., Tůmová, T., Monincová, L., et al. (2012). Toxicity Study of Antimicrobial Peptides from Wild Bee Venom and Their Analogs toward Mammalian normal and Cancer Cells. *Peptides* 33, 18–26. doi:10.1016/j.peptides.2011.11.002
- Song, W., Gui, M., Wang, X., and Xiang, Y. (2018). Cryo-EM Structure of the SARS Coronavirus Spike Glycoprotein in Complex with its Host Cell Receptor ACE2. *Plos Pathog.* 14, e1007236–19. doi:10.1371/journal.ppat.1007236
- V. K., P., Rath, S. P., and Abraham, P. (2021). Computational Designing of a Peptide that Potentially Blocks the Entry of SARS-CoV, SARS-CoV-2 and MERS-CoV. *PLoS ONE* 16, e0251913. doi:10.1371/journal.pone.0251913
- Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004). Development and Testing of a General Amber Force Field. *J. Comput. Chem.* 25, 1157–1174. doi:10.1002/jcc.20035
- Weng, G., Wang, E., Wang, Z., Liu, H., Zhu, F., Li, D., et al. (2019). HawkDock: a Web Server to Predict and Analyze the Protein-Protein Complex Based on Computational Docking and MM/GBSA. *Nucleic Acids Res.* 47, W322–W330. doi:10.1093/nar/gkz397
- Wild, C. T., Shugars, D. C., Greenwell, T. K., McDaniel, C. B., and Matthews, T. J. (1994). Peptides Corresponding to a Predictive Alpha-Helical Domain of Human Immunodeficiency Virus Type 1 Gp41 Are Potent Inhibitors of Virus Infection. *Proc. Natl. Acad. Sci.* 91, 9770–9774. doi:10.1073/pnas.91.21.9770
- Willett, P., and Glen, C. (1995). Molecular Recognition of Receptor Sites Using a Genetic Algorithm with a Description of Desolvation. *J. Mol. Biol.* 245, 43–53. doi:10.1016/s0022-2836(95)80037-9
- Wong, S. K., Li, W., Moore, M. J., Choe, H., and Farzan, M. (2004). A 193-Amino Acid Fragment of the SARS Coronavirus S Protein Efficiently Binds Angiotensin-Converting Enzyme 2. *J. Biol. Chem.* 279, 3197–3201. doi:10.1074/jbc.C300520200
- Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C.-L., Abiona, O., et al. (2020). Cryo-EM Structure of the 2019-nCoV Spike in the Prefusion Conformation. *Science* 367, 1260–1263. doi:10.1126/science.aax090210.1126/science.abb2507
- Wu, A., Peng, Y., Huang, B., Ding, X., Wang, X., Niu, P., et al. (2020). Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host Microbe* 27, 325–328. doi:10.1016/j.chom.2020.02.001
- Xiao, X., Chakraborti, S., Dimitrov, A. S., Gramatikoff, K., and Dimitrov, D. S. (2003). The SARS-CoV S Glycoprotein: Expression and Functional Characterization. *Biochem. Biophys. Res. Commun.* 312, 1159–1164. doi:10.1016/j.bbrc.2003.11.054
- Yan, Y., Zhang, D., Zhou, P., Li, B., and Huang, S.-Y. (2017). HDock: A Web Server for Protein-Protein and Protein-DNA/RNA Docking Based on a Hybrid Strategy. *Nucleic Acids Res.* 45, W365–W373. doi:10.1093/nar/gkx407
- Yang, J., Zhang, Z., Yang, F., Zhang, H., Wu, H., Zhu, F., et al. (2021). Computational Design and Modeling of Nanobodies toward SARS-CoV-2 Receptor Binding Domain. *Chem. Biol. Drug Des.* 98, 1–18. doi:10.1111/cbdd.13847
- Yin, C. (2020). Genotyping Coronavirus SARS-CoV-2: Methods and Implications. *Genomics* 112, 3588–3596. doi:10.1016/j.ygeno.2020.04.016
- Yuan, Y., Cao, D., Zhang, Y., Ma, J., Qi, J., Wang, Q., et al. (2017). Cryo-EM Structures of MERS-CoV and SARS-CoV Spike Glycoproteins Reveal the Dynamic Receptor Binding Domains. *Nat. Commun.* 8, 1–9. doi:10.1038/ncomms15092
- Yuan, M., Liu, H., Wu, N. C., and Wilson, I. A. (2021). Recognition of the SARS-CoV-2 Receptor Binding Domain by Neutralizing Antibodies. *Biochem. Biophys. Res. Commun.* 538, 192–203. doi:10.1016/j.bbrc.2020.10.012
- Zhang, Y.-Z., and Holmes, E. C. (2020). A Genomic Perspective on the Origin and Emergence of SARS-CoV-2. *Cell* 181, 223–227. doi:10.1016/j.cell.2020.03.035
- Zhou, P., Jin, B., Li, H., and Huang, S.-Y. (2018). HPEPDOCK: A Web Server for Blind Peptide-Protein Docking Based on a Hierarchical Algorithm. *Nucleic Acids Res.* 46, W443–W450. doi:10.1093/nar/gky357
- Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., et al. (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* 382, 727–733. doi:10.1056/nejmoa2001017

Conflict of Interest: The authors declare 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.

Copyright © 2022 Biswas, Mahmud, Mita, Afrose, Hasan, Sultana Shimu, Saleh, Mostafa-Hedeab, Alqarni, Obaidullah and Batiha. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.



Phytochemicals of *Euphorbia hirta* L. and Their Inhibitory Potential Against SARS-CoV-2 Main Protease

Ruel Cayona* and Evelyn Creencia*

Department of Chemistry, College of Science and Mathematics, Mindanao State University—Iligan Institute of Technology, Iligan, Philippines

OPEN ACCESS

Edited by:

Arvind Ramanathan,
Argonne National Laboratory (DOE),
United States

Reviewed by:

Debsindhu Bhowmik,
Oak Ridge National Laboratory (DOE),
United States

Lorane Hage-Melim,
Universidade Federal do Amapá,
Brazil

*Correspondence:

Ruel Cayona
ruel.cayona@g.msuiit.edu.ph
Evelyn Creencia
evelyn.creencia@g.msuiit.edu.ph

Specialty section:

This article was submitted to
Biological Modeling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 25 October 2021

Accepted: 30 December 2021

Published: 04 February 2022

Citation:

Cayona R and Creencia E (2022)
Phytochemicals of *Euphorbia hirta* L.
and Their Inhibitory Potential Against
SARS-CoV-2 Main Protease.
Front. Mol. Biosci. 8:801401.
doi: 10.3389/fmolb.2021.801401

Euphorbia hirta L. is a medicinal plant widely used in the Philippines and across tropical Asia against various diseases, including respiratory disorders. In this study, the phytochemical components of *E. hirta* were investigated *in silico* for their potential to inhibit the severe acute respiratory syndrome-coronavirus-2 main protease (SARS-CoV-2 Mpro), a coronavirus disease 2019 (COVID-19) drug target that plays a critical role in the infection process of SARS-CoV-2. Phytochemical mining in tandem with virtual screening (PM-VS) was the strategy implemented in this study, which allows efficient preliminary *in silico* assessment of the COVID-19 therapeutic potential of the reported phytochemicals from the plant. The main rationale for considering *E. hirta* in the investigation was its reported efficacy against respiratory disorders. It is very promising to investigate the phytochemicals of *E. hirta* for their potential efficacy against diseases, such as COVID-19, that also target the respiratory system. A total of 298 *E. hirta* phytochemicals were comprehensively collected from the scientific literature. One hundred seventy of these phytochemicals were computed through molecular docking and were shown to have comparable or better binding properties (promising inhibitors) toward SARS-CoV-2 Mpro than known *in vitro* inhibitors. In connection to our previous work considering different medicinal plants, antiviral compounds were also rediscovered from the phytochemical composition of *E. hirta*. This finding provides additional basis for the potential of the plant (or its phytochemicals) as a COVID-19 therapeutic directly targeting drug targets such as SARS-CoV-2 Mpro and/or addressing respiratory-system-related symptoms. The study also highlights the utility of PM-VS, which can be efficiently implemented in the preliminary steps of drug discovery and development.

Keywords: *Euphorbia hirta*, COVID-19, molecular docking, phytochemical mining, medicinal plant, Philippine medicinal plant, SARS-CoV-2 Mpro, virtual screening

1 INTRODUCTION

Euphorbia hirta L. (Euphorbiaceae) is a medicinal plant widely used in the Philippines and across tropical Asia, and it is commonly known by the following names: “asthma plant” (English), “tawa-tawa” (Filipino), and “mangagaw” (Cebuano). The extract of *E. hirta* is taken orally as an aqueous decoction for most of its folkloric uses. As its English common name suggests, the plant has been used for asthma and other respiratory difficulties (Ekpo and Pretorius, 2007; Ogunlesi et al., 2009; Rao et al., 2017). In addition, available studies conclusively suggest its potential against dengue

(Guzman et al., 2016; Perera et al., 2018; Suganthi and Ravi, 2018); however, additional studies are required to validate the results (Perera et al., 2018). Nevertheless, the studies reveal *E. hirta* as a pool for compounds with interesting biological activities.

E. hirta is one of the medicinal plants currently being investigated in the Philippines for its potential against coronavirus (CoV) disease 2019 (COVID-19) (Luci-Atienza, 2021a, Luci-Atienza, 2021b; Tawa-Tawa Clinical Trial on COVID-19, 2021). The goal is to develop a formulation utilizing the plant as an adjuvant treatment for mild to moderate COVID-19. A recently published review article identified *E. hirta* as one of the Philippine medicinal plants with immunomodulatory effects and potential against severe acute respiratory syndrome-CoV-2 (SARS-CoV-2) (Dayrit et al., 2021), the virus responsible for COVID-19. In this connection, a parallel and complementary *in silico* study was conducted to investigate the potential of its phytochemicals against a specific COVID-19 drug target, SARS-CoV-2 main protease (Mpro). Mpro is seen as an important COVID-19 drug target because of the role it plays in the regulation of viral replication (Di Micco et al., 2021).

It was the reported activities of *E. hirta* or its phytochemicals against respiratory-related ailments that serve as the primary basis for considering it as a subject of the present investigation. This study was conducted in line with the ongoing effort to discover potential COVID-19 therapeutic chemicals from medicinal plants, starting first with those found in the Philippines (Philippine medicinal plants). A strategy called phytochemical mining in tandem with virtual screening (PM-VS) was implemented. PM-VS refers to the systematic and comprehensive collection of medicinal plant phytochemicals reported in the scientific literature (phytochemical mining) and subsequent *in silico* assessment of the potential efficacy of the phytochemicals against specific or multiple drug target(s) (virtual screening). PM-VS and its rationale have been elaborated elsewhere (Cayona and Creencia, 2021a; Cayona and Creencia, 2021b, Cayona and Creencia, 2022). Specifically focused in this study is *E. hirta* and automated targeted molecular docking as the medicinal plant and virtual screening tool, respectively. It is argued that PM-VS can be efficiently implemented in the preliminary steps of drug discovery and development.

2 MATERIALS AND METHODS

2.1 Phytochemical Data Collection

The method implemented in this study is adapted from the method described in our previous papers (Cayona and Creencia, 2021a; Cayona and Creencia, 2021b) with slight modification. The Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) (Moher et al., 2009) protocol was implemented throughout the systematic data collection process. The sources of phytochemicals were peer-reviewed research and review articles from scientific journals deposited in the MEDLINE database by the US National Institutes of Health National Library of Medicine (<https://pubmed.ncbi.nlm.nih.gov/>).

The Google Scholar (<https://scholar.google.com/>) search engine was utilized to find additional literature but other search engines were also consulted (i.e., Microsoft Academic and Semantic Scholar), similarly applying relevant search keys and filters when applicable.

The identified sources were then compared against each other to check for multiple entries and reference-checked to retrieve additional sources unintentionally omitted in the first part of literature gathering. Articles deposited in restricted repositories and which were not written in English were not included. Thereafter, the phytochemicals reported in every literature reference were trimmed down to unique chemical identities only (because one compound may have multiple reported names). For simplicity, the common names of the compounds were taken in cases where ambiguity does not manifest; otherwise, the International Union of Pure and Applied Chemistry (IUPAC) nomenclature was adopted. The study strictly adhered to the data collection protocol described in PRISMA (see **Supplementary Materials**).

2.2 Phytochemical Classification

The collected phytochemicals from *E. hirta* were classified according to the ClassyFire (Djoumbou Feunang et al., 2016) algorithm of chemical classification. This was done to gain insight that might be helpful in assessing the basic structure–activity relationship. The hierarchy of chemical taxonomic classification can be found in the **Supplementary Materials**.

2.3 Preparation of Ligands

Three-dimensional (3D) structure-data files (SDFs) of phytochemicals included in the final list were either conveniently collected from PubChem or manually generated whenever they are unavailable in the database. Hydrogen atoms were explicitly added to the structures. In some cases, two-dimensional (2D) SDFs were used but only for 2D compounds (linear or flat). In preparation for virtual screening and for future convenience, the SDFs of all the structures of the phytochemicals (the ligands) were combined into a single SDF using OpenBabel 2.4.1 (O'Boyle et al., 2011) to facilitate automated importing of the multiple structures into the virtual screening tool. The same preparation was done for the control compounds.

2.4 Receptor Preparation

The crystal structure at 2.16 Å of the SARS-CoV-2 Mpro (PDB ID: 6LU7) in complex with the *in vitro* inhibitor N3 (Jin et al., 2020) was downloaded from the Protein Data Bank (<http://www.rcsb.org/>) in a PDB file format. The noninteracting atoms (e.g., water and buffer molecules) were removed, and hydrogen atoms were explicitly added to the enzyme and the native ligand.

The active site was taken as the region of the SARS-CoV-2 Mpro volume where the *in vitro* inhibitor N3 was attached. From the SARS-CoV-2 Mpro–N3 complex, the search space for the targeted molecular docking was then assigned with the help of BIOVIA Discovery Studio Visualizer v20.1.0.19295 (DSV, 2020). The interacting and the pocket amino acids (AAs) that lie within the 3.5 Å distance from the closest N3 atom were identified by

visual inspection. The residues found within this region totaled 25 AAs. The interacting AAs were H41, M49, F140, N142, G143, H164, M165, E166, L167, P168, H172, Q189, T190, and T191; and the pocket AAs were T24, T25, T26, L27, Y54, L141, S144, C145, H163, D187, R188, and Q192. From this list of AAs, the H41–C145 catalytic dyad can be found (Wang YC. et al., 2020; Hakmi et al., 2020; Ullrich and Nitsche, 2020).

2.5 Virtual Screening Through Automated Molecular Docking

2.5.1 Molecular Docking Tools

The phytochemical ligands were virtually screened against SARS-CoV-2 Mpro (6LU7-neat) using PyRx0.8 (Dallakyan and Olson, 2015), a virtual screening tool that allows automated molecular docking of multiple ligands (or libraries) against target receptor (s). PyRx0.8 utilizes the enabling capabilities of AutoDock tools for receptor and ligand preparation just as in AutoDock 4 (Morris et al., 2009) and the earlier versions; AutoDock Vina for molecular docking (Trott and Olson, 2010); OpenBabel for file format interconversion (O'Boyle et al., 2011); and other open-source software. To save on computational cost, targeted molecular docking on the active site of Mpro was conducted.

2.5.2 Control Parameters

To enhance the accuracy, control parameters were set in molecular docking against Mpro. In addition to the phytochemical ligands, control ligand samples were also tested. Ten known inhibitors with established *in vitro* half-maximal effective concentration (EC₅₀) against SARS-CoV-2 or half-maximal inhibitory concentration (IC₅₀) against SARS-CoV-2 Mpro were used as positive controls. On the other hand, 10 small molecules that do not possess interesting pharmacological properties were also used as negative controls. The positive controls were N3 (6LU7 native ligand), efonidipine, bedaquiline, tideglusib, manidipine, N3, lercanidipine, boceprevir, shikonin, ebselen, and carmofur, whose inhibitory properties were reported elsewhere (Ghahremanpour et al., 2020; Jin et al., 2020; Ma et al., 2020). The negative controls chosen were anthracene, naphthalene, glycerol, decane, hexanol, benzene, cyclohexane, hexane, ethanol, and water. The positive controls (inhibitors) were expected to give satisfactory binding free energy (BFE) values towards the receptor SARS-CoV-2 Mpro because they are empirically established inhibitors. On this basis, their BFEs were taken as a reference in assigning promising phytochemicals against SARS-CoV-2 Mpro. In contrast, the negative controls should have unsatisfactory computed BFE towards the receptor. The control ligands provide a simple means to assess the reliability and performance of the virtual screening tool.

2.5.3 Automated Molecular Docking

The receptor (6LU7-neat) was loaded onto PyRx0.8 and set into the macromolecule (receptor) in the PDBQT format. The collective SDFs of the phytochemicals and positive and negative controls previously prepared using OpenBabel 2.4.1 were also loaded onto PyRx0.8 and subsequently extracted

automatically to individual structures. The structures were then energy minimized by implementing suitable force fields. For most of the structures, MMFF94 was sufficient in energy minimization; however, UFF and/or Ghemical must be implemented for some ligands whose final structures were distorted under specified UFF minimization parameters. Thereafter, the ligands were converted into a docking-ready PDBQT file format.

Before docking, the search space for the targeted automated molecular docking was set. The interacting and the pocket AA residues of SARS-CoV-2 Mpro that were identified previously were selected, and the search space was adjusted manually in the PyRx0.8 interface so that all of the residues were included in the grid volume of the search space. The resulting grid dimensions are the following: center_x = -10.8864; center_y = 14.0407; center_z = 68.7458; size_x = 21.4856; size_y = 26.7715; size_z = 28.0882. The exhaustiveness of the most stable conformation search was set at 16. Finally, docking was commenced using the Vina (AutoDock Vina) tab in PyRx0.8.

2.5.4 Receptor–Ligand Interaction Analysis

Interactions of the ligands which have BFEs comparable to or better than those of the positive controls were analyzed. Those ligands whose most stable binding conformation (docking RMSD = 0) established interactions with the H41–C145 catalytic dyad (Jin et al., 2020; Khan et al., 2020; Menéndez et al., 2020; Mirza and Froeyen, 2020; Shitrit et al., 2020) of the SARS-CoV-2 Mpro and those with reported antiviral properties were given emphasis. Favorable computed BFE and catalytic dyad interaction(s) were considered as major criteria in identifying promising SARS-CoV-2 Mpro inhibitors.

2.5.5 Assessment of the Reliability of the Tools and Strategies

All the tools and strategies used in the study are well established throughout the scientific literature. The number of citations of the articles that report the tools and strategies partly establish their reputation in the field. For example, Google Scholar queries on PRISMA, ClassyFire, PyRx0.8, AutoDock Vina, and AutoDock 4 will reveal 7,156; 283; 873; 15,275; and 12,351 citations, respectively, as of July 17, 2021.

3 RESULTS

3.1 Phytochemical Mining and Classification

Literature reports indicate that leaves, aerial parts, and whole plants are the sources of *E. hirta* phytochemicals. The relevant data collected from phytochemical mining (PM) *E. hirta* are presented in Table 1. Each phytochemical is provided with its molecular formula (MF), BFE value against SARS-CoV-2 Mpro, and chemical taxonomy grouping levels (ClassyFire Superclass, Class, and Subclass). The chemical structures of all *E. hirta* phytochemicals and the control samples (positive and negative) used in molecular docking can be found in the **Supplementary Materials**. In total, 298 phytochemical

TABLE 1 | Phytochemicals from *E. hirta*.

ID	Phytochemical	MF ^a	BFE ^b
<i>Benzenoids/benzene and substituted derivatives</i>			
1	1-(3-aminophenyl)ethanol (Rautela et al.,2020)	C ₈ H ₁₁ NO	-4.6
2	1-O-butyl 2-O-tetradecyl benzene-1,2-dicarboxylate (Ogunlesi et al.,2009)	C ₂₆ H ₄₂ O ₄	-4.9
3	benzoic acid (Ali et al.,2020)	C ₇ H ₆ O ₂	-4.5
4	benzamide, 3-fluoro-N-butyl-N-ethyl (Rautela et al.,2020)	C ₁₃ H ₁₈ FN	-4.9
5	gallic acid (Bach et al.,2020; Linfang et al., 2012; Mahomoodally et al.,2020; Mekam et al.,2019, Suganthi and Ravi, 2018; Wu et al.,2012)	C ₇ H ₆ O ₅	-5.5
6	ethyl gallate (Mekam et al.,2019)	C ₉ H ₁₀ O ₅	-5.7
7	methyl gallate (Mahomoodally et al.,2020)	C ₈ H ₈ O ₅	-5.6
8	protocatechuic acid (Mahomoodally et al.,2020)	C ₇ H ₆ O ₄	-5.4
9	1-(3-ethoxyphenyl)propan-2-one (Rautela et al.,2020)	C ₁₁ H ₁₄ O ₂	-5.0
10	methyl 3-(3,5-di- <i>tert</i> -butyl-4-hydroxyphenyl)propanoate (Perumal and Mahmud, 2013)	C ₁₈ H ₂₈ O ₃	-6.5
<i>Benzenoid/naphthalene</i>			
11	[6-(4-cyanophenyl)naphthalen-2-yl] hexanoate (Rautela et al.,2020)	C ₂₃ H ₂₁ NO ₂	-6.5
<i>Benzenoids/phenols</i>			
12	benzene-1,2,3-triol (Karki et al.,2019)	C ₆ H ₆ O ₃	-4.9
13	2- <i>tert</i> -butyl-4-methoxyphenol (Rautela et al.,2020)	C ₁₁ H ₁₆ O ₂	-5.1
14	4-ethenyl-2-methoxyphenol (Rautela et al.,2020)	C ₉ H ₁₀ O ₂	-4.7
<i>Benzenoid</i>			
15	1,2-benzenedicarboxylic acid diisooctyl ester (Ogunlesi et al.,2009)	C ₂₄ H ₃₈ O ₄	-5.4
<i>Hydrocarbon/saturated hydrocarbon</i>			
16	Tetradecane (Ogunlesi et al.,2009)	C ₁₄ H ₃₀	-4.2
<i>Hydrocarbon/unsaturated hydrocarbon</i>			
17	(<i>E</i>)-pentatriacont-17-ene (Rautela et al.,2020)	C ₃₅ H ₇₀	-4.4
<i>Lignans, neolignans, and related compounds/aryltetralin lignans</i>			
18	Isolintetralin (Zhang et al.,2020)	C ₂₃ H ₂₈ O ₆	-7.1
19	Lintetralin (Zhang et al.,2020)	C ₂₃ H ₂₈ O ₆	-7.3
20	Phyltetralin (Zhang et al.,2020)	C ₂₄ H ₃₂ O ₆	-7.0
21	Hypophyllanthin (Zhang et al.,2020)	C ₂₄ H ₃₀ O ₇	-6.9
<i>Lignans, neolignans, and related compounds/dibenzylbutane lignans</i>			
22	Niranthin (Zhang et al.,2020)	C ₂₄ H ₃₂ O ₇	-6.3
23	5-demethoxyniranthin (Zhang et al.,2020)	C ₂₃ H ₃₀ O ₆	-6.3
24	Phyllanthin (Zhang et al.,2020)	C ₂₄ H ₃₄ O ₆	-5.9
<i>Lignans, neolignans, and related compounds/furanoid lignans</i>			
25	Virgatusin (Zhang et al.,2020)	C ₂₃ H ₂₈ O ₇	-6.4
26	Urinaligran (Zhang et al.,2020)	C ₂₂ H ₂₄ O ₇	-7.4
27	7-hydroxyhinokinin (Zhang et al.,2020)	C ₂₀ H ₁₈ O ₈	-8.2
28	(-)-pinoresinol (Li et al.,2015)	C ₂₀ H ₂₂ O ₆	-7.2
29	(+)-syringaresinol (Li et al.,2015)	C ₂₂ H ₂₆ O ₈	-7.6
<i>Lignans, neolignans, and related compounds/lignan glycosides</i>			
30	(+)-syringaresinol glucoside (Li et al.,2015)	C ₂₈ H ₃₆ O ₁₃	-7.0
31	(-)-pinoresinol glucoside (Li et al.,2015)	C ₂₆ H ₃₂ O ₁₁	-7.6
<i>Lignans</i>			
32	5-methoxyvirgatusin (Zhang et al.,2020)	C ₂₄ H ₃₀ O ₈	-7.2
33	7R-ethoxy-3-methoxyisolintetralin (Zhang et al.,2020)	C ₂₆ H ₃₄ O ₈	-6.7
34	7R-ethoxyisolintetralin (Zhang et al.,2020)	C ₂₅ H ₃₂ O ₇	-6.8
35	7S-ethoxyisolintetralin (Zhang et al.,2020)	C ₂₅ H ₃₂ O ₇	-7.6
36	chebulic acid triethyl ester (Yang et al.,2020)	C ₂₀ H ₂₄ O ₁₁	-6.2
37	euphorhirtin A (Yang et al.,2020; Zhang et al.,2020)	C ₁₉ H ₂₀ O ₁₁	-6.5
38	euphorhirtin B (Yang et al.,2020; Zhang et al.,2020)	C ₁₉ H ₂₀ O ₁₁	-6.6
39	euphorhirtin C (Yang et al.,2020; Zhang et al.,2020)	C ₁₈ H ₁₈ O ₁₁	-6.6
40	euphorhirtin D (Yang et al.,2020; Zhang et al.,2020)	C ₁₈ H ₁₈ O ₁₁	-6.8
41	hirtacoumaroflavonoside (Sheliya et al.,2015)	C ₄₁ H ₄₄ O ₁₇	-8.7
42	hirtacoumaroflavonoside B (Sheliya et al.,2015)	C ₃₁ H ₃₆ O ₁₂	-8.4
43	Neonirtetralin (Zhang et al.,2020)	C ₂₀ H ₂₂ O ₇	-6.7
44	3,5-O-dicafeoylquinic acid (Mekam et al.,2019)	C ₂₅ H ₂₄ O ₁₂	-9.2
<i>Lipids and lipid-like molecules/fatty acyls</i>			
45	2-(dimethylamino)ethyl 3-cyclopentylpropanoate (Rautela et al.,2020)	C ₁₂ H ₂₃ NO ₂	-4.8
46	3-octadecoxypentyl (Z)-octadec-9-enoate (Karki et al.,2019)	C ₃₉ H ₇₆ O ₃	-4.2
47	ethyl hexadecanoate (Sharma et al.,2014)	C ₁₈ H ₃₆ O ₂	-4.4
48	ethyl octadecanoate (Sharma et al.,2014)	C ₂₀ H ₄₀ O ₂	-4.4
49	methyl (11E,14E,17E)-icosa-11,14,17-trienoate (Karki et al.,2019)	C ₂₁ H ₃₆ O ₂	-4.9
50	methyl 9-octadecanoate (Olaoluwa et al.,2018)	C ₁₉ H ₃₆ O ₂	-4.3
51	methyl hexadecanoate (Perumal and Mahmud, 2013; Olaoluwa et al.,2018; Karki et al.,2019; Rautela et al.,2020)	C ₁₇ H ₃₄ O ₂	-4.3

(Continued on following page)

TABLE 1 | (Continued) Phytochemicals from *E. hirta*.

ID	Phytochemical	MF ^a	BFE ^b
52	citronellyl palmitoleate (Rautela et al.,2020)	C ₂₆ H ₄₈ O ₂	-5.0
53	geranyl linoleate (Rautela et al.,2020)	C ₂₈ H ₄₈ O ₂	-5.2
54	(Z)-3,7-dimethylocta-2,6-dien-1-yl palmitate (Rautela et al.,2020)	C ₂₆ H ₄₈ O ₂	-5.4
55	oleic acid (Ogunlesi et al.,2009)	C ₁₈ H ₃₄ O ₂	-5.0
56	pentadecanoic acid (Sharma et al.,2014)	C ₁₅ H ₃₀ O ₂	-4.8
57	tetradecanoic acid (Sharma et al.,2014)	C ₁₄ H ₂₈ O ₂	-4.5
58	hexadecanoic acid (Ogunlesi et al.,2009; Perumal and Mahmud, 2013; Rautela et al.,2020)	C ₁₆ H ₃₂ O ₂	-4.4
59	methyl 3-hydroxyoctanoate O-beta-D-glucopyranoside (Nomoto et al.,2013)	C ₁₅ H ₂₈ O ₈	
60	N-butyl-1-O-alpha-L-rhamnopyranoside (Mallavadhani and Narasimhan, 2009)	C ₁₀ H ₂₀ O ₅	-5.2
61	N-butyl-1-O-beta-L-rhamnopyranoside (Mallavadhani and Narasimhan, 2009)	C ₁₀ H ₂₀ O ₅	-5.4
62	sodium beta-D-glucopyranosyl 12-hydroxyjasmonate (*acid form was used in docking) (Bach et al.,2020)	C ₁₈ H ₂₈ O ₉	-7.0
63	bumaldoside A (Nomoto et al.,2013)	C ₁₉ H ₃₆ O ₁₀	-7.2
64	byzantionoside B (Nomoto et al.,2013)	C ₁₉ H ₃₂ O ₇	-7.1
65	corchoionoside C (Nomoto et al.,2013)	C ₁₉ H ₃₀ O ₈	-7.2
66	Roseoside (Mekam et al.,2019)	C ₁₉ H ₃₀ O ₈	-7.0
67	(Z)-3-hexenyl-beta-D-glucopyranoside (Nomoto et al.,2013)	C ₁₂ H ₂₂ O ₆	-6.3
68	geranyl acetate (Rautela et al.,2020)	C ₁₂ H ₂₀ O ₂	-5.0
69	neryl acetate (Rautela et al.,2020)	C ₁₂ H ₂₀ O ₂	-4.9
70	(9E,12E,15E)-octadeca-9,12,15-trien-1-ol (Sharma et al.,2014)	C ₁₈ H ₃₂ O	-4.7
71	heptadec-13-yn-1-ol (Ogunlesi et al.,2009)	C ₁₇ H ₃₂ O	-4.4
72	(Z)-octadec-13-enal (Karki et al.,2019)	C ₁₈ H ₃₄ O	-4.1
73	(Z)-tetradec-9-enal (Karki et al.,2019)	C ₁₄ H ₂₆ O	-4.5
74	hexadecanal (Ogunlesi et al.,2009)	C ₁₆ H ₃₂ O	-4.2
75	(Z)-octadec-9-enamide (Olaoluwa et al.,2018)	C ₁₈ H ₃₅ NO	-4.2
76	tetradecanamide (Olaoluwa et al.,2018)	C ₁₄ H ₂₉ NO	-4.5
77	(1',R,5'R)-5-(5'-carboxymethyl-2'-oxocyclopentyl)-3-Z-pentenyl acetate (Chi et al.,2012)	C ₁₄ H ₂₀ O ₅	
78	methyl linolenate (Perumal and Mahmud, 2013; Rautela et al.,2020)	C ₁₉ H ₃₂ O ₂	-5.2
79	methyl linoleate (Rautela et al.,2020; Sharma et al.,2014)	C ₁₉ H ₃₄ O ₂	-4.4
80	glyceryl monolinoleate (Rautela et al.,2020)	C ₂₁ H ₃₈ O ₄	-5.1
81	ethyl linoleate (Rautela et al.,2020; Sharma et al.,2014)	C ₂₀ H ₃₆ O ₂	-4.4
82	linolenic acid (Rautela et al.,2020)	C ₁₈ H ₃₀ O ₂	-4.9
83	linoleic acid (Perumal and Mahmud, 2013)	C ₁₈ H ₃₂ O ₂	-4.6
<i>Lipids and lipid-like molecules/glycerolipids</i>			
84	2,3-dihydroxypropyl octadecanoate (Rautela et al.,2020)	C ₂₁ H ₄₂ O ₄	-4.5
85	2-monopalmitin (Perumal and Mahmud, 2013; Rautela et al.,2020)	C ₁₉ H ₃₈ O ₄	-4.7
86	2-monostearin (Perumal and Mahmud, 2013)	C ₂₁ H ₄₂ O ₄	-4.7
87	triolein (Karki et al.,2019)	C ₅₇ H ₁₀₄ O ₆	-4.5
<i>Lipids and lipid-like molecules/prenol lipids/diterpenoids</i>			
88	(E)-3,7,11,15-tetramethylhexadec-2-en-1-ol (Ogunlesi et al.,2009)	C ₂₀ H ₄₀ O	-5.2
89	phytol (Ogunlesi et al.,2009; Perumal and Mahmud, 2013; Sharma et al.,2014; Olaoluwa et al.,2018; Karki et al.,2019)	C ₂₀ H ₄₀ O	-5.1
90	gibberellin (Mekam et al.,2019)	C ₂₀ H ₂₈ O ₆	-6.2
91	ponicidin (Mekam et al.,2019)	C ₂₀ H ₂₆ O ₆	-7.5
92	albopilosin H (Mekam et al.,2019)	C ₂₀ H ₂₈ O ₄	-6.5
93	kaur-16-ene (Olaoluwa et al.,2018)	C ₂₀ H ₃₂	-6.6
<i>Lipids and lipid-like molecules/prenol lipids/monoterpenoids</i>			
94	(E)-3,7-dimethyl-2,6-octadienoic acid (Rautela et al.,2020)	C ₁₀ H ₁₆ O ₂	-4.7
95	citronellol (Rautela et al.,2020)	C ₁₀ H ₂₀ O	-4.5
96	camphol (Shah et al.,2019)	C ₁₀ H ₁₈ O	-4.3
97	cis-alpha-bergamotene (Rautela et al.,2020)	C ₁₅ H ₂₄	-5.0
98	2,6,6-trimethylbicyclo[3.1.1]heptane-2,3-diol (Rautela et al.,2020)	C ₁₀ H ₁₈ O ₂	-4.9
99	para-menth-3-en-9-ol (Olaoluwa et al.,2018)	C ₁₀ H ₁₈ O	-4.9
100	tricyclo[4.2.2.0.1,5]decan-3-ol (Rautela et al.,2020)	C ₁₀ H ₁₆ O	-4.8
<i>Lipids and lipid-like molecules/prenol lipids/quinone and hydroquinone lipids</i>			
101	gamma-tocopherol (Rautela et al.,2020; Sharma et al.,2014)	C ₂₈ H ₄₈ O ₂	-6.2
102	vitamin E (Perumal and Mahmud, 2013; Rautela et al.,2020)	C ₂₉ H ₅₀ O ₂	-6.7
<i>Lipids and lipid-like molecules/prenol lipids/sesquiterpenoids</i>			
103	isospathulenol (Rautela et al.,2020)	C ₁₅ H ₂₄ O	-5.9
104	beta-elemene (Olaoluwa et al.,2018)	C ₁₅ H ₂₄	-5.0
105	neointermedeol (Rautela et al.,2020)	C ₁₅ H ₂₆ O	-5.5
106	germacren D-4-ol (Rautela et al.,2020)	C ₁₅ H ₂₆ O	-5.6
107	beta-bisabolene (Rautela et al.,2020)	C ₁₅ H ₂₄	-5.7
108	cis-nerolidol (Rautela et al.,2020)	C ₁₅ H ₂₆ O	-5.3
109	alpha-humulene (Rautela et al.,2020)	C ₁₅ H ₂₄	-4.9
110	alpha-farnesene (Rautela et al.,2020)	C ₁₅ H ₂₄	-5.2

(Continued on following page)

TABLE 1 | (Continued) Phytochemicals from *E. hirta*.

ID	Phytochemical	MF ^a	BFE ^b
111	beta-caryophyllene (Olaoluwa et al.,2018); Rautela et al.,2020)	C ₁₅ H ₂₄	-5.1
112	farnesol 1 (Rautela et al.,2020)	C ₁₅ H ₂₆ O	-5.2
113	2,6,10-trimethyltetradecane (Ogunlesi et al.,2009)	C ₁₇ H ₃₆	-4.3
114	neophytadiene (Perumal and Mahmud, 2013; Rautela et al.,2020)	C ₂₀ H ₃₈	-4.6
115	6,10,14-trimethylpentadecan-2-one (Ogunlesi et al.,2009; Perumal and Mahmud, 2013)	C ₁₈ H ₃₆ O	-5.0
116	taraxerol acetate (Li et al.,2015)	C ₃₂ H ₅₂ O ₂	-7.5
117	taraxerone (Ragasa and Cornelio, 2013; Li et al.,2015; Tayone et al.,2020)	C ₃₀ H ₄₈ O	-8.0
118	taraxerol (Perumal and Mahmud, 2013; Prachi and Pradeep, 2014; Li et al.,2015; Amos Samkumar et al.,2019; Bach et al.,2020; Tayone et al.,2020)	C ₃₀ H ₅₀ O	-7.8
<i>Lipids and lipid-like molecule/terpene glycoside</i>			
119	citroside A (Nomoto et al.,2013)	C ₁₉ H ₃₀ O ₈	-6.8
<i>Lipids and lipid-like molecule/triterpenoids</i>			
120	friedelane-3beta,29-diol (Li et al.,2015)	C ₃₀ H ₅₂ O ₂	-7.6
121	psi-taraxastane-3,20-diol (Li et al.,2015)	C ₃₀ H ₅₂ O ₂	-7.4
122	squalene (Perumal and Mahmud, 2013; Sharma et al.,2014)	C ₃₀ H ₅₀	-5.4
123	lanost-8-en-3beta-ol (Rautela et al.,2020)	C ₃₀ H ₅₂ O	-6.1
124	lupeol (Ragasa and Cornelio, 2013; Tayone et al.,2020)	C ₃₀ H ₅₀ O	-7.3
125	friedelan-3beta-ol (Li et al.,2015)	C ₃₀ H ₅₂ O	-7.9
126	friedelin (Linfang et al., 2012; Li et al.,2015)	C ₃₀ H ₅₀ O	-8.2
127	alpha-amyirin (Linfang et al., 2012; Perumal and Mahmud, 2013; Ragasa and Cornelio, 2013)	C ₃₀ H ₅₀ O	-7.9
128	beta-amyirin (Martínez-Vázquez et al.,1999; Perumal and Mahmud, 2013; Ragasa and Cornelio, 2013)	C ₃₀ H ₅₀ O	-7.2
<i>Lipids and lipid-like molecule/steroids and steroid derivatives/cycloartanols and derivatives</i>			
129	(23E)-cycloart-23-en-3beta,25-diol (Tayone et al.,2020)	C ₃₀ H ₅₀ O ₂	-7.0
130	cycloart-23-ene-3beta,25,28-triol (Li et al.,2015)	C ₃₀ H ₅₀ O ₃	-6.8
131	cyclolanostan-3beta-ol (Li et al.,2015)	C ₃₀ H ₅₂ O	-6.7
132	24-hydroperoxycycloart-25-en-3beta-ol (Ragasa and Cornelio, 2013; Tayone et al.,2020)	C ₃₀ H ₅₀ O ₃	-7.3
133	25-hydroperoxycycloart-23-en-3beta-ol (Ragasa and Cornelio, 2013; Tayone et al.,2020)	C ₃₀ H ₅₀ O ₃	-8.0
134	cycloart-23-ene-3beta,25-diol (Li et al.,2015)	C ₃₀ H ₅₀ O ₂	-7.1
135	cycloartenol (Perumal and Mahmud, 2013; Ragasa and Cornelio, 2013; Li et al.,2015)	C ₃₀ H ₅₀ O	-6.9
<i>Lipids and lipid-like molecule/steroids and steroid derivative/ergostane steroids</i>			
136	campesterol (Perumal and Mahmud, 2013; Bach et al.,2020; Rautela et al.,2020)	C ₂₈ H ₄₈ O	-6.9
<i>Lipids and lipid-like molecule/steroids and steroid derivative/stigmastanes and derivatives</i>			
137	stigmasterol (Rautela et al.,2020)	C ₂₉ H ₄₈ O	-7.1
138	gamma-sitosterol (Perumal and Mahmud, 2013; Rautela et al.,2020)	C ₂₉ H ₅₀ O	-6.8
139	beta-sitosterol (Martínez-Vázquez et al.,1999; Mallavadhani and Narasimhan, 2009; Tayone et al.,2020)	C ₂₉ H ₅₀ O	-6.8
140	16alpha,17-dihydroxy-ent-kaurane-3-one (Li et al.,2015)	C ₂₀ H ₃₂ O ₃	-7.9
141	16alpha,17,19-trihydroxy-ent-kaurane (Li et al.,2015)	C ₂₀ H ₃₄ O ₃	-6.5
142	16alpha,19-dihydroxy-ent-kaurane (Yan et al.,2011)	C ₂₀ H ₃₄ O ₂	-6.1
143	16beta,17-dihydroxy-ent-kaurane-3-one (Li et al.,2015)	C ₂₀ H ₃₂ O ₃	-7.0
144	23(E)-25-methoxycycloart-23-en-3beta-ol (Li et al.,2015)	C ₃₁ H ₅₂ O ₂	-7.7
145	24-methylcycloartenol (Martínez-Vázquez et al.,1999)	C ₂₉ H ₅₀ O	-7.1
146	28-hydroxyfriedelin (Li et al.,2015)	C ₃₀ H ₅₀ O ₂	-7.7
147	2beta,16alpha,19-trihydroxy-ent-kaurane (Li et al.,2015; Yan et al.,2011)	C ₂₀ H ₃₄ O ₃	-6.3
148	3beta,16alpha,17-trihydroxy-ent-kaurane (Li et al.,2015)	C ₂₀ H ₃₄ O ₃	-6.9
149	3beta-hydroxy-cycloart-25-ene-24-one (Li et al.,2015)	C ₃₀ H ₄₈ O ₂	-6.5
150	3beta-hydroxyurs-12-ene (Mallavadhani and Narasimhan, 2009)	C ₂₉ H ₄₈ O	-7.7
151	ent-kaur-16-ene-3beta-ol (Li et al.,2015)	C ₂₁ H ₃₄	-6.4
152	isojaponin A (Mekam et al.,2019)	C ₂₁ H ₃₀ O ₆	-7.5
<i>Organic 1,3-dipolar compound/allyl-type 1,3-dipolar organic compound</i>			
153	azidocyclohexane (Rautela et al.,2020)	C ₆ H ₁₁ N ₃	-4.3
<i>Organic acids and derivatives/carboxylic acids and derivatives</i>			
154	ethyl 1-ethylpyrrolidine-2-carboxylate (Rautela et al.,2020)	C ₉ H ₁₇ NO ₂	-4.4
155	phenylalanine (Mekam et al.,2019)	C ₉ H ₁₁ NO ₂	-5.3
156	tyrosine (Mekam et al.,2019)	C ₉ H ₁₁ NO ₃	-5.5
157	2-[[2-amino-3-(4-hydroxyphenyl)propanoyl]amino]pentanedioic acid (Mekam et al.,2019)	C ₁₄ H ₁₈ N ₂ O ₆	-6.8
158	maleic acid (Linfang et al., 2012)	C ₄ H ₄ O ₄	-4.3
159	dehydrochebulic acid triethyl ester (Yang et al.,2020)	C ₂₀ H ₂₂ O ₁₁	-6.7
160	hydroxycitric acid (Mekam et al.,2019)	C ₆ H ₈ O ₈	-5.1
161	citric acid (Mekam et al.,2019)	C ₆ H ₈ O ₇	-5.1
<i>Organic acids and derivatives/hydroxy acids and derivatives</i>			
162	malic acid (Mekam et al.,2019)	C ₄ H ₆ O ₅	-4.8
<i>Organic acids and derivative/organic phosphoric acid and derivative</i>			
163	methyl bis(trimethylsilyl) phosphate (Karki et al.,2019)	C ₇ H ₂₁ O ₄ PSi ₂	NA
164	1,4-digalloylquinic acid (Mahmoodally et al.,2020)	C ₂₁ H ₂₀ O ₁₄	-7.8

(Continued on following page)

TABLE 1 | (Continued) Phytochemicals from *E. hirta*.

ID	Phytochemical	MF ^a	BFE ^b
165	3,5-digalloylquinic acid (Mekam et al.,2019)	C ₂₁ H ₂₀ O ₁₄	-8.1
166	3-hydroxyoctanoic acid O-beta-D-glucopyranoside (Nomoto et al.,2013)	C ₁₄ H ₂₆ O ₈	-6.1
167	hirtionoside A (Nomoto et al.,2013)	C ₂₆ H ₃₄ O ₁₂	-8.7
168	hirtionoside B (Nomoto et al.,2013)	C ₂₆ H ₃₄ O ₁₁	-8.8
169	hirtionoside C (Nomoto et al.,2013)	C ₂₆ H ₃₆ O ₁₁	-8.4
Organohalogen compound/organobromide			
170	1,5-dibromo-3-methylpentane (Rautela et al.,2020)	C ₆ H ₁₂ Br ₂	-3.4
Organohalogen compound/organochloride			
171	1-bromo-6-chlorohexane (Rautela et al.,2020)	C ₆ H ₁₂ BrCl	-3.2
Organoheterocyclic compound/benzofuran			
172	3,6-dimethyl-5,6,7,7a-tetrahydro-4H-1-benzofuran-2-one (Rautela et al.,2020)	C ₁₀ H ₁₄ O ₂	-5.2
Organoheterocyclic compound/coumaran			
173	2,3-dihydrobenzofuran (Rautela et al.,2020)	C ₈ H ₈ O	-4.3
Organoheterocyclic compound/epoxide			
174	13-oxabicyclo[10.1.0]tridecane (Karki et al.,2019)	C ₁₂ H ₂₂ O	-4.7
Organoheterocyclic compounds/indoles and derivatives			
175	1,2,3,4-tetrahydrocyclopenta[b]indole (Rautela et al.,2020)	C ₁₁ H ₁₁ N	-5.4
176	tryptophan (Mekam et al.,2019)	C ₁₁ H ₁₂ N ₂ O ₂	-6.1
Organoheterocyclic compound/oxane			
177	1,3,3-trimethyl-2-oxabicyclo[2.2.2]octan-6-ol (Rautela et al.,2020)	C ₁₀ H ₁₈ O ₂	-5.2
Organoheterocyclic compound/oxepane			
178	3,4-epoxycyclohexylmethyl 3,4-epoxycyclohexanecarboxylate (Karki et al.,2019)	C ₁₄ H ₂₀ O ₄	-6.2
Organoheterocyclic compound/piperidine			
179	1-(2-piperidin-4-ylethyl)pyrrolidin-2-one (Rautela et al.,2020)	C ₁₁ H ₂₀ N ₂ O	-5.2
Organoheterocyclic compound/pyran			
180	3,5-dihydroxy-6-methyl-2,3-dihydropyran-4-one (Sharma et al.,2014; Rautela et al.,2020)	C ₆ H ₈ O ₄	-4.9
181	chelidonic acid (Mekam et al.,2019)	C ₇ H ₄ O ₆	-5.8
Organoheterocyclic compounds/pyrrolidines			
182	1-(3-methyl-3-butenyl)pyrrolidine (Rautela et al.,2020)	C ₉ H ₁₇ N	-4.1
183	2,2-bis(but-3-en-2-yl)pyrrolidine (Karki et al.,2019)	C ₁₂ H ₂₁ N	-4.4
184	1-(1-cyclohexen-1-yl)pyrrolidine (Rautela et al.,2020)	C ₁₀ H ₁₇ N	-4.6
Organometallic compound/organometalloid compound			
185	diethyl-hexoxy-(3-methylbutoxy)silane (Rautela et al.,2020)	C ₁₅ H ₃₄ O ₂ Si	NA
Organic nitrogen compound/organonitrogen compound			
186	nonanenitrile (Rautela et al.,2020)	C ₉ H ₁₇ N	-3.8
Organic oxygen compounds/organooxygen compounds			
187	cis-5-O-(4-coumaroyl)-D-quinic acid (Mekam et al.,2019)	C ₁₆ H ₁₈ O ₈	-7.5
188	trigalloylquinic acid (Mekam et al.,2019)	C ₂₈ H ₂₄ O ₁₈	-9.0
189	cryptochlorogenic acid (Mekam et al.,2019; Mahomoodally et al.,2020)	C ₁₆ H ₁₈ O ₉	-7.2
190	trans-5-O-(4-coumaroyl)-D-quinic acid (Mekam et al.,2019)	C ₁₆ H ₁₈ O ₈	-7.0
191	chlorogenic acid (Mekam et al.,2019; Mahomoodally et al.,2020)	C ₁₆ H ₁₈ O ₉	-7.2
192	cis-chlorogenic acid (Mekam et al.,2019)	C ₁₆ H ₁₈ O ₉	-6.6
193	quinic acid (Mekam et al.,2019; Mahomoodally et al.,2020)	C ₇ H ₁₂ O ₆	-5.4
194	shikimic acid (Mekam et al.,2019)	C ₇ H ₁₀ O ₅	-5.2
195	[2,6,6-trimethyl-4-(3-methylbut-2-enyl)cyclohexen-1-yl]methanol (Olaoluwa et al.,2018)	C ₁₅ H ₂₆ O	-5.8
196	2-pentylcyclohexane-1,4-diol (Karki et al.,2019)	C ₁₁ H ₂₂ O ₂	-4.7
197	quercitol (Linfang et al., 2012; Shah et al.,2019)	C ₆ H ₁₂ O ₅	-5.4
198	(R)-lotaustralin (Nomoto et al.,2013)	C ₁₁ H ₁₉ NO ₆	-6.1
199	benzyl-beta-D-glucopyranoside (Nomoto et al.,2013)	C ₁₃ H ₁₈ O ₆	-6.7
200	rutinoside (Nomoto et al.,2013)	C ₁₂ H ₂₂ O ₁₀	-6.8
201	(2R,3S,4S,5R,6R)-2-(hydroxymethyl)-6-methoxyoxane-3,4,5-triol (Rautela et al.,2020)	C ₇ H ₁₄ O ₆	-5.3
202	ternatoside C (Mekam et al.,2019)	C ₂₄ H ₂₃ N ₃ O ₇	-8.6
203	linocinnamarin (Nomoto et al.,2013)	C ₁₆ H ₂₀ O ₈	-6.5
204	6'-O-galloylsalicin (Mekam et al.,2019)	C ₂₀ H ₂₂ O ₁₁	-8.3
205	syringin (Nomoto et al.,2013)	C ₁₇ H ₂₄ O ₉	-7.0
206	gluconic acid (Mekam et al.,2019)	C ₆ H ₁₂ O ₇	-5.3
207	tartaric acid (Linfang et al., 2012)	C ₄ H ₆ O ₆	-4.8
208	5-hydroxymethyl-2-furancarboxaldehyde (Sharma et al.,2014)	C ₆ H ₆ O ₃	-4.4
209	2-hydroxy-1-(1'-pyrrolidyl)-1-buten-3-one (Rautela et al.,2020)	C ₈ H ₁₃ NO ₂	-4.4
210	xanthoxilin (Yang et al.,2020)	C ₁₀ H ₁₂ O ₄	-5.3
211	megastigmatrienone A (Perumal and Mahmud, 2013)	C ₁₃ H ₁₈ O	-5.7
212	2-(4,4,4-trichlorobutyl)cyclohexan-1-one (Rautela et al.,2020)	C ₁₀ H ₁₅ Cl ₃ O	-4.8
213	2-butoxyethanol (Ogunlesi et al.,2009)	C ₆ H ₁₄ O ₂	-3.7
214	2E,6E-dimethyl-2,6-octadiene-1,8-diol (Rautela et al.,2020)	C ₁₀ H ₁₈ O ₂	-4.8

(Continued on following page)

TABLE 1 | (Continued) Phytochemicals from *E. hirta*.

ID	Phytochemical	MF ^a	BFE ^b
215	2-methylhexadecanol (Ogunlesi et al.,2009)	C ₁₇ H ₃₆ O	-4.8
216	3,5-dipropyl-1,2,4,3,5-triselenadiborolane (Karki et al.,2019)	C ₆ H ₁₄ B ₂ Se ₃	NA
217	feruloyl malate (Mekam et al.,2019)	C ₁₄ H ₁₄ O ₈	-7.0
218	<i>trans</i> - <i>para</i> -coumaric acid (Mekam et al.,2019)	C ₉ H ₈ O ₃	-5.1
219	caffeic acid (Perumal et al.,2015); Mekam et al.,2019)	C ₉ H ₈ O ₄	-5.6
220	ferulic acid (Mekam et al.,2019)	C ₁₀ H ₁₀ O ₄	-5.7
221	4-methoxyfuro[3,2-g]chromen-7-one (Rautela et al.,2020)	C ₁₂ H ₈ O ₄	-5.8
222	isopimpinellin (Rautela et al.,2020)	C ₁₃ H ₁₀ O ₅	-5.9
223	xanthoxin (Rautela et al.,2020)	C ₁₂ H ₈ O ₄	-5.9
224	esculetin (Li et al.,2015)	C ₉ H ₆ O ₄	-6.2
225	phyllanthusin E methyl ester (Yang et al.,2020)	C ₁₄ H ₁₀ O ₈	-7.2
226	phyllanthusin E (Yang et al.,2020)	C ₁₃ H ₈ O ₈	-7.2
227	umbelliferone (Li et al.,2015)	C ₉ H ₆ O ₃	-5.5
228	daphnoretin Li et al. (2015)	C ₁₉ H ₁₂ O ₇	-8.4
229	scopoletin (Wu et al.,2012; Li et al.,2015; Shah et al.,2019)	C ₁₀ H ₈ O ₄	-5.8
230	isoscopoletin (Wu et al.,2012; Li et al.,2015)	C ₁₀ H ₈ O ₄	-5.7
231	6,7,8-trimethoxycoumarin (Sharma et al.,2014)	C ₁₂ H ₁₂ O ₅	-5.6
232	scoparone (Wu et al.,2012; Li et al.,2015; Shah et al.,2019)	C ₁₁ H ₁₀ O ₄	-5.7
233	citropten (Rautela et al.,2020)	C ₁₁ H ₁₀ O ₄	-5.7
234	trigallic acid (Mekam et al.,2019)	C ₂₁ H ₁₄ O ₁₃	-9.2
235	digallic acid (Mekam et al.,2019)	C ₁₄ H ₁₀ O ₉	-8.3
236	tetragalloyl glucose (Mahomoodally et al.,2020)	C ₃₄ H ₂₈ O ₂₂	-8.8
237	epicatechin 3-gallate (Perumal et al.,2015; Perumal et al., 2018)	C ₂₂ H ₁₈ O ₁₀	-8.2
238	leucocyanidol (Shah et al.,2019)	C ₁₅ H ₁₄ O ₇	-7.2
239	epicatechin (Mekam et al.,2019)	C ₁₅ H ₁₄ O ₆	-7.0
240	pinocembrin (Wu et al.,2012; Shah et al.,2019)	C ₁₅ H ₁₂ O ₄	-7.2
241	chrysin (Mekam et al.,2019)	C ₁₅ H ₁₀ O ₄	-7.3
242	luteolin (Wu et al.,2012)	C ₁₅ H ₁₀ O ₆	-7.5
243	dimethoxyquercetin (Sheliya et al.,2015)	C ₁₇ H ₁₄ O ₉	-7.3
244	kaempferol (Linfang et al., 2012; Wu et al.,2012; Rao et al.,2017; Ali et al.,2020)	C ₁₅ H ₁₀ O ₆	-7.8
245	quercetin (Liu et al.,2007; Linfang et al., 2012; Wu et al.,2012; Sheliya et al.,2015; Selin-Rani et al.,2016; Bach et al.,2017; Rao et al.,2017; Suganthi and Ravi, 2018; Mekam et al.,2019; Nugroho et al.,2019; Shah et al.,2019; Ali et al.,2020; Tayone et al.,2020)	C ₁₅ H ₁₀ O ₇	-7.5
246	isovitexin (Nomoto et al.,2013)	C ₂₁ H ₂₀ O ₁₀	-8.0
247	kaempferol-3-O-glucuronide (Mekam et al.,2019)	C ₂₁ H ₁₈ O ₁₂	-8.7
248	quercetin-3-O-glucuronide (Mekam et al.,2019)	C ₂₁ H ₁₈ O ₁₃	-8.0
249	euphorbianin (Shah et al.,2019)	C ₂₉ H ₃₂ O ₁₈	-8.2
250	myricetin-3-O-pentoside (Mekam et al.,2019)	C ₂₀ H ₁₈ O ₁₂	-8.4
251	myricetin-3-O-hexoside (Mekam et al.,2019)	C ₂₁ H ₂₀ O ₁₃	-7.3
252	quercetin 3-O- α -L-arabinofuranoside (Bach et al.,2020)	C ₂₀ H ₁₈ O ₁₁	-8.5
253	quercetin-3-O-pentoside (Mekam et al.,2019)	C ₂₀ H ₁₈ O ₁₁	-8.4
254	kaempferol-3-O-rhamnoside (Mekam et al.,2019)	C ₂₁ H ₂₀ O ₁₀	-7.7
255	narcissin (Mekam et al.,2019)	C ₂₈ H ₃₂ O ₁₆	-8.9
256	nicotiflorin (Mekam et al.,2019)	C ₂₇ H ₃₀ O ₁₅	-8.7
257	afzelin (Liu et al.,2007; Nomoto et al.,2013; Shah et al.,2019; Bach et al.,2020; Mahomoodally et al.,2020)	C ₂₁ H ₂₀ O ₁₀	-8.8
258	astragalin (Bach et al.,2020; Mahomoodally et al.,2020)	C ₂₁ H ₂₀ O ₁₁	-8.3
259	myricetin-3-O-rhamnoside (Liu et al.,2007; Linfang et al., 2012; Nugroho et al.,2019; Shah et al.,2019; Bach et al.,2020; Mahomoodally et al.,2020; Tayone et al.,2020)	C ₂₁ H ₂₀ O ₁₂	-9.0
260	isorhamnetin (Wu et al.,2012; Shah et al.,2019)	C ₂₁ H ₂₀ O ₁₂	-7.3
261	hyperoside (Mekam et al.,2019)	C ₂₁ H ₂₀ O ₁₂	-8.5
262	rutin (Linfang et al., 2012; Bach et al.,2017; Rao et al.,2017; Suganthi and Ravi, 2018; Mekam et al.,2019; Ali et al.,2020; Mahomoodally et al.,2020)	C ₂₇ H ₃₀ O ₁₆	-8.8
263	isoquercitrin (Mekam et al.,2019; Mahomoodally et al.,2020)	C ₂₁ H ₂₀ O ₁₂	-8.0
264	quercetin-3-O-rhamnoside (Gopi et al.,2016; Mekam et al.,2019; Mahomoodally et al.,2020)	C ₂₁ H ₂₀ O ₁₁	-9.0
265	luteolin-7-O- β -D-glucopyranoside (Bach et al.,2020)	C ₂₁ H ₂₀ O ₁₁	-7.9
266	cosmosiin (Mahomoodally et al.,2020)	C ₂₁ H ₂₀ O ₁₀	-7.8
267	scutellarein 6-glucoside (Mekam et al.,2019)	C ₂₁ H ₂₀ O ₁₁	-7.8
268	hymenoxin (Bach et al.,2020)	C ₁₉ H ₁₈ O ₈	-7.0

(Continued on following page)

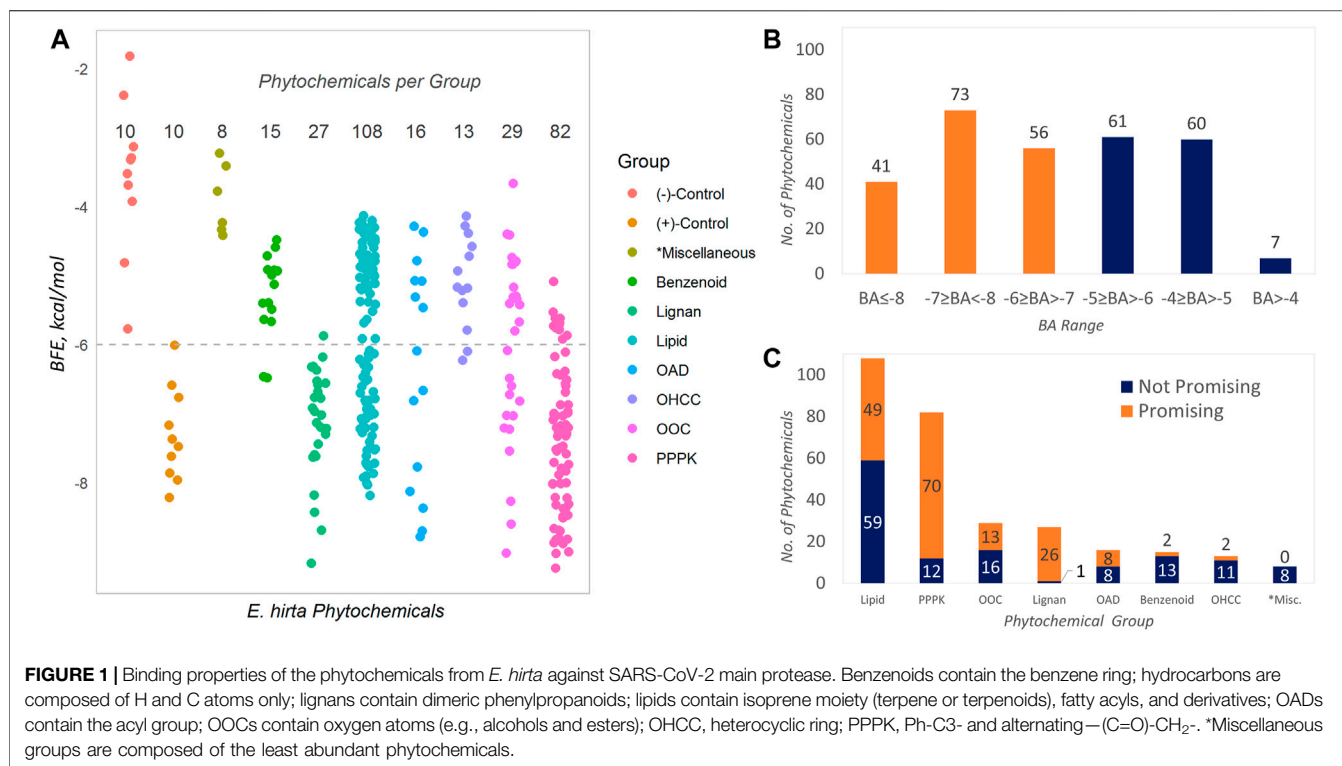
TABLE 1 | (Continued) Phytochemicals from *E. hirta*.

ID	Phytochemical	MF ^a	BFE ^b
<i>Phenylpropanoids and polyketides/isocoumarins and derivatives</i>			
269	brevifolin (Yang et al.,2020)	C ₁₂ H ₈ O ₆	-7.2
270	ethyl brevifolin carboxylate (Yang et al.,2020)	C ₁₅ H ₁₂ O ₈	-7.0
271	brevifolin carboxylic acid (Mahomoodally et al.,2020; Yang et al.,2020)	C ₁₃ H ₈ O ₈	-7.2
272	methyl brevifolin carboxylate (Yang et al.,2020)	C ₁₄ H ₁₀ O ₈	-6.4
<i>Phenylpropanoids and polyketides/tannins</i>			
273	tannic acid (Yang et al.,2020)	C ₇₆ H ₅₂ O ₄₆	-7.1
274	ellagitannin (Yang et al.,2020)	C ₄₄ H ₃₂ O ₂₇	-8.5
275	ellagic acid (Linfang et al., 2012; Mekam et al.,2019; Mahomoodally et al.,2020)	C ₁₄ H ₆ O ₈	-7.3
276	pedunculagin II (Mekam et al.,2019)	C ₃₄ H ₂₆ O ₂₂	-8.9
277	pedunculagin (Mekam et al.,2019)	C ₃₄ H ₂₄ O ₂₂	-8.0
278	corilagin (Mekam et al.,2019; Mahomoodally et al.,2020)	C ₂₇ H ₂₂ O ₁₈	-8.7
279	penta-O-galloylglucose (Mekam et al.,2019; Mahomoodally et al.,2020)	C ₄₁ H ₃₂ O ₂₆	-8.0
<i>Phenylpropanoids and polyketides</i>			
280	(R)-euphorhirtin H (Yang et al.,2020)	C ₁₆ H ₁₂ O ₁₀	-7.7
281	(R)-euphorhirtin I (Yang et al.,2020)	C ₁₅ H ₁₀ O ₁₀	-7.5
282	(R)-euphorhirtin J (Yang et al.,2020)	C ₁₇ H ₁₄ O ₁₀	-7.6
283	(R)-euphorhirtin K (Yang et al.,2020)	C ₁₈ H ₁₆ O ₁₀	-7.5
284	(R)-euphorhirtin L (Yang et al.,2020)	C ₁₈ H ₁₆ O ₁₀	-6.4
285	(R)-euphorhirtin M (Yang et al.,2020)	C ₁₇ H ₁₆ O ₉	-6.4
286	(S)-euphorhirtin H (Yang et al.,2020)	C ₁₆ H ₁₂ O ₁₀	-7.0
287	(S)-euphorhirtin I (Yang et al.,2020)	C ₁₅ H ₁₀ O ₁₀	-7.0
288	(S)-euphorhirtin J (Yang et al.,2020)	C ₁₇ H ₁₄ O ₁₀	-6.9
289	(S)-euphorhirtin K (Yang et al.,2020)	C ₁₈ H ₁₆ O ₁₀	-6.6
290	(S)-euphorhirtin L (Yang et al.,2020)	C ₁₈ H ₁₆ O ₁₀	-7.2
291	(S)-euphorhirtin M (Yang et al.,2020)	C ₁₇ H ₁₆ O ₉	-6.6
292	5-O-feruloylquinic acid (Mekam et al.,2019)	C ₁₇ H ₂₀ O ₈	-7.2
293	chebulic acid-14,15-diethyl ester (Yang et al.,2020)	C ₁₈ H ₂₀ O ₁₁	-6.5
294	euphorhirtin E (Yang et al.,2020)	C ₁₈ H ₁₈ O ₁₁	-6.7
295	euphorhirtin F (Yang et al.,2020)	C ₁₈ H ₂₀ O ₁₁	-6.1
296	euphorhirtin G (Yang et al.,2020)	C ₁₅ H ₁₂ O ₈	-6.9
297	euphorhirtin N (Yang et al.,2020)	C ₂₀ H ₂₁ NO ₉	-7.5
298	feruloylconiferin (Mekam et al.,2019)	C ₂₆ H ₂₈ O ₁₂	-8.5
<i>Negative controls</i>			
N1	anthracene	C ₁₄ H ₁₀	-5.8
N2	naphthalene	C ₁₀ H ₈	-4.8
N3	glycerol	C ₃ H ₈ O ₃	-3.9
N4	decane	C ₁₀ H ₂₂	-3.7
N5	hexanol	C ₆ H ₁₂ O	-3.5
N6	benzene	C ₆ H ₆	-3.3
N7	cyclohexane	C ₆ H ₁₂	-3.3
N8	hexane	C ₆ H ₁₄	-3.1
N9	ethanol	C ₂ H ₆	-2.4
N10	water	H ₂ O	-1.8
<i>Positive controls</i>			
P1	efonidipine	C ₃₄ H ₃₈ N ₃ O ₇ P	-8.2
P2	bedaquiline	C ₃₂ H ₃₁ BrN ₂ O ₂	-8.0
P3	tideglusib	C ₁₉ H ₁₄ N ₂ O ₂ S	-7.9
P4	manidipine	C ₃₅ H ₃₈ N ₄ O ₆	-7.6
P5	N3	C ₃₅ H ₄₈ N ₆ O ₈	-7.5
P6	lercanidipine	C ₃₆ H ₄₁ N ₃ O ₆	-7.4
P7	boceprevir	C ₂₇ H ₄₅ N ₅ O ₅	-7.2
P8	shikonin	C ₁₆ H ₁₆ O ₅	-6.8
P9	ebselen	C ₁₃ H ₉ NOSe	-6.6
P10	carmofur	C ₁₁ H ₁₆ FN ₃ O ₃	-6.0

Notes: a, molecular formula; b, computed BFE in kcal/mol using AutoDock Vina implemented in PyRx0.8. Phytochemicals with NA, indicated for their BFE, contain atoms that are not well parameterized for molecular docking using PyRx0.8. Benzenoids contain the benzene ring; hydrocarbons are composed of H and C atoms only; lignans contain dimeric phenylpropanoids; lipids contain isoprene moiety (terpene or terpenoids), fatty acyls, and derivatives; OADs contain the acyl group; OOCs contain oxygen atoms (e.g., alcohols and esters); OHCC, heterocyclic ring; PPPK, Ph-C3- and alternating-(C = O)-CH₂-. * Miscellaneous groups are composed of the least abundant phytochemicals.

components of *E. hirta* were identified from verified sources. This is by far the most comprehensive data gathering for *E. hirta* phytochemicals. The phytochemicals gathered fall into 13

ClassyFire Superclass levels. Majority are lipids and lipid-like molecules (Lipids) (108, 36.2%); phenylpropanoids and polyketides (PPPKs) (82, 27.5%); organic oxygen compounds



(OOCs) (29, 9.7%); lignan, neolignans, and related compounds (Lignans) (27, 9.1%); organic acids and derivatives (OADs) (16, 5.4%); benzenoids (15, 5.0%); and organoheterocyclic compounds (OHCCs) (13, 4.4%), comprising a total of 97.0%. The rest (*Miscellaneous) of the phytochemicals are hydrocarbons, organic 1,3-dipolar compounds, organic nitrogen, organohalogens, and organic salt.

3.2 Virtual Screening Through Automated Molecular Docking

The data obtained in Table 1 are graphically presented in Figure 1. The BFE values of the phytochemicals are described in 1A, and these are compared to the control compounds (positive and negative). It can be observed that the positive controls obtained more highly negative BFE values (thermodynamically stable receptor–ligand interaction) against SARS-CoV-2 Mpro than the negative controls (see entries in Table 1). The least negative in the group is that of carmofur with -6.0 kcal/mol computed BFE based on the AutoDock Vina docking algorithm. This value (-6.0 kcal/mol) was taken as the threshold for assigning promising inhibitors considering the fact that carmofur and the rest of the positive control compounds are actual *in vitro* inhibitors against SARS-CoV-2 Mpro. Phytochemicals having BFE values of ≤ -6.0 kcal/mol qualify as promising inhibitors. In Figure 1A, these phytochemicals are represented by the points on and below the dashed horizontal line. Over this line are the non-promising inhibitors and the negative controls with less satisfactory BFE values. Overall, 170 (57.0%) of the phytochemicals found in *E. hirta* were identified as

promising inhibitors against SARS-CoV-2 Mpro from a total of 298 phytochemicals.

The distribution of the BFEs is shown in Figure 1B and that of the promising inhibitors is highlighted as orange bars. The BFE range with the most abundant phytochemicals is $-7 \geq \text{BFE} < -8$ with 73 (24.5%) promising inhibitors. It can be seen in both Figure 1A and Figure 1C that the two most abundant groups are lipids (108, 36.2%) and PPPKs (82, 27.5%), collectively comprising 63.7% of the total. Interestingly, PPPKs have the most number of promising phytochemicals per group. There are 70 out of 82 (85.4%) PPPKs that are promising inhibitors. This value is 23.5% of the total number of *E. hirta* phytochemicals. This behavior by the PPPKs has been previously noted (Cayona and Creencia, 2021a, Cayona and Creencia, 2022). The relative numbers of promising and non-promising inhibitors with respect to chemical grouping are given in Figure 1C.

3.3 Antiviral Phytochemicals From *E. hirta*

Virtual screening revealed that *E. hirta* is an abundant source of promising inhibitors of SARS-CoV-2 Mpro. The list of promising inhibitors includes notable compounds with interesting biological and pharmacological properties. At least 12 of the promising inhibitors were established *in vitro* or *in vivo* antiviral compounds against various viruses. These are kaempferol (A), luteolin (B), quercetin (C), isoquercitrin (D), hyperoside (E), rutin (F), myricetin-3-O-rhamnoside (G), daphnoretin (H), digallic acid (I), epicatechin-3-gallate (J), trigallic acid (K), and corilagin (L). The chemical structures of the aforementioned compounds and their overlain conformations on the active site of SARS-CoV-2 Mpro represented by an H-bonding surface are shown in

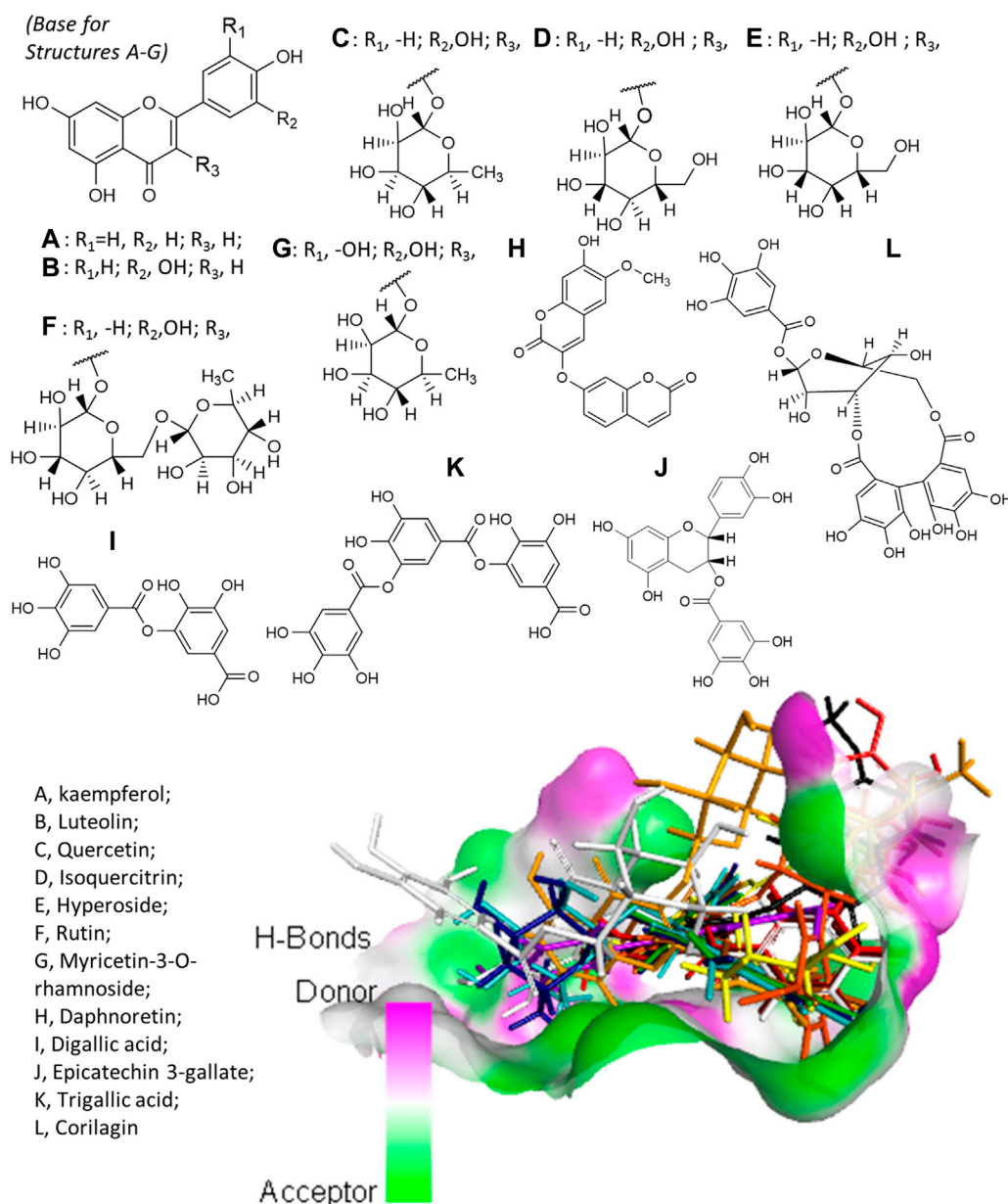


FIGURE 2 | Structures of the antiviral compounds found in *E. hirta* and overlain stick representation of the ligands in complex with SARS-CoV-2 Mpro. **(A)** Brown; **(B)** purple; **(C)** mint green; **(D)** cyan; **(E)** blue; **(F)** gold; **(G)** yellow; **(H)** black; **(I)** maroon; **(J)** orange; **(K)** red; and **(L)** gray.

Figure 2. A–G all have a common molecular skeleton, of which A is the only one without an attached sugar moiety. The skeleton of H is an isomer of A–G skeleton, and I–L are gallic acid derivatives.

The viruses susceptible to compounds A–L are listed in **Table 2** along with relevant details obtained from virtual screening (i.e., BFEs and interacting AAs). The susceptible viruses include herpes simplex virus (HSV), hepatitis, enterovirus, human immunodeficiency virus (HIV), and influenza. Interestingly, specific antivirals are effective against viruses that affect the respiratory tract, such as CoVs and respiratory syncytial viruses (RSVs). This property is clearly

relevant when considering chemical therapy against respiratory tract-related diseases like COVID-19. Kaempferol is active against CoVs (Schwarz et al., 2014), and luteolin (Wang S. et al., 2020) and daphnoretin (Wang S. et al., 2020) are active against RSVs.

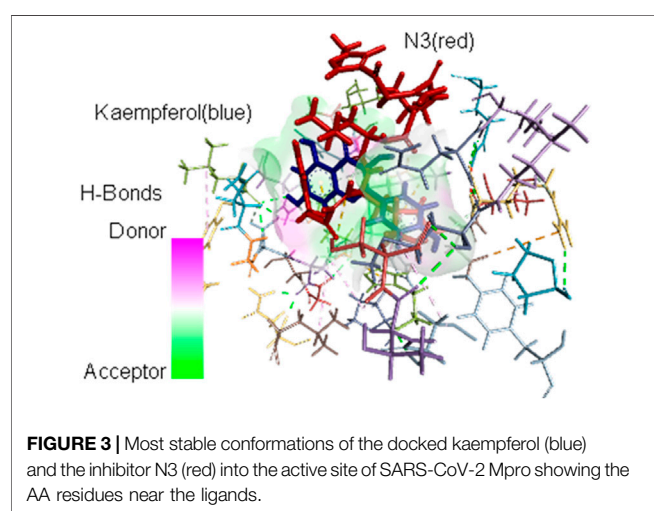
The interacting AAs are obtained from the most stable molecular docking conformation. These AAs are located at least 3.5 Å from the nearest atom of the docked ligands. It can be observed in **Table 2** that H41 and/or C145 (in boldface) catalytic dyad residues in the active site of SARS-CoV-2 Mpro can interact with the promising inhibitors (identified using DSV

TABLE 2 | *In vitro* antiviral phytochemicals rediscovered from the medicinal plants used in this study.

***	Phytochemical	BFE ^a	Interacting AAs ^b	Susceptible viruses ^c
A	Kaempferol	-7.8	H41 , M49, L141, C145 , H163, E166, M165, R188	HSV-1 (Zhu et al.,2018); CoV (Schwarz et al.,2014)
B	Luteolin	-7.5	N142, C145 , M165, R188, T190	RSV (Wang S et al.,2020); HSV (Béládi et al.,1977; Fan et al.,2016; Xu et al.,2000)
C	Quercetin	-7.5	M49, L141, C145 , M165, E166, Q189	HSVs (Gaudry et al.,2018; Kim et al.,2020; Xu et al.,2000)
D	Isoquercitrin	-8.0	H41 , M49, L141, C145 , M165, E166, P168, D187, Q189, T190	HSV (Gaudry et al.,2018; Kim et al.,2020; Xu et al.,2000)
E	Hyperoside	-8.5	M49, L141, C145 , M165, E166, R188, Q189, T190, Q192	Hepa B (Wu et al.,2007)
F	Rutin	-8.8	T26, L141, N142, G143, C145 , H163, M165, E166, R188, Q189, T190	HSVs (Béládi et al.,1977); HIV-1 (Xu et al.,2000); enterovirus (Lin et al.,2012)
G	Myricetin-3-O-rhamnoside	-9.0	M49, L141, N142, S144, C145 , E166	Hepa B (Parvez et al.,2020); influenza A (Motilhatlego et al.,2021); HIV-1 (Ortega et al.,2017)
H	Daphnoretin	-8.4	H41 , G143, C145 , M165	RSV (Ho et al.,2010; Hu et al.,2000)
I	Digallic acid	-8.3	L141, G143, S144, C145 , H163, H164, M165, E166, R188	HIV (Nakane et al.,1990)
J	Epicatechin 3-gallate	-8.2	H41 , F140, L141, N142, C145 , M165, E166, H172	HSV-2 (Alvarez et al.,2012)
K	Trigallic acid	-9.2	T26, L141, G143, S144, C145 , M165, E166, H163, Q189	HIV (Nakane et al.,1990)
***L	Corilagin	-8.7	L141, N142, G143, S144, C145 , H163, E166, P168, T190, Q192	HSV-1 (Guo et al.,2015); Hepa C (Reddy et al.,2018)

Notes: a, computed binding affinity towards SARS-CoV-2, Mpro in kcal/mol; b, interacting AA residues of the most stable conformation of the docked ligands; c, based on reported *in vitro* antiviral activity (HSV, herpes simplex virus; RSV, respiratory syncytial virus; HIV, human immunodeficiency virus; Hepa, hepatitis).

2020); however, molecular dynamics (MD) simulations are necessary to assess the stability of the receptor–ligand complex that can be formed. As stated previously, MD simulations are not covered in the scope of the present study and are reserved for future analyses. Nevertheless, the identification of these dyad residues in close proximity to the docked ligands provides a rationale for further studies. **Figure 3** shows how the most stable docked conformations of kaempferol (one of the promising inhibitors) and N3 (known inhibitor) fit into the active site of SARS-CoV-2 Mpro. The AAs that are in close proximity to the ligands are also shown.



4 DISCUSSION

This study provides the most comprehensive phytochemical gathering for *E. hirta* at present. It is argued in this the study that organized phytochemical composition will generate new information and enable meaningful analyses that may aid in understanding phytochemistry and plant metabolism. It was quite unexpected to discover an abundant cocktail of potential SARS-CoV-2 Mpro inhibitors from a single plant species. Clearly, lipids and PPPKs are among the most diverse groups of phytochemicals in *E. hirta*. These groups are also observed to be significantly more abundant in quantity than other phytochemical groups in *E. hirta* extracts (Sharma et al., 2014; Rao et al., 2017).

The molecular surface representation of the receptor reveals abundant hydrogen donor (purple) and acceptor (green) sites. This partly explains the observation that the ligands that can effectively establish H-bonding generally possess more negative BFEs than those that do not. Careful examination of the individual structures of the phytochemicals tested revealed that the capacity for H-bonding signifies direct correlation to favorable BFE. The ligands represented by the points below the dashed horizontal line in **Figure 1A** can H-bond more effectively and have more negative BFE towards SARS-CoV-2 Mpro than the ones above the line. Hydrocarbons (*Miscellaneous group), benzenoids, and OHCCs are obviously represented in the latter because they cannot (or can poorly) establish H-bonding with the receptor.

The molecular docking behavior of most PPPKs is interesting and deserves further investigation. Their docked conformations like the one presented in **Figure 2** indicate the molecular skeleton that deeply buried and extended through the active site cavity of SARS-CoV-2 Mpro. The phenylpropanoids and their structurally related groups, the lignans, may feature pharmacophoric moieties.

Available preclinical information conclusively reveals that *E. hirta* possesses antiviral properties (Perera et al., 2018). In this study, some of these antiviral phytochemicals with established antiviral properties against various viruses, including those that affect the respiratory tract, were rediscovered through the PM-VS strategy. These properties are relevant in the effort to address a respiratory disease like COVID-19. More importantly, the strategy allowed the identification of many other promising inhibitors of SARS-CoV-2 Mpro despite its simplicity. Further studies are definitely necessary, but preliminary results gathered on the demonstration of the proof-of-principle for PM-VS provide a basis for exhaustive *in silico* investigations and future *in vitro* experiments. PM-VS can be efficiently implemented in the preliminary stages of drug discovery and development with minimal computational cost. Moving forward, other drug targets, not only COVID-19 drug targets, can also be investigated with PM-VS using different medicinal plants.

5 CONCLUSION

A method described as phytochemical mining allowed the systematic collection and organization of phytochemical components from *E. hirta*. A total of 298 *E. hirta* phytochemicals collected from the literature represent the most comprehensive phytochemical data collection for the plant. Virtual screening through molecular docking of the phytochemicals revealed an abundant cocktail of 170 promising inhibitors against SARS-CoV-2 Mpro. Twelve of the promising inhibitors are also prominent natural products with reported antiviral property against diverse viruses including respiratory CoV and RSVs. Finally, PM-VS was successfully

implemented in this study, and the preliminary results obtained so far suggest further investigations.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

The authors (RC and EC) contributed equally to the conceptualization, data gathering, analysis, and manuscript preparation.

ACKNOWLEDGMENTS

The authors acknowledge the Drug Design and Development Related Research Lab (DDDRRL) of the Premier Research Institute of Science and Mathematics (PRISM), MSU-IIT, for supporting this study. RC thanks DOST-SEI for the scholarship grant under ASTHRDP-NSC.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2021.801401/full#supplementary-material>

REFERENCES

- Ali, M. Z., Mehmood, M. H., Saleem, M., and Gilani, A.-H. (2020). The Use of *Euphorbia Hirta* L. (Euphorbiaceae) in Diarrhea and Constipation Involves Calcium Antagonism and Cholinergic Mechanisms. *BMC Complement. Med. Ther.* 20 (1), 14. doi:10.1186/s12906-019-2793-0
- Álvarez, Á. L., Dalton, K. P., Nicieza, I., Diñeiro, Y., Picinelli, A., Melón, S., et al. (2012). Bioactivity-Guided Fractionation of *Phyllanthus Orbicularis* and Identification of the Principal Anti HSV-2 Compounds. *Phytother. Res.* 26 (10), 1513–1520. doi:10.1002/ptr.4608
- Amos Samkumar, R., Premnath, D., and David Paul Raj, R. S. (2019). Strategy for Early Callus Induction and Identification of Anti-snake Venom Triterpenoids from Plant Extracts and Suspension Culture of *Euphorbia Hirta* L. *3 Biotech.* 9 (7), 266. doi:10.1007/s13205-019-1790-9
- Bach, L. T., Dung, L. T., Quetin-Leclercq, J., Kestermont, P., Scippo, M.-L., Phuong, T. N., et al. (2017). The Flavonoid Isolation and Antioxidant Activity of *Euphorbia Hirta* L. Extracts. *Vietnam J. Chem.* 55 (5E34), 568–573. Retrieved from: <https://aquabioactive.ctu.edu.vn/images/upload/publications/bai-bao-chi-bach.pdf>
- Bach, L. T., Hue, B. T. B., Tram, N. T. T., Thu, D. N. A., and Dung, L. T. (2020). Chemical Constituents from N-Hexane and Ethyl Acetate Extracts of *Euphorbia Hirta* L. Grown in Vietnam. *IOP Conf. Ser. Mater. Sci. Eng.* 736, 022083. doi:10.1088/1757-899X/736/2/022083
- Béládi, I., Pusztai, R., Mucsi, I., Bakay, M., and Gábor, M. (1977). Activity of Some Flavonoids against Viruses. *Ann. NY Acad. Sci.* 284 (1 Third Confere), 358–364. doi:10.1111/j.1749-6632.1977.tb21971.x
- Cayona, R., and Creencia, E. (2021a). Discovery of a "Cocktail" of Potential SARS-CoV-2 Main Protease Inhibitors through Virtual Screening of Known Chemical Components of *Vitex Negundo* L. ("Lagundi"). *Med. Chem.* 18 (3), 364381. doi:10.2174/1573406417666210618132003
- Cayona, R., and Creencia, E. (2021b). Phytochemical Mining Of SARS-COV-2 Main Protease Inhibitors From the Philippine Medicinal Plants: *Vitex Negundo*, *Blumea Balsamifera*, *Euphorbia Hirta*, *Catharanthus Roseus*, *Mentha Cordifolia*, *Ehretia Microphylla*, *Combretum Indicum*, *Lagerstroemia Speciosa*, *Peperomia Pellucida*, and *Senna alata* [Dissertation]. Iligan City: Mindanao State University-Iligan Institute of Technology.
- Cayona, R., and Creencia, E. (2022). Phytochemical Mining of Potential SARS-CoV-2 Main Protease Inhibitors from *Blumea Balsamifera* (L.) DC. *Philippine J. Sci.* 151 (1), 235–261. Retrieved from: <https://philjournalsci.dost.gov.ph/111-vol-151-no-1-february-2022/1555-phytochemical-mining-of-potential-sars-cov-2-main-protease-inhibitors-from-blumea-balsamifera-l-dc>
- Chi, S.-M., Wang, Y., Zhao, Y., Pu, J.-X., Du, X., Liu, J.-P., et al. (2012). A New Cyclopentanone Derivative from *Euphorbia Hirta*. *Chem. Nat. Compd.* 48 (4), 577–579. doi:10.1007/s10600-012-0315-0
- Dallakyan, S., and Olson, A. J. (2015). "Small-Molecule Library Screening by Docking with PyRx," in *Chemical Biology: Methods and Protocols*. Editors J. E. Hempel, C. H. Williams, and C. C. Hong (Springer), 243–250. doi:10.1007/978-1-4939-2269-7_19
- Dayrit, F. M., Jr, A. M. G., and Gloriani, N. G. (2021). Philippine Medicinal Plants with Potential Immunomodulatory and Anti-SARS-CoV-2 Activities. *Philippine J. Sci.* 150 (5), 17. Retrieved from: <https://philjournalsci.dost.gov.ph/108-vol-150-no-5-october-2021/1467-philippine-medicinal-plants-with-potential-immunomodulatory-and-anti-sars-cov-2-activities>

- Di Micco, S., Musella, S., Scala, M. C., Sala, M., Campiglia, P., Bifulco, G., et al. (2021). *In Silico* Analysis Revealed Potential Anti-SARS-CoV-2 Main Protease Activity by the Zonulin Inhibitor Larazotide Acetate. *Front. Chem.* 8, 628609. doi:10.3389/fchem.2020.628609
- Djombou Feunang, Y., Eisner, R., Knox, C., Chepelev, L., Hastings, J., Owen, G., et al. (2016). ClassyFire: Automated Chemical Classification with a Comprehensive, Computable Taxonomy. *J. Cheminform.* 8 (1), 61. doi:10.1186/s13321-016-0174-y
- Ekpo, O. E., and Pretorius, E. (2007). Asthma, *Euphorbia hirta* and its Anti-Inflammatory Properties. *South Afr. J. Sci.* 103 (5), 201–203. Retrieved from: https://www.researchgate.net/publication/286838495_Asthma_Euphorbia_hirta_and_its_anti-inflammatory_properties
- Fan, W., Qian, S., Qian, P., and Li, X. (2016). Antiviral Activity of Luteolin against Japanese Encephalitis Virus. *Virus. Res.* 220, 112–116. doi:10.1016/j.virusres.2016.04.021
- Gaudry, A., Bos, S., Viranaicken, W., Roche, M., Krejbich-Trotot, P., Gadea, G., et al. (2018). The Flavonoid Isoquercitrin Precludes Initiation of Zika Virus Infection in Human Cells. *Int. J. Mol. Sci.* 19 (4), 1093. doi:10.3390/ijms19041093
- Ghahremanpour, M. M., Tirado-Rives, J., Deshmukh, M., Ippolito, J. A., Zhang, C.-H., Cabeza de Vaca, I., et al. (2020). Identification of 14 Known Drugs as Inhibitors of the Main Protease of SARS-CoV-2. *ACS Med. Chem. Lett.* 11 (12), 2526–2533. doi:10.1021/acsmchemlett.0c00521
- Gopi, K., Anbarasu, K., Renu, K., Jayanthi, S., Vishwanath, B. S., and Jayaraman, G. (2016). Quercetin-3-O-rhamnoside from *Euphorbia hirta* Protects Against Snake Venom Induced Toxicity. *Biochim. Biophys. Acta (Bba) - Gen. Subjects.* 1860 (7), 1528–1540. doi:10.1016/j.bbagen.2016.03.031
- Guo, Y.-J., Luo, T., Wu, F., Liu, H., Li, H.-R., Mei, Y.-W., et al. (2015). Corilagin Protects Against HSV1 Encephalitis through Inhibiting the TLR2 Signaling Pathways *In Vivo* and *In Vitro*. *Mol. Neurobiol.* 52 (3), 1547–1560. doi:10.1007/s12035-014-8947-7
- Guzman, G., Dacanay, A., Andaya, B., and Alejandro, G. (2016). Ethnopharmacological Studies on the Uses of *Euphorbia hirta* in the Treatment of Dengue in Selected Indigenous Communities in Pangasinan (Philippines). *J. Intercult Ethnopharmacol.* 5 (3), 239–243. doi:10.5455/jice.20160330124637
- Hakmi, M., Bouricha, E. M., Kandoussi, I., Harti, J. E., and Ibrahim, A. (2020). Repurposing of Known Anti-Virals as Potential Inhibitors for SARS-CoV-2 Main Protease Using Molecular Docking Analysis. *Bioinformation.* 16 (4), 301. doi:10.6026/97320630016301
- Ho, W.-S., Xue, J.-Y., Sun, S. S. M., Ooi, V. E. C., and Li, Y.-L. (2010). Antiviral Activity of Daphnoretin Isolated from *Wikstroemia indica*. *Phytother. Res.* 24 (5), 657–661. doi:10.1002/ptr.2935
- Hu, K., Kobayashi, H., Dong, A., Iwasaki, S., and Yao, X. (2000). Antifungal, Antimitotic and Anti-HIV-1 Agents from the Roots of *Wikstroemia indica*. *Planta Med.* 66 (6), 564–567. doi:10.1055/s-2000-8601
- Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., et al. (2020). Structure of Mpro from SARS-CoV-2 and Discovery of its Inhibitors. *Nature.* 582 (7811), 289–293. doi:10.1038/s41586-020-2223-y
- Karki, S., Shrestha, K., Gautam, R., and Narayan, R. (2019). Phytochemical Screening, FT-IR and GC-MS Analysis of *Euphorbia hirta*. *J. Pharmacognosy Phytochemistry.* 9 (1), 1883–1889. Retrieved from: https://www.researchgate.net/publication/349124683_Phytochemical_screening_FT-IR_and_GC-MS_analysis_of_Euphorbia_hirta
- Khan, A., Ali, S. S., Khan, M. T., Saleem, S., Ali, A., Suleman, M., et al. (2020). Combined Drug Repurposing and Virtual Screening Strategies with Molecular Dynamics Simulation Identified Potent Inhibitors for SARS-CoV-2 Main Protease (3CLpro). *J. Biomol. Struct. Dyn.* 39 (0), 4659–4670. doi:10.1080/07391102.2020.1779128
- Kim, C. H., Kim, J.-E., and Song, Y.-J. (2020). Antiviral Activities of Quercetin and Isoquercitrin Against Human Herpesviruses. *Molecules.* 25 (10), 2379. doi:10.3390/molecules25102379
- Li, E.-T., Liu, K.-H., Zang, M.-H., Zhang, X.-L., Jiang, H.-Q., Zhou, H.-L., et al. (2015). Chemical Constituents from *Euphorbia hirta*. *Biochem. Syst. Ecol.* 62, 204–207. doi:10.1016/j.bse.2015.09.007
- Lin, Y.-J., Chang, Y.-C., Hsiao, N.-W., Hsieh, J.-L., Wang, C.-Y., Kung, S.-H., et al. (2012). Fisetin and Rutin as 3C Protease Inhibitors of Enterovirus A71. *J. Virol. Methods.* 182 (1), 93–98. doi:10.1016/j.jviromet.2012.03.020
- Linfang, H., Shilin, C., and Meihua, Y. (2012). *Euphorbia hirta* (Feiyangcao): A Review on its Ethnopharmacology, Phytochemistry and Pharmacology. *J. Med. Plants Res.* 6 (39), 5176–5185. doi:10.5897/JMPR12.206
- Liu, Y., Murakami, N., Ji, H., Abreu, P., and Zhang, S. (2007). Antimalarial Flavonol Glycosides from *Euphorbia hirta*. *Pharm. Biol.* 45 (4), 278–281. doi:10.1080/13880200701214748
- Luci-Atienza, C. (2021a). 68 COVID-19 Patients Already Recruited to Participate in PH's Tawa-Tawa Clinical Trial. Manila Bulletin. Available at: <https://mb.com.ph/2021/07/02/68-covid-19-patients-already-recruited-to-participate-in-phs-tawa-tawa-clinical-trial/>.
- Luci-Atienza, C. (2021b). Recruitment of Volunteers for PH Clinical Trial for Tawa-Tawa as COVID-19 Treatment Still Ongoing—DOST. Manila Bulletin. Available at: <https://mb.com.ph/2021/08/27/recruitment-of-volunteers-for-ph-clinical-trial-for-tawa-tawa-as-covid-19-treatment-still-ongoing-dost/>.
- Ma, C., Hu, Y., Townsend, J. A., Lagarias, P. I., Marty, M. T., Kolocouris, A., et al. (2020). Ebselen, Disulfiram, Carmofur, PX-12, Tideglusib, and Shikonin Are Nonspecific Promiscuous SARS-CoV-2 Main Protease Inhibitors. *ACS Pharmacol. Transl. Sci.* 3 (6), 1265–1277. doi:10.1021/acspsci.0c00130
- Mahmoodally, M. F., Dall'Acqua, S., Sinan, K. I., Sut, S., Ferrarese, I., Etienne, O. K., et al. (2020). Phenolic Compounds Analysis of Three *Euphorbia* Species by LC-DAD-MSn and Their Biological Properties. *J. Pharm. Biomed. Anal.* 189, 113477. doi:10.1016/j.jpba.2020.113477
- Mallavadhani, U. V., and Narasimhan, K. (2009). Two Novel Butanol Rhamnosides from an Indian Traditional Herb, *Euphorbia hirta*. *Nat. Product. Res.* 23 (7), 644–651. doi:10.1080/14786410802214009
- Martínez-Vázquez, M., Apan, T. O. R., Lazcano, M. E., and Bye, R. (1999). Anti-inflammatory Active Compounds from the N-Hexane Extract of *Euphorbia hirta*. *J. Mexican Chem. Soc.* 43 (3,4), 103–105. Retrieved from: <https://www.redalyc.org/pdf/475/47543410.pdf>
- Mekam, P. N., Martini, S., Nguefack, J., Tagliazucchi, D., and Stefani, E. (2019). Phenolic Compounds Profile of Water and Ethanol Extracts of *Euphorbia hirta* L. Leaves Showing Antioxidant and Antifungal Properties. *South Afr. J. Bot.* 127, 319–332. doi:10.1016/j.sajb.2019.11.001
- Menéndez, C. A., Byléhn, F., Perez-Lemus, G. R., Alvarado, W., and de Pablo, J. J. (2020). Molecular Characterization of Ebselen Binding Activity to SARS-CoV-2 Main Protease. *Sci. Adv.* 6 (37), eabd0345. doi:10.1126/sciadv.abd0345
- Mirza, M. U., and Froeyen, M. (2020). Structural Elucidation of SARS-CoV-2 Vital Proteins: Computational Methods Reveal Potential Drug Candidates Against Main Protease, Nsp12 Polymerase and Nsp13 Helicase. *J. Pharm. Anal.* 10 (4), 320–328. doi:10.1016/j.jpba.2020.04.008
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and Group, T. P. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Plos Med.* 6 (7), e1000097. doi:10.1371/journal.pmed.1000097
- Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., et al. (2009). AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* 30 (16), 2785–2791. doi:10.1002/jcc.21256
- Mothlatlego, K. E., Mehrbod, P., Fotouhi, F., Abdalla, M. A., Eloff, J. N., and McGaw, L. J. (2021). Anti-influenza A Virus Activity of Two *Newtonia* Species and the Isolated Compound Myricetin-3-O-Rhamnoside. *BMC Complement. Med. Ther.* 21 (1), 92. doi:10.1186/s12906-021-03250-0
- Nakane, H., Fukushima, M., and Ono, K. (1990). Differential Inhibition of Reverse Transcriptase and Various DNA Polymerases by Digallic Acid and its Derivatives. *J. Nat. Prod.* 53 (5), 1234–1240. doi:10.1021/np50071a015
- Nomoto, Y., Sugimoto, S., Matsunami, K., and Otsuka, H. (2013). Hirtionosides A-C, Gallates of Megastigmane Glucosides, 3-hydroxyoctanoic Acid Glucosides and a Phenylpropanoid Glucoside from the Whole Plants of *Euphorbia hirta*. *J. Nat. Med.* 67 (2), 350–358. doi:10.1007/s11418-012-0692-5
- Nugroho, A., Heryani, H., and Istikowati, W. T. (2020). Quantitative Determination of Quercitrin and Myricitrin in Three Different Parts of *Euphorbia hirta* as Bioflavonoid Source for Functional Food. *IOP Conf. Ser. Earth Environ. Sci.* 443, 012042–012046. doi:10.1088/1755-1315/443/1/012042
- O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., and Hutchison, G. R. (2011). Open Babel: An Open Chemical Toolbox. *J. Cheminform.* 3 (1), 33. doi:10.1186/1758-2946-3-33
- Ogunlesi, M., Okiei, W., Ofor, E., and Osibote, A. E. (2009). Analysis of the Essential Oil from the Dried Leaves of *Euphorbia hirta* Linn (Euphorbiaceae), a

- Potential Medication for Asthma. *Afr. J. Biotechnol.* 8 (24), 7042–7050. doi:10.4314/AJB.V8I24.68792
- Olaoluwa, O., Moronkola, D., Taiwo, O., and Iganboh, P. (2018). Volatile Oil Composition, Antioxidant and Antimicrobial Properties of *Boerhavia Erecta* L. And *Euphorbia Hirta* L. *Trends Phytochem. Res.* 2 (3), 171–178. Retrieved from: http://tpr.iau-shahrood.ac.ir/article_543327.html
- Ortega, J. T., Suárez, A. I., Serrano, M. L., Baptista, J., Pujol, F. H., and Rangel, H. R. (2017). The Role of the Glycosyl Moiety of Myricetin Derivatives in Anti-HIV-1 Activity *In Vitro*. *AIDS Res. Ther.* 14 (1), 57. doi:10.1186/s12981-017-0183-6
- Parvez, M. K., Al-Dosari, M. S., Arbab, A. H., Al-Rehaily, A. J., and Abdelwahid, M. A. S. (2020). Bioassay-Guided Isolation of Anti-Hepatitis B Virus Flavonoid Myricetin-3-O-Rhamnoside Along with Quercetin from *Guiera Senegalensis* Leaves. *Saudi Pharm. Journal* The Official Publ. Saudi Pharm. Soc. 28 (5), 550–559. doi:10.1016/j.jsps.2020.03.006
- Perera, S. D., Jayawardena, U. A., and Jayasinghe, C. D. (2018). Potential Use of *Euphorbia Hirta* for Dengue: A Systematic Review of Scientific Evidence. *J. Trop. Med.* 2018, 7. doi:10.1155/2018/2048530
- Perumal, S., and Mahmud, R. (2013). Chemical Analysis, Inhibition of Biofilm Formation and Biofilm Eradication Potential of *Euphorbia Hirta* L. Against Clinical Isolates and Standard Strains. *BMC Complement. Altern. Med.* 13, 346. doi:10.1186/1472-6882-13-346
- Perumal, S., Mahmud, R., and Mohamed, N. (2018). Combination of Epicatechin 3-Gallate from *Euphorbia Hirta* and Cefepime Promotes Potential Synergistic Eradication Action against Resistant Clinical Isolate of *Pseudomonas aeruginosa*. *Evidence-Based Complement. Altern. Med.* 2018, 1–7. doi:10.1155/2018/5713703
- Perumal, S., Mahmud, R., and Ramanathan, S. (2015). Anti-Infective Potential of Caffeic Acid and Epicatechin 3-gallate Isolated from Methanol Extract of *Euphorbia hirta* (L.) against *Pseudomonas Aeruginosa*. *Nat. Product. Res.* 29 (18), 1766–1769. doi:10.1080/14786419.2014.999242
- Prachi, S., and Pradeep, T. (2014). 13 α -methyl-27-norolean-14-en-3 β -ol, a Triterpene Isolated from the Stem of *Euphorbia Hirta* (Linn) Possess an Anti-asthmatic Properties. *Res. J. Chem. Sci.* 4 (3), 21–26. Retrieved from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1082.2070&rep=rep1&type=pdf>
- Ragasa, C. Y., and Cornelio, K. B. (2013). Triterpenes from *Euphorbia Hirta* and Their Cytotoxicity. *Chin. J. Nat. Medicines*. 11 (5), 528–533. doi:10.1016/S1875-5364(13)60096-5
- Rao, C., Gupta, S., Azmi, L., and Mohapatra, P. (2017). Flavonoids from Whole Plant of *Euphorbia Hirta* and Their Evaluation against Experimentally Induced Gastroesophageal Reflux Disease in Rats. *Phcog Mag.* 13 (Suppl. 1), 127–S134. doi:10.4103/0973-1296.203987
- Rautela, I., Joshi, P., Thapliyal, P., Pant, M., Dheer, P., Bisht, S., et al. (2020). Comparative GC-MS Analysis of *Euphorbia Hirta* and *Euphorbia Milli* for Therapeutic Potential Utilities. *Plant Arch.* 20 (2), 3515–3512. Retrieved from: [http://plantarchives.org/20-2/3515-3522%20\(6354\).pdf](http://plantarchives.org/20-2/3515-3522%20(6354).pdf)
- Reddy, B. U., Mullick, R., Kumar, A., Sharma, G., Bag, P., Roy, C. L., et al. (2018). A Natural Small Molecule Inhibitor Corilagin Blocks HCV Replication and Modulates Oxidative Stress to Reduce Liver Damage. *Antiviral Res.* 150, 47–59. doi:10.1016/j.antiviral.2017.12.004
- Schwarz, S., Sauter, D., Wang, K., Zhang, R., Sun, B., Karioti, A., et al. (2014). Kaempferol Derivatives as Antiviral Drugs Against the 3 α Channel Protein of Coronavirus. *Planta Med.* 80 (2–3), 177–182. doi:10.1055/s-0033-1360277
- Selin-Rani, S., Senthil-Nathan, S., Thanigaivel, A., Vasantha-Srinivasan, P., Edwin, E.-S., Ponsankar, A., et al. (2016). Toxicity and Physiological Effect of Quercetin on Generalist Herbivore, Spodoptera Litura Fab. And a Non-Target Earthworm *Eisenia fetida* Savigny. *Chemosphere*. 165, 257–267. doi:10.1016/j.chemosphere.2016.08.136
- Shah, A. P., Parmar, G. R., Sailor, G. U., and Seth, A. K. (2019). Antimalarial Phytochemicals Identification from *Euphorbia Hirta* against Plasmeppsin Protease: An In Silico Approach. *Folia Med.* 61 (4), 584–593. doi:10.3897/folmed.61.e47965
- Sharma, N., Samarakoon, K., Gyawali, R., Park, Y.-H., Lee, S.-J., Oh, S., et al. (2014). Evaluation of the Antioxidant, Anti-Inflammatory, and Anticancer Activities of *Euphorbia Hirta* Ethanolic Extract. *Molecules*. 19 (9), 14567–14581. doi:10.3390/molecules190914567
- Sheliya, M. A., Rayhana, B., Ali, A., Pillai, K. K., Aeri, V., Sharma, M., et al. (2015). Inhibition of α -glucosidase by New Prenylated Flavonoids from *euphorbia Hirta* L. Herb. *J. Ethnopharmacology*. 176, 1–8. doi:10.1016/j.jep.2015.10.018
- Shitrit, A., Zaidman, D., Kalid, O., Bloch, I., Doron, D., Yarnitzky, T., et al. (2020). Conserved Interactions Required for Inhibition of the Main Protease of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). *Sci. Rep.* 10 (1), 20808. doi:10.1038/s41598-020-77794-5
- Suganthi, A., and Ravi, T. K. (2018). Estimation of Anti-Dengue Phytochemical Markers Gallic Acid, Rutin and Quercetin in Methanolic Extract of *Euphorbia Hirta* (L.) and Tawa-Tawa Capsule Formulation by Validated RP-HPLC Method. *Chem. Methodologies*. 3 (1), 43–54. doi:10.22034/chemm.2018.129381.1051
- Tawa-tawa Clinical Trial on COVID-19 (2021). Philippine Health Research Registry. Available at: <https://registry.healthresearch.ph/index.php/registry?view=research&layout=details&cid=3475>
- Tayone, W. C., Ishida, K., Goto, S., Tayone, J. C., Arakawa, M., Morita, E., et al. (2020). Anti-Japanese Encephalitis Virus (JEV) Activity of Triterpenes and Flavonoids from *Euphorbia Hirta*. *Philippine J. Sci.* 149 (3), 603–613. Retrieved from: <https://philjournalsci.dost.gov.ph/publication/regular-issues/past-issues/98-vol-149-no-3-september-2020/1211-anti-japanese-encephalitis-virus-jev-activity-of-triterpenes-and-flavonoids-from-euphorbia-hirta>
- Trott, O., and Olson, A. J. (2009). AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* 31 (2), 455–461. doi:10.1002/jcc.21334
- Ullrich, S., and Nitsche, C. (2020). The SARS-CoV-2 Main Protease as Drug Target. *Bioorg. Med. Chem. Lett.* 30 (17), 127377. doi:10.1016/j.bmcl.2020.127377
- Wang, S., Ling, Y., Yao, Y., Zheng, G., and Chen, W. (2020). Luteolin Inhibits Respiratory Syncytial Virus Replication by Regulating the MiR-155/SOCS1/STAT1 Signaling Pathway. *Virology*. 17 (1), 187. doi:10.1186/s12985-020-01451-6
- Wang, Y. C., Yang, W. H., Yang, C. S., Hou, M. H., Tsai, C. L., Chou, Y. Z., et al. (2020). Structural Basis of SARS-CoV-2 Main Protease Inhibition by a Broad-Spectrum Anti-coronaviral Drug. *Am. J. Cancer Res.* 10 (8), 2535–2545. Retrieved from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7471349/pdf/ajcr0010-2535.pdf>
- Wu, L.-I., Yang, X.-b., Huang, Z.-m., Liu, H.-z., and Wu, G.-x. (2007). *In Vivo* and *In Vitro* Antiviral Activity of Hyperoside Extracted from *Abelmoschus Manihot* (L.) Medik. *Acta Pharmacologica Sinica*. 28 (3), 404–409. doi:10.1111/j.1745-7254.2007.00510.x
- Wu, Y., Qu, W., Geng, D., Liang, J.-Y., and Luo, Y.-L. (2012). Phenols and Flavonoids from the Aerial Part of *Euphorbia Hirta*. *Chin. J. Nat. Medicines*. 10 (1), 40–42. doi:10.1016/S1875-5364(12)60009-0
- Xu, H.-X., Wan, M., Dong, H., But, P. P.-H., and Foo, L. Y. (2000). Inhibitory Activity of Flavonoids and Tannins against HIV-1 Protease. *Biol. Pharm. Bull.* 23 (9), 1072–1076. doi:10.1248/bpb.23.1072
- Yan, S., Ye, D., Wang, Y., Zhao, Y., Pu, J., Du, X., et al. (2011). Ent-Kaurane Diterpenoids from *Euphorbia hirta*. *Rec. Nat. Prod.* 5 (4), 247–251.
- Yang, Z.-N., Su, B.-J., Wang, Y.-Q., Liao, H.-B., Chen, Z.-F., and Liang, D. (2020). Isolation, Absolute Configuration, and Biological Activities of Chebulic Acid and Brevifolinicarbonylic Acid Derivatives from *Euphorbia Hirta*. *J. Nat. Prod.* 83 (4), 985–995. doi:10.1021/acs.jnatprod.9b00877
- Zhang, L., Wang, X.-L., Wang, B., Zhang, L.-T., Gao, H.-M., Shen, T., et al. (2020). Lignans from *Euphorbia Hirta* L. *Nat. Product. Res.* 36, 26–36. doi:10.1080/14786419.2020.1761358
- Zhu, L., Wang, P., Yuan, W., and Zhu, G. (2018). Kaempferol Inhibited Bovine Herpesvirus 1 Replication and LPS-Induced Inflammatory Response. *Acta Virologica*. 62 (2), 220–225. doi:10.4149/av_2018_206

Conflict of Interest: The authors declare 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.

Copyright © 2022 Cayona and Creencia. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.



SARS-CoV-2 Spike Protein Unlikely to Bind to Integrins *via* the Arg-Gly-Asp (RGD) Motif of the Receptor Binding Domain: Evidence From Structural Analysis and Microscale Accelerated Molecular Dynamics

OPEN ACCESS

Edited by:

Mattia Falconi,
University of Rome Tor Vergata, Italy

Reviewed by:

Subodh Samrat,
University of Arizona, United States
Richard J. Bingham,
University of Huddersfield,
United Kingdom

*Correspondence:

Houcemeddine Othman
houcemeddine.othman@wits.ac.za
Najet Srairi-Abid
najet.abid@pasteur.ms.tn

Specialty section:

This article was submitted to
Biological Modeling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 13 December 2021

Accepted: 12 January 2022

Published: 14 February 2022

Citation:

Othman H, Messaoud HB,
Khamessi O, Ben-Mabrouk H,
Ghedira K, Bharuthram A, Treurnicht F,
Achilonu I, Sayed Y and Srairi-Abid N
(2022) SARS-CoV-2 Spike Protein
Unlikely to Bind to Integrins *via* the Arg-
Gly-Asp (RGD) Motif of the Receptor
Binding Domain: Evidence From
Structural Analysis and Microscale
Accelerated Molecular Dynamics.
Front. Mol. Biosci. 9:834857.
doi: 10.3389/fmolb.2022.834857

Houcemeddine Othman^{1,2*}, Haifa Ben Messaoud³, Oussema Khamessi⁴,
Hazem Ben-Mabrouk², Kais Ghedira⁵, Avani Bharuthram⁶, Florette Treurnicht⁶,
Ikechukwu Achilonu⁷, Yasien Sayed⁷ and Najet Srairi-Abid^{2*}

¹Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa, ²Laboratory of Biomolecules, Venoms and Theranostic Applications, LR20IPT01, Institut Pasteur de Tunis, University of Tunis El Manar, Tunis, Tunisia, ³National Gene Bank of Tunisia, Boulevard du Leader Yesser Arafet, Tunis, Tunisia, ⁴Université de Tunis El Manar, Institut Pasteur de Tunis, LR11IPT08 Venins et Biomolécules Thérapeutiques, Tunis, Tunisie, ⁵Laboratory of Bioinformatics, Biomathematics and Biostatistics (BIMS), Institut Pasteur de Tunis (IPT), University of Tunis El Manar, Tunis, Tunisia, ⁶Department of Virology, National Health Laboratory Services and the School of Pathology, University of the Witwatersrand, Johannesburg, South Africa, ⁷Protein Structure-Function Research Unit, School of Molecular and Cell Biology, University of Witwatersrand, Johannesburg, South Africa

The Receptor Binding Domain (RBD) of SARS-CoV-2 virus harbors a sequence of Arg-Gly-Asp tripeptide named RGD motif, which has also been identified in extracellular matrix proteins that bind integrins as well as other disintegrins and viruses. Accordingly, integrins have been proposed as host receptors for SARS-CoV-2. However, given that the microenvironment of the RGD motif imposes a structural hindrance to the protein-protein association, the validity of this hypothesis is still uncertain. Here, we used normal mode analysis, accelerated molecular dynamics microscale simulation, and protein-protein docking to investigate the putative role of RGD motif of SARS-CoV-2 RBD for interacting with integrins. We found, that neither RGD motif nor its microenvironment showed any significant conformational shift in the RBD structure. Highly populated clusters of RBD showed no capability to interact with the RGD binding site in integrins. The free energy landscape revealed that the RGD conformation within RBD could not acquire an optimal geometry to allow the interaction with integrins. In light of these results, and in the event where integrins are confirmed to be host receptors for SARS-CoV-2, we suggest a possible involvement of other residues to stabilize the interaction.

Keywords: SARS-CoV-2, COVID-19, integrin, RBD, molecular dynamics

1 INTRODUCTION

The molecular mechanism of human infection with SARS-CoV-2 has been studied extensively (Shang et al., 2020a; Harrison et al., 2020; Li et al., 2021). Alveolar epithelial cells are thought to be the main target for the virus. Indeed, in pioneering work, Chu *et al.* (Chu et al., 2020), studied the tropism of SARS-CoV-2 by inoculating it into 24 cell lines covering seven organs and tracts. They found that the virus most efficiently replicates on lung-type cell lines. Other organs can also be targeted including intestinal tracts, liver, and kidney (*idem*). At the molecular level, the interaction with the host cell involves primarily the homotrimeric spike protein (S protein) expressed on the virus surface. Prior to cell attachment, the spike protein arranges its three Receptor Binding Domains (RBD) in a laying-down configuration, which could help to evade the immune system (Berry et al., 2004). Human viruses frequently use mammalian cell surface receptors to attach and to enter host cells (Sheppard, 2003). During the interaction process with the host cell, the spike protein switches one of the RBD domains to a standing-up configuration, thus exposing the Receptor Binding Motif (RBM) to the interaction surface of the Angiotensin-Converting Enzyme 2 (ACE2) receptor. ACE2 is widely regarded as the main entry point for the virus to the cellular machinery of the host (Othman et al., 2020; Wan et al., 2020). However, evidence suggests the possibility of other receptors and co-receptors that might be as relevant as ACE2. The proteomic analysis that helped to establish the interactome map, suggested the putative implication of more than 300 host proteins in the interaction with SARS-CoV-2 (Gordon et al., 2020). While many of these proteins are expected to be false-positive hits, other studies have pointed out the critical role of specific host proteins and macromolecules as co-receptors (Zamorano Cuervo and Grandvaux, 2020), such as neuropilin-1 (Cantuti-Castelvetri et al., 2020), heparan sulfate (Clausen et al., 2020), sialic acids (Qing et al., 2020), CD147 (Aguiar et al., 2020) and GRP78 (Ibrahim et al., 2020). Recently, Sigrist et al. (2020) have identified an Arg-Gly-Asp (RGD) motif in the sequence of the spike RBD which is found to be exposed at the surface of the interaction domain. This motif was originally identified within the extracellular matrix proteins, including fibronectin, fibrinogen, vitronectin, and laminin that mediate cell attachment. Integrins are membrane proteins that act as receptors for these cell adhesion molecules via the RGD motif (Hamidi et al., 2016). Three main integrins expressed on airway epithelial cells were described to play an important role in virus infection (Isberg and Tran Van Nhieu, 1994). $\alpha_2\beta_1$, a collagen and laminin receptor, play a critical role in cell infection by echovirus (Eisner, 1992). Based on these findings, Sigrist et al. (2020) concluded that integrins can also interact with the spike protein. Several other studies have built on this hypothesis to support the role of integrins as spike protein receptors (Luan et al., 2020; Beddingfield et al., 2021; Dakal, 2021) and to exploit the property for potential therapeutic applications (Yan et al., 2020). Moreover, Beddingfield et al. (2021) showed, by *in vitro* analysis, that

the interaction with integrins is a plausible hypothesis. Integrins are heterodimeric receptors that interact favorably with the extracellular molecules by forming a cleft at the protein-protein interface between the beta-propeller and a betal domains from the alpha and beta subunits (Xiao et al., 2004). The cleft contains the Metal Ion-Dependent Adhesion Site (MIDAS) harboring an Mg^{2+} ion. Differential expression of $\alpha_2\beta_1$, $\alpha_3\beta_1$, $\alpha_4\beta_1$, $\alpha_5\beta_1$, $\alpha_7\beta_1$, $\alpha_6\beta_4$, $\alpha_9\beta_1$, $\alpha_v\beta_5$, $\alpha_v\beta_6$, $\alpha_v\beta_8$ integrins was revealed in human lung cells (Weinacker et al., 1995; Cambier et al., 2000; Bazan-Socha et al., 2005). Indeed, $\alpha_2\beta_1$, $\alpha_3\beta_1$, $\alpha_6\beta_4$, $\alpha_9\beta_1$, $\alpha_v\beta_5$, $\alpha_v\beta_6$ and $\alpha_v\beta_8$ are expressed in airway epithelial cells, which are the main target of coronavirus (Ravindra et al., 2020). Among these, only $\alpha_v\beta_5$, $\alpha_v\beta_6$ and $\alpha_v\beta_8$ can recognize RGD motif while $\alpha_5\beta_1$ integrin was not shown to be expressed in healthy epithelial cells (Sheppard, 2003). The activity of integrins can be inhibited by disintegrin peptides purified from animals such as snakes, scorpions and insects. The majority of these disintegrins incorporate an RGD motif in their sequences (Gasmi et al., 2001; Olfa et al., 2005; Bazaa et al., 2007; Assumpcao et al., 2012; Ben-Mabrouk et al., 2016). Most of the arguments about the validity of the RGD motif in SARS-CoV-2 RBD as an interacting segment with integrins are supported by sequence-based and structural-based analysis. However, the microenvironment of RGD imposes a critical steric hindrance that could prevent the RBD from optimally interacting with integrins. To investigate the extent of such effect on the RGD/RBD conformational and binding properties, we conducted a computational study involving microscale accelerated molecular dynamics simulation and protein-protein docking.

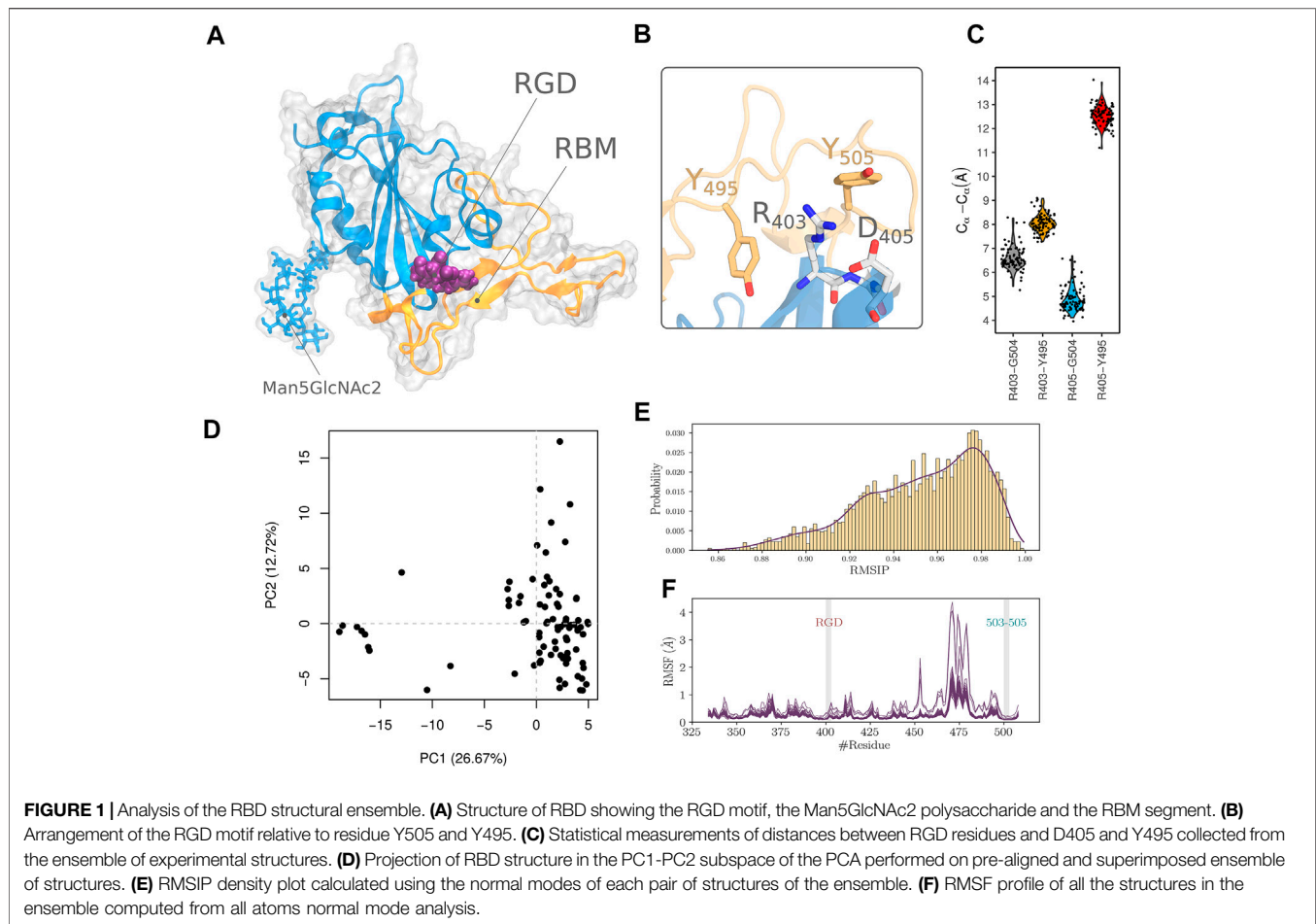
2 METHODS

2.1 Structural Data

All the structures with complete 3D coordinates of the RBD were explored. They include X-ray crystallography and the cryo-electron microscopy structures. The coordinates of the RBD domain were extracted from the entries of the complete spike protein. In total, we obtained 90 Protein Data Bank (PDB) files (**Supplementary Data S1**).

2.2 Normal Mode Analysis

The normal mode analysis (NMA) approach represents an efficient and powerful tool for predicting and characterizing the large-scale conformational transitions in protein structures around their equilibrium fluctuation. For this study, the Bio3D package in R (version 2.4-1.9000) was utilized to conduct a comparative NMA analysis of a large ensemble of structures (Fuglebak et al., 2012). All atoms low-frequency normal modes were calculated under the coarse-grained Elastic Network Model (ENM). Prior to the calculation, structures were aligned to an invariant region of RBD residues. Root Mean Squared Inner Product (RMSIP) was computed from the corresponding eigenvectors of the normal modes to calculate a score quantifying the overlap between modes. The RMSIP was calculated between all the pairs of RBD structures from the collected ensemble of PDB files.



2.3 Accelerated Molecular Dynamics

Accelerated molecular dynamics (aMD) enhances the sampling of a protein conformational space by lowering energy barriers of the energy landscape (Hamelberg et al., 2004). A bias term is added to the potential energy $V(r)$ when the value falls below a certain threshold as follows:

$$V^*(r) = V(r) + \Delta V(r)$$

$$\Delta V(r) = \begin{cases} 0 & \text{if } V(r) > E \\ \frac{(E - V(r))^2}{\alpha + E - V(r)} & \text{if } V(r) < E \end{cases}$$

where $\Delta V(r)$ is the bias; $V(r)$ is the potential energy calculated from the vector of coordinates r of all the atoms in the system; E is the threshold value of the energy, and α is the acceleration factor (Wang et al., 2011). We used the crystal structure of SARS-CoV2 RBD in complex with H11-D4 antibody (PDB code 6YZ5) at a resolution of 1.8 Å to conduct the simulations. Parameters from the ff14SB force field (Maier et al., 2015) were assigned to the atoms of the system using AMBER molecular dynamics simulation package, version 18 (Case et al., 2018). After removing the antibody and the heteroatoms from the structure, we built a Oligomannose-5 glycan (Man5GlcNAc2) type polysaccharide structure and linked it covalently to residue

N343 of the RBD (Figure 1A). The topology of the glycan was identified to be the major form for this amino acid (Watanabe et al., 2020). The system was then neutralized, and TIP3P water molecules were added to a truncated octahedron simulation box where the edges are at a minimum distance of 12 Å for any atom of the solute. Three stages of energy minimization were used to clean the geometry of the atoms and to relax the system. First, we used 5,000 steps of steepest descent minimization followed by 15,000 steps of conjugate gradient minimization while restraining both water and protein atoms at their initial positions using a force constant of 100 kcal/mol/Å² and a non-bonded contact cutoff of 12 Å. We then applied the same minimization series with 400 steps of the steepest descent algorithm and 9,600 steps of the conjugate gradient algorithm while applying the constraining force on the protein atoms only. At the final stage, we ran the same cycle and we only lowered the constraining force constant to 0.1 kcal/mol/Å² applied to the protein atoms. To further relax the system, we applied a heating stage of molecular dynamics by increasing the temperature from 50 to 300 K while maintaining a force constant of 10 kcal/mol/Å² on the heavy atoms of the RBD. A Langevin thermostat with a collision frequency of 5 ps⁻¹ was applied to control the temperature fluctuation. Following the heating stage, we lifted the constraining forces gradually by an increment of 1 over 11

intervals of 100 ps. The restrained molecular dynamics were run in the NPT ensemble by maintaining the pressure at 1 atm using a relaxation time of 2 ps. The SHAKE method was applied for all the stages of the simulation to constrain the bonds involving hydrogen atoms which allowed an integration time of 100 fs. The Particle Mesh Ewald method was applied to calculate the electrostatic forces. The production phases were run under the NVT conditions. To calculate the different parameters for the aMD simulation, we first run classical molecular dynamics for a total time of 100 ns. From there, we estimated the values of the parameters to calculate the boosting term. The aMD simulation was run in 3 independent replicates for a total time of 1 μ s each. An extra boost to the torsional space was added, and the trajectory was constructed by collecting the snapshots at every 10 ps of the running simulation.

2.4 Molecular Dynamics Data Analysis

The crystal structure was set as a reference conformation. Analysis of the molecular dynamics trajectory was made with an in-house python code. Principal Component Analysis (PCA) (David and Jacobs, 2014) was calculated for all heavy atoms in the protein, which allowed the detection of dynamical patterns with functional relevance. The translational and rotational related dynamic was first removed by fitting the ensemble of snapshots to the crystal structure of RBD. The low dimension components were calculated to return the corresponding eigenvalues and eigenvectors as well as the projection of the atomic coordinates into the lower-dimensional subspace. Clustering analysis was executed using a hierarchical algorithm embedded in the “cpptraj” analysis tool implemented by AMBER. In this regard an ϵ cutoff of 2 \AA was used. To assess the convergence of the simulation, the cumulative number of clusters (CNC) as a function of time and the evolution of informational entropy (H) were calculated. The informational entropy is defined by the following formula.

$$H = - \sum_{i=1}^n p_i \log(p_i)$$

p_i is the probability of the i th found cluster, as a function of simulation time. To recover the unbiased free energy landscape from the ensemble of conformations sampled by aMD, we reweighted the probability sampling landscape according to the following equation.

$$v_i = k_b T \ln \left[\frac{P(x_i)}{P_{\max}(x)} \right]^2$$

k_b is the Boltzmann constant, T was set to 298 K, $P(x_i)$ estimates the probability of a conformational event obtained by binning along the reaction coordinate using the histogram method. The number of bins was set to 50. $P_{\max}(x)$ is the maximum probability of the discrete state.

2.5 Protein-Protein Docking

Protein-protein docking was run using the prediction interface of HADDOCK2.2 web server (van Zundert et al., 2016). Integrin structures of $\alpha_5\beta_1$, $\alpha_{IIb}\beta_3$, and $\alpha_V\beta_3$, corresponding to PDB entries 3VI4, 3ZDY and 6UJC respectively, were defined as receptors. The

structure of integrins is in a bound state with an RGD binding segment which was removed before running the docking. All residues within a 7 \AA distance from the bound RGD in the integrin structure were used to define the active residues of the receptor. Multiple conformations of RBD, compiled from the molecular dynamics simulation, were employed as ligand structures to run the cross-docking. The amino acids of the RGD motif (in position 403–405) were used to define the active residues of the ligand structures. All other parameters of HADDOCK2.2 were kept to their default settings. The structure of the most populated cluster for each docking run was selected for analysis.

3 RESULTS

We explored the crystal structure of RBD (PDB code 6YZ5). The RGD motif extends over residues 403–405. R403 is located at the C-terminal end of the fourth β -strand of the RBD, while both G404 and D405 are part of its α -helix (Figure 1A). We noticed that only D405 and the guanidinium group of the R403 side chain are solvent-exposed (Figure 1B). RGD motif shows a considerable kink defined by the main chain atoms and the C_β atoms of R403 and D405. Such configuration leads to the close contact between the RGD motif charged groups with a distance of 4.1 \AA . This conformation is different from the optimized configuration of integrin interacting RGDs that adopt an extensive or a slightly kinked configuration (Kapp et al., 2017). The conformation might be imposed, in part, by the tight interactions with nearby amino acids of the RBD that include Y495 and Y504 (Figure 1B). Both residues are part of the receptor-binding motif with ACE2 (Shang et al., 2020b). We, therefore, hypothesized that in order to come to an integrin-compatible conformation for RGD, the nearby segments incorporating Y495 and Y504, have to move outwardly relative to the motif. We first attempted to detect such an event in the collected dataset, by assuming that a functionally relevant conformation, could be sampled in the large number of RBD experimentally solved structures. We thought that measuring the distance between reference amino acids in the RGD segment, i.e. R403 and D405, and other residues in the nearby RBM amino acids (Y495 and G504) might be a good proxy to evaluate the extent of the outward movement of the latter segment relative to RGD. The results of this analysis are reported in (Figure 1C). The median distances are 6.4, 8.0, 4.7, and 12.5 \AA , corresponding respectively to R403-Y504, R403-Y495, D405-Y504, and D405-Y495 pairs of residues. The distances also show low variability with a maximum difference between the upper and lower values of 2.7 \AA noticed for the D405-Y504 pair of residues.

3.1 Normal Mode Analysis

Previous work (Bahar et al., 2010; Bende, 2010) showed all-atoms elastic network normal mode analysis to be successful in describing the collective dynamics of a wide range of biomolecular systems. We therefore analyzed the ensemble of experimental RBD structures to verify the extent of conformational remodelling that can be adopted and whether it can lead to a better configuration of the RGD atoms in order to be able to interact with integrins. We performed a PCA on the

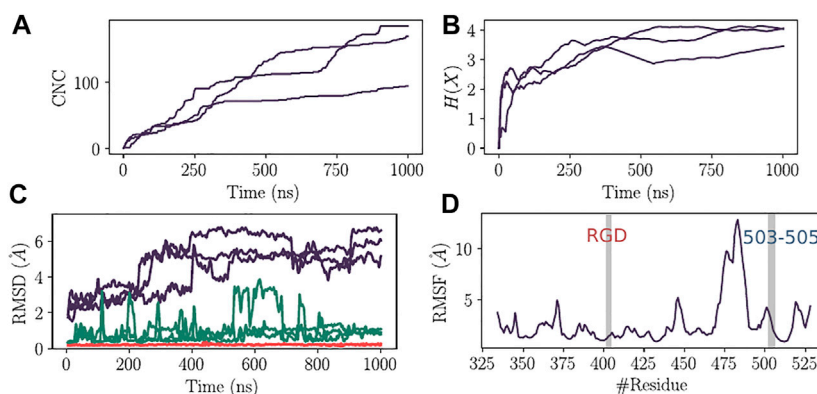


FIGURE 2 | Convergence analysis of aMD and structural deviation of RBD. **(A)** Cumulative number of clusters as a function of time for the three replicates of aMD trajectories. **(B)** Evolution of the Shannon's entropy ($H(X)$) for the three replicates of aMD trajectories. **(C)** Root Mean Square Deviation of RBD structure (Purple), C1 cluster of residues (Green) and the RGD motif (Red). **(D)** Root Mean Square Fluctuation of RBD residues calculated for the C_{α} atoms from the combined aMD trajectories.

pre-aligned and superimposed ensemble of structures. Data along PC1 and PC2 are relatively clustered in the lower right corner of **Figure 1D**, except for a few structures that showed the highest values of PC2 or the lowest values of PC1. Particularly for these structures, this might indicate divergent structural properties compared to the other members of the dataset. To proceed with a quantitative and more objective comparison, we calculated the RMSIP to assess the degree of overlap of the normal modes between the members of the constructed ensemble as proposed in related work (Yao et al., 2016). A score of 0.70 is considered a good correspondence, while a score of 0.50 is considered fair (Amadei et al., 1999). We found that the RMSIP values are ranged from 0.86 to 1 (**Figure 1E**) which shows a high level of similarity and agrees with the results from the PCA calculated from the normal modes. We also evaluated the structural deformation adopted by the RBD in terms of Root Mean Square Fluctuation (RMSF) calculated from the projection of the normal modes (**Figure 1F**). The structures of the ensemble show an overall similar profile of residue fluctuations in almost all except for some, where increasing flexibility by the amino-acids of the RBM segment is noticed. Furthermore, we noticed that the lateral chains of segment 503-505 residues (we refer to this cluster of residues as C1) are the closest residues from RBM that interact with the RGD segment. This was also detected from the distance calculation shown in **Figure 1C**. We thought that these residues are critical in controlling the conformational properties of the RGD segment. However, The RMSF profile revealed limited flexibility for both RGD motif and 503-505 segment showing a maximum displacement of 0.2 Å.

3.2 Accelerated Molecular Dynamics Shows Local Flexibility Mainly in the Receptor Binding Motif Segment but Not in RGD Microenvironment

Three independent aMD simulations were conducted for a total simulation time of 3 μ s. This allows for efficient sampling of the

energy landscape for SARS-CoV-2 RBD. The utility of aMD has been previously shown in many macromolecular systems including G-protein coupled receptors, bovine pancreatic trypsin inhibitor, and α -1-Antitrypsin (Duan et al., 2019). The main goal of this analysis was to identify the most populated conformations that the RBD can take to exert its function of interacting with the host receptors. In the event that the virus binds to integrins via the RGD motif, we would be able to detect a conformational state adapted for such interaction within the set of the sampled aMD snapshots. First, to assess the convergence of the different independent simulations, we calculated the cumulative number of the detected conformational clusters as well as the evolution of Shannon's entropy (**Figures 2A,B**). We found that, except for one run, all the trajectories show adequate convergence starting from 300 ns in terms of CNCs. The entropy value also converged for all the replicates around 300 ns (**Figure 2B**). The coverage of the conformational landscape for RBD was therefore reasonable in the context of our research question. We then verified the conformational drift from the initial structure of RBD for the total C_{α} atoms, the C_{α} of the RGD segment, and those of both RGD motif and the C1 cluster that harbors the Y504 residue (**Figure 2C**). The latter was included given its proximity to RGD as well as the presumed role that it may play to control the structural properties of the motif. Based on all residue Root Mean Square Deviation (RMSD) values, that can exceed 6 Å, RBD might adopt a significant conformational arrangement. However, the RGD motif does not seem to share this property as the range of RMSDs is less than 0.5 Å. In addition, the C1 residues also did not show a large conformational drift compared to the crystal structure since the corresponding RMSD values are mostly below 2.5 Å. This indeed can also be seen from the RMSF profile of the C_{α} atoms of RBD (**Figure 2D**). The region that shows the largest flexibility corresponds roughly to the RBM residues. The RGD motif shows RMSF values of less than 2 Å while the loop 503-505 has a maximum value of 3.1 Å.

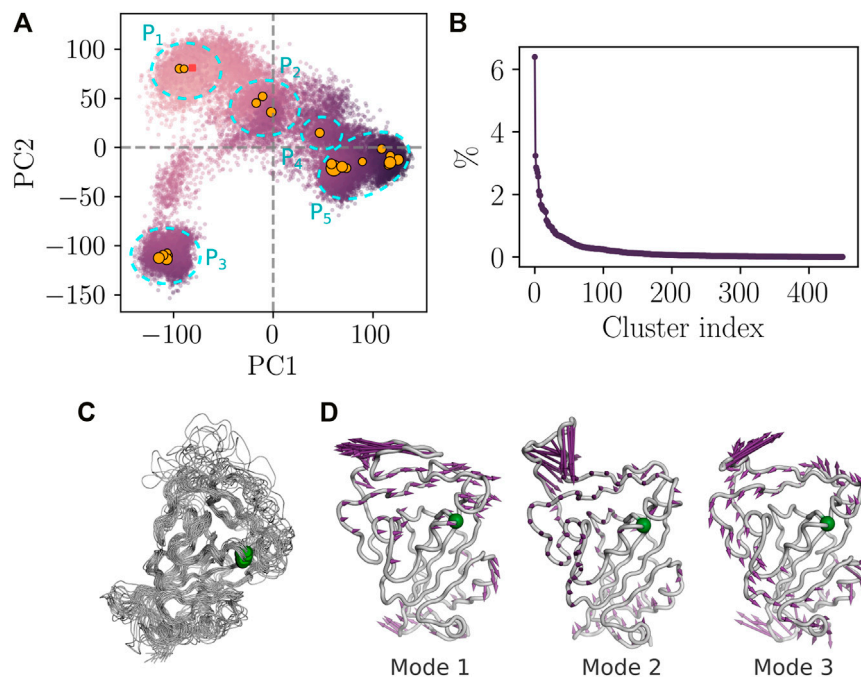


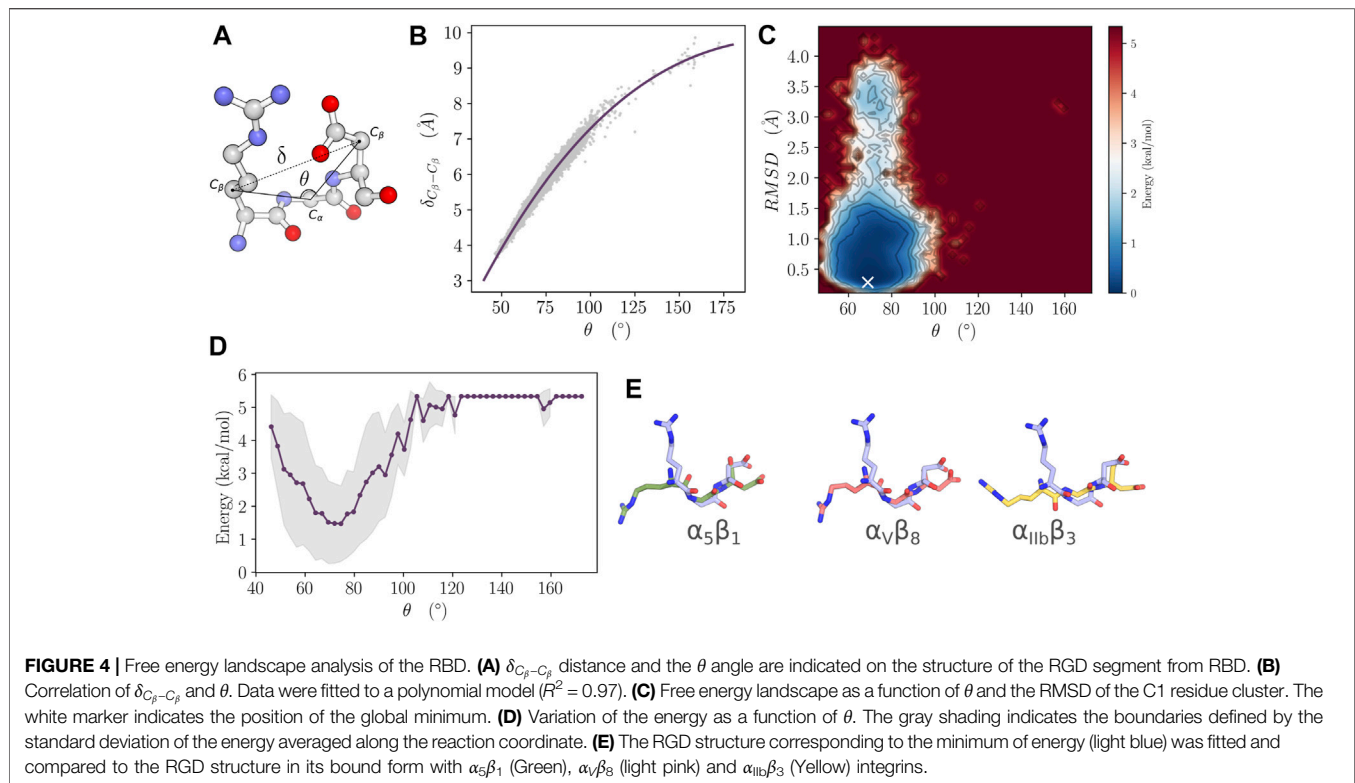
FIGURE 3 | Essential dynamics of RBD from aMD simulation. **(A)** PCA analysis from the combined replicas. The color of the dots varies as a function of the structural deviation (RMSD) to the crystal structure of RBD. i.e., light purple color indicates lesser deviation and dark purple indicates higher values of RMSD. The square point corresponds to the projection of the crystal structure onto the first and the second subspaces. Orange circles correspond to the centroids of the highly populated clusters and the size of the circles is proportional to the occupancy of the cluster. **(B)** Occupancy of RBD structural clusters. **(C)** Structural alignment of the highly populated clusters (occupancy >1%). Green spheres indicate the position of the RGD motif. **(D)** Porcupine plot corresponding to projections of C_{α} atoms onto the first three non-rotational and non-translational normal modes.

3.3 Principal Component Analysis and Clustering Analysis Show No Major Conformational Change in RGD and Its Microenvironment

We have conducted a principal component analysis using the total set of conformations from the three combined independent trajectories. The protein heavy-atom coordinates were projected onto the subspaces defined by the first and the second components. The aMD simulation was capable of capturing different states of the RBD. We noticed that the structure drifted considerably from the initial crystal structure (red rectangle in **Figure 3A**), thus demonstrating the convenient sampling of the RBD phase space that allows ascending the energy barriers. Clustering analysis focused on the clusters showing more than 1% of occupancy. Twenty one major clusters were detected of which the highest-ranked member shows the occupancy of 6.3% (**Figure 3B**).

Essentially, the PCA plot can be subdivided into five different partitions according to the density of the major conformational clusters (**Figure 3A**). P1 partition consists of the structures that are close to the bound conformation of RBD. Partitions P2 and P4 correspond to transition states with lower occupancies compared to the other partitions. P3 and P5 correspond to highly populated

partitions where the density of the projected atom coordinates is high as shown from the large number of major clusters agglomerated together in the PCA plot. Highly populated partitions, i.e. P1, P3, and P5, may describe the three relevant discrete functional states of RBD corresponding to the bound, up and down states (Henderson et al., 2020). However, we were unable to verify this, given that the experimental structure of these states lack the atomic details in some RBD segment regions and those at close proximity to subdomain-1 of the spike protein. Nevertheless, the free energy landscape based on PC1 and PC2, as reaction coordinates established after correcting for the biased sampling of aMD, shows indeed that P2, P3, and P5 correspond to minimum energy wells on the one hand and confirms that P2 and P4 partitions describe transition states on the other hand (**Supplementary Data S2**). Superposition of the representative structures of the highly populated clusters revealed a rigid core of the RBD that harbors the RGD motif of low flexibility (**Figure 3C**). Porcupine plots, depicting the direction and the amplitude of motion across the three non-rotational and non-translational normal modes, also highlight the location of the RGD motif within a rigid core of the RBD, characterized by a low amplitude displacement vector (**Figure 3D**). Moreover, the RGD motif is rigid in modes 2 and 3, while it moves in the same direction of the segment 503-505 in mode 1.



3.4 Favorable Geometrical Features for the Interaction Between RGD and Integrins Are Not Sampled in the Receptor Binding Domain Ensemble

Previous research using RGD peptide analogs suggested that extended conformation, spanning the atoms of the aliphatic side chain of Arg and Asp residues as well as the atoms of the main chain of RGD, has to take place to be capable of interacting with integrins (Civera et al., 2017; Kapp et al., 2017). Moreover, the distance between the C_β atoms of Arg and Asp must be within a range of 7 Å to 9 Å. To examine if these properties occurred during aMD simulation, we calculated the angle described by the C_β , C_α , C_β of R403, G404, and D405 residues, respectively, allowing to assess the level of extension (Figure 4A). We also calculated the distance between the C_β atoms of R403 and D405. $\delta_{C\beta-C\beta}$ and θ describe a wide range of values of 3.6 Å to 9.8 Å and 46° to 172°, respectively (Figure 4B). However, the data are skewed towards the lower end of the value ranges. Roughly, θ has more density in the 46° to 110° range, while the proportion of $\delta_{C\beta-C\beta}$ is ranging in higher values of 3.8 Å to 7.7 Å. A strong correlation was also noticed between $\delta_{C\beta-C\beta}$ and θ with an R^2 value of 0.97 when we fitted the data to a polynomial model. Therefore, we choose the θ angle and the RMSD of the C1 cluster of residues as reaction coordinates (Figure 4C). The FEL has a single highly populated minimum where the values of θ roughly span a range of 58° to 83° while the RMSD is low and does not exceed 1.5 Å. Averaging the energy over the binned values of θ shows a depth in the energy well of around 3 kcal/mol

(Figure 4D). It also reveals that the more extended θ is in the less favorable energy. Indeed the conformation with the lowest energy value shows a significant divergence compared to the states of the RGD motif in its bound form with $\alpha_5\beta_1$, $\alpha_{IIb}\beta_3$ and $\alpha_V\beta_8$ integrins (Figure 4E). θ and $\delta_{C\beta-C\beta}$ for the lowest energy conformation were measured to 67° and 5.4 Å, respectively. The RGD motif however, clearly adopts an extended conformation in its bound form as revealed by θ values of 146°, 173° and 145° and $\delta_{C\beta-C\beta}$ values of 8.9, 9.6 and 8.9 Å for $\alpha_5\beta_1$, $\alpha_{IIb}\beta_3$ and $\alpha_V\beta_8$ respectively.

3.5 Protein-Protein Docking Shows the Inability of RGD Motif to Interact With Integrins

We used 22 structures of the highly populated cluster centers obtained from the molecular dynamics simulation to conduct a protein-protein docking. The analysis was conducted by restraining the sampling space to include the RGD motif of RBD and the native binding site on $\alpha_5\beta_1$, $\alpha_{IIb}\beta_3$, and $\alpha_V\beta_8$ integrins (Figure 5). These integrins have been chosen mainly for their high-quality crystal structures in a bound state with an RGD motif. Of note, the homology relationship with RGD-binding integrins expressed in airway epithelial cells; namely $\alpha_V\beta_5$, $\alpha_V\beta_6$, and $\alpha_V\beta_8$, is confirmed, implying a conserved 3D fold. Moreover, $\alpha_{IIb}\beta_3$ was included to assess the putative binding of SARS-CoV-2 to platelets as suggested by previous studies (Koupenova and Freedman, 2020; Zaid et al., 2020; Zhang et al., 2020). Our results show that RBD has not been able to

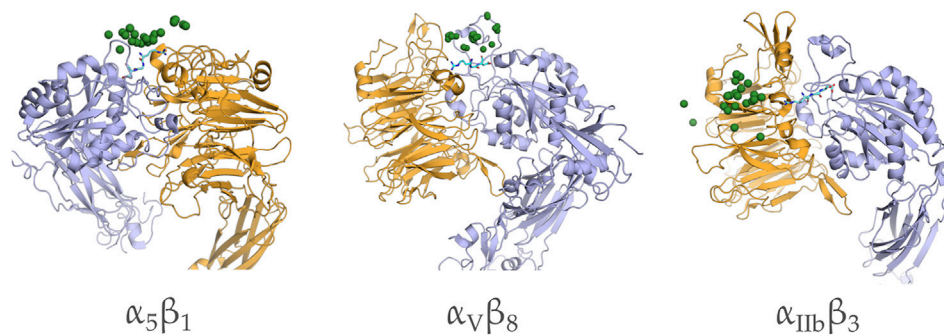


FIGURE 5 | Distribution of the candidate complexes of RBD docked to $\alpha_5\beta_1$, $\alpha_{IIb}\beta_3$, and $\alpha_V\beta_8$ integrins. The positions of G405 of the RBD motif are shown in green spheres and the native bound configuration of RBD from the crystal structure is shown in cyan sticks.

interact favorably with any of the studied integrins. Indeed, RGD motif was not capable of reaching its native binding site in any given structural state.

4 DISCUSSION

The optimal interaction of the RGD motif with integrin involves the establishment of a minimal set of contacts with the MIDAS interaction site and the nearby amino acid residues. Experimental structures of RGD in the bound form with integrins show that the motif is laid extensively, crossing the interface cleft between the alpha and beta integrin subunits. The carboxylic and guanidine groups of RGD act as electrostatic clamps with the MIDAS site and the acidic residues of the alpha subunit respectively. However, when we superposed the RGD motif from the RBD domain of SARS-CoV-2 with its corresponding sequence on the cilengitide molecule co-crystallized with the integrin (data not shown), we found that severe clashes persist in this mode of interaction. Following this observation, we hypothesized that the RBD must undergo structural adaptation to allow for the favorable interaction with integrins.

The RMSIP distribution demonstrated that the conformational space sampled from the analysis of all the experimental structures are relatively homogeneous, given the observed low variance in the data. Therefore, it is expected that the normal mode properties are linked directly to the conformational behaviour of the RBD. Both normal mode analysis and molecular dynamics simulation are supportive of the relative rigidity of the RGD motif, compared to the RBM amino acids. Therefore, the motif is highly unlikely to undergo a significant structural rearrangement to increase its exposure to the solvent and allow the interaction with integrins. The RGD motif in the structure of different disintegrins, like triflavin, schistatin, echistatin, decorsin and salmosin is located at the tip of a hairpin-like structure that allows an easy fitting with the integration head cleft without steric hindrance (Matsui et al., 2010). The same type of structure was observed in $\alpha_V\beta_6$ integrin interacting with the capsid protein VP1 of the foot-and-mouth disease virus (Kotecha et al., 2017). In the case of SARS-CoV-2 RBD, the RGD motif did not show any structural similarities with

disintegrins, and the steric hindrance imposed by the segments close to the motif, seems to be maintained in all the functionally relevant conformational states.

Microscale aMD allowed for an extensive sampling of the conformational phase of RBD where we have detected three highly populated states that could correspond to the bound, up and down configurations of the domain. However, potential integrin-binding conformations were not detected. The free energy landscape also confirmed that the geometrical features of the RGD binding to integrins are unfavorable. Moreover, protein-protein docking showed the inability of all the highly populated conformations to reach the depth of the interaction site of integrins where the electrostatic clamping and the interaction with MIDAS must happen to maintain a stable association.

Most of the former works have relied on sequence conservation and motif detection analysis to conclude on the implication of RGD motif in SARS-CoV-2 RBD in the interaction with integrins (Luan et al., 2020; Sigrist et al., 2020; Carvacho and Piesche, 2021; Dakal, 2021). However, few of them have considered the structural features to reinforce or confirm the hypothesis with details, as presented in this study. Indeed, Sigrist et al. (2020) and Luan et al. (2020), stated the solvent exposure of RGD as the single argument supporting its involvement in integrin binding, but they did not consider the geometrical features of the motif that must be fulfilled nor the steric hindrance that can be imposed by the surrounding segments. Mészáros et al. (2021) and Makowski et al. (2021) proposed that the surrounding residues of RGD are flexible and, therefore, allow the interaction with integrins. Nevertheless, our results from molecular dynamics simulation and normal mode analysis are congruent in showing that the level of plasticity of these segments is not sufficient to eliminate sterical hindrance that prevents the association with integrins. Moreover, we were not able to detect any hairpin-like structure as observed in disintegrins and VP1 protein of the foot-and-mouth disease virus, despite the extensive sampling of the conformational space. Computational analysis by Dakal (2021) concluded that the RGD could bind favorably to $\alpha_5\beta_1$ and $\alpha_5\beta_6$ integrins. However, in his study the author used only the β -propeller head of the alpha subunit for the protein-protein docking, which is not adequate to infer physiological

binding properties. On the other hand, Beddingfield et al. (2021) showed that the protein-protein complex between integrins and S protein, obtained from docking, does not show a favorable fitting in the RGD binding site, which is in agreement with what we have observed from the constrained protein-protein docking analysis.

Among integrins expressed in airway epithelial cells, and that could be potential SARS-CoV-2 receptor, $\alpha_2\beta_1$, a collagen and laminin receptor, plays a critical role in cell infection by echovirus (Eisner, 1992). The $\alpha_2\beta_1$ integrin is known to be a non-RGD binding receptor, and therefore, it is unlikely that it binds to the 403–405 segment of RBD. The second receptor $\alpha_V\beta_5$ is well known to be an adenovirus receptor (Wickham et al., 1994), but is not expressed on the luminal surface (Grubb et al., 1994) which makes it difficult to be involved in the infection by coronavirus. $\alpha_V\beta_6$, an RGD receptor, was described to be implicated in infection by foot and mouth disease virus (Jackson et al., 2000). $\alpha_V\beta_6$ is the only one known to be expressed on the mucosal epithelial cells that are the primary site of infection by respiratory viruses (Sheppard, 2003). However, studies using developed antibodies show that $\alpha_V\beta_6$ is poorly expressed in lung epithelium cells and is constitutively expressed at low levels in uninjured epithelia (Breuss et al., 1995; Weinacker et al., 1995). Furthermore, the expression pattern of RGD-binding integrins is very differentiated between healthy and unhealthy pulmonary cells. Indeed, many integrins are not seen on healthy adult airway epithelium cells especially $\alpha_5\beta_1$ and $\alpha_9\beta_1$ (Pilewski et al., 1997; Sheppard, 2003). On the other hand, the other expressed RGD dependent integrins have a distinct functional, spatial and chronological expression (Pilewski et al., 1997). $\alpha_V\beta_5$, $\alpha_V\beta_6$ and $\alpha_V\beta_8$ are constitutively expressed at low levels on healthy lung cells (Breuss et al., 1995; Sheppard, 2003), recognize many ligands that are not expressed on healthy epithelial basement membranes, and are only involved in cases of lung inflammation and injury (Nishimura et al., 1994; Nishimura et al., 1994; Mu et al., 2002). Nader et al. (2021) have conducted experiments to assess the binding of SARS-CoV-2 spike protein to integrin $\alpha_V\beta_3$, to Human Aortic Endothelial Cell or to Caco-2 endothelial cells. Their result shows indeed a direct interaction with the integrin. However, a competition assay with Cilengitide, an RGD binding peptide, was only conducted for the cell binding assay. It is therefore difficult to assert whether the observed effect in their work is the result of the direct interaction of the RBD-RGD motif with $\alpha_V\beta_3$ or if it is the outcome of a modulation effect. It is worthy to note, that the Cilengitide can induce the downregulation of the *ITGAV* gene which encodes the α_V subunit (Wang et al., 2014). On another note, Schimmel et al. demonstrated that primary endothelial cells can not be infected with SARS-CoV 2 *in vivo* nor *in vitro* (Schimmel et al., 2021). Moreover, another study has concluded that the incubation with the integrin inhibitor ATN-161, had no effect on the infection capacity of SARS-CoV-2 with Caco-2 endothelial cells (Zech et al., 2021). All these studies are not contradicting our results. In fact, we are not excluding integrins as putative receptors for SARS-CoV-2. We, however, postulate that RGD from RBD is unlikely to be the interacting motif with integrin. This implies that other motifs could be involved in such interaction. Our claim, is also sustained with the recent study by Beaudoin et al. (2021). All this information, consolidated by our above-cited results,

emphasize the need for more evidence to confirm the role of integrins in the physiopathology of SARS-CoV-2.

5 CONCLUSION

Based on the evidence provided in this paper, we suggest that the RGD motif from the RBD of SARS-CoV-2 is unlikely to interact with integrins. That, however, does not imply that integrins are not host receptors for the virus. Thus, in light of our results, as well as previous works, the potential interaction of the RGD motif from the RBD of SARS-CoV-2 with integrins should be revised extensively. Consequently, potential involvement of other segments belonging to the spike protein, is more likely to take place if integrins are confirmed to be host receptors for SARS-CoV-2.

DATA AVAILABILITY STATEMENT

Molecular dynamics trajectories, raw data and the code used to make the analysis and figures of this paper are available online from the Zenodo repository “Dry trajectories of SARS-CoV-2 RBD from accelerated molecular dynamics simulation” <https://doi.org/10.5281/zenodo.5775514>.

AUTHOR CONTRIBUTIONS

HO and NS-A coordinated and designed the study and co-led the writing of the paper. HB-Ma performed the normal mode analysis study. OK performed the protein-protein docking. HB-Ma, KG, AB, FT, IA, and YS provided the critical analysis of the paper and contributed to writing. All authors contributed to writing the manuscript, and read and approved the final manuscript.

FUNDING

The present work was partially supported by the European project PHINDaccess: Strengthening Omics data analysis capacities in pathogen-host interaction (Grant Agreement ID: 811034).

ACKNOWLEDGMENTS

The authors acknowledge the Centre for High Performance Computing (CHPC), South Africa, for providing computational resources to this research project. We thank Fatma Zahra Guerfali from Institut Pasteur de Tunis for providing critical reviews of the paper.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2022.834857/full#supplementary-material>

REFERENCES

- Aguiar, J. A., Tremblay, B. J., Mansfield, M. J., Woody, O., Lobb, B., Banerjee, A., et al. (2020). Gene Expression and *In Situ* Protein Profiling of Candidate SARS-CoV-2 Receptors in Human Airway Epithelial Cells and Lung Tissue. *Eur. Respir. J.* 56, 2001123. doi:10.1183/13993003.01123-2020
- Amadei, A., Ceruso, M. A., and Di Nola, A. (1999). On the Convergence of the Conformational Coordinates Basis Set Obtained by the Essential Dynamics Analysis of Proteins' Molecular Dynamics Simulations. *Proteins* 36, 419–424. doi:10.1002/(sici)1097-0134(19990901)36:4<419:aid-prot5>3.0.co;2-u
- Assumpcao, T. C. F., Ribeiro, J. M. C., and Francischetti, I. M. B. (2012). Disintegrins from Hematophagous Sources. *Toxins* 4, 296–322. doi:10.3390/toxins4050296
- Bahar, I., Lezon, T. R., Bakan, A., and Shrivastava, I. H. (2010). Normal Mode Analysis of Biomolecular Structures: Functional Mechanisms of Membrane Proteins. *Chem. Rev.* 110, 1463–1497. doi:10.1021/cr900095e
- Bazaa, A., Juárez, P., Marrakchi, N., Lasfer, Z. B., Ayeb, M. E., Harrison, R. A., et al. (2007). Loss of Introns along the Evolutionary Diversification Pathway of Snake Venom Disintegrins Evidenced by Sequence Analysis of Genomic DNA from *Macrovipera lebetina* Transmediterranea and *Echis Ocellatus*. *J. Mol. Evol.* 64, 261–271. doi:10.1007/s00239-006-0161-4
- Bazan-Socha, S., Bukiej, A., Marcinkiewicz, C., and Musial, J. (2005). Integrins in Pulmonary Inflammatory Diseases. *Cpd* 11, 893–901. doi:10.2174/1381612053381710
- Beaudoin, C. A., Hamaia, S. W., Huang, C. L.-H., Blundell, T. L., and Jackson, A. P. (2021). Can the SARS-CoV-2 Spike Protein Bind Integrins Independent of the RGD Sequence? *Front. Cel. Infect. Microbiol.* 11, 765300. doi:10.3389/fcimb.2021.765300
- Beddingfield, B. J., Iwanaga, N., Chapagain, P. P., Zheng, W., Roy, C. J., Hu, T. Y., et al. (2021). The Integrin Binding Peptide, ATN-161, as a Novel Therapy for SARS-CoV-2 Infection. *JACC: Basic Translational Sci.* 6, 1–8. doi:10.1016/j.jacbs.2020.10.003
- Ben-Mabrouk, H., Zouari-Kessentini, R., Montassar, F., Koubaa, Z. A., Messaadi, E., Guillonneau, X., et al. (2016). CC5 and CC8, Two Homologous Disintegrins from Cerastes Cerastes Venom, Inhibit *In Vitro* and *Ex Vivo* Angiogenesis. *Int. J. Biol. Macromolecules* 86, 670–680. doi:10.1016/j.ijbiomac.2016.02.008
- Bende, A. (2010). Hydrogen Bonding in the Urea Dimers and Adenine-Thymine DNA Base Pair: Anharmonic Effects in the Intermolecular H-Bond and Intramolecular H-Stretching Vibrations. *Theor. Chem. Acc.* 125, 253–268. doi:10.1007/s00214-009-0645-6
- Berry, J. D., Jones, S., Drebrot, M. A., Andonov, A., Sabara, M., Yuan, X. Y., et al. (2004). Development and Characterisation of Neutralising Monoclonal Antibody to the Sars-Coronavirus. *J. Virol. Methods* 120, 87–96. doi:10.1016/j.jviromet.2004.04.009
- Bruss, J. M., Gallo, J., DeLisser, H. M., Klimanskaya, I. V., Folkesson, H. G., Pittet, J. F., et al. (1995). Expression of the Beta 6 Integrin Subunit in Development, Neoplasia and Tissue Repair Suggests a Role in Epithelial Remodeling. *J. Cel Sci* 108 (Pt 6), 2241–2251. doi:10.1242/jcs.108.6.2241
- Cambier, S., Mu, D. Z., O'Connell, D., Boylen, K., Travis, W., Liu, W. H., et al. (2000). A Role for the Integrin Alpha5beta1 in the Negative Regulation of Epithelial Cell Growth. *Cancer Res.* 60, 7084–7093.
- Cantuti-Castelvetri, L., Ojha, R., Pedro, L. D., Djannatian, M., Franz, J., Kuivanen, S., et al. (2020). Neuropilin-1 Facilitates Sars-Cov-2 Cell Entry and Provides a Possible Pathway into the central Nervous System. *bioRxiv*. doi:10.1101/2020.06.07.137802
- Carvacho, I., and Piesche, M. (2021). RGD-binding Integrins and TGF- β in SARS-CoV-2 Infections - Novel Targets to Treat COVID-19 Patients? *Clin. Transl Immunol.* 10, e1240. doi:10.1002/cti2.1240
- Case, D. A., Ben-Shalom, I., Brozell, S., Cerutti, D., Cheatham, T. E., Cruzeiro, V., et al. (2018). *AMBER 2018*. San Francisco: University of California.
- Chu, H., Chan, J. F.-W., Yuen, T. T.-T., Shuai, H., Yuan, S., Wang, Y., et al. (2020). Comparative Tropism, Replication Kinetics, and Cell Damage Profiling of SARS-CoV-2 and SARS-CoV with Implications for Clinical Manifestations, Transmissibility, and Laboratory Studies of COVID-19: an Observational Study. *The Lancet Microbe* 1, e14–e23. doi:10.1016/s2666-5247(20)30004-5
- Civera, M., Arosio, D., Bonato, F., Manzoni, L., Pignataro, L., Zanella, S., et al. (2017). Investigating the Interaction of Cyclic RGD Peptidomimetics with $\alpha V\beta 6$ Integrin by Biochemical and Molecular Docking Studies. *Cancers (Basel)* 9, 128. doi:10.3390/cancers9100128
- Clausen, T. M., Sandoval, D. R., Spliid, C. B., Pihl, J., Painter, C. D., Thacker, B. E., et al. (2020). SARS-CoV-2 Infection Depends on Cellular Heparan Sulfate and ACE2. *Cell* 183, 1043. doi:10.1016/j.cell.2020.09.033
- Dakal, T. C. (2021). SARS-CoV-2 Attachment to Host Cells Is Possibly Mediated via RGD-Integrin Interaction in a Calcium-dependent Manner and Suggests Pulmonary EDTA Chelation Therapy as a Novel Treatment for COVID 19. *Immunobiology* 226, 152021. doi:10.1016/j.imbio.2020.152021
- David, C. C., and Jacobs, D. J. (2014). Principal Component Analysis: a Method for Determining the Essential Dynamics of Proteins. *Methods Mol. Biol.* 1084, 193–226. doi:10.1007/978-1-62703-658-0_11
- Duan, L., Guo, X., Cong, Y., Feng, G., Li, Y., and Zhang, J. Z. H. (2019). Accelerated Molecular Dynamics Simulation for Helical Proteins Folding in Explicit Water. *Front. Chem.* 7, 540. doi:10.3389/fchem.2019.00540
- Eisner, R. (1992). Finding Out How a Viral Hitchhiker Snags a Ride. *Science* 255, 1647. doi:10.1126/science.1553554
- Fuglebak, E., Echave, J., and Reuter, N. (2012). Measuring and Comparing Structural Fluctuation Patterns in Large Protein Datasets. *Bioinformatics* 28, 2431–2440. doi:10.1093/bioinformatics/bts445
- Gasmi, A., Srairi, N., Guerhazi, S., Dkhil, H., Karoui, H., El Ayeb, M., et al. (2001). Amino Acid Structure and Characterization of a Heterodimeric Disintegrin from *Vipera Lebetina* Venom. *Biochim. Biophys. Acta (Bba) - Protein Struct. Mol. Enzymol.* 1547, 51–56. doi:10.1016/s0167-4838(01)00168-6
- Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., et al. (2020). A SARS-CoV-2 Protein Interaction Map Reveals Targets for Drug Repurposing. *Nature* 583, 459–468. doi:10.1038/s41586-020-2286-9
- Grubb, B. R., Pickles, R. J., Ye, H., Yankaskas, J. R., Vick, R. N., Engelhardt, J. F., et al. (1994). Inefficient Gene Transfer by Adenovirus Vector to Cystic Fibrosis Airway Epithelia of Mice and Humans. *Nature* 371, 802–806. doi:10.1038/371802a0
- Hamelberg, D., Mongan, J., and McCammon, J. A. (2004). Accelerated Molecular Dynamics: a Promising and Efficient Simulation Method for Biomolecules. *J. Chem. Phys.* 120, 11919–11929. doi:10.1063/1.1755656
- Hamidi, H., Pietilä, M., and Ivaska, J. (2016). The Complexity of Integrins in Cancer and New Scopes for Therapeutic Targeting. *Br. J. Cancer* 115, 1017–1023. doi:10.1038/bjc.2016.312
- Harrison, A. G., Lin, T., and Wang, P. (2020). Mechanisms of SARS-CoV-2 Transmission and Pathogenesis. *Trends Immunol.* 41, 1100–1115. doi:10.1016/j.it.2020.10.004
- Henderson, R., Edwards, R. J., Mansouri, K., Janowska, K., Stalls, V., Gobeil, S. M. C., et al. (2020). Controlling the SARS-CoV-2 Spike Glycoprotein Conformation. *Nat. Struct. Mol. Biol.* 27, 925–933. doi:10.1038/s41594-020-0479-4
- Ibrahim, I. M., Abdelmalek, D. H., Elshahat, M. E., and Elfiky, A. A. (2020). COVID-19 Spike-Host Cell Receptor GRP78 Binding Site Prediction. *J. Infect.* 80, 554–562. doi:10.1016/j.jinf.2020.02.026
- Isberg, R. R., and Tran Van Nhieu, G. (1994). Binding and Internalization of Microorganisms by Integrin Receptors. *Trends Microbiol.* 2, 10–14. doi:10.1016/0966-842x(94)90338-7
- Jackson, T., Sheppard, D., Denyer, M., Blakemore, W., and King, A. M. Q. (2000). The Epithelial Integrin $\alpha V\beta 6$ Is a Receptor for Foot-And-Mouth Disease Virus. *J. Virol.* 74, 4949–4956. doi:10.1128/jvi.74.11.4949-4956.2000
- Kapp, T. G., Rechenmacher, F., Neubauer, S., Maltsev, O. V., Cavalcanti-Adam, E. A., Zarka, R., et al. (2017). A Comprehensive Evaluation of the Activity and Selectivity Profile of Ligands for RGD-Binding Integrins. *Sci. Rep.* 7, 39805. doi:10.1038/srep39805
- Kotecha, A., Wang, Q., Dong, X., Ilca, S. L., Ondiviela, M., Zihe, R., et al. (2017). Rules of Engagement between $\alpha V\beta 6$ Integrin and Foot-And-Mouth Disease Virus. *Nat. Commun.* 8, 15408. doi:10.1038/ncomms15408
- Koupenova, M., and Freedman, J. E. (2020). Platelets and COVID-19. *Circ. Res.* 127, 1419–1421. doi:10.1161/circresaha.120.318218
- Li, S., Li, S., Disoma, C., Zheng, R., Zhou, M., Razzaq, A., et al. (2021). Sars-cov-2: Mechanism of Infection and Emerging Technologies for Future Prospects. *Rev. Med. Virol.* 31, e2168. doi:10.1002/rmv.2168
- Luan, J., Lu, Y., Gao, S., and Zhang, L. (2020). A Potential Inhibitory Role for Integrin in the Receptor Targeting of SARS-CoV-2. *J. Infect.* 81, 318–356. doi:10.1016/j.jinf.2020.03.046

- Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theor. Comput.* 11, 3696–3713. doi:10.1021/acs.jctc.5b00255
- Makowski, L., Olson-Sidford, W., and W-Weisel, J. (2021). Biological and Clinical Consequences of Integrin Binding via a Rogue RGD Motif in the SARS CoV-2 Spike Protein. *Viruses* 13, 146. doi:10.3390/v13020146
- Matsui, T., Hamako, J., and Titani, K. (2010). Structure and Function of Snake Venom Proteins Affecting Platelet Plug Formation. *Toxins (Basel)* 2, 10–23. doi:10.3390/toxins2010010
- Mészáros, B., Sámáno-Sánchez, H., Alvarado-Valverde, J., Čalyševa, J., Martínez-Pérez, E., Alves, R., et al. (2021). Short Linear Motif Candidates in the Cell Entry System Used by SARS-CoV-2 and Their Potential Therapeutic Implications. *Sci. Signal.* 14, abd0334. doi:10.1126/scisignal.abd0334
- Mu, D., Cambier, S., Fjellbirkeland, L., Baron, J. L., Munger, J. S., Kawakatsu, H., et al. (2002). The Integrin $\alpha\beta 8$ Mediates Epithelial Homeostasis through MT1-MMP-dependent Activation of TGF- $\beta 1$. *J. Cell Biol* 157, 493–507. doi:10.1083/jcb.200109100
- Nader, D., Fletcher, N., Curley, G. F., and Kerrigan, S. W. (2021). SARS-CoV-2 Uses Major Endothelial Integrin $\alpha\beta 3$ to Cause Vascular Dysregulation *In Vitro* during COVID-19. *PLoS One* 16, e0253347. doi:10.1371/journal.pone.0253347
- Nishimura, S. L., Sheppard, D., and Pytela, R. (1994). Integrin Alpha V Beta 8. Interaction with Vitronectin and Functional Divergence of the Beta 8 Cytoplasmic Domain. *J. Biol. Chem.* 269, 28708–28715. doi:10.1016/s0021-9258(19)61963-0
- Olfa, K.-Z., José, L., Salma, D., Amine, B., Najet, S. A., Nicolas, A., et al. (2005). Lebestatin, a Disintegrin from Macrovipera Venom, Inhibits Integrin-Mediated Cell Adhesion, Migration and Angiogenesis. *Lab. Invest.* 85, 1507–1516. doi:10.1038/labinvest.3700350
- Othman, H., Bouslama, Z., Brandenburg, J.-T., Da Rocha, J., Hamdi, Y., Ghedira, K., et al. (2020). Interaction of the Spike Protein Rbd from Sars-Cov-2 with Ace2: Similarity with Sars-Cov, Hot-Spot Analysis and Effect of the Receptor Polymorphism. *Biochem. biophysical Res. Commun.* 527, 702–708. doi:10.1016/j.bbrc.2020.05.028
- Pilewski, J. M., Latoche, J. D., Arcasoy, S. M., and Albelda, S. M. (1997). Expression of Integrin Cell Adhesion Receptors during Human Airway Epithelial Repair *In Vivo*. *Am. J. Physiology-Lung Cell Mol. Physiol.* 273, L256–L263. doi:10.1152/ajplung.1997.273.1.L256
- Qing, E., Hantak, M., Perlman, S., and Gallagher, T. (2020). Distinct Roles for Sialoside and Protein Receptors in Coronavirus Infection. *mBio* 11. doi:10.1128/mBio.02764-19
- Ravindra, N. G., Alfajaro, M. M., Gasque, V., Habet, V., Wei, J., Filler, R. B., et al. (2020). Single-cell Longitudinal Analysis of SARS-CoV-2 Infection in Human Bronchial Epithelial Cells. *bioRxiv*.
- Schimmel, L., Chew, K. Y., Stocks, C. J., Yordanov, T. E., Essebier, P., Kulasinghe, A., et al. (2021). Endothelial Cells Are Not Productively Infected by SARS-CoV-2. *Clin. Transl Immunol.* 10, e1350. doi:10.1002/cti.1350
- Shang, J., Wan, Y., Luo, C., Ye, G., Geng, Q., Auerbach, A., et al. (2020a). Cell Entry Mechanisms of SARS-CoV-2. *Proc. Natl. Acad. Sci. USA* 117, 11727–11734. doi:10.1073/pnas.2003138117
- Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., et al. (2020b). Structural Basis of Receptor Recognition by SARS-CoV-2. *Nature* 581, 221–224. doi:10.1038/s41586-020-2179-y
- Sheppard, D. (2003). Functions of Pulmonary Epithelial Integrins: from Development to Disease. *Physiol. Rev.* 83, 673–686. doi:10.1152/physrev.00033.2002
- Sigris, C. J., Bridge, A., and Le Mercier, P. (2020). A Potential Role for Integrins in Host Cell Entry by SARS-CoV-2. *Antiviral Res.* 177, 104759. doi:10.1016/j.antiviral.2020.104759
- van Zundert, G. C. P., Rodrigues, J. P. G. L. M., Trellet, M., Schmitz, C., Kastiris, P. L., Karaca, E., et al. (2016). The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* 428, 720–725. doi:10.1016/j.jmb.2015.09.014
- 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. *J. Virol.* 94, e00127. doi:10.1128/JVI.00127-20
- Wang, J. T., Liu, Y., Kan, X., Liu, M., and Lu, J. G. (2014). Cilengitide, a Small Molecule Antagonist, Targeted to Integrin αv Inhibits Proliferation and Induces Apoptosis of Laryngeal Cancer Cells *In Vitro*. *Eur. Arch. Otorhinolaryngol.* 271, 2233–2240. doi:10.1007/s00405-014-2918-5
- Wang, Y., Harrison, C. B., Schulten, K., and McCammon, J. A. (2011). Implementation of Accelerated Molecular Dynamics in NAMD. *Comput. Sci. Discov.* 4, 015002. doi:10.1088/1749-4699/4/1/015002
- Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S., and Crispin, M. (2020). Site-specific Analysis of the Sars-Cov-2 Glycan Shield. *BioRxiv*.
- Weinacker, A., Ferrando, R., Elliott, M., Hogg, J., Balmes, J., and Sheppard, D. (1995). Distribution of Integrins Alpha V Beta 6 and Alpha 9 Beta 1 and Their Known Ligands, Fibronectin and Tenascin, in Human Airways. *Am. J. Respir. Cell Mol Biol* 12, 547–556. doi:10.1165/ajrcmb.12.5.7537970
- Wickham, T. J., Filardo, E. J., Cheres, D. A., and Nemerow, G. R. (1994). Integrin Alpha V Beta 5 Selectively Promotes Adenovirus Mediated Cell Membrane Permeabilization. *J. Cell Biol* 127, 257–264. doi:10.1083/jcb.127.1.257
- Xiao, T., Takagi, J., Coller, B. S., Wang, J.-H., and Springer, T. A. (2004). Structural Basis for Allostery in Integrins and Binding to Fibrinogen-Mimetic Therapeutics. *Nature* 432, 59–67. doi:10.1038/nature02976
- Yan, S., Sun, H., Bu, X., and Wan, G. (2020). New Strategy for COVID-19: An Evolutionary Role for RGD Motif in SARS-CoV-2 and Potential Inhibitors for Virus Infection. *Front. Pharmacol.* 11, 912. doi:10.3389/fphar.2020.00912
- Yao, X.-Q., Skjærven, L., and Grant, B. J. (2016). Rapid Characterization of Allosteric Networks with Ensemble normal Mode Analysis. *J. Phys. Chem. B* 120, 8276–8288. doi:10.1021/acs.jpcc.6b01991
- Zaid, Y., Puhm, F., Allaey, I., Naya, A., Oudghiri, M., Khalki, L., et al. (2020). Platelets Can Associate with Sars-Cov-2 Rna and Are Hyperactivated in Covid-19. *Circ. Res.* 127, 1404–1418. doi:10.1161/circresaha.120.317703
- Zamorano Cuervo, N., and Grandvaux, N. (2020). ACE2: Evidence of Role as Entry Receptor for SARS-CoV-2 and Implications in Comorbidities. *Elife* 9, e61390. doi:10.7554/eLife.61390
- Zech, F., Schniertshauer, D., Jung, C., Herrmann, A., Cordsmeier, A., Xie, Q., et al. (2021). Spike Residue 403 Affects Binding of Coronavirus Spikes to Human ACE2. *Nat. Commun.* 12, 6855. doi:10.1038/s41467-021-27180-0
- Zhang, S., Liu, Y., Wang, X., Yang, L., Li, H., Wang, Y., et al. (2020). SARS-CoV-2 Binds Platelet ACE2 to Enhance Thrombosis in COVID-19. *J. Hematol. Oncol.* 13, 120. doi:10.1186/s13045-020-00954-7

Conflict of Interest: The authors declare 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.

Copyright © 2022 Othman, Messaoud, Khamessi, Ben-Mabrouk, Ghedira, Bharuthram, Treurnicht, Achilonu, Sayed and Srairi-Abid. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.



Computational Repurposing of Drugs and Natural Products Against SARS-CoV-2 Main Protease (M^{pro}) as Potential COVID-19 Therapies

Sakshi Piplani^{1,2†}, Puneet Singh^{1,2†}, Nikolai Petrovsky^{1,2*‡} and David A. Winkler^{3,4,5*‡}

OPEN ACCESS

Edited by:

James Leland Olds,
George Mason University,
United States

Reviewed by:

Yu Kang,
Zhejiang University, China
Debsindhu Bhowmik,
Oak Ridge National Laboratory (DOE),
United States
Lance Hellman,
Nevada State College, United States
Richard J. Bingham,
University of Huddersfield,
United Kingdom

*Correspondence:

Nikolai Petrovsky
nikolai.petrovsky@flinders.edu.au
David A. Winkler
d.winkler@latrobe.edu.au

[†]These authors share first authorship

[‡]These authors share senior authorship

Specialty section:

This article was submitted to
Biological Modeling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 22 September 2021

Accepted: 28 January 2022

Published: 14 March 2022

Citation:

Piplani S, Singh P, Petrovsky N and
Winkler DA (2022) Computational
Repurposing of Drugs and Natural
Products Against SARS-CoV-2 Main
Protease (M^{pro}) as Potential COVID-
19 Therapies.
Front. Mol. Biosci. 9:781039.
doi: 10.3389/fmolb.2022.781039

¹College of Medicine and Public Health, Flinders University, Bedford, SA, Australia, ²Vaxine Pty Ltd., Warradale, SA, Australia, ³Department of Biochemistry and Chemistry, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, VIC, Australia, ⁴Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, VIC, Australia, ⁵School of Pharmacy, University of Nottingham, Nottingham, United Kingdom

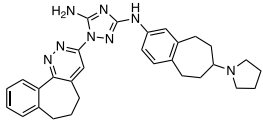
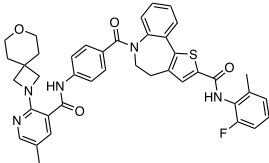
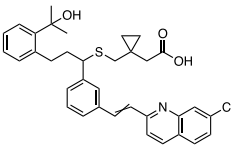
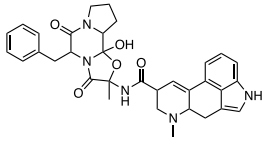
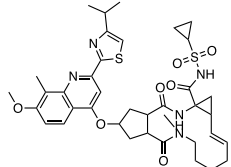
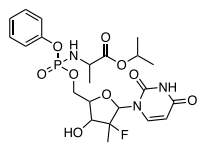
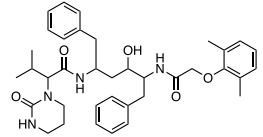
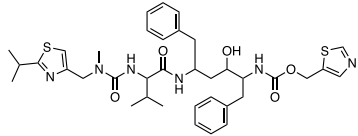
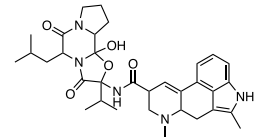
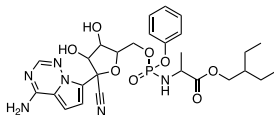
We urgently need to identify drugs to treat patients suffering from COVID-19 infection. Drugs rarely act at single molecular targets. Off-target effects are responsible for undesirable side effects and beneficial synergy between targets for specific illnesses. They have provided blockbuster drugs, e.g., Viagra for erectile dysfunction and Minoxidil for male pattern baldness. Existing drugs, those in clinical trials, and approved natural products constitute a rich resource of therapeutic agents that can be quickly repurposed, as they have already been assessed for safety in man. A key question is how to screen such compounds rapidly and efficiently for activity against new pandemic pathogens such as SARS-CoV-2. Here, we show how a fast and robust computational process can be used to screen large libraries of drugs and natural compounds to identify those that may inhibit the main protease of SARS-CoV-2. We show that the shortlist of 84 candidates with the strongest predicted binding affinities is highly enriched ($\geq 25\%$) in compounds experimentally validated *in vivo* or *in vitro* to have activity in SARS-CoV-2. The top candidates also include drugs and natural products not previously identified as having COVID-19 activity, thereby providing leads for experimental validation. This predictive *in silico* screening pipeline will be valuable for repurposing existing drugs and discovering new drug candidates against other medically important pathogens relevant to future pandemics.

Keywords: SARS-CoV-2, binding affinity, main protease, 3CL, computational chemistry, docking, molecular dynamics

INTRODUCTION

The devastating impact of the COVID-19 pandemic caused by SARS coronavirus-2 (SARS-CoV-2) has stimulated unprecedented international activity to discover effective drugs for this and other pathogenic coronaviruses such as SARS and MERS CoV (Ciotti et al., 2019; Zhang J. et al., 2020; Berkley, 2020; Zhang J.-J. et al., 2020; Zhang T. et al., 2020; Jiang, 2020; Lu, 2020; Mendes, 2020; Olsen et al., 2020; Rosa and Santos, 2020; Rosales-Mendoza et al., 2020; Sanders et al., 2020; Schlagenhauf et al., 2020; Sohrabi et al., 2020; Thanh Le et al., 2020; Whitworth, 2020; Yavuz and Unal, 2020). Computational methods are useful, fast approaches to determine the

TABLE 1 | Binding energies of 10 top ranked small-molecule ligands for SARS-CoV-2 M^{pro}.

ID	Structure	Description	ΔG_{MMPBSA} (ΔG_{bind}) (kcal/mol)
C3809489 bemcentinib		Inhibitor of the kinase domain of AXL receptor	-34.7 ± 2.6 (-30.7)
C4291143 PC786		Respiratory syncytial virus (RSV) L protein polymerase inhibitor	-33.1 ± 0.3 (-29.2)
C787 Montelukast		Leukotriene receptor antagonist used with cortico-steroids for asthma therapy	-32.7 ± 0.2 (-20.6)
C442 Ergotamine		Alpha-1 selective adrenergic agonist used in migraine treatment	-31.5 ± 0.3 (-28.7)
D06290 simeprevir		Hepatitis C virus (HCV) NS3/4A protease inhibitor	-31.4 ± 0.2 (-29.2)
D08934 sofosbuvir		Nucleotide prodrug and HCV NS5B polymerase inhibitor	-31.0 ± 0.5 (-22.8)
D01601 lopinavir		Antiretroviral protease inhibitor for treatment of HIV-1	-30.7 ± 0.3 (-20.4)
D00503 ritonavir		Peptidomimetic inhibitor of HIV-1 and HIV-2 proteases	-30.5 ± 0.5 (-21.3)
C2105887 Mergocriptine		Synthetic ergot derivative, dopamine receptor agonist	-30.0 ± 0.3 (-17.9)
D14761 remdesivir		Viral RNA-dependent RNA polymerase inhibitor	-30.0 ± 0.2 (-27.1)

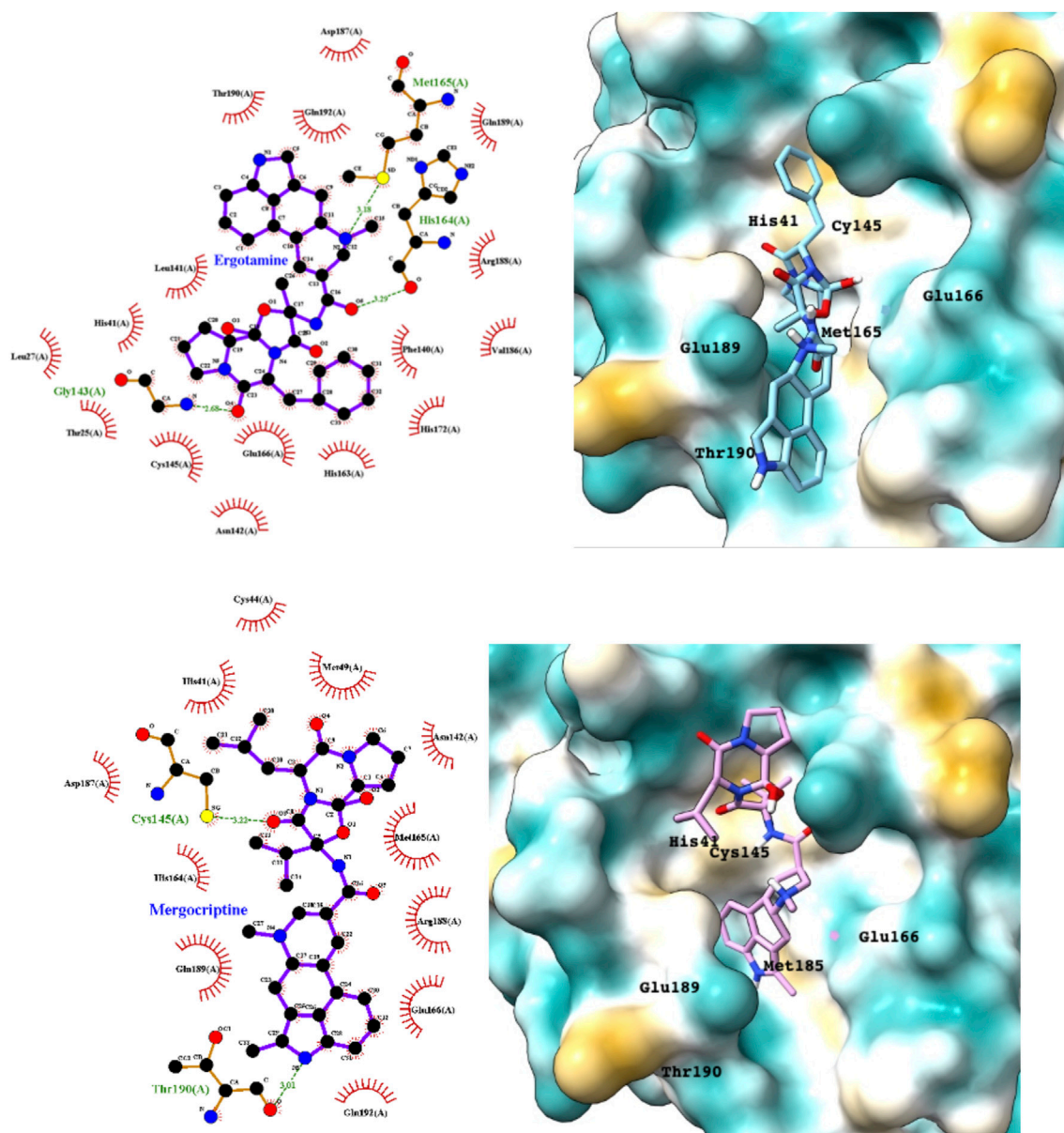


FIGURE 2 | LigPlot (left) and hydrophobic protein surface representation (right) of the main interactions between M^{pro} and ergotamine (top) and mergocriptine (bottom). The molecular surface denotes hydrophobicity of the pockets (blue hydrophilic, yellow/brown hydrophobic). Key binding site residues are labeled to help orient the viewer.

RESULTS

Molecular dynamics calculations were used to predict the optimal binding poses and binding energies for 84 of the top hits from docking-based virtual screening of ~12,000 drug candidates against the SARS-CoV-2 M^{pro}. The docking protocols were validated by redocking ligands from 10 x-ray structures. The top candidates were ranked for COVID-19 repurposing based on binding affinity and novelty. Conspicuously, we found that ~30% of the computationally repurposed drug candidates have experimentally validated activity against the M^{pro} target

protein, the SARS-CoV-2 virus, or both. Several of the drugs we identified are currently in clinical trials for COVID-19.

The binding energies of the 84 top ranked ligands from the docking calculations are listed in **Supplementary Table S1**. Note that calculating accurate *absolute* binding energies is difficult, and the approach we have taken provides good estimates of the *relative* binding energies of repurposing candidates. The ten drugs with the tightest binding to M^{pro} are summarized in **Table 1**, together with their GMXPBSA binding energies. The binding energies of several of the antiviral drugs, namely, simeprevir, sofosbuvir, lopinavir,

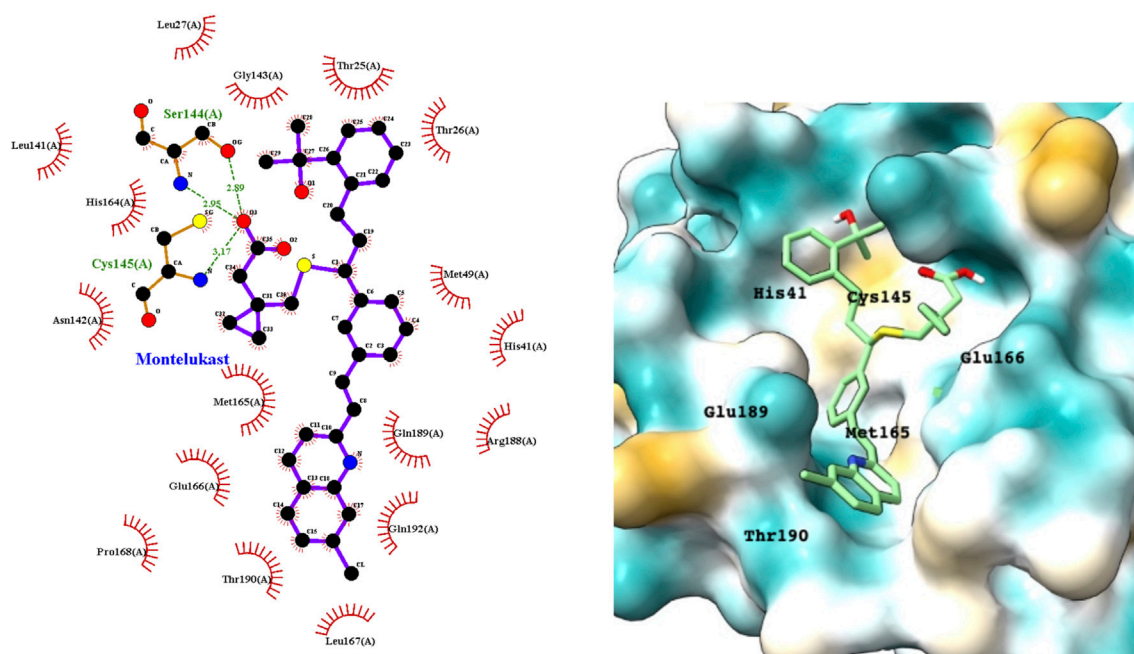


FIGURE 3 | LigPlot (left) and hydrophobic protein surface representation (right) of the main interactions between M^{PRO} and montelukast. The molecular surface denotes hydrophobicity of the pockets (blue hydrophilic, yellow/brown hydrophobic). Key binding site residues are labeled to help orient the viewer.

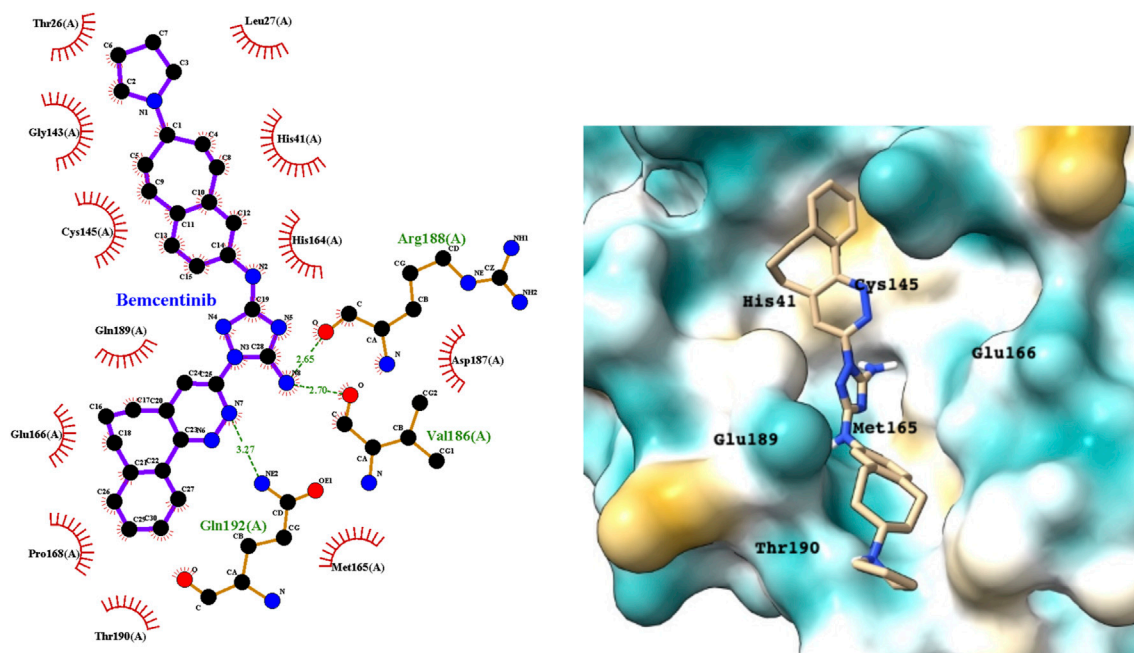


FIGURE 4 | LigPlot (left) and hydrophobic protein surface representation (right) of the main interactions between M^{PRO} and bemcentinib. The molecular surface denotes hydrophobicity of the pockets (blue hydrophilic, yellow/brown hydrophobic). Key binding site residues are labeled to help orient the viewer.

and ritonavir, are very similar, within the uncertainties in calculated energies. Some of the antivirals were also identified in other *in silico* docking studies or wet-lab SARS-CoV-2

activity studies, as we discuss below. This, together with a subsequent extensive search of the literature for experimental data, provides strong validation of the utility of our

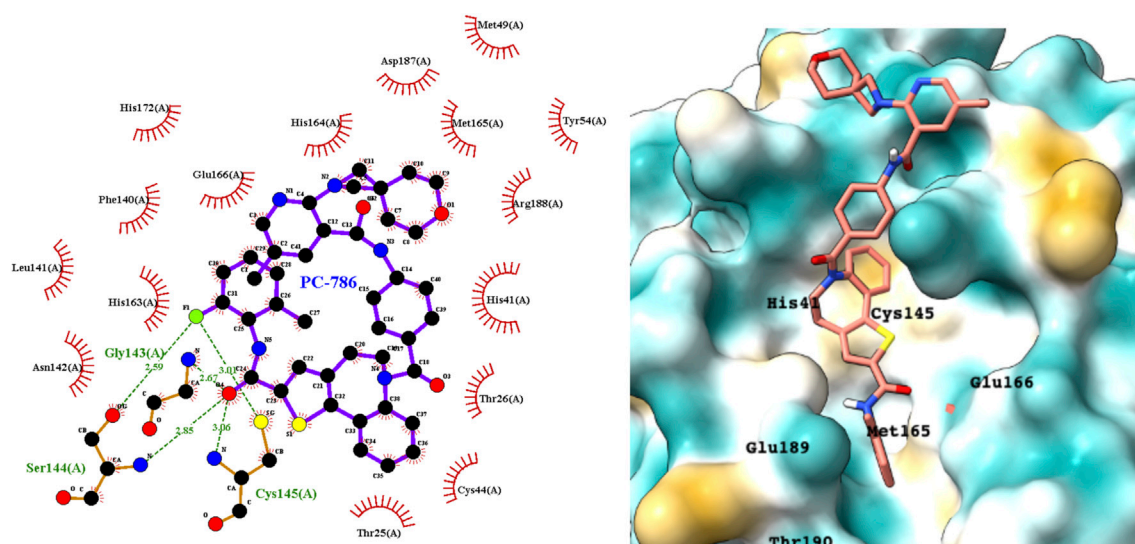


FIGURE 5 | LigPlot (left) and hydrophobic protein surface representation (right) of the main interactions between M^{PRO} and PC786. The molecular surface denotes hydrophobicity of the pockets (blue hydrophilic, yellow/brown hydrophobic). Key binding site residues are labeled to help orient the viewer. Other novel putative M^{PRO} inhibitors from the short list of 84 drugs.

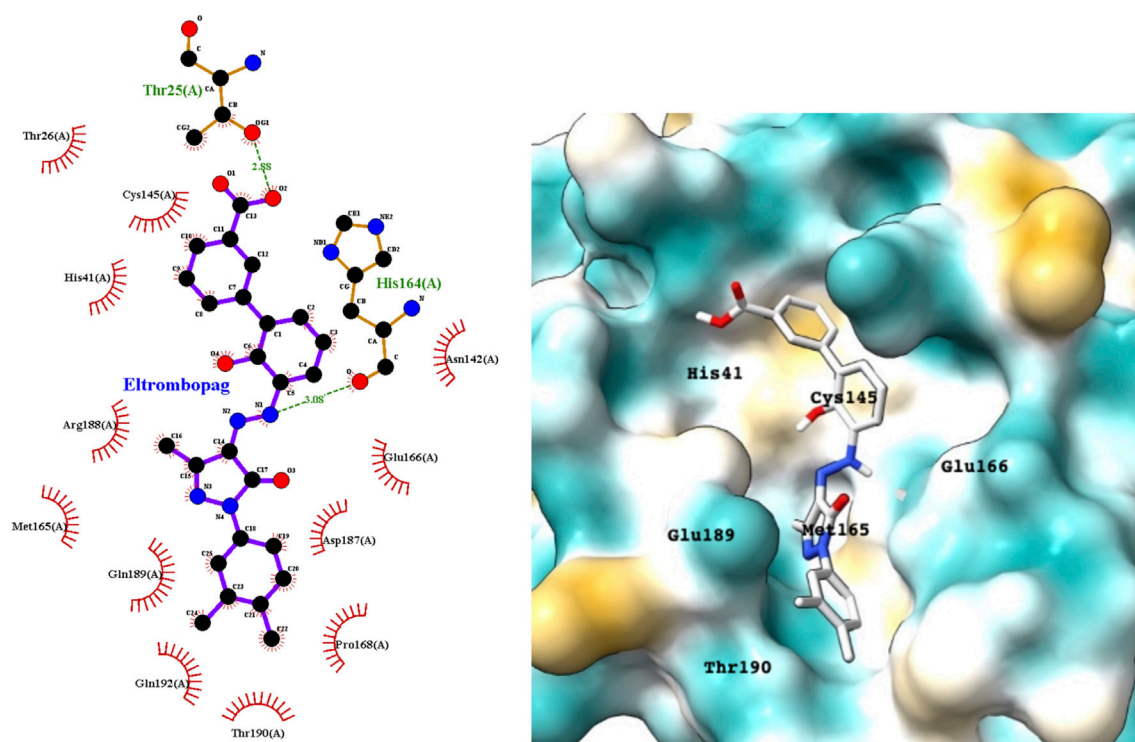


FIGURE 6 | LigPlot (left) and hydrophobic M^{PRO} protein surface representation (right) of the main interactions between M^{PRO} and eltrombopag and M^{PRO}. The molecular surface denotes hydrophobicity of the pockets (blue hydrophilic, yellow/brown hydrophobic).

computational methods to find leads consistent with other studies that also inhibit SARS-CoV-2 or the relevant protein target. It strongly suggests that the computational protocols we

have adopted are very capable of generating a list of repurposing candidates, many of which are likely to exhibit useful experimental *in vitro* activity at least.

In this paper, we have focused particularly on tightly binding drugs with novel structures, such as ergot compounds, bemcentinib, PC786, and montelukast.

Although the main focus of the paper is to show that appropriate computational methods can make useful predictions of the repurposing potential of drugs, we also provide a preliminary analysis of the binding of candidate drugs to the M^{Pro} active site. M^{Pro} achieves protein cleavage *via* the catalytic dyad His41 and Cys145. The main active site residues that have previously been implicated in drug binding are His41, Gly143, Cys145, His163, Glu166, and Glu166. All of the drugs whose interactions with the M^{Pro} binding site are summarized below interact with these six residues (see **Supplementary Table S2**). Most form strong hydrogen bonds to one or more of Gly143, Cys145, and His163. All docked and MD simulated structures of M^{Pro} with the repurposed drug candidates were also deposited in open access data archives. **Supplementary Figure S2** shows a superimposition of the top 10 drugs bound to the M^{Pro} site.

Ergotamine and ergocriptine, a synthetic long-acting ergot derivative, are α 1 selective adrenergic agonist vasoconstrictors and an agonist of dopamine receptors, respectively. **Figure 2** shows a LigPlot representation of the interactions of key functional groups in ergotamine and ergocriptine with protease active site residues. These, together with the accompanying M^{Pro} binding site molecular surface plots encoded for lipophilicity, illustrate how these drugs bind in the protease binding site. The specific interactions between these drugs and the residues in the binding site are summarized in **Supplementary Table S2**. Both drugs make strong and multiple interactions with 20 active site residues, notably hydrogen bonds with Gly143, His164, Met165, Cys145, and Thr190.

Montelukast is a cysteinyl leukotriene receptor antagonist used to treat asthma and allergic rhinitis. It reduces pulmonary responses to antigen, tissue eosinophilia and IL-5 expression in inflammatory cells and decreases elevated levels of IL-1 β and IL8 in viral upper respiratory tract infections (Almerie and Kerrigan, 2020). **Figure 3** shows a LigPlot representation of the interactions of key functional groups in montelukast with protease active site residues and a representation of how this drug binds in the active site of M^{Pro}. The specific interactions between montelukast and the active site residues are also summarized in **Supplementary Table S2**. The drug interacts extensively with the active site, binding to 23 residues, forming strong hydrogen bonds with Ser144 and Cys145. Montelukast spans the relatively broad binding pocket of the enzyme.

Bemcentinib selectively inhibits AXL kinase activity, which blocks viral entry and enhances the antiviral type I interferon response. **Figure 4** provides a LigPlot representation of the interactions of key functional groups in bemcentinib with protease active site residues, which are also summarized in detail in **Supplementary Table S2**. It forms strong hydrogen-bonding interactions with Val186, Arg188, and Gln192. **Figure 4** also shows the docking pose of bemcentinib in the protease active site after simulation by MD. The hydrophobic

benzocycloheptapyridazine moiety occupies a relative hydrophobic pocket, while the hydrophilic triazolyldiamine moiety binds strongly to the polar pocket formed partially by Asp187 and Arg188.

PC786 targets the respiratory syncytial virus (RSV) L protein and is designed to be a topical inhalation treatment, a likely route of infection for SARS-CoV-2. **Figure 5** shows a LigPlot representation of the interactions of key functional groups in PC786 with protease active site residues, with the specific interactions listed in **Supplementary Table S2**. It forms a hydrogen bond network with Gly143, Ser144, and Cys145. **Figure 5** illustrates the binding pose of PC786 in the M^{Pro} binding site after MD simulations based on the structure obtained from Vina docking calculations. The hydrophobic phenyl ring of the benzazepine moiety projects into a hydrophobic pocket formed partly by Thr25 and Thr26.

The predicted binding energies of the 84 drugs in the short list are summarized in **Supplementary Table S1**, along with details of any experiments to determine their activities against M^{Pro} or SARS-CoV-2 *in vitro* or *in vivo*. This suggests that our screening and MD simulation methods are sufficiently robust and accurate to identify drugs for repurposing against SARS-CoV-2 and, more broadly, other coronaviruses. The 33% of drugs in the hit list that have not been reported before are clearly of potential interest as novel drugs for treating COVID-19. We discuss below some of the more interesting and novel hit compounds with stronger binding affinities.

Eltrombopag is a thrombopoietin (TPO) receptor agonist that acts at the transmembrane domain of its cognate receptor C-Mpl *via* a histidine residue that occurs only in humans and apes. It scored highly in the docking studies, suggesting that it could inhibit the M^{Pro} and exhibit antiviral activity. **Figure 6** shows a LigPlot representation of the interactions of key functional groups in eltrombopag with protease active site residues. The binding pose of eltrombopag in the active site of M^{Pro} from the MD simulations is also shown in **Figure 6**. Close analysis of the binding mode shows that eltrombopag occupies the main part of the M^{Pro} binding pocket, with the hydrophilic biphenyl moiety binding to the hydrophobic pocket formed partly by Cys145. The hydrophilic pyrazolone lies in a polar cleft bounded by Glu189 and Glu166, with the terminal dimethyl phenyl ring undergoing a hydrophobic interaction with Thr190.

Eltrombopag is of particular interest as an M^{Pro} inhibitor lead because it is novel and is also a member of a large class of small molecular TPO receptor agonists that may also exhibit activity against the viral protease, and potentially the spike protein and human ACE2 (Tarasova and Winkler, 2009). However, given the clotting disorders that SARS-CoV-2 generates, the TPOR agonist activities would need to be minimized to prevent platelet enhancement, while retaining or enhancing the antiviral activities.

Apart from the drugs discussed above, several other drugs in **Supplementary Table S1** are of interest. There are several other ergot derivatives with good predicted binding affinities to M^{Pro}. Metergotamine and dihydroergocristine were predicted to have ΔG_{bind} of -29 and -24 kcal/mol, respectively.

DISCUSSION

Our virtual screening approach, using Autodock Vina and MD simulation in tandem to calculate binding poses and energies for repurposed drugs, identified 84 compounds with potential for treating COVID-19. The top hits from our study consisted of a mixture of antiviral agents, natural products and drugs developed for other applications and that have additional models of action. We now discuss the results of our computational screening in the context of other computational studies of M^{Pro} in the literature.

Relevant Computational Drug Repurposing Modeling Studies

We reviewed the literature for other *in silico* studies that also identified some of these hit compounds as potential M^{Pro} inhibitors and SARS-CoV-2 antiviral agents. Many drugs on the list in **Supplementary Table S1** are predicted by published computational studies to be potential inhibitors of SARS-CoV-2 target proteins, largely M^{Pro} but also RNA-dependent RNA polymerase (RdRp), spike, helicase, 2'-O-methyltransferase, nsp16/nsp10 complex, nsp1, PL^{Pro}, nsp3, and nsp12, and human angiotensin converting enzyme 2 (ACE2). Satisfyingly, those with the best predicted binding affinity from our study have also been of greatest interest clinically, with a larger number of *in vitro* assay results and clinical trials for drugs with the highest binding affinities (see below).

Modeling Studies Related to the Top 10 Predicted Drugs for Repurposing

Simeprevir was reported to be an inhibitor of the M^{Pro} by Abhithaj et al. (2020). They used a pharmacophore search followed by grid-based ligand docking (GLIDE, Schrodinger) and binding energy estimates from the MMGBSA method of -81.7 kcal/mol. However, they did not use MD to simulate the interaction of simeprevir in the M^{Pro} binding site. Similarly, sofosbuvir was reported to be a strong inhibitor of the protease by Lo et al. (2021).

The potential protease inhibition properties of lopinavir and ritonavir were reported by Bolcato et al., who used supervised MD to calculate the trajectories of the ligands in the protease binding site (Bolcato et al., 2020). Muralidharan et al. also used AutoDock (another docking program similar to Vina produced by the Scripps group) followed by MD simulations using the Generalized Amber Force Field (GAFF) in Amber16 to screen for repurposed drugs (Muralidharan et al., 2021). They reported AutoDock binding energies for lopinavir, oseltamivir, and ritonavir of -4.1 kcal/mol, -4.65 kcal/mol, and -5.11 kcal/mol, respectively, but did not provide the binding energies from the MD calculations. The best-known antiviral drug, which has been the subject of several clinical trials for COVID-19, is remdesivir (Hendaus, 2021). The potential inhibition of M^{Pro} by this drug has been reported in several computational screening studies. For example, Al-Khafaji and colleagues reported a combined computational docking and MD study of a range of antiviral drugs to the viral protease (Al-Khafaji et al., 2021). They calculated a binding energy for remdesivir of -65.19 kcal/mol

from a GROMACS simulation and a MMGBSA binding energy calculation. Beck et al. reported a K_d for binding of remdesivir to 3CLPro of 113 nM using a deep learning model.

Novel potential M^{Pro} inhibitors that emerged from our study included the ergot alkaloids ergotamine, mergocriptine, the thrombopoietin receptor agonist eltrombopag (ranked 13 with $\Delta G_{MMPBSA} = -28.2$ kcal/mol, see **Supplementary Table S1**), bemcentinib, PC786, and montelukast. These drugs were predicted to have better binding energies than the antiviral drugs discussed above and were not previously known to be antiviral.

Gurung et al. reported potential binding of ergotamine to the SARS-CoV-2 main protease in a preprint (Gurung et al., 2020). They employed AutoDock Vina but without subsequent MD simulation of the complex. They reported the binding energy as -9.4 kcal/mol for dihydroergotamine and -9.3 kcal/mol for ergotamine. Mevada et al. also reported *in silico* estimates of the binding of ergotamine to the protease using AutoDock Vina for the virtual screening (Mevada et al., 2020). They found the drug bound with an energy of -10.2 kcal/mol, calculated using Vina (no subsequent MD simulation). Gul et al. used a similar docking approach, this time with MD simulation, and identified ergotamine and its derivatives, dihydroergotamine and bromocriptine, as having high binding affinity to SARS-CoV-2 M^{Pro}. Ergotamine is an α -1 selective adrenergic agonist and vasoconstrictor and exhibited a favorable docking binding energy against SARS-CoV-2 M^{Pro} of -8.6 kcal/mol. Dihydroergotamine, the 9,10- α -dihydro derivative of ergotamine, showed a similar high affinity of -8.6 kcal/mol, and bromocriptine had a high affinity of -9.2 kcal/mol. Ergotamine has also been predicted to bind tightly to the SARS-CoV-2 spike (S) protein (Qiao et al., 2020).

Montelukast has been shown to inhibit at least one other protease, eosinophil protease (Langlois et al., 2006). Mansoor and colleagues deduced that it may bind to M^{Pro} on the basis of a simple molecular docking study (Mansoor et al., 2020). Wu et al. also reported putative binding of montelukast to M^{Pro} in a computational study using the same Internal Coordinate Mechanics modeling methods (Wu et al., 2020). No accurate binding affinities were reported in either study.

There is very little published work on the PC786 SARS-CoV-2 efficacy or predicted binding affinity to M^{Pro}. Panda and coworkers reported a binding energy ΔG_{bind} for PC786 of -179.79 , tighter binding than calculated for lopinavir (-131.49 kJ/mol), using a combined docking and MD approach (Panda et al., 2020). Like our study, they employed Autodock Vina to dock a molecular library into the active site of M^{Pro}, followed by MD simulation using GROMACS.

Relevant Modeling Studies of the Drugs From the List of 84 Drugs

Several *in silico* screening studies have identified eltrombopag as a potential SARS-CoV-2 drug. Feng et al.'s study suggested that eltrombopag bound not only to the M^{Pro} active site but also to the viral spike protein and to human ACE2 (Feng et al., 2020). This potential synergistic polypharmacy could be particularly beneficial for treating COVID-19. Eltrombopag

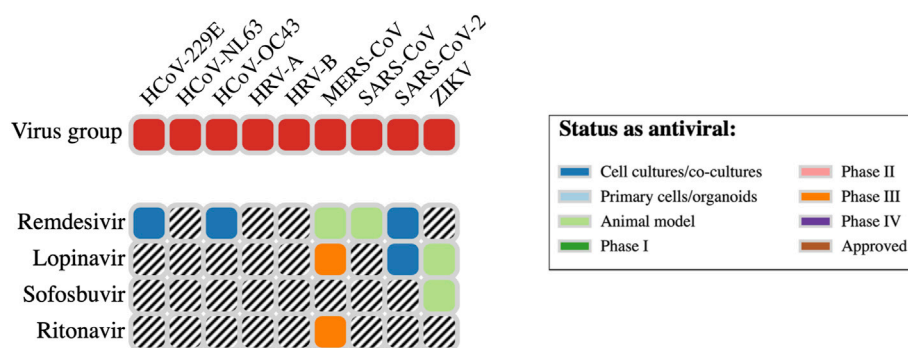


FIGURE 7 | Spectrum of antiviral activity and nature of assessment for four antiviral hit drugs.

has also been proposed as a useful drug against SARS-CoV-2 spike protein on the basis of its predicted strong binding to a pocket in the fusion cores of S2 domain (Feng et al., 2020). Eltrombopag was identified as having a high binding affinity to human ACE2, the primary binding site for the SARS-CoV-2 spike protein. This virtual screening study also used Autodock Vina, but no subsequent MD simulation was used for the top hit compounds from the screen. Surface Plasmon Resonance (SPR) was used to assess the binding of the drug to M^{Pro}.

Other drugs with binding energies stronger than -25 kcal/mol include galicaftor (in clinical trial for cystic fibrosis), rolitetracycline (a broad spectrum antibiotic), disoglucide (a natural product from *Dioscorea nipponica* Makino that reduces liver chronic inflammation and fibrosis), zafirlukast (a leukotriene receptor antagonist for asthma), diosmin (a natural flavone for treating venous disease), AZD-5991 (in clinical trial for relapsed or refractory hematologic malignancies), and ruzasvir (in clinical trials for treatment of hepatitis C). Li et al. also reported predicted M^{Pro} binding for galicaftor (Li et al., 2020). These drugs and natural products merit assessment in SAR-CoV-2 assays and M^{Pro} inhibition experiments.

As we stated in the introduction, few studies have used computational docking followed by MD simulation of the best repurposing candidates, and fewer still have reported experimental validation of the computational predictions. Here, we report a comprehensive review of experimentally determined protein target, *in vitro*, or *in vivo* activities of our 84 top binding drug candidates.

Experimental Validation of Biological Activity of Computational Repurposing Candidates

Clearly, blind computational predictions of likely M^{Pro} activity and, potentially, SARS-CoV-2 activities are of limited use if the predictions are not validated experimentally. **Supplementary Table S1** shows that 70% of the top 10 hit compounds have confirmed experimental activity against SARS-CoV-2, and 25% of the 84 entries in this table also have confirmed experimental activity either against SARS-CoV-2 or M^{Pro}. Of the remaining

75% of putative repurposed drugs, most have not been studied experimentally so they may have relevant antiviral activity, at least *in vitro*. The relatively high experimental validation rate of compounds predicted to be strong binders to M^{Pro} suggests that our computational paradigm is useful for selecting drugs for repurposing against SARS-CoV-2. Clearly, computational studies are investigating molecular interactions at the target, which will be most useful for identifying candidates with strong target activity, M^{Pro} in this instance. The high experimental validation rate also strongly suggests that the drugs not yet experimentally tested should at least be screened in an *in vitro* antiviral assay.

Validation of SARS-CoV-2 Activity of Top 10 Predicted Drugs for Repurposing

The website DrugVirus.info provides a concise picture of the broad-spectrum antiviral activity of a range of drugs. A summary for four of the top 10 antiviral hits (**Table 1**) from our *in silico* screens is provided in **Figure 7**. Here, we discuss the experimental SARS-CoV-2 or molecular target activity of seven of the top ten repurposing drug candidates identified by our computational studies—bemcentinib, montelukast, simeprevir, sofosbuvir, lopinavir, ritonavir, and remdesivir.

Bemcentinib selectively inhibits AXL kinase activity, which blocks viral entry and enhances the antiviral type I interferon response. Its *in vitro* activity against SARS-CoV-2 has been assessed by several groups. In a Vero cell assay, Liu et al. reported 10–40% protection at 50 μ M (Liu et al., 2020). However, in an alternative assay using human Huh7.5 cells (Dittmar et al., 2021), bemcentinib exhibited an IC₅₀ of 100 nM and CC₅₀ of 4.7 μ M. These authors also developed an assay in Vero cells and reported an IC₅₀ of 470 nM and CC₅₀ of 1.6 μ M, considerably higher activity than that reported by Liu et al. As a result, it is an investigational treatment for COVID-19 (www.clinicaltrialsregister.eu), with a phase 2 trial underway (Wilkinson et al., 2020). Dittmar and co-workers also reported an ED₅₀ for bemcentinib of 0.1 μ M (Huh7.5 cells), 0.47 μ M (Vero cells), and 2.1 μ M (Calu3 cells) (Dittmar et al., 2021). Six of the top 10 drugs (**Table 1**) are antiviral agents. Using a Vero E6 cellular infection model, they also reported that simeprevir was the only drug among their prioritized candidates that suppressed

SARS-CoV-2 replication at below 10 μM . Dose–response studies showed that simeprevir had an EC_{50} of 4 μM and a CC_{50} of 20 μM , similar to remdesivir in their experiments. Simeprevir had an experimental *in vitro* EC_{50} activity of 4.08 μM . Ma et al. developed a fluorescence resonance energy transfer (FRET)-based enzymatic assay for the SARS-CoV-2 M^{pro} and used it to screen a library of M^{pro} inhibitors (Ma et al., 2020). In their assay, simeprevir exhibited an IC_{50} of $14 \pm 3 \mu\text{M}$.

Sofosbuvir was reported to be a strong inhibitor of the main protease by Lo et al. (2021). It has *in vitro* antiviral EC_{50} values of 6.2 and 9.5 μM (Sacramento et al., 2021). Lopinavir exhibits an antiviral *in vitro* EC_{50} of 5.7 μM in one study (Yamamoto et al., 2020), and an EC_{50} of 26.6 μM in another study (Choy et al., 2020). It is also the subject of multiple single-agent and combination human trials (e.g., Cao et al., 2020; Costanzo et al., 2020). Ritonavir has an experimental *in vitro* EC_{50} of 8.6 μM (Yamamoto et al., 2020), and it too is being assessed in multiple single-agent and combination human trials (e.g., Cao et al., 2020; Verdugo-Paiva et al., 2020). Costanzo and colleagues likewise reported high protease binding for these two antiviral drugs (Costanzo et al., 2020). They also reported updates on experimental drugs successfully employed in the treatment of the disease caused by SARS-CoV-2 coronavirus. Patient recovery has been reported after treatment with lopinavir/ritonavir (used to treat HIV infection) in combination with the anti-flu drug oseltamivir.

Remdesivir has also been assessed in multiple human trials (e.g., Wang Y. et al., 2020; Olender et al., 2021), and it has reported antiviral *in vitro* EC_{50} values of 23.2 μM (Choy et al., 2020) and 0.77 μM and a $\text{CC}_{50} > 100 \mu\text{M}$ in another study (Wang M. et al., 2020). It had a SARS-CoV-2 EC_{50} in Vero cells of 6.6 μM and $\text{CC}_{50} > 100 \mu\text{M}$ (Pirzada et al., 2021). Beck et al. reported a K_d for binding of remdesivir to M^{pro} of 113 nM using a deep learning model. Liu et al. reported an *in vitro* assay that exploited the pronounced cytopathic effects of SARS-CoV-2 on Vero cells and the ability of a range of antiviral drugs to protect cells against the virus (Liu et al., 2020). In their assay, remdesivir exhibited an IC_{50} of 2.5 μM and a CC_{50} of 175 μM , while sofosbuvir, lopinavir, and ritonavir were inactive.

Montelukast has been shown to produce a significant reduction in SARS-CoV-2 infection in the treated elderly asthmatic patients. Kumar et al. also reported *in vitro* SARS-CoV-2 inhibition, with an IC_{50} of 18.8 μM and $\text{CC}_{50} > 20 \mu\text{M}$ (Bozek and Winterstein, 2020; Kumar et al., 2021).

Other Novel Putative SARS-CoV-2 Drugs From the List of 84 Drugs

We also highlight some novel and interesting drugs for repurposing in the list of 74 (Supplementary Table S1) with weaker predicted binding affinity than the top 10 listed in Table 1. As stated above, 25% of the drugs in Supplementary Table S1 have reported experimental data that support the validity of predictions from our computational experiments.

Eltrombopag, a thrombopoietin receptor agonist used to treat thrombocytopenia, has a reported IC_{50} 8.3 μM for SARS-CoV-2 infection in Vero and Calu-3 cells (Ko et al., 2021). Recently, Vogel et al. reported direct inhibition of cytomegalovirus (CMV) by therapeutic doses of eltrombopag used to treat thrombocytopenia (Vogel et al., 2019). They showed that

TABLE 2 | RMSD errors for redocking small molecule ligands into the binding site of M^{pro} .

Crystal ligand	Docking score (kcal/mol)	RMSD (Å)
5R7Y	−5.0	0.44
5R81	−5.0	0.51
5RE4	−4.2	0.62
5REJ	−5.3	0.52
5RG0	−4.5	0.55
6LU7	−6.6	0.44
6W63	−7.2	0.39
5R7Z	−5.3	0.57
5REL	−5.6	0.51
5R83	−5.4	0.57

eltrombopag inhibits the late stages of the HCMV replication cycle and reduces virus titers by 1.8×10^4 -fold at 10 μM and by 15-fold at 500 nM.

Saquinavir, an HIV protease inhibitor used in combination with other antiretroviral agents for the treatment of HIV-1, displays an *in vitro* EC_{50} of 8.8 μM (Yamamoto et al., 2020); zafirlukast, a leukotriene receptor antagonist used for prophylaxis and chronic treatment of asthma, exhibits an *in vitro* SARS-CoV-2 IC_{50} value of 3.6 μM (Zeng et al., 2021). Zhu and coworkers also measured the SARS-CoV-2 and M^{pro} inhibition of zafirlukast (Zhu et al., 2020). The IC_{50} for M^{pro} was 24 μM and the EC_{50} for the virus $> 20 \mu\text{M}$.

Eravacycline, a tetracycline antibiotic used to treat complicated intra-abdominal infections, has a reported *in vitro* activity against recombinant SARS-CoV-2, SARS-CoV, and MERS-CoV main proteases, with IC_{50} values of 1.7, 10.0, and 16.4 μM , respectively. It also inhibits SARS-CoV-2 infection in VeroE6 cells with an $\text{IC}_{50} = 30.6 \mu\text{M}$ (Reig and Shin, 2020). Umifenovir (Arbidol), exerts antiviral effects through multiple pathways that see its use against a variety of enveloped and non-enveloped RNA and DNA viruses. It inhibits coronavirus OC43 with an IC_{50} of 4.4 μM and SARS-CoV-2 with an IC_{50} of 10 μM (Xiao et al., 2020). It has also been reported to inhibit SARS-CoV-2 infection at 10–30 μM *in vitro* (Vafaei et al., 2020). Multiple clinical trials show a larger negative rate of PCR on day 14 in adult COVID-19 patients (Huang et al., 2021), and it shortens the viral shedding interval (Huang et al., 2020). Atazanavir, an antiretroviral drug of the protease inhibitor (PI) class, displays SARS-CoV-2 inhibition (in Vero cells) with an EC_{50} of 2 μM and a CC_{50} of 312 μM . It is also active against SARS-CoV-2 infection in a human epithelial pulmonary cell line (A549) with an EC_{50} of 0.22 μM (Fintelman-Rodrigues et al., 2020). Zhu and corksers measured the SARS-CoV-2 and M^{pro} inhibition of zafirlukast (Zhu et al., 2020). The IC_{50} for M^{pro} was 24 μM and the EC_{50} for the virus is $> 20 \mu\text{M}$.

The protease binding of rolitetracycline has been reported by Durdagi (2020) and Gul et al. (2020). The potential of the natural product diosmin as an antiviral agent targeting M^{pro} has also been reported in several recent computational studies (Arun et al., 2020; Ngo et al., 2020; Peterson, 2020b; Peterson, 2020a). Chakraborti et al. reported the potential of ruzasvir as a drug against SARS-CoV-2, although no data were provided (Chakraborti et al., 2020).

The prognostic value of our computational approach has been demonstrated by the fact that it identified a diverse range of drugs that have been reported in other computational studies or that exhibit useful SARS-CoV-2 antiviral effects *in vitro*. The antiviral drugs simeprevir, sofosbuvir, lopinavir, ritonavir, and remdesivir exhibit strong antiviral properties, and several are in clinical trial or used against SARS-CoV-2. These drugs have also been reported as binding to M^{Pro} by numerous virtual screening studies, and by *in vitro* assays. The more interesting and least studied hit drugs among our candidate list, bemcentinib, PC786, montelukast, ergotamine, and mergocriptine, were predicted to have binding affinities equal to or greater than the antiviral drugs, and have also been shown to have *in vitro* antiviral activity against SARS-CoV-2. A few computational studies mostly using less rigorous methods than those employed here have also suggested that these drugs may bind to M^{Pro}.

This high validation success rate strongly suggests that this type of virtual screening approach is capable of identifying compounds with potentially useful activity against SARS-CoV-2 and, by analogy, other coronaviruses. In particular, the 28 drugs for which no SARS-CoV-2 activity has been yet reported may be of particular interest for *in vitro* screening. The results of the current drug repurposing study provide information that could be useful to identify additional candidate drugs for testing for use in the current pandemic, as well as a rational computational paradigm for identifying therapeutic agents for future viral pandemics.

MATERIALS AND METHODS

Protein Structure Preparation and Grid Preparation

The crystal structure of the COVID-19 M^{Pro} was downloaded from the RCSB PDB (<http://www.rcsb.org>; refcode 6Y2F) (Zhang L. et al., 2020).

Protein preparation and removal of non-essential and non-bridging water molecules for docking studies were performed using the UCSF Chimera package (<https://www.cgl.ucsf.edu/chimera/>) (Pettersen et al., 2004) AutoDock Tools (ADT) software was used to prepare the required files for Autodock Vina by assigning hydrogen polarities, calculating Gasteiger charges to protein structures and converting protein structures from the pdb file format to pdbqt format (Forli et al., 2016). The surface area of the 3CLPro binding pocket is 335 Å², and the volume is 364.101 Å³ (Tian et al., 2018).

As recommended by Llanos et al., the ability of Vina to redock known ligands from x-ray structures was assessed to determine the reliability of the algorithm for this target. **Table 2** shows the RMSD values for redocking the ligands for 10 experimental structures of M^{Pro} with bound ligands. The relatively low RMSD values show that Vina can recapitulate the experimental binding poses well.

Screening Databases

Drugs were downloaded from the DrugBank database (Wishart et al., 2018) and ChEMBL database (FDA approved) (Gaulton et al., 2017). A total of 8,773 and 13,308 drugs were retrieved from DrugBank and ChEMBL database, respectively.

The drugs were downloaded in sdf format and converted to pdbqt format using Raccoon (Forli et al., 2016).

Docking Methodology Small-molecule ligand structures were docked against protein structure using the AutoDock Vina (version 1.1.3) package (Forli et al., 2016). AutoDock Vina employs gradient-based conformational search approach and an energy-based empirical scoring function that includes an approximate correction for ligand conformational entropy. AutoDock Vina is also flexible, easily scripted, and extensively validated in many published studies with a variety of proteins and ligands and takes advantage of large multi-CPU or -GPU machines to run many calculations in parallel. The code has also been employed very successfully to dock millions of small-molecule drug candidates into a series of protein targets to discover new potent drug leads. The package includes useful scripts for generating modified pdb files required for grid calculations and for setting up the grid calculations around each protein automatically. The software requires the removal of hydrogens, addition of polar hydrogens, setting of the correct atom types, and calculation of atom charges compatible with the AutoGrid code. The algorithm generates a grid around each protein and calculates the interaction energy of a probe noble gas atom at each grid position outside and within internal cavities of the protein. The grid resolution was set to 1 Å, the maximum number of binding modes to output was fixed at 10, and the exhaustiveness level (controlling the number of independent runs performed) was set at 8. The docking employed a genetic algorithm to optimize the binding conformations of the ligands during docking to the protease site. Drugs were docked individually to the active site of M^{Pro} (3CLPro, refcode 6Y2F) with the grid coordinates (grid center) and grid boxes of appropriate sizes generated by the bash script `vina_screen.sh` (**Supplementary Information**). The top scored compounds were identified with a python script `1.py` (**Supplementary Information**) and subjected to molecular dynamic simulation. The docked structures were analyzed using UCSF Chimera (Pettersen et al., 2004) and LigPlot + software (Laskowski and Swindells, 2011) to illustrate hydrogen-bond and hydrophobic interactions. A total of fifty top compounds were selected from each of the DrugBank and ChEMBL compounds. Sixteen compounds were common to both database top hits. Molecular dynamics studies were conducted on the unique set of eighty-four compounds from both sets.

Molecular Dynamics Simulation The top screened compound complexes with protease were minimized with CHARMM force field. The topology files of the ligands were prepared from Swissparam (<http://www.swissparam.ch/>) (Zoete et al., 2011) and minimized in Gromacs 2020 (<http://www.gromacs.org/>) (Abraham et al., 2015). Docked complexes of ligands and COVID-19 M^{Pro} protein were used as starting geometries for MD simulations. Simulations were carried out using the GPU accelerated version of the program with the CHARMM force field I periodic boundary conditions in ORACLE server. Docked complexes were immersed in a truncated octahedron box of TIP3P water molecules. The solvated box was further neutralized with Na⁺ or Cl⁻ counter ions using the `tleap` program. Particle Mesh Ewald

(PME) was employed to calculate the long-range electrostatic interactions. The cutoff distance for the long-range van der Waals (VDW) energy term was 12.0 Å. The whole system was minimized without any restraint. The above steps applied 2,500 cycles of steepest descent minimization followed by 5,000 cycles of conjugate gradient minimization. After equilibration at 300 K using Langevin thermostat NVT ensemble for 50 ps, the system was then equilibrated at 1 atm pressure using Berendsen thermostat NPT ensemble for 50 ps. After the system was fully equilibrated at the desired temperature and pressure (NVT/NPT ensembles), we used Parrinello–Rahman pressure coupling to run MD for data collection. Duplicate production runs starting with different random seeds were also run to allow estimates of binding energy uncertainties to be determined. Finally, a production run of 20 ns of MD simulation was performed.

During the MD procedure, the SHAKE algorithm was applied for the constraint of all covalent bonds involving hydrogen atoms. The time step was set to 2 fs. The structural stability of the complex was monitored by the RMSD and RMSF values of the backbone atoms of the entire protein. Calculations were also performed for up to 100 ns on few compounds to ensure that 20 ns is sufficiently long for convergence. We checked the RMSD of M^{Pro} and drug during this time and it was within the range of 1.5 Å. The RMSF graph revealed minimal fluctuations and relatively stable conformations of SARS-CoV-2 M^{Pro} bound to screened drugs.

The protein–ligand binding affinities were evaluated in two ways. One calculates the energies of solvated SARS-CoV-2 protease and small-molecule ligands and the other calculates that of the bound complex and derive the binding energy by subtraction.

$$\Delta E(bind) = \Delta E(complex) - (\Delta E(protein) + \Delta E(ligand)) \quad (1)$$

We also calculated binding free energies using the molecular mechanics Poisson Boltzmann surface area (MM/PBSA) tool in GROMACS that uses the nonbonded interaction energies of the complex. The method is also a widely used method for binding free energy calculations (Spiliotopoulos et al., 2012). However, accurate calculation of absolute binding energies requires very extensive sampling, so the methods we employed provide accurate relative binding energies of ligands that are useful for ranking them, as we have done in this work.

We used GMXPBSA2.1 program to perform MM/PBSA calculations on selected docked complexes derived from GROMACS trajectories (Paissoni et al., 2015). It is a suite of Bash/Perl scripts for streamlining MM/PBSA calculations on structural ensembles derived from GROMACS trajectories and to automatically calculate binding free energies for protein–protein or ligand–protein. GMXPBSA 2.1, which provides the freedom to calculate free binding energy of complexes with any force field, calculates diverse MM/PBSA energy contributions from molecular mechanics (MM) and electrostatic contribution to solvation (PB) and non-polar

contribution to solvation (SA). This tool combines the capability of MD simulations (GROMACS) and the Poisson–Boltzmann equation (APBS) for calculating solvation energy (Baker et al., 2001). The g_mmpbsa tool in GROMACS was used after molecular dynamics simulations, and the output files obtained were used to post-process binding free energies by the single-trajectory MMPBSA method. In the current study, we considered 100 frames at equal distance from 20-ns trajectory files.

Specifically, for a non-covalent binding interaction in the aqueous phase, the binding free energy, $\Delta G(bind, aq)$, is:

$$\Delta G(bind, aq) = \Delta G(bind, vac) + \Delta G(bind, solv) \quad (2)$$

where $\Delta G(bind, vac)$ is the binding free energy in vacuum, and $\Delta G(bind, solv)$ is the solvation free energy change upon binding:

$$\Delta G(bind, solv) = \Delta G(R:L, solv) - \Delta G(R, solv) - \Delta G(L, solv) \quad (3)$$

where $\Delta G(R:L, solv)$, $\Delta G(R, solv)$, and $\Delta G(L, solv)$ are solvation free energies of complex, receptor, and ligand, respectively.

Method Note Added in Proof

Guterres and Im recently showed how substantial improvements in protein–ligand docking results could be achieved using high-throughput MD simulations (Guterres and Im, 2020). As with our study, they also employed AutoDock Vina for docking, followed by MD simulation using CHARMM. The MD parameters they advocated were very similar to those used in our study. Proteins were solvated in a box of TIP3P water molecules extending 10 Å beyond the proteins and the particle-mesh Ewald method was used for electrostatic interactions. Nonbonded interactions over 10 and 12 Å were truncated. Their systems were minimized for 5,000 steps using the steepest descent method followed by 1 ns of equilibration with an NVT setting. For each protein–ligand complex, they ran 3×100 -ns production runs from the same initial structure using different initial velocity random seeds and an integration step size of 2 fs. Over 56 protein targets (of seven different protein classes) and 560 ligands, this shows 22% improvement in the area under the receiver operating characteristic curve, from an initial value of 0.68 using AutoDock Vina alone to a final value of 0.83 when the Vina results were refined by MD.

DATA AVAILABILITY STATEMENT

The data have been deposited in the OPAL repository at La Trobe University and are available at DOI 10.26181/19235004.

AUTHOR CONTRIBUTIONS

NP conceived the project, analyzed data, and contributed to the manuscript; SP and PS performed the computations, analyzed data, and contributed to the manuscript; DW analyzed data and contributed to the manuscript.

FUNDING

We would like to thank Oracle for providing their Cloud computing resources for the modeling studies described herein.

ACKNOWLEDGMENTS

We would like to thank Harinda Rajapaksha for assistance in optimizing GROMACS for this project. We would also like to thank Oracle for providing their Cloud computing resources for the modeling studies described herein. In particular, we wish to

thank Peter Winn, Dennis Ward, and Alison Derbenwick Miller from Oracle in facilitating these studies. The opinions expressed herein are solely those of the individual authors and should not be inferred to reflect the views of their affiliated institutions, funding bodies, or Oracle corporation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2022.781039/full#supplementary-material>

REFERENCES

- Abhithaj, J., Dileep, F., Sharanya, C. S., Arun, K. G., Sadasivan, C., and Jayadevi, V. (2022). Repurposing Simeprevir, Calpain Inhibitor IV and a Cathepsin F Inhibitor Against SARS-CoV-2 and Insights into Their Interactions With Mpro. *J. Biomol. Struct. Dyn.* 40 (1), 325–336. doi:10.1080/07391102.2020.1813200
- Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., et al. (2015). GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* 1–2, 19–25. doi:10.1016/j.softx.2015.06.001
- 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. *J. Biomol. Struct. Dyn.* 39, 3387–3395. doi:10.1080/07391102.2020.1764392
- Almerie, M. Q., and Kerrigan, D. D. (2020). The Association between Obesity and Poor Outcome after COVID-19 Indicates a Potential Therapeutic Role for Montelukast. *Med. Hypotheses* 143, 109883. doi:10.1016/j.mehy.2020.109883
- Arun, K. G., Sharanya, C. S., Abhithaj, J., Francis, D., and Sadasivan, C. (2020). Drug Repurposing against SARS-CoV-2 Using E-Pharmacophore Based Virtual Screening, Molecular Docking and Molecular Dynamics with Main Protease as the Target. *J. Biomol. Struct. Dyn.*, 1–12. doi:10.1080/07391102.2020.1779819
- Berkley, S. (2020). COVID-19 Needs a Big Science Approach. *Science* 367, 1407. doi:10.1126/science.abb8654
- Bolcato, G., Bissaro, M., Pavan, M., Sturlese, M., and Moro, S. (2020). Targeting the Coronavirus SARS-CoV-2: Computational Insights into the Mechanism of Action of the Protease Inhibitors Lopinavir, Ritonavir and Nelfinavir. *Sci. Rep.* 10, 20927. doi:10.1038/s41598-020-77700-z
- Bozek, A., Winterstein, J., Galuszka, B., and Jarzab, J. (2020). Different Development Forms of Local Allergic Rhinitis towards Birch. *Biomed. Res. Int.* 2020, 1–9. doi:10.1155/2020/3408561
- Cao, B., Wang, Y., Wen, D., Liu, W., Wang, J., Fan, G., et al. (2020). A Trial of Lopinavir-Ritonavir in Adults Hospitalized with Severe Covid-19. *N. Engl. J. Med.* 382, 1787–1799. doi:10.1056/nejmoa2001282
- Chakraborti, S., Bheemireddy, S., and Srinivasan, N. (2020). Repurposing Drugs against the Main Protease of SARS-CoV-2: Mechanism-Based Insights Supported by Available Laboratory and Clinical Data. *Mol. Omics* 16, 474–491. doi:10.1039/d0mo00057d
- Choy, K.-T., Wong, A. Y.-L., Kaewpreedee, P., Sia, S. F., Chen, D., Hui, K. P. Y., et al. (2020). Remdesivir, Lopinavir, Emetine, and Homoharringtonine Inhibit SARS-CoV-2 Replication *In Vitro*. *Antiviral Res.* 178, 104786. doi:10.1016/j.antiviral.2020.104786
- Ciotti, M., Angeletti, S., Minieri, M., Giovannetti, M., Benvenuto, D., Pascarella, S., et al. (2019). COVID-19 Outbreak: An Overview. *Chemotherapy* 64, 215–223. doi:10.1159/000507423
- Costanzo, M., De Giglio, M. A. R., and Roviello, G. N. (2020). SARS-CoV-2: Recent Reports on Antiviral Therapies Based on Lopinavir/Ritonavir, Darunavir/Umifenovir, Hydroxychloroquine, Remdesivir, Favipiravir and Other Drugs for the Treatment of the New Coronavirus. *Curr. Med. Chem.* 27, 4536–4541. doi:10.2174/0929867327666200416131117
- 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, 1331–1335. doi:10.1126/science.abb4489
- Dittmar, M., Lee, J. S., Whig, K., Segrist, E., Li, M., Kamalia, B., et al. (2021). Drug Repurposing Screens Reveal Cell-Type-Specific Entry Pathways and FDA-Approved Drugs Active Against SARS-Cov-2. *Cell Rep.* 35 (1), 108959. doi:10.1016/j.celrep.2021.108959
- Durdagi, S. (2020). Virtual Drug Repurposing Study against SARS-CoV-2 TMPRSS2 Target. *Turk J. Biol.* 44, 185–191. doi:10.3906/biy-2005-112
- Feng, S., Luan, X., Wang, Y., Wang, H., Zhang, Z., Wang, Y., et al. (2020). Eltrombopag Is a Potential Target for Drug Intervention in SARS-CoV-2 Spike Protein. *Infect. Genet. Evol.* 85, 104419. doi:10.1016/j.meegid.2020.104419
- Fintelman-Rodrigues, N., Sacramento, C. Q., Ribeiro Lima, C., Souza Da Silva, F., Ferreira, A. C., Mattos, M., et al. (2020). Atazanavir, Alone or in Combination with Ritonavir, Inhibits SARS-CoV-2 Replication and Proinflammatory Cytokine Production. *Antimicrob. Agents Chemother.* 64, e00825–00820. doi:10.1128/AAC.00825-20
- Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., and Olson, A. J. (2016). Computational Protein-Ligand Docking and Virtual Drug Screening with the AutoDock Suite. *Nat. Protoc.* 11, 905–919. doi:10.1038/nprot.2016.051
- Gaulton, A., Hersey, A., Nowotka, M., Bento, A. P., Chambers, J., Mendez, D., et al. (2017). The ChEMBL Database in 2017. *Nucleic Acids Res.* 45 (D1), D945–D954. doi:10.1093/nar/gkw1074
- Gul, S., Ozcan, O., Asar, S., Okyar, A., Baris, I., and Kavakli, I. H. (2020). In Silico identification of Widely Used and Well-Tolerated Drugs as Potential SARS-CoV-2 3C-like Protease and Viral RNA-dependent RNA Polymerase Inhibitors for Direct Use in Clinical Trials. *J. Biomol. Struct. Dyn.*, 39(17):6772–6791. doi:10.1080/07391102.2020.1802346
- Gurung, A. B., Ali, M. A., Lee, J., Abul Farah, M., and Al-Anazi, K. M. (2020). In Silico screening of FDA Approved Drugs Reveals Ergotamine and Dihydroergotamine as Potential Coronavirus Main Protease Enzyme Inhibitors. *Saudi J. Biol. Sci.* 27, 2674–2682. doi:10.1016/j.sjbs.2020.06.005
- Guterres, H., and Im, W. (2020). Improving Protein-Ligand Docking Results with High-Throughput Molecular Dynamics Simulations. *J. Chem. Inf. Model.* 60, 2189–2198. doi:10.1021/acs.jcim.0c00057
- Hendaus, M. A. (2021). Remdesivir in the Treatment of Coronavirus Disease 2019 (COVID-19): a Simplified Summary. *J. Biomol. Struct. Dyn.* 39, 3787–3792. doi:10.1080/07391102.2020.1767691
- Huang, D., Yu, H., Wang, T., Yang, H., Yao, R., and Liang, Z. (2021). Efficacy and Safety of Umifenovir for Coronavirus Disease 2019 (COVID-19): A Systematic Review and Meta-analysis. *J. Med. Virol.* 93, 481–490. doi:10.1002/jmv.26256
- Huang, H., Guan, L., Yang, Y., Le Grange, J. M., Tang, G., Xu, Y., et al. (2020). Chloroquine, Arbidol (Umifenovir) or Lopinavir/ritonavir as the Antiviral Monotherapy for COVID-19 Patients: a Retrospective Cohort Study. *Research Square*. doi:10.21203/rs.3.rs-24667/v1
- Jiang, S. (2020). Don't rush to Deploy COVID-19 Vaccines and Drugs without Sufficient Safety Guarantees. *Nature* 579, 321. doi:10.1038/d41586-020-00751-9
- Ko, M., Jeon, S., Ryu, W. S., and Kim, S. (2021). Comparative Analysis of Antiviral Efficacy of FDA-approved Drugs against SARS-CoV-2 in Human Lung Cells. *J. Med. Virol.* 93, 1403–1408. doi:10.1002/jmv.26397

- Kumar, S., Singh, B., Kumari, P., Kumar, P. V., Agnihotri, G., Khan, S., et al. (2021). Identification of Multipotent Drugs for COVID-19 Therapeutics with the Evaluation of Their SARS-CoV2 Inhibitory Activity. *Comput. Struct. Biotechnol. J.* 19, 1998–2017. doi:10.1016/j.csbj.2021.04.014
- Langlois, A., Ferland, C., Tremblay, G. M., and Laviolette, M. (2006). Montelukast Regulates Eosinophil Protease Activity through a Leukotriene-independent Mechanism. *J. Allergy Clin. Immunol.* 118, 113–119. doi:10.1016/j.jaci.2006.03.010
- Laskowski, R. A., and Swindells, M. B. (2011). LigPlot+: Multiple Ligand-Protein Interaction Diagrams for Drug Discovery. *J. Chem. Inf. Model.* 51, 2778–2786. doi:10.1021/ci200227u
- Li, Y., Zhang, Y., Han, Y., Zhang, T., and Du, R. (2020). Prioritization of Potential Drugs Targeting the SARS-CoV-2 Main Protease. *ChemRxiv*. doi:10.26434/chemrxiv.12629858.v1
- Liu, S., Lien, C. Z., Selvaraj, P., and Wang, T. T. (2020). Evaluation of 19 Antiviral Drugs against SARS-CoV-2 Infection. *bioRxiv*. doi:10.1101/2020.04.29.067983
- Llanos, M. A., Gantner, M. E., Rodriguez, S., Alberca, L. N., Bellera, C. L., Talevi, A., et al. (2021). Strengths and Weaknesses of Docking Simulations in the SARS-CoV-2 Era: the Main Protease (Mpro) Case Study. *J. Chem. Inf. Model.* 61, 3758–3770. doi:10.1021/acs.jcim.1c00404
- Lo, H. S., Hui, K. P., Lai, H.-M., Khan, K. S., Kaur, S., Huang, J., et al. (2021). Simeprevir Potently Suppresses SARS-CoV-2 Replication and Synergizes with Remdesivir. *ACS Central Science* 7 (5), 792–802. doi:10.1021/acscentsci.0c01186
- Lu, S. (2020). Timely Development of Vaccines against SARS-CoV-2. *Emerging Microb. Infect.* 9, 542–544. doi:10.1080/22221751.2020.1737580
- Ma, C., Sacco, M. D., Hurst, B., Townsend, J. A., Hu, Y., Szeto, T., et al. (2020). Boceprevir, GC-376, and Calpain Inhibitors II, XII Inhibit SARS-CoV-2 Viral Replication by Targeting the Viral Main Protease. *Cell Res* 30, 678–692. doi:10.1038/s41422-020-0356-z
- Mansoor, S., Saadat, S., Amin, A., Ali, I., Ghaffar, M. T., Amin, U., et al. (2020). A Case for Montelukast in COVID-19: The Use of Computational Docking to Estimate the Effects of Montelukast on Potential Viral Main Protease Catalytic Site. *Research Square*. [Preprint]. doi:10.21203/rs.3.rs-27079/v1
- Mendes, A. (2020). Research towards Treating COVID-19. *Br. J. Community Nurs.* 25, 204–205. doi:10.12968/bjcn.2020.25.4.204
- Mengist, H. M., Fan, X., and Jin, T. (2020). Designing of Improved Drugs for COVID-19: Crystal Structure of SARS-CoV-2 Main Protease Mpro. *Sig Transduct Target. Ther.* 5, 67. doi:10.1038/s41392-020-0178-y
- Mevada, V., Dudhagara, P., Gandhi, H., Vaghamsi, N., Beladiya, U., and Patel, R. (2020). Drug Repurposing of Approved Drugs Elbasvir, Ledipasvir, Paritaprevir, Velpatasvir, Antrafenine and Ergotamine for Combating COVID19. *ChemRxiv*. doi:10.26434/chemrxiv.12115251.v2
- Muralidharan, N., Sakthivel, R., Velmurugan, D., and Gromiha, M. M. (2021). Computational Studies of Drug Repurposing and Synergism of Lopinavir, Osetamivir and Ritonavir Binding with SARS-CoV-2 Protease against COVID-19. *J. Biomol. Struct. Dyn.* 39, 2673–2678. doi:10.1080/07391102.2020.1752802
- Ngo, S. T., Quynh Anh Pham, N., Thi Le, L., Pham, D.-H., and Vu, V. V. (2020). Computational Determination of Potential Inhibitors of SARS-CoV-2 Main Protease. *J. Chem. Inf. Model.* 60, 5771–5780. doi:10.1021/acs.jcim.0c00491
- Olender, S. A., Perez, K. K., Go, A. S., Balani, B., Price-Haywood, E. G., Shah, N. S., et al. (2021). Remdesivir for Severe COVID-19 versus a Cohort Receiving Standard of Care. *Clin. Infect. Dis.* 73 (11), e4166–e4174. ciae1041. doi:10.1093/cid/ciaa1041
- Olsen, M., Cook, S. E., Huang, V., Pedersen, N., and Murphy, B. G. (2020). Perspectives: Potential Therapeutic Options for SARS-CoV-2 Patients Based on Feline Infectious Peritonitis Strategies: Central Nervous System Invasion and Drug Coverage. *Int. J. Antimicrob. Agents* 55, 105964. doi:10.1016/j.ijantimicag.2020.105964
- Paisoni, C., Spiliotopoulos, D., Musco, G., and Spitaleri, A. (2015). GROMACS 2.1: A GROMACS Tool to Perform MM/PBSA and Computational Alanine Scanning. *Computer Phys. Commun.* 186, 105–107. doi:10.1016/j.cpc.2014.09.010
- Panda, P. K., Arul, M. N., Patel, P., Verma, S. K., Luo, W., Rubahn, H. G., et al. (2020). Structure-based Drug Designing and Immunoinformatics Approach for SARS-CoV-2. *Sci. Adv.* 6, eabb8097. doi:10.1126/sciadv.abb8097
- Peterson, L. (2020a). COVID-19 and Flavonoids: In Silico Molecular Dynamics Docking to the Active Catalytic Site of SARS-CoV and SARS-CoV-2 Main Protease. *SSRN Electronic Journal*. doi:10.2139/ssrn.3599426
- Peterson, L. (2020b). In Silico Molecular Dynamics Docking of Drugs to the Inhibitory Active Site of SARS-CoV-2 Protease and Their Predicted Toxicology and ADME. *SSRN Electronic Journal*. doi:10.2139/ssrn.3580951
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al. (2004). UCSF Chimera?A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* 25, 1605–1612. doi:10.1002/jcc.20084
- Pirzada, R. H., Haseeb, M., Batool, M., Kim, M., and Choi, S. (2021). Remdesivir and Ledipasvir Among the FDA-Approved Antiviral Drugs Have Potential to Inhibit SARS-CoV-2 Replication. *Cells* 10, 1052. doi:10.3390/cells10051052
- Qiao, Z., Zhang, H., Ji, H.-F., and Chen, Q. (2020). Computational View toward the Inhibition of SARS-CoV-2 Spike Glycoprotein and the 3CL Protease. *Computat.* 8, 53. doi:10.3390/computation8020053
- Rathnayake, A. D., Zheng, J., Kim, Y., Perera, K. D., Mackin, S., Meyerholz, D. K., et al. (2020). 3C-like Protease Inhibitors Block Coronavirus Replication *In Vitro* and Improve Survival in MERS-CoV-Infected Mice. *Sci. Transl. Med.* 12, eabc5332. doi:10.1126/scitranslmed.abc5332
- Reig, N., and Shin, D.-H. (2020). 560. Repurposing Eravacycline for the Treatment of SARS-CoV-2 Infections. *Open Forum Infect. Dis.* 7, S345. doi:10.1093/ofid/ofaa439.754
- Rosa, S. G. V., and Santos, W. C. (2020). Clinical Trials on Drug Repositioning for COVID-19 Treatment. *Rev. Panam Salud Publica* 44, e40. doi:10.26633/RPSP.2020.40
- Rosales-Mendoza, S., Márquez-Escobar, V. A., González-Ortega, O., Nieto-Gómez, R., and Arévalo-Villalobos, J. I. (2020). What Does Plant-Based Vaccine Technology Offer to the Fight against COVID-19? *Vaccines* 8, 183. doi:10.3390/vaccines8020183
- Sacramento, C. Q., Fintelman-Rodrigues, N., Temerozo, J. R., de Paula Dias Da Silva, A., Da Silva Gomes Dias, S., dos Santos da Silva, C., et al. (2021). In vitro Antiviral Activity of the Anti-HCV Drugs Daclatasvir and Sofosbuvir Against SARS-CoV-2, the Aetiological Agent of COVID-19. *J. Antimicrob. Chemother.* 76 (7), 1874–1885. doi:10.1093/jac/dkab072
- Sanders, J. M., Monogue, M. L., Jodlowski, T. Z., and Cutrell, J. B. (2020). Pharmacologic Treatments for Coronavirus Disease 2019 (COVID-19): A Review. *JAMA* 323, 1824–1836. doi:10.1001/jama.2020.6019
- Saul, S., and Einav, S. (2020). Old Drugs for a New Virus: Repurposed Approaches for Combating COVID-19. *ACS Infect. Dis.* 6, 2304–2318. doi:10.1021/acscinfdis.0c00343
- Schlagenhauf, P., Grobusch, M. P., Maier, J. D., and Gautret, P. (2020). Repurposing Antimalarials and Other Drugs for COVID-19. *Trav. Med. Infect. Dis.* 34, 101658. doi:10.1016/j.tmaid.2020.101658
- Sheik Amamuddy, O., Verkhivker, G. M., and Tastan Bishop, Ö. (2020). Impact of Early Pandemic Stage Mutations on Molecular Dynamics of SARS-CoV-2 Mpro. *J. Chem. Inf. Model.* 60, 5080–5102. doi:10.1021/acs.jcim.0c00634
- Sohrabi, C., Alsafi, Z., O'Neill, N., Khan, M., Kerwan, A., Al-Jabir, A., et al. (2020). World Health Organization Declares Global Emergency: A Review of the 2019 Novel Coronavirus (COVID-19). *Int. J. Surg.* 76, 71–76. doi:10.1016/j.ijsu.2020.02.034
- Spiliotopoulos, D., Spitaleri, A., and Musco, G. (2012). Exploring PHD Fingers and H3K4me0 Interactions with Molecular Dynamics Simulations and Binding Free Energy Calculations: AIRE-PHD1, a Comparative Study. *PLoS One* 7, e46902. doi:10.1371/journal.pone.0046902
- Tarasova, A., and Winkler, D. A. (2009). Modelling Atypical Small-Molecule Mimics of an Important Stem Cell Cytokine, Thrombopoietin. *ChemMedChem* 4, 2002–2011. doi:10.1002/cmdc.200900340
- Thanh Le, T., Andreadakis, Z., Kumar, A., Gómez Román, R., Tollefsen, S., Saville, M., et al. (2020). The COVID-19 Vaccine Development Landscape. *Nat. Rev. Drug Discov.* 19, 305–306. doi:10.1038/d41573-020-00073-5
- Tian, W., Chen, C., Lei, X., Zhao, J., and Liang, J. (2018). CASTp 3.0: Computed Atlas of Surface Topography of Proteins. *Nucleic Acids Res.* 46, W363–W367. doi:10.1093/nar/gky473
- Vafaie, S., Razmi, M., Mansoori, M., Asadi-Lari, M., and Madjd, Z. (2020). Spotlight of Remdesivir in Comparison with Ribavirin, Favipiravir,

- Oseltamivir and Umifenovir in Coronavirus Disease 2019 (COVID-19) Pandemic. *SSRN Electronic Journal*. doi:10.2139/ssrn.3569866
- Verdugo-Paiva, F., Izcovich, A., Ragusa, M., and Rada, G. (2020). Lopinavir-ritonavir for COVID-19: A Living Systematic Review. *Medwave* 20, e7967. doi:10.5867/medwave.2020.06.7966
- Vogel, J.-U., Schmidt, S., Schmidt, D., Rothweiler, F., Koch, B., Baer, P., et al. (2019). The Thrombopoietin Receptor Agonist Eltrombopag Inhibits Human Cytomegalovirus Replication via Iron Chelation. *Cells* 9, 31. doi:10.3390/cells9010031
- Wang, M., Cao, R., Zhang, L., Yang, X., Liu, J., Xu, M., et al. (2020a). Remdesivir and Chloroquine Effectively Inhibit the Recently Emerged Novel Coronavirus (2019-nCoV) *In Vitro*. *Cell Res* 30, 269–271. doi:10.1038/s41422-020-0282-0
- Wang, Y., Zhang, D., Du, G., Du, R., Zhao, J., Jin, Y., et al. (2020b). Remdesivir in Adults with Severe COVID-19: a Randomised, Double-Blind, Placebo-Controlled, Multicentre Trial. *Lancet* 395, 1569–1578. doi:10.1016/s0140-6736(20)31022-9
- Whitworth, J. (2020). COVID-19: a Fast Evolving Pandemic. *Trans. R. Soc. Trop. Med. Hyg.* 114, 241–248. doi:10.1093/trstmh/traa025
- Wilkinson, T., Dixon, R., Page, C., Carroll, M., Griffiths, G., Ho, L.-P., et al. (2020). ACCORD: A Multicentre, Seamless, Phase 2 Adaptive Randomisation Platform Study to Assess the Efficacy and Safety of Multiple Candidate Agents for the Treatment of COVID-19 in Hospitalised Patients: A Structured Summary of a Study Protocol for a Randomised Controlled Trial. *Trials* 21, 691. doi:10.1186/s13063-020-04584-9
- Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., et al. (2018). DrugBank 5.0: A Major Update to the DrugBank Database for 2018. *Nucleic Acids Res.* 46 (D1), D1074–D1082. doi:10.1093/nar/gkx1037
- Wu, C., Liu, Y., Yang, Y., Zhang, P., Zhong, W., Wang, Y., et al. (2020). Analysis of Therapeutic Targets for SARS-CoV-2 and Discovery of Potential Drugs by Computational Methods. *Acta Pharmaceut. Sin.* 10, 766–788. doi:10.1016/j.apsb.2020.02.008
- Xiao, X., Wang, C., ChangWang, D., Wang, Y., Dong, X., Jiao, T., et al. (2020). Identification of Potent and Safe Antiviral Therapeutic Candidates against SARS-CoV-2. *Front. Immunol.* 11, 586572. doi:10.3389/fimmu.2020.586572
- Yamamoto, N., Matsuyama, S., Hoshino, T., and Yamamoto, N. (2020). Nelfinavir Inhibits Replication of Severe Acute Respiratory Syndrome Coronavirus 2 in Vitro. *bioRxiv*. doi:10.1101/2020.04.06.026476
- Yavuz, S. S., and Ünal, S. (2020). Antiviral Treatment of Covid-19. *Turk. J. Med. Sci.* 50, 611–619. doi:10.3906/sag-2004-145
- Zeng, J., Weissmann, F., Bertolin, A. P., Posse, V., Canal, B., Ulferts, R., et al. (2021). Identifying SARS-CoV-2 Antiviral Compounds by Screening for Small Molecule Inhibitors of Nsp13 Helicase. *Biochem. J.* 478, 2405–2423. doi:10.1042/bcj20210201
- Zhang, J.-J., Shen, X., Yan, Y.-M., Wang, Y., and Cheng, Y.-X. (2020b). Discovery of Anti-SARS-CoV-2 Agents from Commercially Available Flavor via Docking Screening. doi:10.31219/osf.io/vjch2
- Zhang, J., Zeng, H., Gu, J., Li, H., Zheng, L., and Zou, Q. (2020a). Progress and Prospects on Vaccine Development against SARS-CoV-2. *Vaccines* 8, 153. doi:10.3390/vaccines8020153
- Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., et al. (2020c). Crystal Structure of SARS-CoV-2 Main Protease Provides a Basis for Design of Improved α -ketoamide Inhibitors. *Science* 368, 409–412. doi:10.1126/science.abb3405
- Zhang, T., He, Y., Xu, W., Ma, A., Yang, Y., and Xu, K.-F. (2020d). Clinical Trials for the Treatment of Coronavirus Disease 2019 (COVID-19): A Rapid Response to Urgent Need. *Sci. China Life Sci.* 63, 774–776. doi:10.1007/s11427-020-1660-2
- Zhu, W., Xu, M., Chen, C. Z., Guo, H., Shen, M., Hu, X., et al. (2020). Identification of SARS-CoV-2 3CL Protease Inhibitors by a Quantitative High-Throughput Screening. *ACS Pharmacol. Transl. Sci.* 3 (5), 1008–1016. doi:10.1021/acspsci.0c00108
- Zoete, V., Cuendet, M. A., Grosdidier, A., and Michielin, O. (2011). SwissParam: a Fast Force Field Generation Tool for Small Organic Molecules. *J. Comput. Chem.* 32, 2359–2368. doi:10.1002/jcc.21816

Conflict of Interest: SP, PS, and NP were employed by Vaxine Pty Ltd.

The remaining 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.

Copyright © 2022 Piplani, Singh, Petrovsky and Winkler. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.



Antigen–Antibody Complex-Guided Exploration of the Hotspots Conferring the Immune-Escaping Ability of the SARS-CoV-2 RBD

Kit-Man Fung^{1,3}, Shu-Jung Lai², Tzu-Lu Lin³ and Tien-Sheng Tseng^{3*}

¹Academia Sinica, Institute of Biological Chemistry, Taipei, Taiwan, ²Graduate Institute of Biomedical Sciences, China Medical University, Taichung, Taiwan, ³Institute of Molecular Biology, National Chung Hsing University, Taichung, Taiwan

OPEN ACCESS

Edited by:

James Leland Olds,
George Mason University,
United States

Reviewed by:

Jie Hu,
Chongqing Medical University, China
Rajat Desikan,
Certara United Kingdom Limited,
United Kingdom

*Correspondence:

Tien-Sheng Tseng
emersonseng@
dragon.nchu.edu.tw

Specialty section:

This article was submitted to
Biological Modeling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 18 October 2021

Accepted: 02 March 2022

Published: 22 March 2022

Citation:

Fung K-M, Lai S-J, Lin T-L and
Tseng T-S (2022) Antigen–Antibody
Complex-Guided Exploration of the
Hotspots Conferring the Immune-
Escaping Ability of the SARS-CoV-
2 RBD.
Front. Mol. Biosci. 9:797132.
doi: 10.3389/fmolb.2022.797132

The COVID-19 pandemic resulting from the spread of SARS-CoV-2 spurred devastating health and economic crises around the world. Neutralizing antibodies and licensed vaccines were developed to combat COVID-19, but progress was slow. In addition, variants of the receptor-binding domain (RBD) of the spike protein confer resistance of SARS-CoV-2 to neutralizing antibodies, nullifying the possibility of human immunity. Therefore, investigations into the RBD mutations that disrupt neutralization through convalescent antibodies are urgently required. In this study, we comprehensively and systematically investigated the binding stability of RBD variants targeting convalescent antibodies and revealed that the RBD residues F456, F490, L452, L455, and K417 are immune-escaping hotspots, and E484, F486, and N501 are destabilizing residues. Our study also explored the possible modes of actions of emerging SARS-CoV-2 variants. All results are consistent with experimental observations of attenuated antibody neutralization and clinically emerging SARS-CoV-2 variants. We identified possible immune-escaping hotspots that could further promote resistance to convalescent antibodies. The results provide valuable information for developing and designing novel monoclonal antibody drugs to combat emerging SARS-CoV-2 variants.

Keywords: SARS-CoV-2, COVID-19, binding stability, hotspots, neutralization, convalescent antibody, immunity

INTRODUCTION

The coronavirus disease 2019 (COVID-19) pandemic resulting from the spread of severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2) led to social, health, and economic crises worldwide (Lai et al., 2020). By early June 2021, SARS-CoV-2 had already infected 176.4 million people and caused 3.8 million deaths. SARS-CoV-2 infections presents through common symptoms such as a hacking cough, sore throat, nasal stuffiness, diarrhea, respiratory illness, and fever (Baj et al., 2020; Pal et al., 2020). Additionally, severe acute respiratory syndrome, pneumonia, and kidney failure were observed in severe cases and were fatal (Lai et al., 2020). Coronaviruses are a family of RNA virus with a genome comprising approximately 30 kilobases and can be divided into four genera (Fehr and Perlman, 2015). SARS-CoV-1 and SARS-CoV-2 belong to the β -genus. In SARS-CoV-2, the structural proteins, spike (S), nucleocapsid (N), envelope (E), and membrane (M), are encoded by its RNA genome (Malik, 2020; Mousavizadeh and Ghasemi, 2021; V'Kovski et al., 2021). The S protein consists of N-terminal signal peptide S1 and S2 subunits (Lan et al., 2020; Tai et al., 2020); the S1 subunit contains the receptor-binding domain (RBD) and N-terminal domain, and the

S2 subunit is composed of a cytoplasm domain, TM domain, heptapeptide repeat sequence 1 (HR1), heptapeptide repeat sequence 2 (HR2), and fusion peptide (FP; (Xia et al., 2020). Notably, the FP of the S2 subunit is essential for facilitating viral membrane fusion with the host cell membrane (Millet and Whittaker, 2018). In addition, the HR1 and HR2 of the S2 subunit form the six-helix bundle (6-HB), which is critical for membrane fusion and viral entry (Chambers et al., 1990; Xia et al., 2020). In the native state, the S protein exists as an inactive precursor. During viral infection, a host cell protease (TMPRSS2) cleaves the S protein into the S1 and S2 subunits (Bertram et al., 2013; Hoffmann et al., 2020), after which the angiotensin-converting enzyme 2 (ACE2) receptor of the host cell is recognized and bound by the RBD of the S1 subunit (Wrapp et al., 2020). Following binding of the RBD to ACE2, the S2 subunit undergoes conformational changes. The FP enters the host cell membrane, and the prehairpin coiled-coil of the HR1 domain is exposed (Huang et al., 2020). Consequently, the 6-HB triggers HR1 and HR2, drawing the viral and host cellular membrane close enough for fusion (Xia et al., 2018). Remarkably, as a surface-exposed protein essential for entry into the host cell, the RBD is regarded as a first-line therapeutic target for developing vaccines and antiviral agents.

Rapid action and the development of vaccines protecting against COVID-19 are urgently required. Some vaccine candidates are being evaluated in preclinical models, and others are being investigated through human clinical trials (World Health Organization, 2021). Currently, only four vaccines are licensed (Gomez et al., 2021). The pharmaceutical companies Janssen and Oxford University/AstraZeneca developed the Ad26.COV2.S (Sadoff et al., 2021) and AZD1222 (Voysey et al., 2021) vaccines, respectively, based on nonreplicating adenoviruses. The Janssen vaccine was developed from modified adenovirus serotype 26, which generates prefusion S protein to elicit immunity and has an efficacy of 66.9%. The monovalent AstraZeneca vaccine contains a replication-deficient chimpanzee adenovirus (ChAdOx1) that encodes the S protein. When administrated, the expressed S protein stimulates cellular immune response with an efficacy of 63.09%. Moderna (mRNA-1273; (Baden et al., 2021; Oronsky et al., 2021); and Pfizer/BioNTech (BNT162b2; (Polack et al., 2020); are two other vaccines that were developed based on mRNA. The genetic fragment encoding the prefusion form of the S protein enhancing uptake in the immune cells of the host are packed in these vaccines. Thus, the S protein can be produced through the transcription and translation machinery of the host cell, further eliciting adaptive immunity against COVID-19. Two regimens of BNT162b2 and mRNA-1273 exerted 95% (Polack et al., 2020) and 94.1% (Baden et al., 2021; Oronsky et al., 2021) protective efficacy, respectively. Although all of these vaccines can provide a level of protection against SARS-CoV-2, the rates of mutations evident in emerging variants that confer resistance to and immune-escaping ability toward neutralizing antibodies continue to grow (Geers et al., 2021; Gomez et al., 2021).

The devastating effects of the COVID-19 pandemic have increased the urgency to disrupt the spread and replication of SARS-CoV-2. Natural infection or vaccination can elicit

neutralizing antibodies of adaptive immunity against viruses (Rydyznski Moderbacher et al., 2020; Bettini and Locci, 2021; Sette and Crotty, 2021). Alternatively, passive immunity can be conferred when antibodies are administrated in the form of recombinant proteins or convalescent plasma (Casadevall, 2002). SARS-CoV-2 infection can reportedly elicit neutralizing antibodies that potentially recognize the RBD (Barnes et al., 2020; Gavor et al., 2020; Liu et al., 2020). Nevertheless, COVID-19 convalescents exhibit low levels of plasma neutralizing activity, indicating that B-cells' generation of high-titer neutralizing antibodies through natural infection is insufficient (Weisblum et al., 2020b). The RBD of the SARS-CoV-2 S protein binds to ACE2 with high affinity, with neutralizing antibodies targeting the RBD exerting protective effects against infection in both animal models and humans (Dong et al., 2020). However, antigenic evolution has been observed eroding the immunity of neutralizing antibodies (Greaney A. J. et al., 2021; Eguia et al., 2021). The positive selection for mutations of RBD drives the antigenic evolution (Jaroszewski et al., 2020; Velazquez-Salinas et al., 2020). Furthermore, several researchers have demonstrated that mutations of the SARS-CoV-2 S protein contribute to its immune-escaping ability from monoclonal antibodies as well as polyclonal human sera (Weisblum et al., 2020a; Greaney A. J. et al., 2021; Chen et al., 2021; Harvey et al., 2021). Moreover, RBD mutants are emerging and were observed in SARS-CoV-2 pandemic isolates, presenting a considerable challenge for the combat against COVID-19. Thus, to understand the viral evolution, a comprehensive examination of the hotspots of the SARS-CoV-2 RBD that confers the ability to escape convalescent antibodies' recognition is timely and necessary.

In this study, we comprehensively and systematically investigated the RBD variants that markedly destabilize binding to six convalescent neutralizing antibodies, namely CV07-270, B38, CT-P59, CA1-B12, CA1-B3, and 47D1. We utilized the complex structures of the RBD antibody in convalescents to conduct in-depth mutational scanning to estimate their binding stability. During calculation, each residue of the RBD contributing to antibody interaction was replaced with distinct amino acids to examine the hotspots that impair binding. Our results demonstrated that mutations at the residues of F456, F490, G416, G502, K417, L452, L455, N487, R403, Y449, and Y489 were unfavorable for RBD binding to most of the convalescent antibodies. Consistently, the K417 T/N, L452R, and E484K/Q mutants were identified in emerging SARS-CoV-2 variants. In addition, the emerging N501Y variant was identified in several viral lineages (P.1, C.37, B.1.1.7, and B.1.351). Moreover, mutations at residues F456, F486, F490, G485, and L455 located near the RBD's receptor-binding ridge reportedly have considerable antigenic effects. All of the aforementioned experimental and clinical evidence corroborates our results. Evidently, the RBD hotspots (F456, F490, G416, G502, K417, L452, L455, N487, R403, Y449, and Y489) explored in this study would benefit from the development of potent therapeutics to combat against emerging SARS-CoV-2 variants.

MATERIALS AND METHODS

Preparation of Protein Structures

The complex structure of the RBD antibody was obtained using Protein Data Bank (PDB IDs: 6XKP, 7BZ5, 7CM4, 7KFV, 7KFW, and 7MF1). In addition, the CHARMM polar H forcefield was applied to all structures in advance for computation.

Calculation of Mutational Binding Stability

The binding stability of the RBD variants in complex with convalescent antibodies was determined using MutaBind2 (Weisblum et al., 2020b), mCSM-PPI2 (Rodrigues et al., 2019), FoldX (Schymkowitz et al., 2005), and Discovery Studio (DS) 3.5 (Accelrys, San Diego, CA, United States). The calculations in mCSM-PPI2 and MutaBind2 were performed according to the online instructions. In the FoldX prediction, the binding stability was estimated using the method previously reported (Teng et al., 2021). The Calculate Mutation Energy (Binding) protocol of DS 3.5 was employed to evaluate the changes in binding stability upon mutation. The complex structure of the RBD antibody and RBD moiety were selected as the “input typed molecule” and “ligand chain” parameters, respectively. Additionally, the “mutation sites” parameter was set to a single mutation with all other 20 amino acids. Moreover, the solvent dielectric constant, maximum number of mutants, dielectric constant, and maximum structures to save were set to 80, 25, 10, and 10, respectively; the remaining parameters were default settings. The values of mutational energies denoted destabilized (positive) and stabilized (negative) binding in the predictions of DS, FoldX, and MutaBind2 but were reversed in mCSM-PPI2.

Generation of the Heatmap of Mutational Binding Energy

Excel (Microsoft Office 2013; Microsoft, Redmond, WA, United States) was used to create the heatmap based on the output binding energy values from single mutations. In the heatmap, the *x*-axis represented the types of amino acids for single mutations, and the *y*-axis represented the mutated residues of the RBD. Conditional formatting was applied on the obtained post mutation binding energy values. Green and purple were employed to create the two-color gradient scales that designate the values of the binding energies.

LIGPLOT Analyses

The LIGPLOT program (Wallace et al., 1995; Laskowski and Swindells, 2011) was used to analyze the molecular interactions within the complex structure of the RBD antibody. During the analysis, the “antibody” LIGPLOT module was employed to explore domain–domain and protein–protein interactions.

Molecular Dynamic Simulation and Binding Affinity Prediction

The complex structures of RBD-CA1-B3 (PDB ID: 7KFW), RBD-CA1-B12 (PDB ID: 7KFV), RBD-CT-P59 (PDB ID:

7CM4) and RBD-B38 (PDB ID: 7BZ5) were employed to perform molecular dynamics (MD) simulations. The L455 and F456 variants of RBD complexed with convalescent antibodies were generated by Build and Edit Protein module of Discovery Studio (DS) 3.5 (Accelrys, San Diego, CA, United States). Additionally, the covalently attached carbohydrates of glycosylated residues of RBD were removed before solvation step. The generated structures were firstly subjected to Solvation (Discovery Studio 3.5) with orthorhombic cell shape under CHARMM forcefield (Supplementary Table S1). Subsequently, the complex structures of RBD-convalescent antibodies were solvated with waters molecules (14,170–15,280), sodium atoms (40–41), and chloride atoms (49–52). Consequently, the Standard Dynamics Cascade (Discovery Studio 3.5) were conducted for each solvated complex structure for 5 ns simulation times with 2 ps as save results interval (detailed parameter setting was shown in Supplementary Table S2). Furthermore, the Analyze Trajectory tool (Discovery Studio 3.5) was employed to plot the total energy changes as functions of simulation time. After MD simulations, the binding affinities of RBD variants to convalescent antibodies were determined by CSM-AB, a machine learning method capable of predicting antibody–antigen binding affinity by modelling interaction interfaces as graph-based signatures (Myung et al., 2021).

RESULTS

Deviations in the Binding Stability of RBD Variants Targeting Convalescent Antibody CV07-270

To investigate the mutational effects of the RBD on its interactions with convalescent antibodies, we obtained the three-dimensional structures of the RBD antibody complex from Protein Data Bank. The convalescent antibody CV07-270 isolated from patients with COVID-19 has high affinity to the SARS-CoV-2 S protein ($IC_{50} = 82.3$ ng/ml; K_D is not applicable) (Kreye et al., 2020). The complex structure of RBD–CV07-270 is depicted in Figure 1A, with the LIGPLOT of its molecular interactions presented in Figure 2A. The RBD residues R346, S349, K444, G447, N448, Y449, N450, and E484 primarily form hydrogen bonds, and residues Y351, G446, Y451, L452, T470, F490, L492, and S494 achieve hydrophobic contact with CV07-270. Next, these residues were replaced with distinct amino acids to evaluate their mutational effects on targeting the CV07-270 antibody. The output binding energies upon mutation were used to generate the heatmap, on which the *x*-axis and *y*-axis indicate the amino acid types for various mutations and mutated residues, respectively. The values of the binding energies are colored purple (destabilized) and green (stabilized) within a gradient range. The results demonstrated that single-point mutation at residues E484, F490, G447, L452, L492, N448, N450, S439, S494, Y449, and Y451 were largely unfavorable for binding to antibodies (Figure 2B and Table 1).

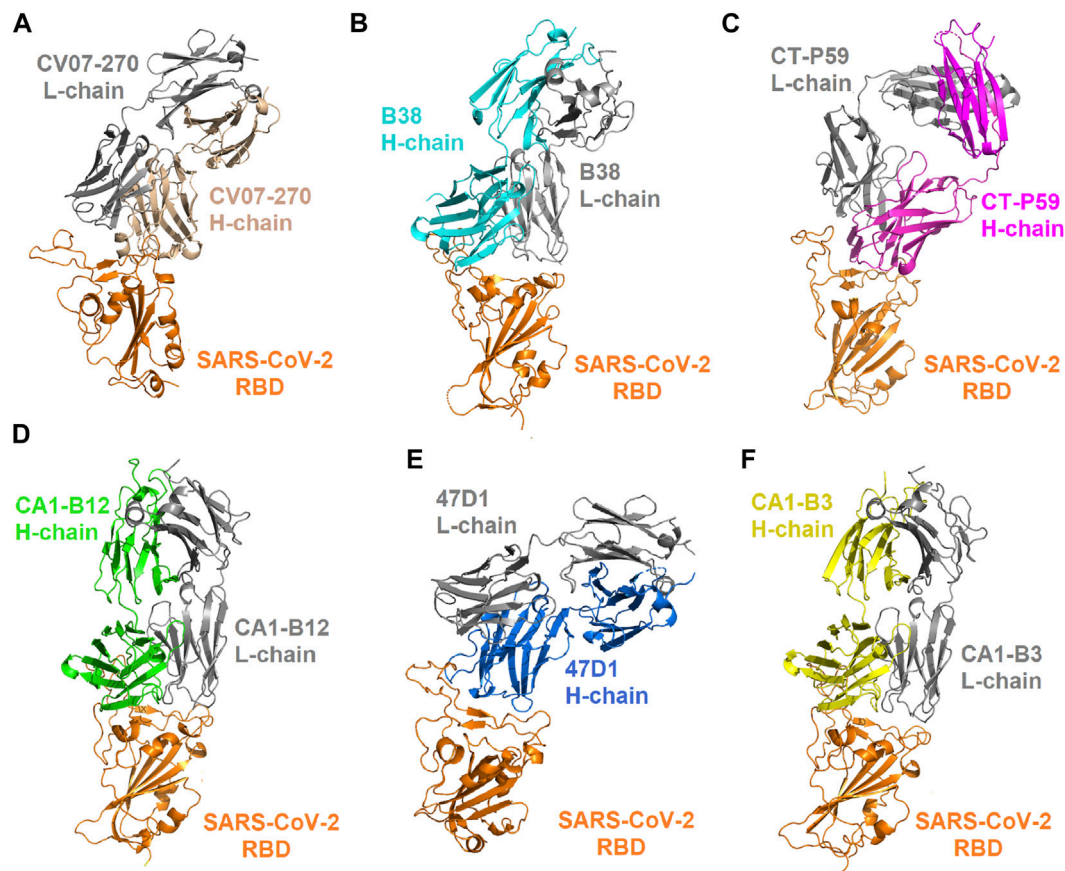


FIGURE 1 | Complex structures of the SARS-CoV-2 RBD bound with convalescent antibodies. **(A)** The complex structure of the RBD–CV07-270 antibody (PDB ID: 6XKP). **(B)** The structure of the RBD in complex with the B38 antibody (PDB ID: 7BZ5). **(C)** The structure of the RBD bound with the CT-P59 antibody (PDB ID: 7CM4). **(D)** The structure of the CA1-B12 antibody in complex with the RBD (PDB ID: 7KFV). **(E)** The structure of the 47D1 antibody bound with the RBD (PDB ID: 7MF1). **(F)** The structure of the CA1-B3 antibody in complex with the RBD (PDB ID: 7KFW).

Mutational Scanning of the RBD Targeting Convalescent Antibody B38

Distinct complex structures were employed to analyze the effects of mutations on RBD binding to convalescent antibodies. The structure of the RBD in complex with a human-origin monoclonal antibody B38 ($K_D = 14.3$ nM) (Wu et al., 2020) from a convalescent patient is illustrated in **Figure 1B** (Wu et al., 2020). **Figure 3** depicts the LIGPLOT of the RBD residues T415, R403, K417, D420, Y421, L455, R457, K458, Y473, A475, N487, Y495, G496, Q498, N501, G502, and Y505 that form hydrogen bonds and that of D405, G416, F456, N460, G476, F486, Y489, Q493, and T500 that made hydrophobic contact with B38 (**Figure 3**). These functionally essential residues were mutated to evaluate the effects on RBD binding to B38, with the results demonstrating that mutations at the residues R403, G416, K417, D420, Y421, L455, F456, Y473, F486, N487, Y489, Y495, G496, N501, G502, and Y505 were unfavorable for binding to the B38 convalescent antibody (**Figure 4** and **Table 1**).

Mutational Binding Stability of RBD Variants Targeting Convalescent Antibody CT-P59

The CT-P59 antibody neutralizes SARS-CoV-2 isolates through blocking of the interaction regions of the RBD ($K_D = 0.05$ nM) for ACE2 (Kim et al., 2021; Ryu et al., 2021). We assessed the destabilizing ability of RBD variants in binding to convalescent antibody CT-P59. The structure of the CT-P59 antibody in complex with the RBD is illustrated in **Figure 1C**. The molecular interactions analyzed using LIGPLOT demonstrated that hydrogen bond interactions were formed between the RBD residues E484, F486, N450, Q493, R403, S494, and Y453 and CT-P59 (**Figure 5A**). Additionally, the residues F456, F490, G485, L452, L455, L492, K417, Y449, Y489, and Y505 exhibited hydrophobic interaction with CT-P59. Therefore, RBD variants from the single-point mutations of the aforementioned residues were assessed in terms of binding stability. The results indicated that the residues F456, F486, F490, G485, L452, L455, L492, Q493, S494, Y449, and Y453 were unfavorable for binding to CT-P59 postmutation (**Figure 5B** and **Table 1**).

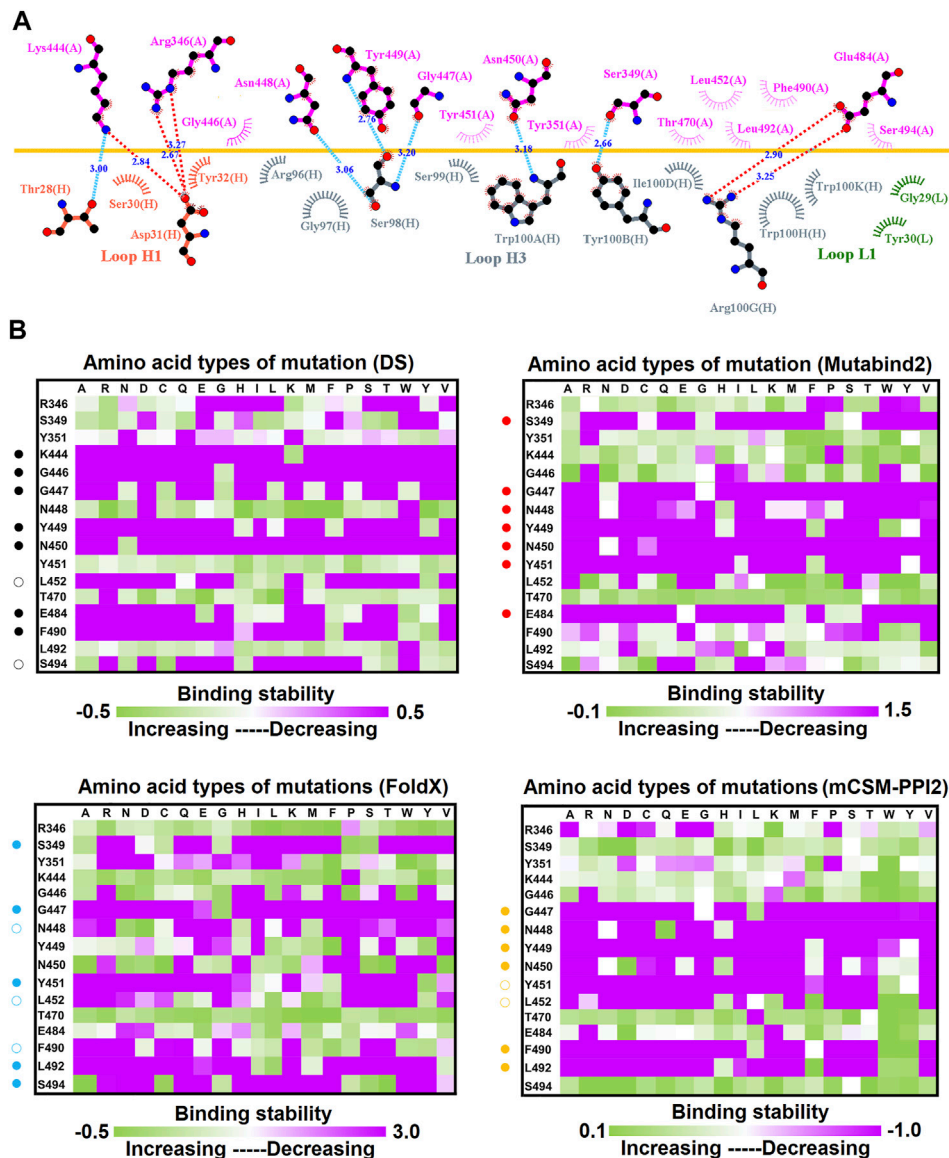


FIGURE 2 | Mutational heatmaps of the RBD targeting antibody CV07-270. **(A)** The LIGPLOT presents the interaction network between CV07-270 and the SARS-CoV-2 RBD. The RBD and light and heavy chains of CV07-270 are labeled A, H, and L, respectively. The golden yellow line denotes the interface of the RBD–CV07-270 complex. The interactive residues of the RBD are labeled and colored magenta. The salt-bridges and hydrogen bonds are indicated by red and cyan dashes, and the hydrophobic interactions are represented by arcs with spokes radiating toward the ligand atoms with which they are in contact. **(B)** The mutational binding stabilities of RBD variants to CV07-270 were estimated using FoldX, MutaBind2, DS 3.5, and mCSM-PP12. The unit of binding stability was the kcal/mole, with the obtained values employed to create the heatmaps. Green (increased stability) and purple (reduced stability) were used to generate a color gradient and were applied in each box. In each heatmap, the moderately and significantly reduced stabilities were labeled with hollow and solid circles, respectively.

Single-Point Mutations of the RBD Interfered With its Binding to Convalescent Antibodies C1A-B12 and C1A-B3

C1A-B12 ($K_D = 4$ nM) and C1A-B3 ($K_D = 70.6$ nM), potent neutralizing antibodies targeting SARS-CoV-2, were isolated from COVID-19 convalescents (Clark et al., 2020). **Figure 1D** depicts the complex structure of RBD–C1A-B12. The detailed molecular interactions of C1A-B12 to RBD were analyzed and are presented in **Figure 6**. Structurally, RBD interacted with

C1A-B12 through the formation of hydrogen bonds (R403, T415, K417, Y421, Y453, L455, R457, Y473, A475, N487, Q493, S494, Y495, G496, Q498, N501, G502, and Y505) and hydrophobic contact (G416, D420, F456, K458, N460, G476, F486, Y489, and T500). The single-point mutations at these interactive RBD residues were examined in terms of their binding stability. As presented in **Figure 7** and **Table 1**, the mutations at residues F456, G416, G496, K417, L455, N487, N501, Q498, R403, Y421, Y453, Y489, Y495, and Y505 did not promote favorable binding to C1A-B12. The mutational effects of the RBD variants on the

TABLE 1 | Functionally essential residues of the RBD mutations that confer the immune-escaping ability of SARS-CoV-2 in relation to neutralizing convalescent antibodies.

Destabilizing mutations on SARS-CoV-2 RBD																					
Predictions validated with experiments											Predictions can be tested by future experiments										
Antibodies	K417*	L452*	E484 ^C	F490 ^{C*}	S494	N501	R403*	G416*	Y421*	Y449*	Y453*	L455 ^{C*}	F456 ^{C*}	F486 ^C	N487*	Y489*	L492	Y495	G496	G502*	Y505
CV07-270		O	O	O	O					O							O				
B38	O					O	O	O	O			O	O	O	O	O		O	O	O	O
CT-P59		O		O	O					O	O	O	O	O			O		O	O	
CA1-B12	O					O	O	O	O		O	O	O		O	O		O	O	O	
CA1-B3	O						O	O	O		O	O	O		O	O				O	O
47D1		O	O	O						O											
Emerging SARS-CoV-2 Variants																					
^b United Kingdom (Nigeria); B.1.525			E484K																		
^b United Kingdom B.1.1.7			E484K		S494P	N501Y															
^b United States (New York); B.1.526			E484K																		
^b United States (California) B.1.427, B.1.429		L452R																			
^b United States (New York); B.1.526.1		L452R																			
^b India B.1.617, B.1.617.1																					
B.1.617.2, B.1.617.3		L452R	E484Q																		
^b Brazil, P.1	K417N		E484K			N501Y															
^b Brazil, P.2			E484K																		
^b South Africa, B.1.351	K417N		E484K			N501Y															
^b Japan, P.1	K417T		E484K			N501Y															
^c Peru, C.37		L452R		F490S		N501Y															

^aHotspots.^b[U.S. Department of Health and Human Services (<https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/variant-surveillance/variant-info.html>)].^c[European Centre for disease Prevention and Control (<https://www.ecdc.europa.eu/en/covid-19/variants-concern>)].^dMutations at other structurally adjacent sites in the RBD's receptor-binding ridge can also have substantial antigenic effects.

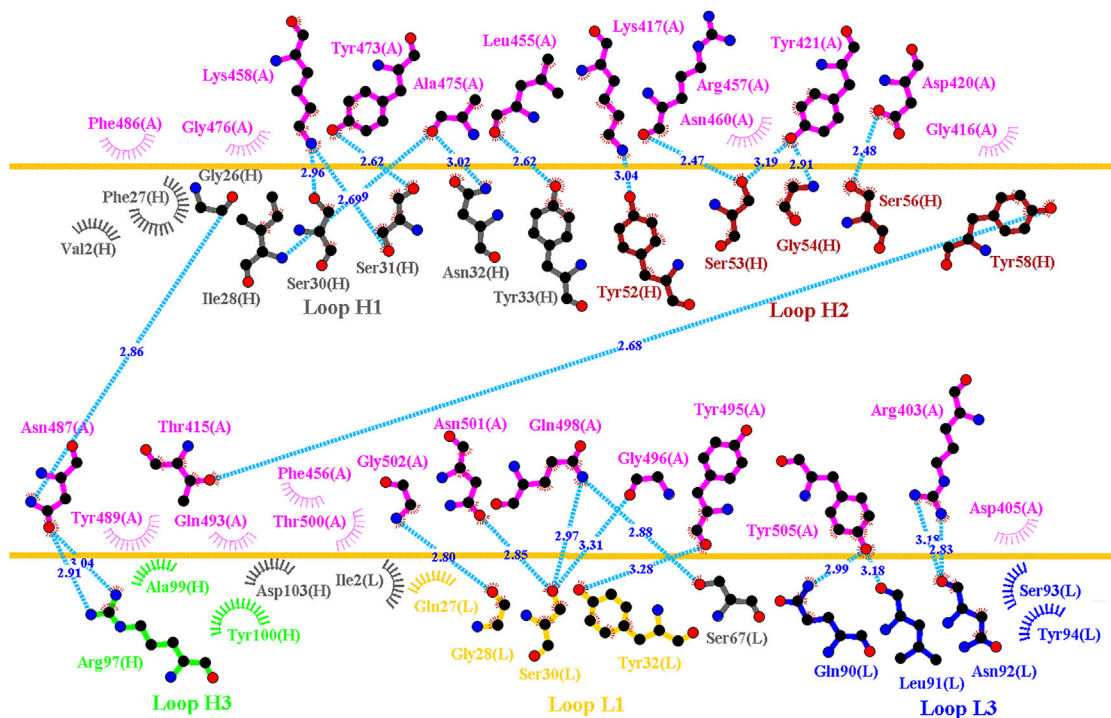


FIGURE 3 | Interactional network of the RBD and B38 antibody analyzed using LIGPLOT. LIGPLOT was employed to examine the interaction network of the SARS-CoV-2 RBD and B38. The RBD and light and heavy chains of B38 are labeled A, H, and L, respectively. The golden yellow line denotes the interface of the RBD–B38 complex. The interactive residues of the RBD are labeled and colored magenta. The hydrogen bonds and hydrophobic interactions are represented by cyan dashes and arcs with spokes radiating toward the ligand atoms with which they are in contact, respectively.

interactions with the C1A-B3 antibody were also investigated. The structure of the C1A-B3 antibody in complex with the RBD is illustrated in **Figure 1E**. As revealed through LIGPLOT analysis (**Figure 8**), the epitope residues of the RBD were A475, D420, G496, G502, K417, L455, N487, Q493, R403, R408, Y421, Y453, and Y473 (forming hydrogen bonds) and residues T415, G416, F456, R457, K458, N460, Q474, G476, Y489, Q498, T500, N501, and Y505 (through hydrophobic interaction). Accordingly, the binding stability of the RBD toward C1A-B3 was systematically analyzed through the individual mutation of these interactive residues. We observed that the single-point mutation at residues R403, G416, K417, Y421, Y453, L455, F456, N487, Y489, G502, and Y505 led to the destabilized binding of C1A-B3 to RBD (**Figure 9** and **Table 1**).

Effect of Mutation on the Binding Stability of the SARS-CoV-2 RBD Targeting Convalescent Antibody 47D1

The neutralization of the convalescent antibody 47D1 (2 nM) was achieved through the hindering of RBD–ACE2 binding (Zhou X. et al., 2021). The crystal structure of RBD–47D1 is presented in **Figure 1E**. The LIGPLOT analysis revealed that the RBD residues T470, G482, E484, and S494 interact with 47D1 through hydrogen bonding, whereas the RBD residues R346, Y351, Y449, N450, L452, I472, V483, N481, and F490 hydrophobically interact with 47D1 (**Figure 10A**). Single-point

mutation of the aforementioned residues was performed, followed by determination of their binding stability. The result indicated that the mutations on residues Y449, L452, I472, G482, V483, E484, and F490 were largely unfavorable for binding to 47D1 (**Figure 10B**).

DISCUSSION

The novel virus SARS-CoV-2 spread rapidly around the globe, leading to unprecedented health and economic crises. The global SARS-CoV-2 pandemic and its devastating effects are ongoing owing to the paucity of effective therapeutics. COVID-19 infection is mainly mediated through binding of the RBD of SARS-CoV-2 to ACE2 in human cells (Walls et al., 2020; Wang et al., 2020; Zhou et al., 2020). Consequently, the functionally essential RBD is a key target in the development of drugs and vaccines. Several neutralizing antibodies working against SARS-CoV-2 have been developed (Jiang et al., 2020; Pinto et al., 2020; Rogers et al., 2020; Xiaojie et al., 2020; Zhou X. et al., 2021; Lau et al., 2021; Lu et al., 2021), and neutralizing antibodies targeting SARS-CoV-2 have been isolated from convalescent patients (Clark et al., 2020; Hoffmann et al., 2020; Kreye et al., 2020; Wang P. et al., 2021; Zhou D. et al., 2021; Zhou X. et al., 2021; Kim et al., 2021; Planas et al., 2021; Yuan et al., 2021). Most of the antibodies targeting the RBD prevent the entry and replication of SARS-CoV-2 (Huo et al., 2020; Ju et al., 2020; Pinto et al., 2020;

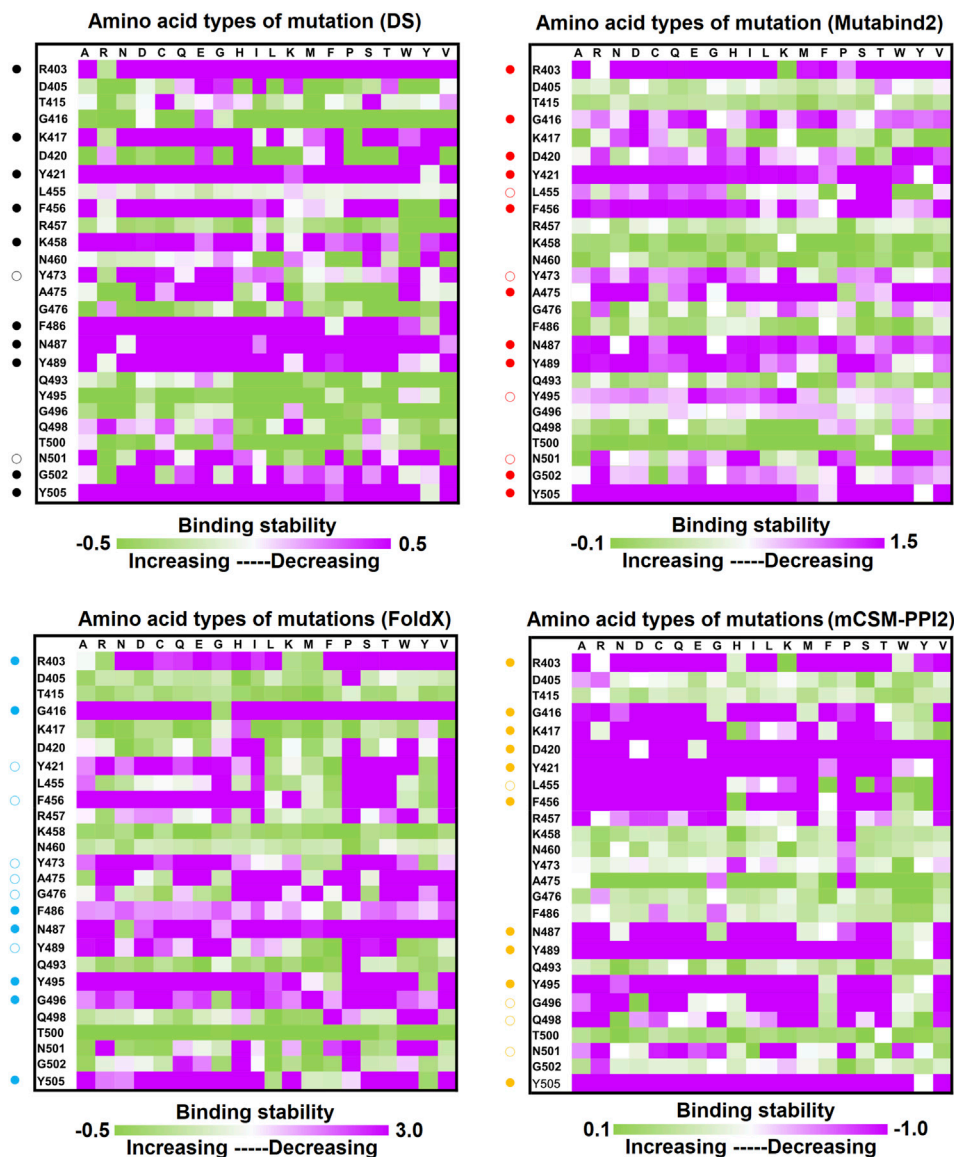


FIGURE 4 | Mutational heatmaps of the RBD targeting antibody B38. The mutational binding stabilities of RBD variants to B38 were estimated using FoldX, MutaBind2, DS 3.5, and mCSM-PPI2. The unit of binding stability was kcal/mole, and the obtained values were employed to create the heatmaps. Green (increased stability) and purple (reduced stability) were used to generate a color gradient and were applied in each box. In each heatmap, the moderately and significantly reduced stabilities are labeled with hollow and solid circles, respectively.

Rogers et al., 2020; Xiaojie et al., 2020; Zhou X. et al., 2021; Lu et al., 2021; Rapp et al., 2021). Moreover, the binding level of the RBD to antibodies in the serum is correlated with the neutralizing activity in patients with COVID-19 (Wang P. et al., 2021). However, the evolving rate of SARS-CoV-2 in the current pandemic equates to one to two mutations per month (Callaway, 2020). RBD mutations in particular allow SARS-CoV-2 and circulating strains to escape antibody neutralization (Andreano et al., 2020; Weisblum et al., 2020a; Huo et al., 2020; Starr et al., 2020; Greaney A. J. et al., 2021; Garcia-Beltran et al., 2021b; Greaney A. J. et al., 2021; Eguia et al., 2021). Thus, to successfully combat these immune-escaping

variants, the potential RBD mutations that destabilize binding with convalescent antibodies must be investigated.

In our previous study, we made efforts to identify the important residue mutations on SARS-CoV-2 RBD which play critical roles in eroding the neutralizing immunities through computational analyses (Tsai et al., 2021). We analyzed putative mutational effects of RBD on binding to two developed nanobodies (H11-D4 (PDB ID: 6YZ5) and VH1-2-15 (PDB ID: 7L5B)), two synthetic nanobodies (MR17 (PDB ID: 7C8W) and SR4 (PDB ID: 7C8V)) and one Fab (P2B-2F6 (PDB ID: 7BWJ)). We found that the interactive residues of RBD (Y449, L452, L455, E484, Y489, F490, L492, Q493, and S494) can be

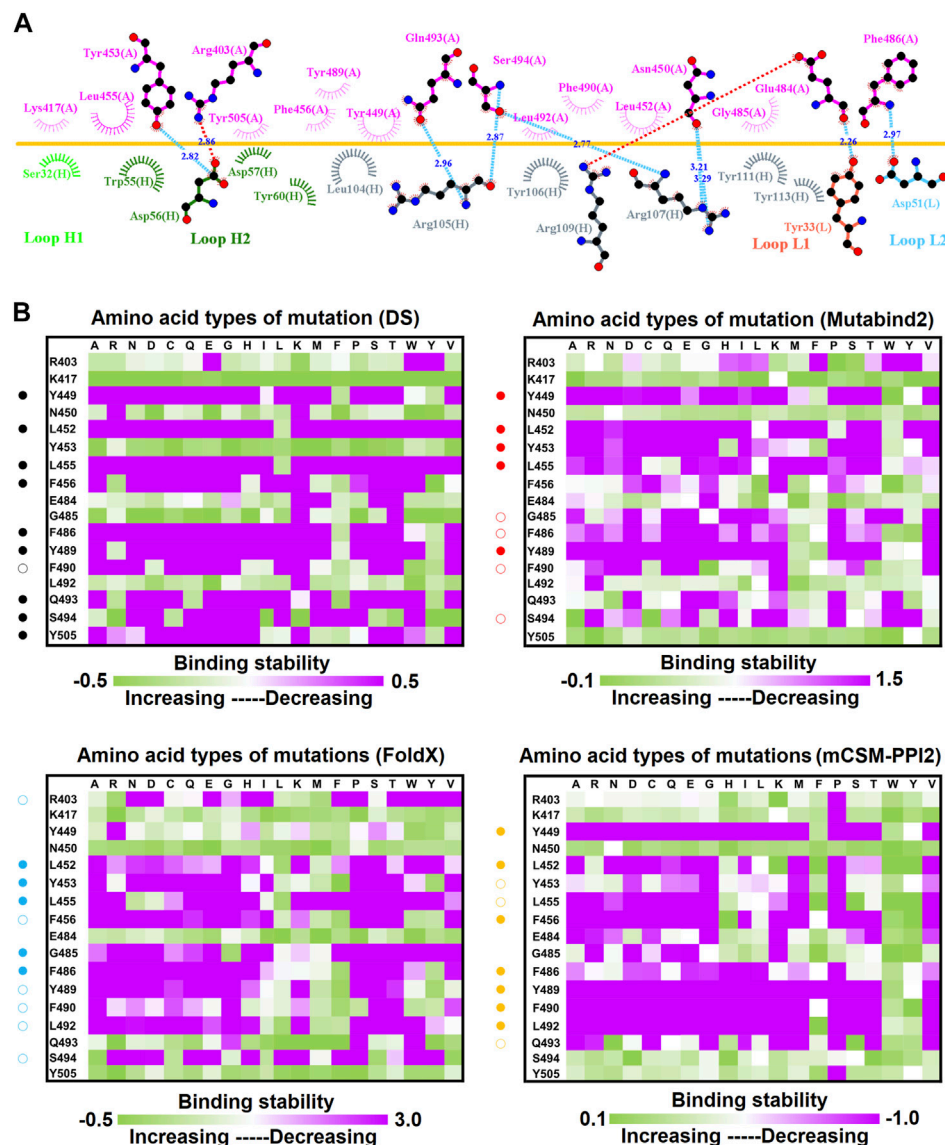


FIGURE 5 | Mutational heatmaps of the RBD targeting antibody CT-P59. **(A)** LIGPLOT was employed to examine the interaction network of CT-P59 and SARS-CoV-2 RBD. The RBD and light and heavy chains of CT-P59 were labeled A, H, and L, respectively. The golden yellow line denotes the interface of the RBD-CT-P59 complex. The interactive residues of the RBD are labeled and colored magenta. The salt-bridges, hydrogen bonds and hydrophobic interactions are represented by red, cyan dashes and arcs with spokes radiating toward the ligand atoms with which they contact. **(B)** The mutational binding stabilities of RBD variants to CT-P59 were estimated using FoldX, MutaBind2, DS 3.5, and mCSM-PPI2. The unit of binding stability was kcal/mole, and the obtained values were employed to create the heatmaps. Green (increased stability) and purple (reduced stability) were used to generate a color gradient and were applied in each box. In each heatmap, the moderately and significantly reduced stabilities were labeled with hollow and solid circles, respectively.

hotspots conferring the ability to neutralizing antibody escape. These results provide valuable information of the mutational effects of RBD variants on interacting with developed and synthetic SARS-CoV-2-neutralizing antibodies. In addition to these man-made neutralizing antibodies, there are antibodies determined from COVID-19 convalescent patients (Piccoli et al., 2020; Greaney A. J. et al., 2021; Rosati et al., 2021; Tada et al., 2021). Especially, several potent SARS-CoV-2-neutralizing monoclonal antibodies were isolated from convalescent individuals (Clark et al., 2020; Kreye et al., 2020; Wu et al.,

2020; Zhou X. et al., 2021; Kim et al., 2021; Xie et al., 2021). These monoclonal antibodies target the RBD of SARS-CoV-2 and compete with its binding to ACE2 preventing virus entry and replication. However, growing evidence show that escape mutations, reducing the neutralizing activity of antibodies in the convalescent plasma of COVID-19 patients, could weaken the effectiveness of antibodies and vaccines under developments (Weisblum et al., 2020a; Greaney A. J. et al., 2021; Wang Z. et al., 2021; Wibmer et al., 2021). To develop potent antiviral prophylaxis combating the SARS-CoV-2 variants which could

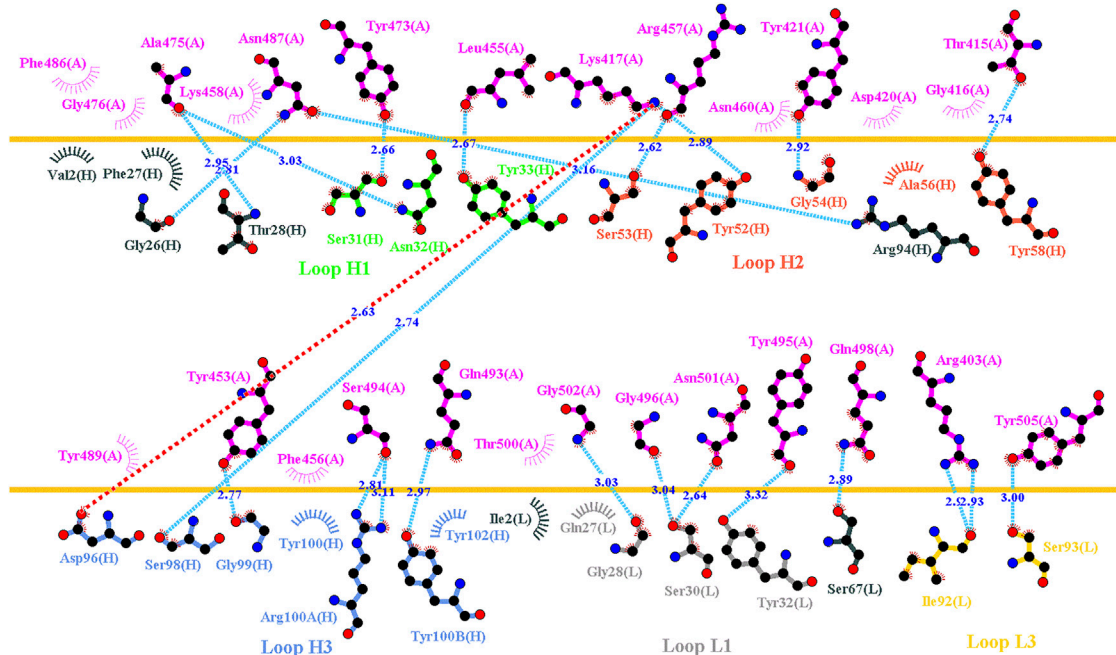


FIGURE 6 | Interaction network of the RBD and CA1-B12 antibody analyzed using LIGPLOT. LIGPLOT was employed to examine the interaction network of the SARS-CoV-2 RBD and CA1-B12. The RBD and light and heavy chains of CA1-B12 were labeled A, H, and L, respectively. The golden yellow line denotes the interface of the RBD–CA1-B12 complex. The interactive residues of the RBD are labeled and colored magenta. The salt-bridges, hydrogen bonds and hydrophobic interactions are represented by red, cyan dashes and arcs with spokes radiating toward the ligand atoms with which they are contact, respectively.

lead neutralization escape from the convalescent antibodies, it is imperative to investigate the mutational effects of SARS-CoV-2 RBD on binding to these convalescent antibodies. Nowadays, the complex structures of RBD-convalescent antibodies are available; this make it feasible to explore the mutations of RBD rendering the capability of neutralizing antibody escape. Therefore, here we aim to identify the possible immune escape hotspots of RBD targeting convalescent antibodies by using similar methodology of our previous work.

In this study, we systematically and comprehensively investigated the changes in the binding stability of RBD variants to convalescent antibodies. The results revealed that single-point mutations at the RBD residues E484, F490, G447, L452, L492, N448, N450, S439, S494, Y449, and Y451 considerably destabilized the interactions with the CV07-270 convalescent antibody (**Figure 2** and **Supplementary Tables S3, S9, S15, S21**). Furthermore, the RBD residues S349, G447, N448, Y449, N450, and E484 mainly interacted with CV07-270 through hydrogen bonding; L452, F490, L492, and S494 interacted with the antibody through hydrophobic contact. Similarly, the binding stability to the convalescent antibody B38 was disrupted when RBD residues F456, F486, Y489 (contributing to hydrophobic contact), R403, K417, D420, Y421, L455, Y473, N487, G496, N501, G502, and Y505 (forming hydrogen bonds) were mutated through replacement with other amino acids (**Figure 4** and **Supplementary Tables S4, S10, S16, S22**). Some RBD mutations interfered with its interaction to the CT-P59 convalescent antibody (**Figure 5**

and **Supplementary Tables S5, S11, S17, S23**); the residues Y453, F486, and Q493 were mostly connected through hydrogen bonding, whereas Y449, L452, F456, G485, F490, and L492 were connected through hydrophobic interaction. Moreover, the amino acid replacements impaired RBD binding to the C1A-B12 convalescent antibody. The mutations at G416, F456, 489 (contributing to hydrophobic contact), R403, K417, Y421, Y453, L455, N487, Y495, G496, N501, and Y505 (forming hydrogen bonds) significantly disrupted the binding stability (**Figure 6** and **Supplementary Tables S6, S12, S18, S24**). The mutations at R403, K417, Y421, Y453, L455, N487, G502 (essentially hydrogen bonds), G416, F456, Y489, and Y505 (mostly hydrophobic interactions) disrupted the interactions between the RBD and C1A-B3 convalescent antibody (**Figure 9** and **Supplementary Tables S7, S13, S19, S25**). We also investigated the changes in the binding of the convalescent antibody 47D1 to the RBD variants. The RBD residues G482 and E484 interacting with 47D1 primarily through hydrogen bonding, and residues Y449, L452, I472, V483, and F490 connecting with 47D1 through hydrophobic interaction disrupted the binding to the antibody when subjected to single-point mutations (**Figure 10** and **Supplementary Tables S8, S14, S20, S26**). Furthermore, we analyzed and integrated the key RBD residues with single-point mutations that destabilize binding to most of the convalescent antibodies. We determined that the RBD residues E484, F456, F486, F490, G416, G496, G502, K417, L452, L455, L492, N487, N501, R403, S494, Y421, Y449, Y489, Y495, and Y505 were prone to disrupt interaction with

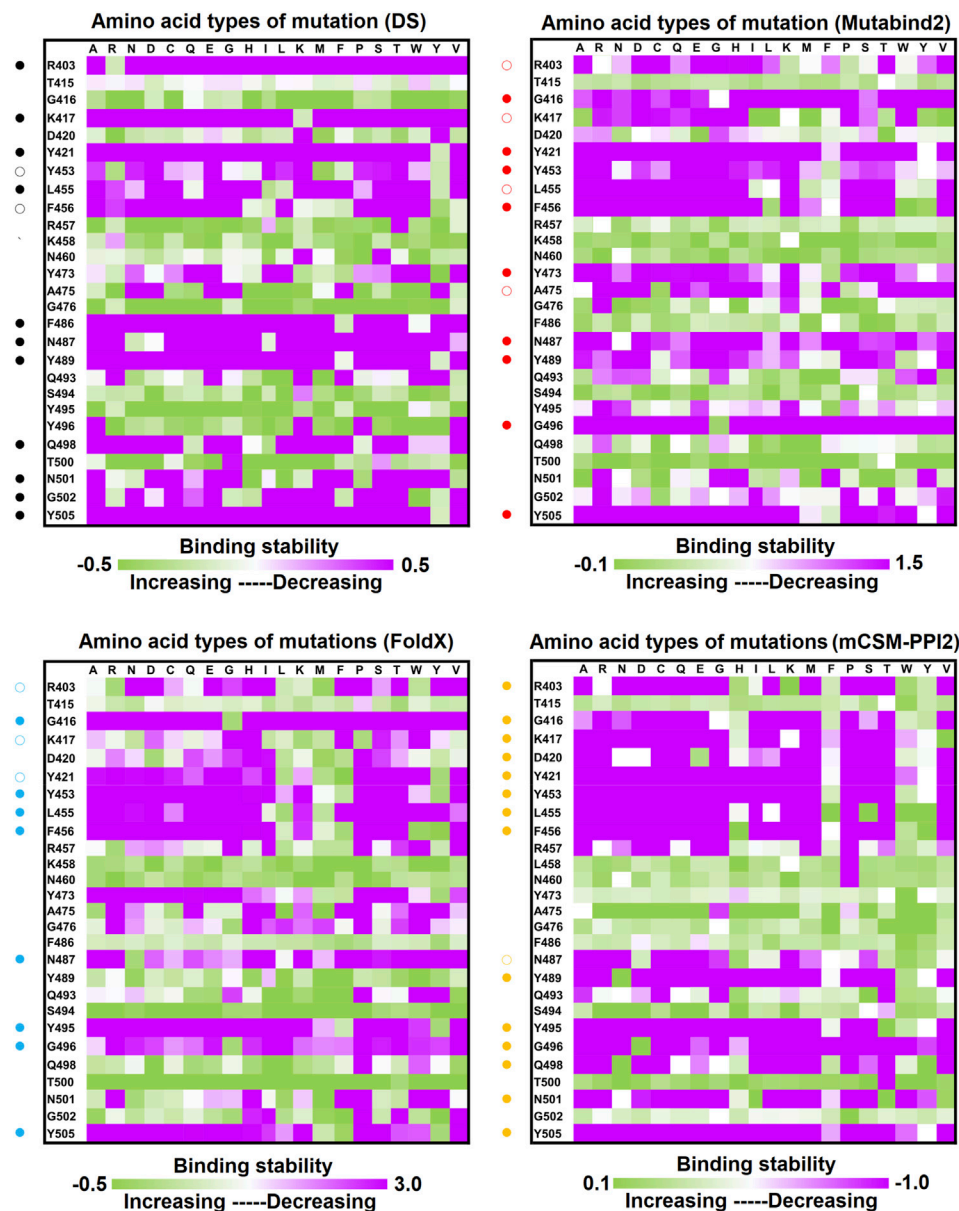


FIGURE 7 | Mutational heatmaps of RBD targeting antibody CA1-B12. The mutational binding stabilities of RBD variants to CA1-B12 were estimated using FoldX, MutaBind2, DS 3.5, and mCSM-PPI2. The unit of binding stability was kcal/mole, and the obtained values were employed to create the heatmaps. Green (increased stability) and purple (reduced stability) were used to generate a color gradient and were applied in each box. In each heatmap, the moderately and significantly reduced stabilities were labeled with hollow and solid circles, respectively.

convalescent antibodies when mutated (Table 1). These residues were aromatic and/or hydrophobic, with the exceptions of E484, K417, N487, N501, R403, and S494. Notably, mutations at F490, F456, G416, G502, K417, L452, L455, Y449, N487, R403, and Y489 were concurrently observed to be unfavorable for the binding of RBD to most of the convalescent antibodies. This indicated that the immune-escaping ability of RBD may be attributable to the hotspots F456, F490, G416, G502, K417, L452, L455, N487, R403, Y449, and Y489, which disrupt interaction with convalescent

antibodies, further interfering or even attenuating the immune responses.

We therefore placed our findings among the emerging variants of SARS-CoV-2 to verify the effects of the determined immune-escape hotspots R403, G416, K417, Y449, L452, L455, F456, N487, Y489, F490, and G502. The K417T variant was first observed in Japan (lineage: P.1), with K417N later identified in lineages B.1.351 and P.1 in South Africa and Brazil, respectively (Table 1). The K417N and K417T variants of the RBD destabilized interaction with convalescent antibodies B38,

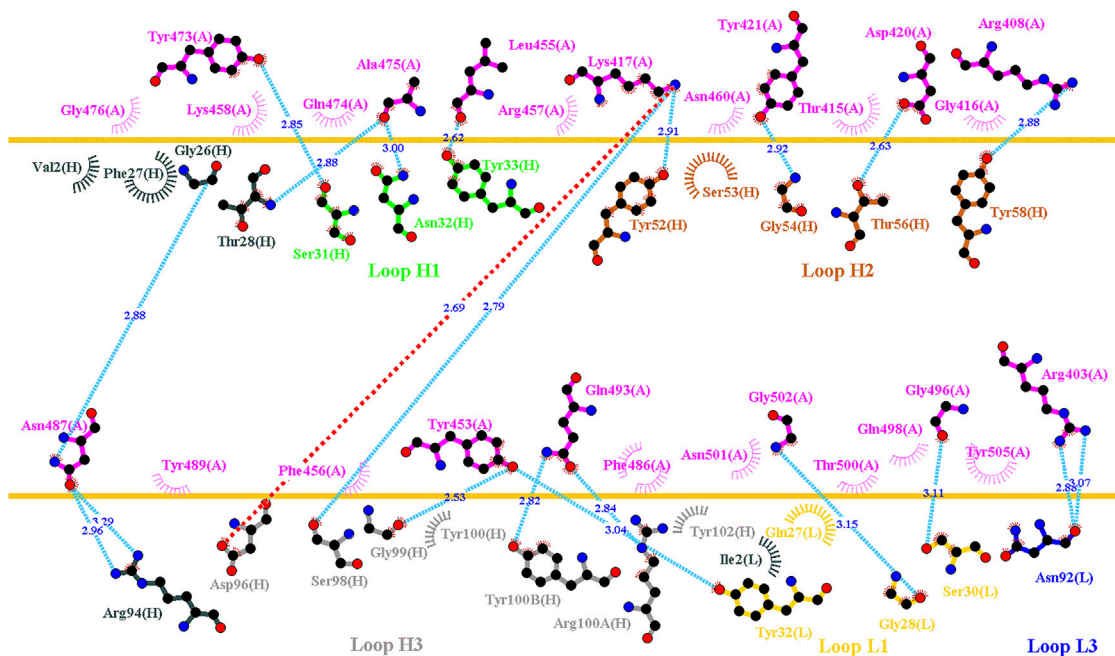
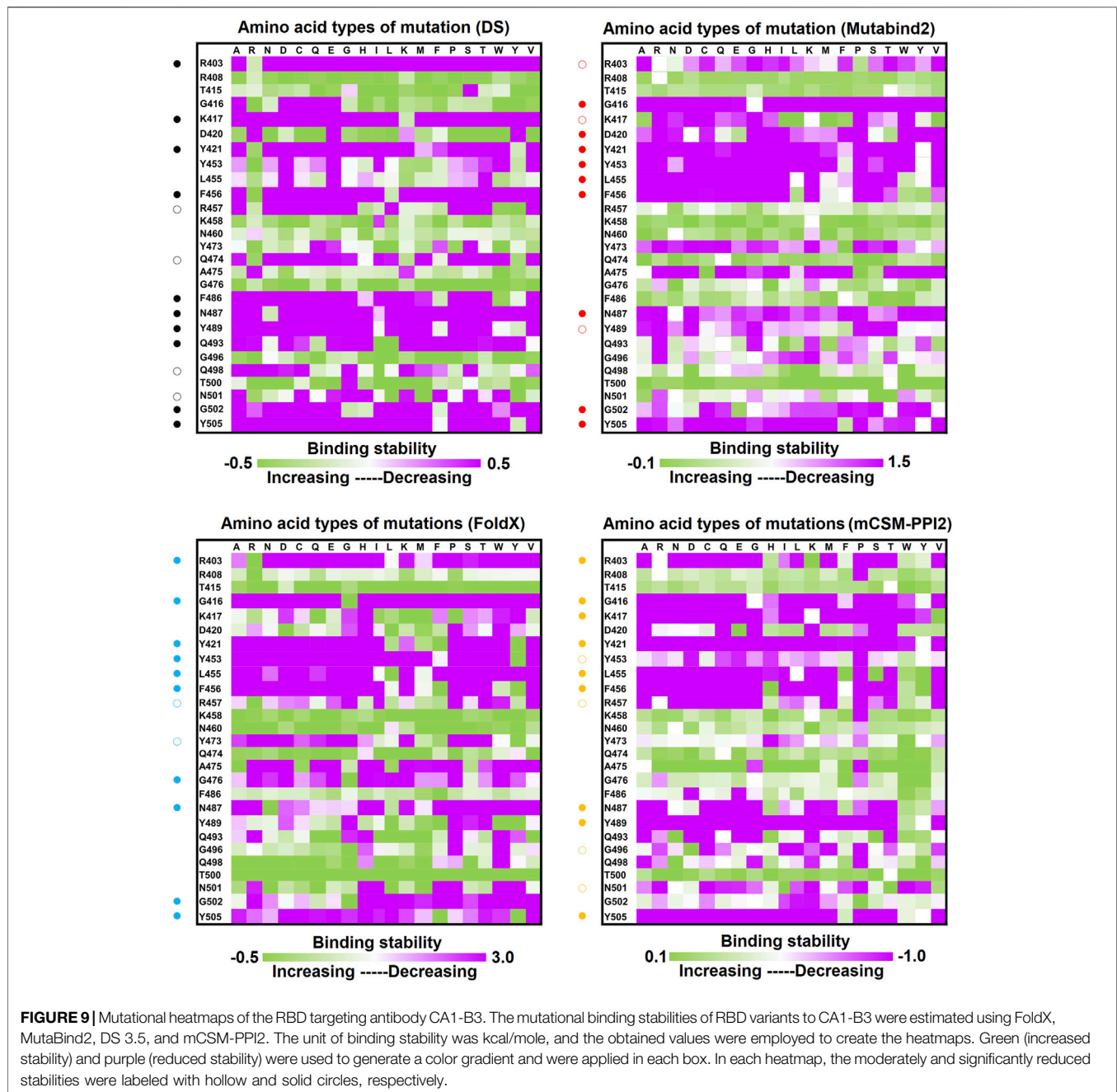


FIGURE 8 | Interactional network of the RBD and CA1-B3 antibody analyzed using LIGPLOT. LIGPLOT was employed to examine the interaction network of the SARS-CoV-2 RBD and CA1-B3. The RBD and light and heavy chains of CA1-B3 are labeled A, H, and L, respectively. The golden yellow line denotes the interface of the RBD–CA1-B3 complex. The interactive residues of the RBD were labeled and colored magenta. The salt-bridges, hydrogen bonds and hydrophobic interactions are represented by red, cyan dashes and arcs with spokes radiating toward the ligand atoms with which they are contact, respectively.

C1A-B12, and C1A-B3 in our study. Furthermore, we observed that the variants of L452 conspicuously decreased the binding stability of the RBD to convalescent antibodies CV07-270, CT-P59, and 47D1 in our predictions. This variant (L452R) was identified in both the United States and India (**Table 1**). Also, E484 variants reportedly escape from neutralizing antibodies, especially the E484K variant that emerged in several SARS-CoV-2 lineages (**Table 1**) and the E484Q variant in India (Weisblum et al., 2020a; Chen et al., 2021). Our computational analysis revealed that the variants of E484 significantly disrupted RBD binding to the convalescent antibodies CV07-270 and 47D1. The variants of F490 and S494 also exerted destabilizing effects on RBD binding to convalescent antibodies in our study, which is consistent with the SARS-CoV-2 variants F490S and S494P observed in the United States and Peru (**Table 1**). Contraction of the highly infectious and potentially lethal variant N501Y leads to a high chance of hospitalization (Leung et al., 2021), with this variant identified in the lineages of C.37, P.1, B.1.1.7, and B.1.351 (Singh et al., 2021; Tang et al., 2021; Tegally et al., 2021). This report corroborates our finding that variants of N501 destabilized the binding stability of RBD to convalescent antibodies B38 and CA1-B12. The RBD variants conferring the immune-escaping ability were also investigated using experimental methods (Andreano et al., 2020; Weisblum et al., 2020a; Huo et al., 2020; Starr et al., 2020; Greaney A. J. et al., 2021; Garcia-Beltran et al., 2021b; Greaney A. J. et al., 2021). The reporter virus VSV/SARS-CoV-2 was examined in terms of the mutational effects of RBD variants, with the results revealing that E484K, F490L, and Q493 K/R exert resistance to neutralizing antibodies

(Weisblum et al., 2020b). Li et al. also reported that L452R and F490L are resistant to some neutralizing antibodies and sera from convalescent patients (Andreano et al., 2020). Moreover, variants E484K, K417N, K417T, and N501Y were able to escape antibody-induced neutralization, as observed in an emerging SARS-CoV-2 isolate in South Africa (B.1.351; Garcia-Beltran et al., 2021a). Another study conducted yeast display and deep mutational scanning of the SARS-CoV-2 RBD and demonstrated that single-point mutations at sites E484, F456, F486, F490, K417, L452, L455, N450, and Q493 escaped neutralizing antibodies (Greaney A. J. et al., 2021). Our results indicated that the RBD residues F456, F490, G502, G416, K417, L452, L455, N487, R403, Y489, and Y449 were essential sites; following mutation, their variants profoundly disrupted interaction with neutralizing antibodies. In particular, the results of the hotspots F456, F490, L452, L455, and K417 and destabilizing residues E484, F486, and N501 identified in this study are corroborated through reports on emerging SARS-CoV-2 variants. This robust collection of evidence strongly supports the identification of these hotspots and the precision and reliability of our computational study. Therefore, these results can assist in the further development and application of potent antibody and vaccine therapeutics to combat emerging SARS-CoV-2 variants.

The detailed interactions between RBD variants and convalescent antibodies were analyzed to determine possible modes of actions relating to the analyzed hotspots. The RBD residue K417 interacted with Y52 and/or D96 of the convalescent antibodies B38, CA1-B12, and CA1-B3 through hydrogen bonding (**Figures 11A–F**). However, the hydrogen bonds were



disrupted when K417 was replaced with threonine or arginine amino acids, which further destabilized the binding stability of the RBD to the antibodies. The RBD residue L452 interacted with W100 of CV07-270 and Y52 and D96 of CT-P59 (Figures 11G,H) through hydrophobic contact. These hydrophobic interactions were disturbed when L452 was mutated to an arginine amino acid, which further weakened the binding stability of RBD to CV07-270 and CT-P59. Structurally, the charge-charge interactions often make considerable contributions to the binding stability of protein complexes. The residue E484 electrostatically interacted with R100 and R98 of CV07-270 and 47D1, respectively, but this interaction

was disrupted when E484 was substituted with lysine or glutamine residues, interfering the convalescent antibody-induced neutralization of RBD (Figures 11I,J). In addition, the aromatic RBD residue F490 exerted hydrophobic and cation- π interactions contributing to binding to the antibodies (Figures 11M,N). These interactions do not occur in the complex structure of the F490S variant, which therefore explains the reduced binding stability of RBD to the antibodies. Furthermore, the RBD residue S494 interacting with R105 and R107 of CT-P59 through hydrogen bonding was disrupted when S494 was mutated to a proline amino acid, thus weakening the binding affinity (Figures 11O,P). In addition, the N501Y variant

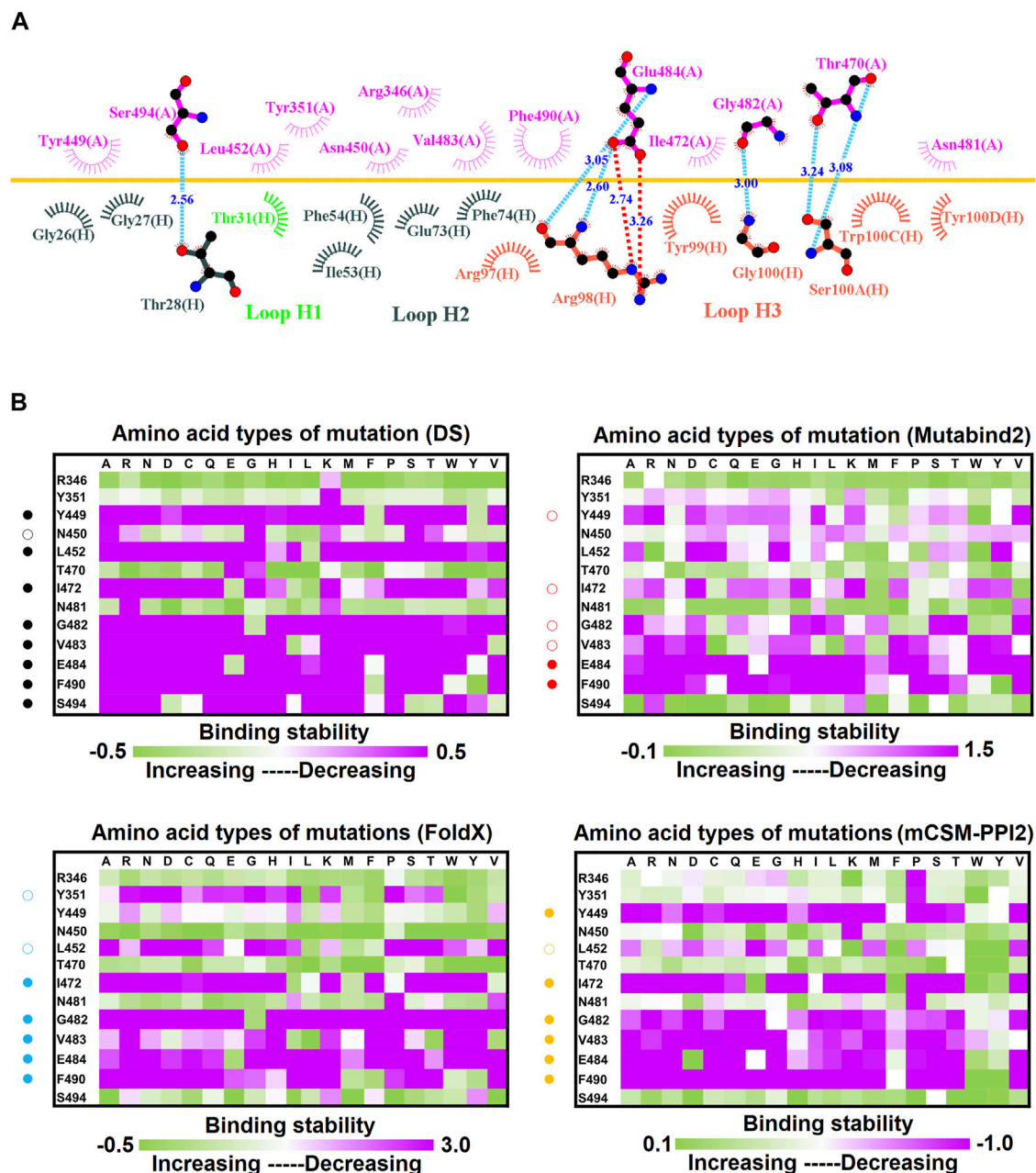


FIGURE 10 | Mutational heatmaps of the RBD targeting antibody 47D1. **(A)** LIGPLOT was employed to examine the interaction network of the SARS-CoV-2 RBD and 47D1. The RBD and light and heavy chains of 47D1 were labeled A, H, and L, respectively. The golden yellow line denotes the interface of the RBD–47D1 complex. The interactive residues of the RBD are labeled and colored magenta. The hydrogen salt-bridges, bonds and hydrophobic interactions are represented by red, cyan dashes and arcs with spokes radiating toward the ligand atoms with which they are contact, respectively. **(B)** The mutational binding stabilities of RBD variants to 47D1 were estimated using FoldX, MutaBind2, DS 3.5, and mCSM-PPI2. The unit of binding stability was kcal/mole, and the obtained values were employed to create the heatmaps. Green (increased stability) and purple (reduced stability) were used to generate a color gradient and were applied in each box. In each heatmap, the moderately and significantly reduced stabilities were labeled with hollow and solid circles, respectively.

exhibited considerable concurrency in several countries. In the complex structure of RBD–B38 and RBD–CA1–B12, N501 forms a hydrogen bond with S30. Nevertheless, the bulky aromatic ring of N510Y collides with S30 and other residues at the interface, hence decreasing the binding stability of RBD to convalescent antibodies B38 and CA1–B12 (**Figures 11Q–T**).

In this study, we employed four programs/methods (FoldX, Mutabind2, mMCS-PPI2, and DS) to simultaneously evaluate the mutational binding energy of variants of RBD targeting distinct convalescent antibodies. Notably, mutations could change protein conformations in the binding interface, and this may have significant effects on antibody binding. While the used four

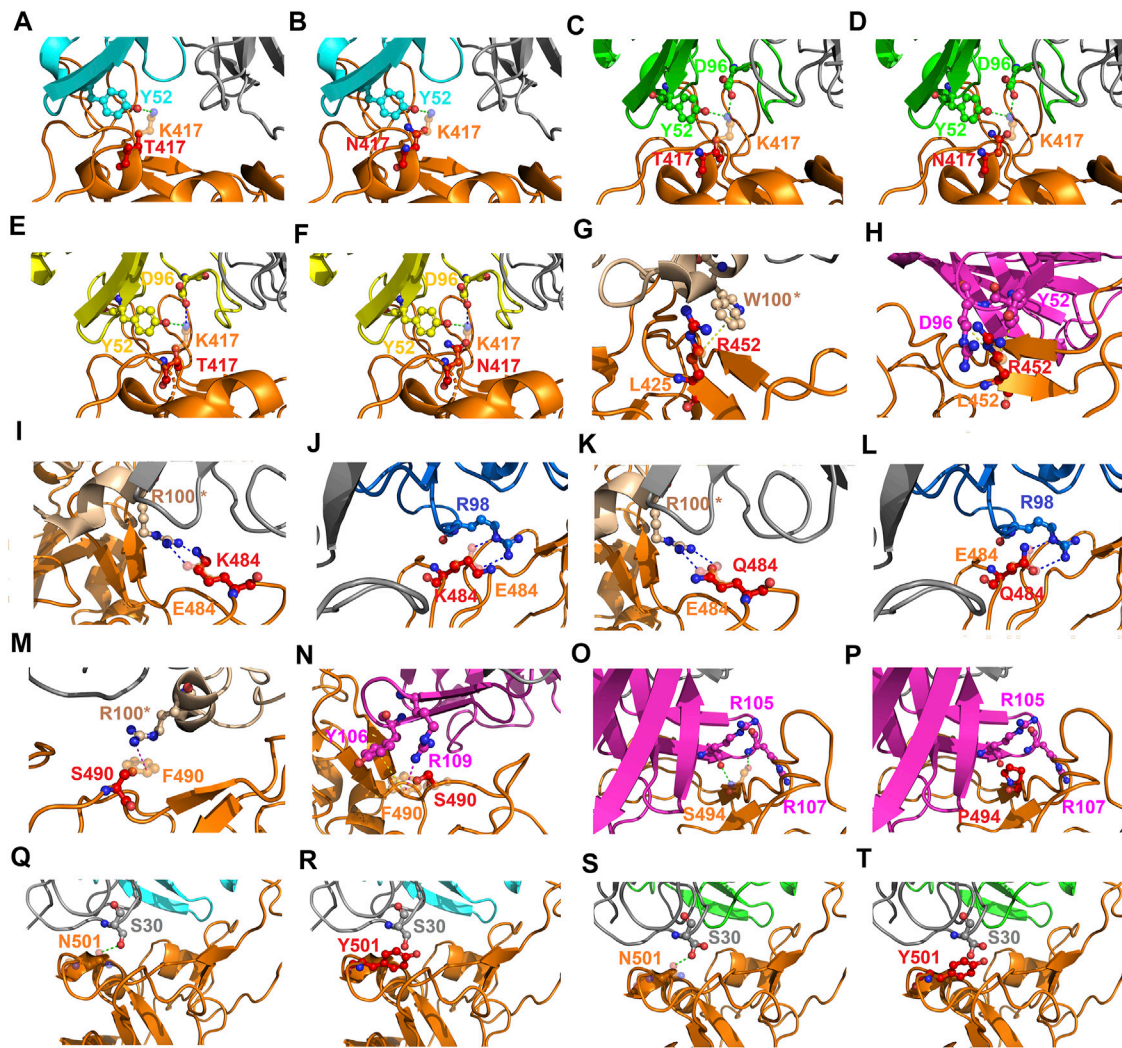
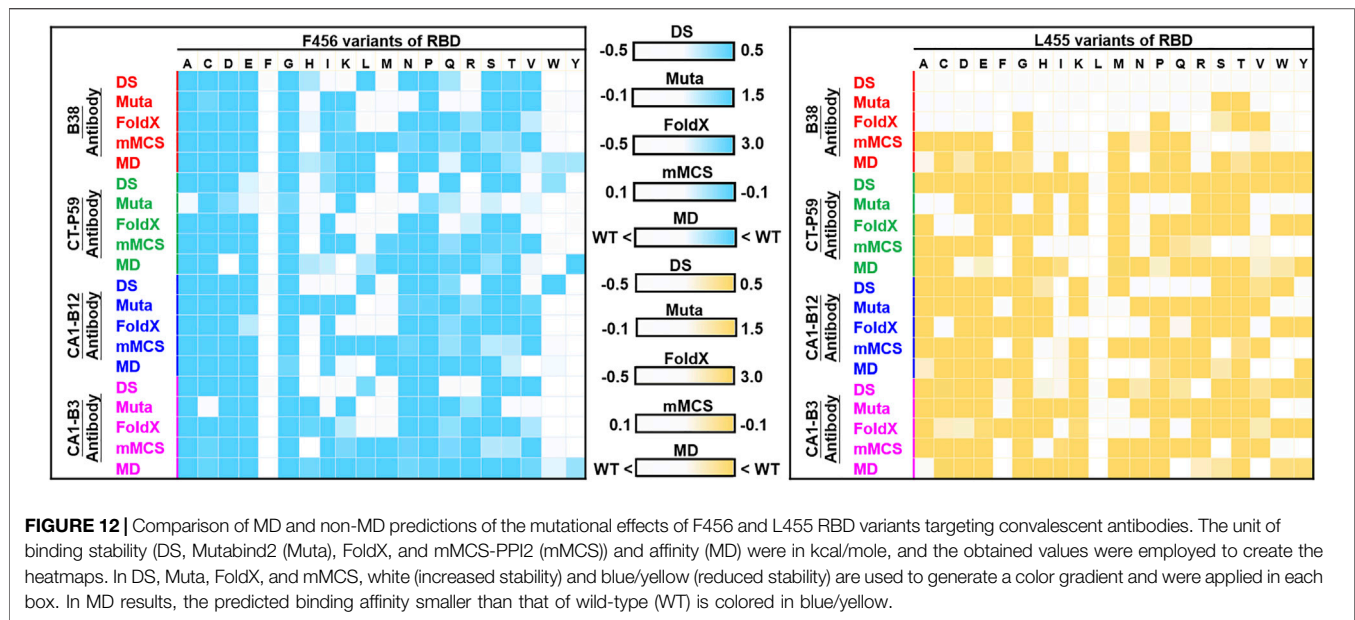


FIGURE 11 | Mutational effects contributing to the modes of action of emerging SARS-CoV-2 RBD variants. **(A)** An overlapped view of the intermolecular interactions of the B38 antibody to K417 (orange chain) and mutant T417 (red chain) of the RBD. **(B)** A superimposition of K417 (orange chain) and mutant N417 (red chain) of the RBD bound with antibody B38. **(C)** The distinct interactions of K417 (orange chain) and mutant T417 (red chain) of the RBD targeting the CA1-B12 antibody. **(D)** A superimposition of K417 (orange chain) and mutant N417 (red chain) of the RBD interacting with the CA1-B12 antibody. **(E)** An overlapped view of K417 (orange chain) and mutant T417 (red chain) of the RBD bound with CA1-B3. **(F)** A superimposition of K417 (orange chain) and mutant N417 (red chain) of the RBD targeting the CA1-B3 antibody. **(G)** A comparison of L452 (orange chain) and mutant R452 (red chain) of the RBD binding to the CV07-270 antibody. **(H)** The distinct interactions of L452 (orange chain) and mutant R452 (red chain) of the RBD targeting the CT-P59 antibody. **(I)** A superimposed view of E484 (orange chain) and mutant K484 (red chain) of the RBD interacting with the CV07-270 antibody. **(J)** An overlapped view of E484 (orange chain) and mutant K484 (red chain) of the RBD binding to 47D1. **(K)** A comparison of E484 (orange chain) and mutant Q484 (red chain) of the RBD in complex with the CV07-270 antibody. **(L)** The differences between E484 (orange chain) and mutant Q484 (red chain) of the RBD interacting with the 47D1 antibody. **(M)** The deviations between F490 (orange chain) and mutant S490 (red chain) of the RBD targeting the CV07-270 antibody. **(N)** A superimposition of F490 (orange chain) and mutant S490 (red chain) of the RBD bound with the CT-P59 antibody. **(O)** The interactions of S494 (orange chain) and **(P)** mutant P494 (red chain) of the RBD binding to the CT-P59 antibody. **(Q)** The interactions of N501 (orange chain) and **(R)** mutant Y501 (red chain) of the RBD targeting the B38 antibody. **(S)** The interactions of N501 (orange chain) and **(T)** mutant Y501 (red chain) of the RBD bound to the CA1-B12 antibody. The structures of proteins are presented with ribbons and dashed lines colored green, blue, red, and yellow denoting a hydrogen bond, cation- π interaction, hydrophobic interaction, and electrostatic interaction, respectively.

programs do not appear to consider protein conformational flexibility, ensemble predictions, and explicit solvent effects. The molecular dynamic (MD) simulations to complexes can incorporate protein flexibility, and explicit solvent effects which are important for electrostatics and hydrogen bonding. Although the energy minimization is included during the calculations of the used program, such as Mutabind2, the

standard MD simulation is not incorporated into the predictions of these methods. Hence, to confirm the reliability and precision of the used strategy and methodology in this study, we further employed MD simulations for the strongest mutation hotspots L455 and F456 to again estimate their mutational effects on RBD's binding to convalescent antibodies. The corresponding mutant and wild-type structures of RBD complexed with



convalescent antibodies (B38, CT-P59, CA1-B3 and CA1-B12) were subjected to MD simulations. All the trajectory profiles of total energy changes as functions of MD simulation time were shown in **Supplementary Figures S1-S16**. Besides, the initial and final conformations before and after MD simulations were compared and shown in **Supplementary Figures S17-S32**. Furthermore, the final conformations of each complex structure after the MD simulations were subjected to CSM-AB to calculate the binding affinity. The results showed that over 80% of the L455 and F456 variants of RBD displayed decreased binding affinity to convalescent antibodies when compare to that of the wild-type RBD (**Supplementary Tables S27,S28**). These observations are consistent with the predicted results of the four used programs (FoldX, Mutabind2, mMCS-PPI2, and DS) (**Figure 12**), providing evidence to support the proposed immune escape hotspots and demonstrating the reliability of the predictions in this study.

In addition to our research, there are approaches conducted to predict the possible mutations in SARS-CoV-2. Rodriguez-Rivas et al. used the epistatic models to predict the mutable sites of proteins and epitopes in SARS-CoV-2 (Rodriguez-Rivas et al., 2022). In their sequence-based predictions, the predicted mutability of SARS-CoV-2 RBD was observed to be well correlated with experimentally determined protein stability. Also, their study identified mutable positions (K417, L452, S477, T478, V483, E484, and N501) of RBD in which current variants of concern are highly overrepresented. Consistently, residues, K417, L452, E484, and N501 were predicted as possible immune escape hotspots in our study. On top of that, Taft and colleagues have performed the predictive profiling of SARS-CoV-2 variants based on protein sequences (Joseph M. Taft, 2021). They built a deep machine learning to integrate and analyze a huge sequence space of SARS-CoV-2 variants, computationally estimating their effects on antibody neutralization. Their DML study can efficiently predict new

and possible SARS-CoV-2 variants, but the sequence-based prediction without consideration of structural properties, prime determinants of protein-protein binding affinity, could erode the reliability. In our study, we systematically conducted *in silico* mutagenesis on the structure of RBD complexed with convalescent antibodies to comprehensively predicted the mutational impacts of its variants on binding stability. The results of our structure-based predictions, consistent with experimental measures and some circulating variants, can be complementary to those of sequence-based predictions. Still, the used method of our study is restricted to give insight into the conformation-destabilizing mutations of allosteric residues that may also result in escape from neutralizing antibodies. In this study, the functionally important residues of RBD which interact and bind to convalescent antibody were mainly subjected to single point mutations. The binding stabilities of the obtained RBD variants in complex with convalescent antibody were predicted to evaluate the mutational effects. We found that single point mutation at specific residues of RBD indeed apparently disrupted the binding of RBD to antibodies. These residues of RBD are of great potential to be hotspots which could escape the recognition by convalescent antibodies. Our study mainly aims to identify the immune-escaping hotspot from the binding interface of RBD to convalescent antibody, while the allosteric residues of RBD that result in escape from neutralizing antibodies could be another research scope for future work. In summary, we explored and revealed the probable immune-escaping hotspots from the interface residues of RBD interacting with convalescent antibodies. Single-point mutations at these residues that could considerably impair the interaction between the SARS-CoV-2 RBD and convalescent antibodies. Our findings are immensely beneficial for designing and developing therapeutics to combat emerging SARS-CoV-2 variants.

CONCLUSION

In this study, we determined the potential residues that, after mutation, can confer the ability to escape convalescent antibody-induced neutralization on SARS-CoV-2. The complex structure-based analysis revealed that specific variants of RBD significantly impair the binding stability to convalescent antibodies. The single-point mutations at hotspots G502, G416, F456, F490, K417, L452, L455, N487, R403, Y449, and Y489 markedly destabilized binding to convalescent antibodies. The results of the immune-escaping hotspots (K417, L452, L455, F456, and F490) and destabilizing residues (E484, F486, and N501) corroborate experimental observations and clinically emerging SARS-CoV-2 variants. Our study provides insight into the structural hotspots that confer the immune-escaping ability of SARS-CoV-2 RBD in relation to convalescent antibodies with distinct modes of action. The results of our study can assist in the development of new antiviral agents to protect against the emerging variants of SARS-CoV-2.

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.

REFERENCES

- Andreano, E., Piccini, G., Licastro, D., Casalino, L., Johnson, N. V., Paciello, I., et al. (2020). SARS-CoV-2 Escape *In Vitro* from a Highly Neutralizing COVID-19 Convalescent Plasma. *bioRxiv*.
- Baden, L. R., El Sahly, H. M., Essink, B., Kotloff, K., Frey, S., Novak, R., et al. (2021). Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *N. Engl. J. Med.* 384, 403–416. doi:10.1056/nejmoa2035389
- Baj, J., Karakula-Juchnowicz, H., Teresiński, G., Buszewicz, G., Ciesielka, M., Sitarz, E., et al. (2020). COVID-19: Specific and Non-specific Clinical Manifestations and Symptoms: The Current State of Knowledge. *J. Clin. Med.* 9, 1753. doi:10.3390/jcm9061753
- Barnes, C. O., Jette, C. A., Abernathy, M. E., Dam, K.-M. A., Esswein, S. R., Gristick, H. B., et al. (2020). SARS-CoV-2 Neutralizing Antibody Structures Inform Therapeutic Strategies. *Nature* 588, 682–687. doi:10.1038/s41586-020-2852-1
- Bertram, S., Dijkman, R., Habjan, M., Heurich, A., Gierer, S., Glowacka, I., et al. (2013). TMPRSS2 Activates the Human Coronavirus 229E for Cathepsin-independent Host Cell Entry and Is Expressed in Viral Target Cells in the Respiratory Epithelium. *J. Virol.* 87, 6150–6160. doi:10.1128/jvi.03372-12
- Bettini, E., and Locci, M. (2021). SARS-CoV-2 mRNA Vaccines: Immunological Mechanism and beyond. *Vaccines (Basel)* 9, 147. doi:10.3390/vaccines9020147
- Callaway, E. (2020). The Coronavirus Is Mutating - Does it Matter? *Nature* 585, 174–177. doi:10.1038/d41586-020-02544-6
- Casadevall, A. (2002). Passive Antibody Administration (Immediate Immunity) as a Specific Defense against Biological Weapons. *Emerg. Infect. Dis.* 8, 833–841. doi:10.3201/eid0808.010516
- Chambers, P., Pringle, C. R., and Easton, A. J. (1990). Heptad Repeat Sequences Are Located Adjacent to Hydrophobic Regions in Several Types of Virus Fusion Glycoproteins. *J. Gen. Virol.* 71 (Pt 12), 3075–3080. doi:10.1099/0022-1317-71-12-3075
- Chen, R. E., Zhang, X., Case, J. B., Winkler, E. S., Liu, Y., Vanblargan, L. A., et al. (2021). Resistance of SARS-CoV-2 Variants to Neutralization by Monoclonal

AUTHOR CONTRIBUTIONS

T-ST and K-MF designed the experiments. T-ST, K-MF, T-LL, and S-JL performed the experiments. T-ST, K-MF, T-LL, and S-JL analyzed the data. T-ST and K-MF wrote the manuscript. All authors have read and agreed to the published version of the manuscript. All authors contributed to the article and have approved the submitted version.

FUNDING

This study was supported by the Ministry of Science and Technology, Taiwan, ROC (MOST 109-2320-B-005-006-MY2).

ACKNOWLEDGMENTS

We thank National Center for High-performance Computing (NCHC) for providing computational and storage resources.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2022.797132/full#supplementary-material>

- and Serum-Derived Polyclonal Antibodies. *Nat. Med.* 27, 717. doi:10.1038/s41591-021-01294-w
- Clark, S. A., Clark, L. E., Pan, J., Coscia, A., McKay, L. G. A., Shankar, S., et al. (2020). Molecular Basis for a Germline-Biased Neutralizing Antibody Response to SARS-CoV-2. *bioRxiv*.
- Dong, Y., Dai, T., Wei, Y., Zhang, L., Zheng, M., and Zhou, F. (2020). A Systematic Review of SARS-CoV-2 Vaccine Candidates. *Sig Transduct Target. Ther.* 5, 237. doi:10.1038/s41392-020-00352-y
- Eguia, R. T., Crawford, K. H. D., Stevens-Ayers, T., Kelnhofer-Millevolte, L., Greninger, A. L., Englund, J. A., et al. (2021). A Human Coronavirus Evolves Antigenically to Escape Antibody Immunity. *Plos Pathog.* 17, e1009453. doi:10.1371/journal.ppat.1009453
- Fehr, A. R., and Perlman, S. (2015). Coronaviruses: an Overview of Their Replication and Pathogenesis. *Methods Mol. Biol.* 1282, 1–23. doi:10.1007/978-1-4939-2438-7_1
- Garcia-Beltran, W. F., Lam, E. C., St. Denis, K., Nitido, A. D., Garcia, Z. H., Hauser, B. M., et al. (2021a). Multiple SARS-CoV-2 Variants Escape Neutralization by Vaccine-Induced Humoral Immunity. *Cell* 184, 2523. doi:10.1016/j.cell.2021.04.006
- Garcia-Beltran, W. F., Lam, E. C., St. Denis, K., Nitido, A. D., Garcia, Z. H., Hauser, B. M., et al. (2021b). Multiple SARS-CoV-2 Variants Escape Neutralization by Vaccine-Induced Humoral Immunity. *Cell* 184, 2372–2383. doi:10.1016/j.cell.2021.03.013
- Gavor, E., Choong, Y. K., Er, S. Y., Sivaraman, H., and Sivaraman, J. (2020). Structural Basis of SARS-CoV-2 and SARS-CoV Antibody Interactions. *Trends Immunol.* 41, 1006–1022. doi:10.1016/j.it.2020.09.004
- Geers, D., Shamier, M. C., Bogers, S., Den Hartog, G., Gommers, L., Nieuwkoop, N. N., et al. (2021). SARS-CoV-2 Variants of Concern Partially Escape Humoral but Not T-Cell Responses in COVID-19 Convalescent Donors and Vaccinees. *Sci. Immunol.* 6, 1750. doi:10.1126/sciimmunol.abj1750
- Gomez, C. E., Perdiguer, B., and Esteban, M. (2021). Emerging SARS-CoV-2 Variants and Impact in Global Vaccination Programs against SARS-CoV-2/covid-19. *Vaccines (Basel)* 9, 243. doi:10.3390/vaccines9030243

- Greaney, A. J., Loes, A. N., Crawford, K. H. D., Starr, T. N., Malone, K. D., Chu, H. Y., et al. (2021a). Comprehensive Mapping of Mutations in the SARS-CoV-2 Receptor-Binding Domain that Affect Recognition by Polyclonal Human Plasma Antibodies. *Cell Host & Microbe* 29, 463–476. doi:10.1016/j.chom.2021.02.003
- Greaney, A. J., Starr, T. N., Barnes, C. O., Weisblum, Y., Schmidt, F., Caskey, M., et al. (2021b). Mutational Escape from the Polyclonal Antibody Response to SARS-CoV-2 Infection Is Largely Shaped by a Single Class of Antibodies. *bioRxiv*.
- Harvey, W. T., Carabelli, A. M., Jackson, B., Gupta, R. K., Thomson, E. C., Harrison, E. M., et al. (2021). SARS-CoV-2 Variants, Spike Mutations and Immune Escape. *Nat. Rev. Microbiol.* 19, 409. doi:10.1038/s41579-021-00573-0
- Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., et al. (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 181, 271–280. doi:10.1016/j.cell.2020.02.052
- Huang, Y., Yang, C., Xu, X.-f., Xu, W., and Liu, S.-w. (2020). Structural and Functional Properties of SARS-CoV-2 Spike Protein: Potential Antivirus Drug Development for COVID-19. *Acta Pharmacol. Sin* 41, 1141–1149. doi:10.1038/s41401-020-0485-4
- Huo, J., Le Bas, A., Ruza, R. R., Duyvesteyn, H. M. E., Mikolajek, H., Malinauskas, T., et al. (2020). Neutralizing Nanobodies Bind SARS-CoV-2 Spike RBD and Block Interaction with ACE2. *Nat. Struct. Mol. Biol.* 27, 846–854. doi:10.1038/s41594-020-0469-6
- Jaroszewski, L., Iyer, M., Alisoltani, A., Sedova, M., and Godzik, A. (2020). The Interplay of SARS-CoV-2 Evolution and Constraints Imposed by the Structure and Functionality of its Proteins. *bioRxiv*.
- Jiang, S., Hillyer, C., and Du, L. (2020). Neutralizing Antibodies against SARS-CoV-2 and Other Human Coronaviruses. *Trends Immunol.* 41, 355–359. doi:10.1016/j.it.2020.03.007
- Joseph M. Taft, C. R. W., Gao, B., Ehling, R. A., Han, J., Frei, L., Metcalfe, S. W., et al. (2021). Predictive Profiling of SARS-CoV-2 Variants by Deep Mutational Learning. *bioRxiv*.
- Ju, B., Zhang, Q., Ge, J., Wang, R., Sun, J., Ge, X., et al. (2020). Human Neutralizing Antibodies Elicited by SARS-CoV-2 Infection. *Nature* 584, 115–119. doi:10.1038/s41586-020-2380-z
- Kim, C., Ryu, D.-K., Lee, J., Kim, Y.-I., Seo, J.-M., Kim, Y.-G., et al. (2021). A Therapeutic Neutralizing Antibody Targeting Receptor Binding Domain of SARS-CoV-2 Spike Protein. *Nat. Commun.* 12, 288. doi:10.1038/s41467-020-20602-5
- Kreye, J., Reincke, S. M., Kornau, H.-C., Sánchez-Sendin, E., Corman, V. M., Liu, H., et al. (2020). A Therapeutic Non-self-reactive SARS-CoV-2 Antibody Protects from Lung Pathology in a COVID-19 Hamster Model. *Cell* 183, 1058–1069. doi:10.1016/j.cell.2020.09.049
- Lai, C.-C., Shih, T.-P., Ko, W.-C., Tang, H.-J., and Hsueh, P.-R. (2020). Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) and Coronavirus Disease-2019 (COVID-19): The Epidemic and the Challenges. *Int. J. Antimicrob. Agents* 55, 105924. doi:10.1016/j.ijantimicag.2020.105924
- Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., et al. (2020). Structure of the SARS-CoV-2 Spike Receptor-Binding Domain Bound to the ACE2 Receptor. *Nature* 581, 215–220. doi:10.1038/s41586-020-2180-5
- Laskowski, R. A., and Swindells, M. B. (2011). LigPlot+: Multiple Ligand-Protein Interaction Diagrams for Drug Discovery. *J. Chem. Inf. Model.* 51, 2778–2786. doi:10.1021/ci200227u
- Lau, E. H. Y., Tsang, O. T. Y., Hui, D. S. C., Kwan, M. Y. W., Chan, W.-h., Chiu, S. S., et al. (2021). Neutralizing Antibody Titres in SARS-CoV-2 Infections. *Nat. Commun.* 12, 63. doi:10.1038/s41467-020-20247-4
- Leung, K., Shum, M. H., Leung, G. M., Lam, T. T., and Wu, J. T. (2021). Early Transmissibility Assessment of the N501Y Mutant Strains of SARS-CoV-2 in the United Kingdom, October to November 2020. *Euro Surveill.* 26, 2106. doi:10.2807/1560-7917.es.2020.26.1.2002106
- Liu, L., Wang, P., Nair, M. S., Yu, J., Rapp, M., Wang, Q., et al. (2020). Potent Neutralizing Antibodies against Multiple Epitopes on SARS-CoV-2 Spike. *Nature* 584, 450–456. doi:10.1038/s41586-020-2571-7
- Lu, Q., Zhang, Z., Li, H., Zhong, K., Zhao, Q., Wang, Z., et al. (2021). Development of Multivalent Nanobodies Blocking SARS-CoV-2 Infection by Targeting RBD of Spike Protein. *J. Nanobiotechnol.* 19, 33. doi:10.1186/s12951-021-00768-w
- Malik, Y. A. (2020). Properties of Coronavirus and SARS-CoV-2. *Malays J. Pathol.* 42, 3–11.
- Millet, J. K., and Whittaker, G. R. (2018). Physiological and Molecular Triggers for SARS-CoV Membrane Fusion and Entry into Host Cells. *Virology* 517, 3–8. doi:10.1016/j.virol.2017.12.015
- Mousavizadeh, L., and Ghasemi, S. (2021). Genotype and Phenotype of COVID-19: Their Roles in Pathogenesis. *J. Microbiol. Immunol. Infect.* 54, 159–163. doi:10.1016/j.jmii.2020.03.022
- Myung, Y., Pires, D. E. V., and Ascher, D. B. (2021). CSM-AB: Graph-Based Antibody-Antigen Binding Affinity Prediction and Docking Scoring Function. *Bioinformatics* 2021, btab762. doi:10.1093/bioinformatics/btab762
- Organization, W. H. (2021). COVID-19 Vaccine Tracker and Landscape [Online]. Available at: <https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines> (Accessed).
- Oronsky, B., Gruber, H. E., Reiners, W., and Reid, T. R. (2021). A Short Discussion about the SARS-CoV-2 mRNA-1273 Vaccine. *Int. J. Infect. Dis.* 104, 532–533. doi:10.1016/j.ijid.2021.01.048
- Pal, M., Berhanu, G., Desalegn, C., and Kandi, V. (2020). Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2): An Update. *Cureus* 12, e7423. doi:10.7759/cureus.7423
- Piccoli, L., Park, Y.-J., Tortorici, M. A., Czudnochowski, N., Walls, A. C., Beltramello, M., et al. (2020). Mapping Neutralizing and Immunodominant Sites on the SARS-CoV-2 Spike Receptor-Binding Domain by Structure-Guided High-Resolution Serology. *Cell* 183, 1024–1042. doi:10.1016/j.cell.2020.09.037
- Pinto, D., Park, Y.-J., Beltramello, M., Walls, A. C., Tortorici, M. A., Bianchi, S., et al. (2020). Cross-neutralization of SARS-CoV-2 by a Human Monoclonal SARS-CoV Antibody. *Nature* 583, 290–295. doi:10.1038/s41586-020-2349-y
- Planas, D., Bruel, T., Grzelak, L., Guivel-Benhassine, F., Staropoli, I., Porrot, F., et al. (2021). Sensitivity of Infectious SARS-CoV-2 B.1.1.7 and B.1.351 Variants to Neutralizing Antibodies. *Nat. Med.* 27, 917–924. doi:10.1038/s41591-021-01318-5
- Polack, F. P., Thomas, S. J., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., et al. (2020). Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N. Engl. J. Med.* 383, 2603–2615. doi:10.1056/nejmoa2034577
- Rapp, M., Guo, Y., Reddem, E. R., Yu, J., Liu, L., Wang, P., et al. (2021). Modular Basis for Potent SARS-CoV-2 Neutralization by a Prevalent VH1-2-Derived Antibody Class. *Cel Rep.* 35, 108950. doi:10.1016/j.celrep.2021.108950
- Rodrigues, C. H. M., Myung, Y., Pires, D. E. V., and Ascher, D. B. (2019). mCSM-PPI2: Predicting the Effects of Mutations on Protein-Protein Interactions. *Nucleic Acids Res.* 47, W338–W344. doi:10.1093/nar/gkz383
- Rodriguez-Rivas, J., Croce, G., Muscat, M., and Weigt, M. (2022). Epistatic Models Predict Mutable Sites in SARS-CoV-2 Proteins and Epitopes. *Proc. Natl. Acad. Sci. U S A* 119, 8119. doi:10.1073/pnas.2113118119
- Rogers, T. F., Zhao, F., Huang, D., Beutler, N., Burns, A., He, W.-t., et al. (2020). Isolation of Potent SARS-CoV-2 Neutralizing Antibodies and protection from Disease in a Small Animal Model. *Science* 369, 956–963. doi:10.1126/science.abc7520
- Rosati, M., Terpos, E., Ntanasios-Stathopoulos, I., Agarwal, M., Bear, J., Burns, R., et al. (2021). Sequential Analysis of Binding and Neutralizing Antibody in COVID-19 Convalescent Patients at 14 Months after SARS-CoV-2 Infection. *Front. Immunol.* 12, 793953. doi:10.3389/fimmu.2021.793953
- Rydzynski Moderbacher, C., Ramirez, S. I., Dan, J. M., Grifoni, A., Hastie, K. M., Weiskopf, D., et al. (2020). Antigen-Specific Adaptive Immunity to SARS-CoV-2 in Acute COVID-19 and Associations with Age and Disease Severity. *Cell* 183, 996–1012. doi:10.1016/j.cell.2020.09.038
- Ryu, D.-K., Song, R., Kim, M., Kim, Y.-I., Kim, C., Kim, J.-I., et al. (2021). Therapeutic Effect of CT-P59 against SARS-CoV-2 South African Variant. *Biochem. Biophysical Res. Commun.* 566, 135–140. doi:10.1016/j.bbrc.2021.06.016
- Sadoff, J., Gray, G., Vandebosch, A., Cárdenas, V., Shukarev, G., Grinsztejn, B., et al. (2021). Safety and Efficacy of Single-Dose Ad26.COV2.S Vaccine against Covid-19. *N. Engl. J. Med.* 384, 2187–2201. doi:10.1056/nejmoa2101544
- Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., and Serrano, L. (2005). The FoldX Web Server: an Online Force Field. *Nucleic Acids Res.* 33, W382–W388. doi:10.1093/nar/gki387
- Sette, A., and Crotty, S. (2021). Adaptive Immunity to SARS-CoV-2 and COVID-19. *Cell* 184, 861–880. doi:10.1016/j.cell.2021.01.007

- Singh, J., Samal, J., Kumar, V., Sharma, J., Agrawal, U., Ehtesham, N. Z., et al. (2021). Structure-Function Analyses of New SARS-CoV-2 Variants B.1.1.7, B.1.351 and B.1.1.28.1: Clinical, Diagnostic, Therapeutic and Public Health Implications. *Viruses* 13, 439. doi:10.3390/v13030439
- Starr, T. N., Greaney, A. J., Hilton, S. K., Ellis, D., Crawford, K. H. D., Dingens, A. S., et al. (2020). Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* 182, 1295–1310. e1220. doi:10.1016/j.cell.2020.08.012
- Tada, T., Dcosta, B. M., Samanovic, M. I., Herati, R. S., Cornelius, A., Zhou, H., et al. (2021). Convalescent-Phase Sera and Vaccine-Elicited Antibodies Largely Maintain Neutralizing Titer against Global SARS-CoV-2 Variant Spikes. *mBio* 12, e0069621. doi:10.1128/mbio.00696-21
- Tai, W., He, L., Zhang, X., Pu, J., Voronin, D., Jiang, S., et al. (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. *Cell Mol Immunol* 17, 613–620. doi:10.1038/s41423-020-0400-4
- Tang, J. W., Tambyah, P. A., and Hui, D. S. (2021). Emergence of a New SARS-CoV-2 Variant in the UK. *J. Infect.* 82, e27–e28. doi:10.1016/j.jinf.2020.12.024
- Tegally, H., Wilkinson, E., Giovanetti, M., Iranzadeh, A., Fonseca, V., Giandhari, J., et al. (2021). Detection of a SARS-CoV-2 Variant of Concern in South Africa. *Nature* 592, 438. doi:10.1038/s41586-021-03402-9
- Teng, S., Sobitan, A., Rhoades, R., Liu, D., and Tang, Q. (2021). Systemic Effects of Missense Mutations on SARS-CoV-2 Spike Glycoprotein Stability and Receptor-Binding Affinity. *Brief Bioinform* 22, 1239–1253. doi:10.1093/bib/bbaa233
- Tsai, K.-C., Lee, Y.-C., and Tseng, T.-S. (2021). Comprehensive Deep Mutational Scanning Reveals the Immune-Escaping Hotspots of SARS-CoV-2 Receptor-Binding Domain Targeting Neutralizing Antibodies. *Front. Microbiol.* 12, 698365. doi:10.3389/fmicb.2021.698365
- V'kovski, P., Kratzel, A., Steiner, S., Stalder, H., and Thiel, V. (2021). Coronavirus Biology and Replication: Implications for SARS-CoV-2. *Nat. Rev. Microbiol.* 19, 155–170.
- Velazquez-Salinas, L., Zarate, S., Eberl, S., Gladue, D. P., Novella, I., and Borca, M. V. (2020). Positive Selection of ORF1ab, ORF3a, and ORF8 Genes Drives the Early Evolutionary Trends of SARS-CoV-2 during the 2020 COVID-19 Pandemic. *Front. Microbiol.* 11, 550674. doi:10.3389/fmicb.2020.550674
- Voysey, M., Clemens, S. A. C., Madhi, S. A., Weckx, L. Y., Folegatti, P. M., Aley, P. K., et al. (2021). Safety and Efficacy of the ChAdOx1 nCoV-19 Vaccine (AZD1222) against SARS-CoV-2: an Interim Analysis of Four Randomised Controlled Trials in Brazil, South Africa, and the UK. *Lancet* 397, 99–111. doi:10.1016/S0140-6736(20)32661-1
- Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995). LIGPLOT: a Program to Generate Schematic Diagrams of Protein-Ligand Interactions. *Protein Eng. Des. Sel* 8, 127–134. doi:10.1093/protein/8.2.127
- Walls, A. C., Park, Y.-J., Tortorici, M. A., Wall, A., McGuire, A. T., and Veesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* 181, 281–292. doi:10.1016/j.cell.2020.02.058
- Wang, P., Nair, M. S., Liu, L., Iketani, S., Luo, Y., Guo, Y., et al. (2021a). Antibody Resistance of SARS-CoV-2 Variants B.1.351 and B.1.1.7. *Nature* 593, 130–135. doi:10.1038/s41586-021-03398-2
- Wang, Q., Zhang, Y., Wu, L., Niu, S., Song, C., Zhang, Z., et al. (2020). Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* 181, 894–904. doi:10.1016/j.cell.2020.03.045
- Wang, Z., Schmidt, F., Weisblum, Y., Muecksch, F., Barnes, C. O., Finklin, S., et al. (2021b). mRNA Vaccine-Elicited Antibodies to SARS-CoV-2 and Circulating Variants. *Nature* 592, 616–622. doi:10.1038/s41586-021-03324-6
- Weisblum, Y., Schmidt, F., Zhang, F., Dasilva, J., Poston, D., Lorenzi, J. C., et al. (2020a). Escape from Neutralizing Antibodies by SARS-CoV-2 Spike Protein Variants. *Elife* 9, e61312. doi:10.7554/eLife.61312
- Weisblum, Y., Schmidt, F., Zhang, F., Dasilva, J., Poston, D., Lorenzi, J. C. C., et al. (2020b). Escape from Neutralizing Antibodies by SARS-CoV-2 Spike Protein Variants. *bioRxiv*.
- Wibmer, C. K., Ayres, F., Hermanus, T., Madzivhandila, M., Kgagudi, P., Oosthuysen, B., et al. (2021). SARS-CoV-2 501Y.V2 Escapes Neutralization by South African COVID-19 Donor Plasma. *Nat. Med.* 27, 622–625. doi:10.1038/s41591-021-01285-x
- Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C. L., Abiona, O., et al. (2020). Cryo-EM Structure of the 2019-nCoV Spike in the Prefusion Conformation. *bioRxiv*.
- Wu, Y., Wang, F., Shen, C., Peng, W., Li, D., Zhao, C., et al. (2020). A Noncompeting Pair of Human Neutralizing Antibodies Block COVID-19 Virus Binding to its Receptor ACE2. *Science* 368, 1274–1278. doi:10.1126/science.abc2241
- Xia, S., Xu, W., Wang, Q., Wang, C., Hua, C., Li, W., et al. (2018). Peptide-Based Membrane Fusion Inhibitors Targeting HCoV-229E Spike Protein HR1 and HR2 Domains. *Int. J. Mol. Sci.* 19, 487. doi:10.3390/ijms19020487
- Xia, S., Zhu, Y., Liu, M., Lan, Q., Xu, W., Wu, Y., et al. (2020). Fusion Mechanism of 2019-nCoV and Fusion Inhibitors Targeting HR1 Domain in Spike Protein. *Cel Mol Immunol* 17, 765–767. doi:10.1038/s41423-020-0374-2
- Xiaojie, S., Yu, L., Lei, Y., Guang, Y., and Min, Q. (2020). Neutralizing Antibodies Targeting SARS-CoV-2 Spike Protein. *Stem Cel Res* 50, 102125. doi:10.1016/j.scr.2020.102125
- Xie, J., Ding, C., He, J., Zhang, Y., Ni, S., Zhang, X., et al. (2021). Novel Monoclonal Antibodies and Recombined Antibodies against Variant SARS-CoV-2. *Front. Immunol.* 12, 715464. doi:10.3389/fimmu.2021.715464
- Yuan, M., Huang, D., Lee, C. D., Wu, N. C., Jackson, A. M., Zhu, X., et al. (2021). Structural and Functional Ramifications of Antigenic Drift in Recent SARS-CoV-2 Variants. *bioRxiv*.
- Zhou, D., Dejnirattisai, W., Supasa, P., Liu, C., Mentzer, A. J., Ginn, H. M., et al. (2021a). Evidence of Escape of SARS-CoV-2 Variant B.1.351 from Natural and Vaccine-Induced Sera. *Cell* 184, 2348–2361. doi:10.1016/j.cell.2021.02.037
- Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., et al. (2020). A Pneumonia Outbreak Associated with a New Coronavirus of Probable Bat Origin. *Nature* 579, 270–273. doi:10.1038/s41586-020-2012-7
- Zhou, X., Ma, F., Xie, J., Yuan, M., Li, Y., Shaabani, N., et al. (2021b). Diverse Immunoglobulin Gene Usage and Convergent Epitope Targeting in Neutralizing Antibody Responses to SARS-CoV-2. *Cel Rep.* 35, 109109. doi:10.1016/j.celrep.2021.109109

Conflict of Interest: The authors declare 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.

Copyright © 2022 Fung, Lai, Lin and Tseng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.



Potential of Natural Alkaloids From Jadwar (*Delphinium denudatum*) as Inhibitors Against Main Protease of COVID-19: A Molecular Modeling Approach

Anuj Kumar^{1,2†}, Mansi Sharma^{1,2†}, Christopher D. Richardson² and David J. Kelvin^{1,2*}

OPEN ACCESS

Edited by:

Mattia Falconi,
University of Rome Tor Vergata, Italy

Reviewed by:

Sai Ganesan,
University of California, San Francisco,
United States
Subhash Chandra,
Kumaun University, India

*Correspondence:

David J. Kelvin
David.Kelvin@dal.ca

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Biological Modeling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 18 March 2022

Accepted: 08 April 2022

Published: 10 May 2022

Citation:

Kumar A, Sharma M, Richardson CD
and Kelvin DJ (2022) Potential of
Natural Alkaloids From Jadwar
(*Delphinium denudatum*) as Inhibitors
Against Main Protease of COVID-19: A
Molecular Modeling Approach.
Front. Mol. Biosci. 9:898874.
doi: 10.3389/fmolb.2022.898874

¹Laboratory of Immunity, Shantou University Medical College, Shantou, China, ²Department of Microbiology and Immunology, Canadian Centre for Vaccinology CCFV, Faculty of Medicine, Dalhousie University, Halifax, Canada

The ongoing pandemic coronavirus disease (COVID-19) caused by a novel corona virus, namely, severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), has had a major impact on global public health. COVID-19 cases continue to increase across the globe with high mortality rates in immunocompromised patients. There is still a pressing demand for drug discovery and vaccine development against this highly contagious disease. To design and develop antiviral drugs against COVID-19, the main protease (M^{pro}) has emerged as one of the important drug targets. In this context, the present work explored Jadwar (*Delphinium denudatum*)-derived natural alkaloids as potential inhibitors against M^{pro} of SARS-CoV-2 by employing a combination of molecular docking and molecular dynamic simulation-based methods. Molecular docking and interaction profile analysis revealed strong binding on the M^{pro} functional domain with four natural alkaloids viz. panicutine (−7.4 kcal/mol), vilmorrianone (−7.0 kcal/mol), denudatine (−6.0 kcal/mol), and condelphine (−5.9 kcal/mol). The molecular docking results evaluated by using the MD simulations on 200 nanoseconds confirmed highly stable interactions of these compounds with the M^{pro}. Additionally, mechanics/generalized Born/Poisson–Boltzmann surface area (MM/G/P/BSA) free energy calculations also affirmed the docking results. Natural alkaloids explored in the present study possess the essential drug-likeness properties, namely, absorption, distribution, metabolism, and excretion (ADME), and are in accordance with Lipinski's rule of five. The results of this study suggest that these four bioactive molecules, namely, condelphine, denudatine, panicutine, and vilmorrianone, might be effective candidates against COVID-19 and can be further investigated using a number of experimental methods.

Keywords: COVID-19, alkaloids, Jadwar, molecular docking, molecular dynamics simulations and free energy

INTRODUCTION

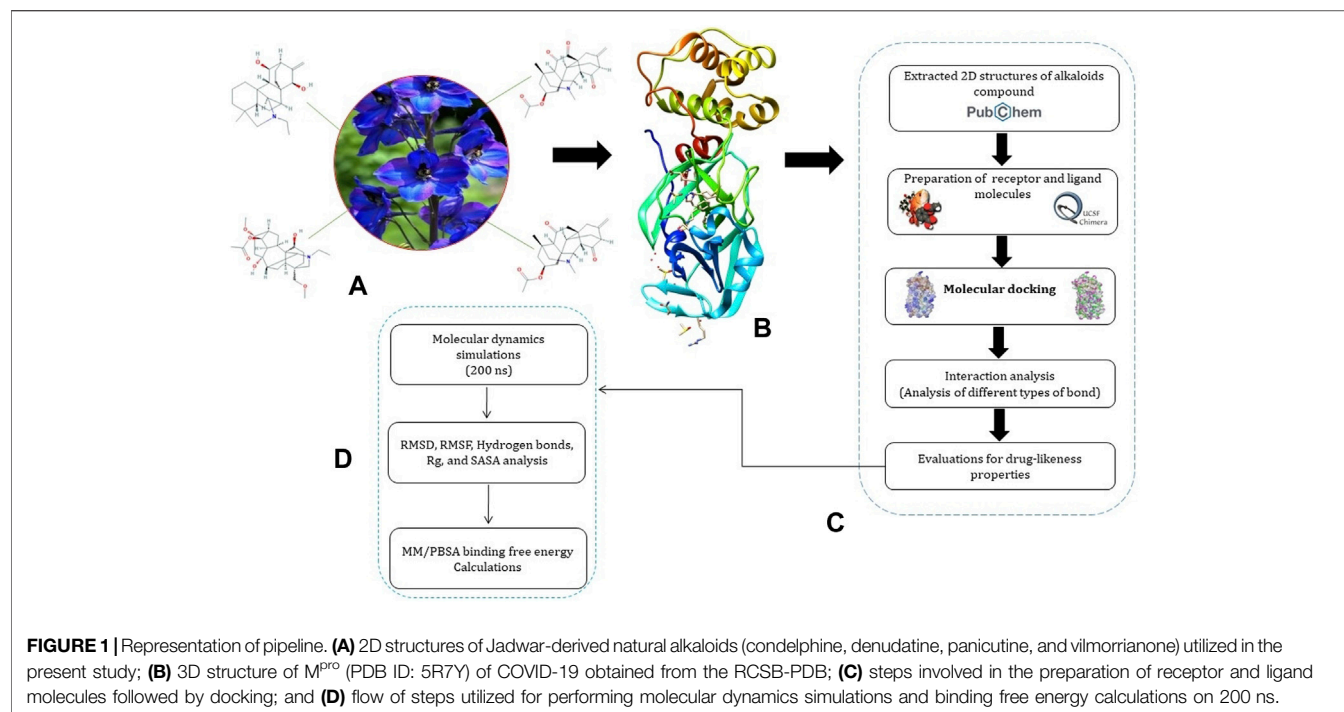
The unprecedented pandemic of coronavirus disease (COVID-19) was caused by a novel coronavirus, namely, severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), which appeared in Wuhan, Hubei Province, China, at the end of December 2019 (Jamwal et al., 2020). A few months later, this viral disease spread to 219 nations and territories across the globe. On 30 January 2020, the World Health Organization (WHO) declared this contagious disease as a Public Health Emergency of International Concern (PHEIC) and based on the recommendations of the Emergency Committee announced it to be a pandemic on 11 March 2020 (Shi et al., 2020; Wilder-Smith and Osman 2020; Yu and Yu 2020). The ongoing pandemic eruption adversely affected the global economy and financial markets (Pak et al., 2020). As of 3 March 2022, COVID-19 has led to more than 438,968,263 confirmed cases and 5,969,439 deaths (<https://covid19.who.int/>), with high mortality rates in immunocompromised and elderly patients. A large number of Canadians were infected with SARS-CoV-2 during the different waves of the ongoing pandemic. As of 3 March 2022, the total caseload in Canada has soared to 3,296,503 with 36,638 fatalities (<https://covid19.who.int/region/amro/country/ca>). Based on infection, morbidity, and mortality, this respiratory infectious disease has greatly superseded previous outbreaks of SARS and the Middle East respiratory syndrome (MERS) (de Wit et al., 2016; Wang et al., 2020; Yuan et al., 2020). Previous SARS and MERS outbreaks possessed fatality rates of 10 and 35%, respectively (Lee et al., 2004; Cheng et al., 2007). It has been reported that COVID-19 is associated with disorders in the respiratory and digestive tracts of the body (Chen et al., 2020; Pal et al., 2020; Tang et al., 2020).

Morphologically, the coronaviruses (CoVs) are a highly diverse family of enveloped positive-sense single-strand RNA viruses (Fehr and Perlman 2015). The Coronavirinae are classified into two subfamilies: Coronavirinae and Torovirinae. Based on the molecular structure and biological functions, the Coronavirinae are further divided into four genera: alpha- (α -), beta- (β -), gamma- (γ -), and delta-coronavirus (δ -CoV) (Hulswit et al., 2016; Payne 2017). To date, six human coronavirus species have been identified, namely, HCoV-NL63, HCoV-229E, HCoV-OC43, HCoV-HKU1, SARS-CoV, and MERS-CoV (Arden et al., 2005; Su et al., 2016; Zhang et al., 2018). The novel strain SARS-CoV-2 has been reported as the seventh CoV known to infect humans (Andersen et al., 2020) in the genus *Betacoronavirus* (<https://talk.ictvonline.org/>; Helmy et al., 2020; Wang et al., 2020).

The single-stranded positive RNA genome of the SARS-CoV-2 virus is ~29.9 kb in size (Wu et al., 2020). The genome sequence for the Wuhan-Hu-1 strain of SARS-CoV-2 is available from the GenBank with the accession number MN908947 (~29,903 nucleotides) (Wu et al., 2020). The SARS-CoV-2 genome contains 14 open reading frames (ORFs) encoding 27 proteins (Alsobaie 2021). The 5' untranslated region (UTR) end harbors ORF1a/b that produces a polyprotein which is posttranslationally cleaved into 16 different nonstructural proteins (nsp1–nsp16). These form the replicase/transcriptase complex (RTC). They

include a papain-like protease (nsp3), main protease (M^{pro} , 3CL pro , nsp5), nsp7–nsp8 primase complex, primary RNA-dependent RNA polymerase (RdRp; nsp12), helicase/triphosphatase (nsp13), exoribonuclease (nsp14), endonuclease (nsp15), and N7- and 2'-O-methyltransferases (nsp10/nsp16). The 3'-end of the SARS-CoV-2 contains ORFs which encode the four structural proteins, namely, E (envelope protein), M (membrane protein), N (nucleocapsid protein), and S (spike protein), as well as nine putative accessory factors (Rastogi et al., 2020; Alsobaie 2021; Shamsi et al., 2021). The SARS-CoV-2 M^{pro} (nsp5) is encoded by the major ORF1ab following posttranslational cleavage in the cytosol (Ullrich and Nitsche 2020; Mengist et al., 2021). Based on its key role in mediating viral replication and transcription, M^{pro} has been considered as one of the promising drug targets against SARS-CoV-2 (Jin et al., 2020; Ullrich and Nitsche 2020; Mengist et al., 2021). M^{pro} consists of 306 amino acids yielding a molecular mass of 33,797 Da (Khan et al., 2020). M^{pro} is composed of three functional domains: domain I (8–101 aa), domain II (102–184 aa), and domain III (201–306 aa) (Jin et al., 2020; Kumar et al., 2021). Based on structure topology, it has been well reported that an antiparallel β -barrel structure is present in domains I and II. While domain III was found to possess a set of five α -helices organized as a large antiparallel cluster. Domains II and III were connected to each other with the help of a 15-residue-long loop region (185–200 residues). Numerous *in vitro*, *in vivo*, and *in silico* studies have been performed to screen the candidate natural compounds as potential inhibitors of SARS-CoV M^{pro} (Benarba and Pandiella 2020; Chikhale et al., 2020; Joshi et al., 2020; Kumar et al., 2020; Raj et al., 2020; Sharma et al., 2020; Tripathi et al., 2020; Yadav et al., 2020; Chowdhury 2021; Mishra et al., 2021; Mukherjee et al., 2021; Patel et al., 2021; Rangsinth et al., 2021; Teli et al., 2021). By now, many FDA-approved known inhibitors of viral protease such as HIV-1 (atazanavir, amprenavir, darunavir, nelfinavir, tipranavir, lopinavir, saquinavir, indinavir, and ritonavir) and hepatitis C virus (ritonavir, boceprevir, telaprevir, paritaprevir, asunaprevir, grazoprevir, glecaprevir, voxilaprevir, and sofosbuvir) have been proposed for the treatment of COVID-19 (Lv et al., 2015; de Leuw and Stephan 2017; Abdelli et al., 2020; Das et al., 2020; Khan et al., 2020; and Mengist et al., 2021). Chloroquine, an FDA-approved antimalarial drug has also been explored as a potential inhibitor of M^{pro} (Ou et al., 2021). In a recent study, two drugs, namely, rifampicin and letermovir have been repurposed as inhibitors of M^{pro} based on investigations of their molecular docking (Pathak et al., 2021). So far, no effective method has been developed for the treatment of this contagious disease, therefore, there is an urgent need to design targeted therapeutic agents to prevent and treat COVID-19.

Delphinium denudatum Wall (Ranunculaceae), also known as Jadwar, is an annual or perennial ornamental shrub that grows to a height of 40–80 cm. This plant grows at high altitudes ranging from 2,400 to 36,500 m on the outer ranges of the western Himalayas from Kashmir to Uttarakhand (Nizami and Jafri 2006; Kumar et al., 2018; Singh et al., 2018). Different portions of *D. denudatum* have been used medicinally for centuries (Pelletier 1996; Nizami and Jafri 2006). The extracts



of Jadwar have been shown to exhibit neuroprotective and cardioprotective properties (Kumar et al., 2018; Singh KP, Kumar A, 2018). Its roots have many diverse uses such as analgesic, antipyretic, antiseptic, anti-inflammatory, aphrodisiac, antidote (for the snake's venom), cardiotonic, diuretic, exhilarant, sedative, and solvent applications (Abid et al., 2017; Singh KP, Kumar A, 2018). It is well documented that Jadwar is traditionally used for the treatment of various diseases such as fungal infections, cardiac diseases, cholera, epilepsy, jaundice, mania, migraine, paralysis, pain, snake bite, scorpion sting, toothache, etc. (Atta-ur-Rahman et al., 1997; Atta-ur-Rahman and Choudhary 1999; Raza et al., 2003; Ahmad et al., 2017; Kumar et al., 2018). Also, the Jadwar root is used in morphine de-addiction therapy (Zafar et al., 2001; Rahman et al., 2002). A plethora of natural compounds belong to flavonoids, triterpenoids, and alkaloids, such as delphocurarine, staphisagrine, delphine, condelphine, and denudatin. A diterpenoid alkaloid identical to condelphine is exclusively found in Jadwar (Singh and Chopra 1962; Jain et al., 2021). Jadwar-derived isotalatazidine hydrate has demonstrated cholinesterase inhibitory potential and can be used as the target drug for Alzheimer's disease (Ahmad et al., 2017). Despite the rich pharmacological properties of Jadwar, natural compounds exclusively found in this important medicinal herb have not yet been explored for the treatment of SARS-CoV-2. Available open-source platforms and molecular modeling algorithms can be utilized to explore the potential of Jadwar-derived natural alkaloids as potential inhibitors against targets of COVID-19.

In the present study, we have employed a molecular modeling approach to assess the potential of Jadwar-derived natural alkaloids (panicutine, vilmorrianone, denudatine, and condelphine) as inhibitors of M^{Pro} from SARS-CoV-2. These

bioactive compounds were subjected to molecular docking analysis with M^{Pro} enzyme. Docking complexes were further evaluated for conformational stability using molecular dynamics simulations of 200 ns followed by mechanics/generalized Born/Poisson-Boltzmann surface area (MM/G/P/BSA) free energy calculations.

MATERIALS AND METHODS

A systematic approach of the molecular modeling used in this study is presented in **Figure 1**.

Protein Structure Retrieval and Preparation

We retrieved the crystal structure of SARS-CoV-2 M^{Pro} in complex with an inhibitor Z45617795 (PDB ID: 5R7Y, at a resolution of 1.65 Å, R-Value Free of 0.237, and R-Value Work of 0.179) from the RCSB Protein Data Bank (RCSB-PDB) (<https://www.rcsb.org/>) in PDB format (Berman 2000; Burley et al., 2018). The protein structure was prepared using AutoDockTools version 1.5.6, UCSF Chimera, and Discovery Studio. Prior to molecular docking, water and hetero atoms were removed, polar hydrogen was added, and Kollman charges were assigned as well on the receptor protein. Amber force field ff14SB embedded in UCSF Chimera was applied for protein structure optimization and energy minimization. The side chain correction was executed using the clean geometry algorithm monitored in the Discovery Studio platform.

Ligand Structure Retrieval and Preparation

In a search for a potential drug candidate against COVID-19, the four Jadwar-derived natural alkaloids, namely, panicutine,

vilmorrianone, denudatine, and condelphine, that have been reported to have therapeutic potential were selected based on an extensive literature survey. The comprehensive PubChem repository was utilized to extract the 2D structures of these alkaloids—denudatine (CID_161515), condelphine (CID_441720), vilmorrianone (CID_44566629), and panicutine (CID_44566630)—in the SDF format (Kim et al., 2020). 3D and geometry optimizations with energy minimization for each molecule were performed using the UCSF Chimera program. The Open Babel toolbox (O'Boyle et al., 2011) which is available in the PyRx package was utilized to convert these molecules into the PDBQT format. All four compounds were prepared by adding the polar hydrogens and Gasteiger charges as previously described in Kumar et al. (2021).

Molecular Docking

To predict the molecular interactions between M^{Pro} of SARS-CoV-2 and the four natural alkaloid compounds—denudatine, condelphine, vilmorrianone, and panicutine—we have performed molecular docking to identify the interaction between the chemical molecules and target proteins. Molecular docking was done with AutoDock v4.2 (Morris et al., 2009), and the binding affinity score was calculated for the docking complexes. Eleven amino acid residues, namely, Thr24, Thr26, Asn119, Phe140, Gly143, Cys145, His163, His164, Glu166, Gln189, and Thr190, were used as the active sites of the receptor protein. These active site residues were considered based on the previous reports by Khan et al. (2020) and Kumar et al. (2021). During the molecular docking process, the M^{Pro} was fixed, while the ligand molecules were flexible. A grid box was created with dimensions 60 Å × 60 Å × 60 Å centered at the coordinates X = -10, Y = 13, and Z = 70, with 100 conformations for each molecule based on the Lamarckian genetic algorithm (LGA) (Fuhrmann et al., 2010). The representative binding position for the ligand molecules was selected based on the negative binding energy and binding interactions with the receptor protein.

Drug Likeness Properties

In terms of absorption, distribution, metabolism, and excretion, the ADME and drug-likeness of all four alkaloid compounds were predicted using the SWISS-ADME server (Daina et al., 2017). During the drug-likeness prediction process, all four alkaloids were analyzed based on the Lipinski's rule of five (Lipinski 2004), using Veber's rule (Veber et al., 2002), polar surface area (TPSA), bioavailability, and solubility potential (Daina et al., 2017).

Molecular Dynamics Simulations

To elucidate the behavior of the natural alkaloids (panicutine, vilmorrianone, denudatine, and condelphine) binding to the M^{Pro} of COVID-19 and monitor the conformational changes the docking complexes undergo over a stipulated time interval, the docking assemblies were subjected to molecular dynamics simulations for 200-ns time frame. All-atom additive protein force field, CHARMM36, available in the GROMACS 2021 package installed on a Linux-based system was utilized to perform the MD simulations (Abraham and Gready 2011;

Huang and Mackerell 2013; Kutzner et al., 2019). The topology files of ligands were prepared using the ACPYPE (AnteChamber PYthon Parser interfacE) server (Sousa Da Silva and Vranken 2012). The docking complexes were contained in a triclinic simulation box and solvated with a TIP3P water model. Counter Na⁺ and Cl⁻ ions were added to neutralize the solvated system, followed by the quick energy minimization of the system with the help of the LINCS constraint algorithm and the steepest descent algorithm. The process of equilibrium was divided into two steps. In the first step, equilibration was established with a constant number of particles, volume, and temperature (NVT), with a 500-ps timestep, while the second step was completed with a constant number of particles, pressure, and temperature (NPT) with the ensemble at 300 K. The particle mesh Ewald (PME) method was used to calculate the long-range electrostatic interactions (Abraham and Gready 2011). Prior to the production run, different thermodynamics properties (pressure, density, potential energy, and temperature) of the system were carefully monitored to verify adequate equilibration. The v-rescale, Berendsen temperature coupling method was employed to regulate the temperature inside the box. After completion of the pressure and temperature equilibration of the system, a production run of 200 ns was run with each step of 2 fs

Molecular Dynamics Trajectory Analysis

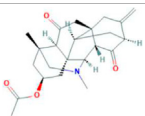
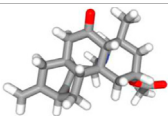
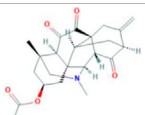
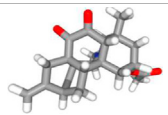
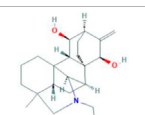
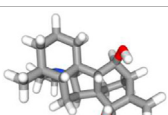
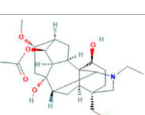
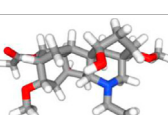
After the successful completion of the 200 ns MD simulation run, different factors of MD, namely, the RMSD (root mean square deviations), RMSF (root mean square fluctuations), number of hydrogen bonds, Rg (the radius of gyration), and SASA (solvent accessible surface area) were calculated using a set of tools embedded in the GROMACS package. The RMSD plot of all complexes was calculated using the gRMS tool, while the RMSF was generated using the gRMSF module of the GROMACS. The Rg, SASA, and hydrogen bonds were estimated using gyrate, gmxsasa, and gh bond tools, respectively.

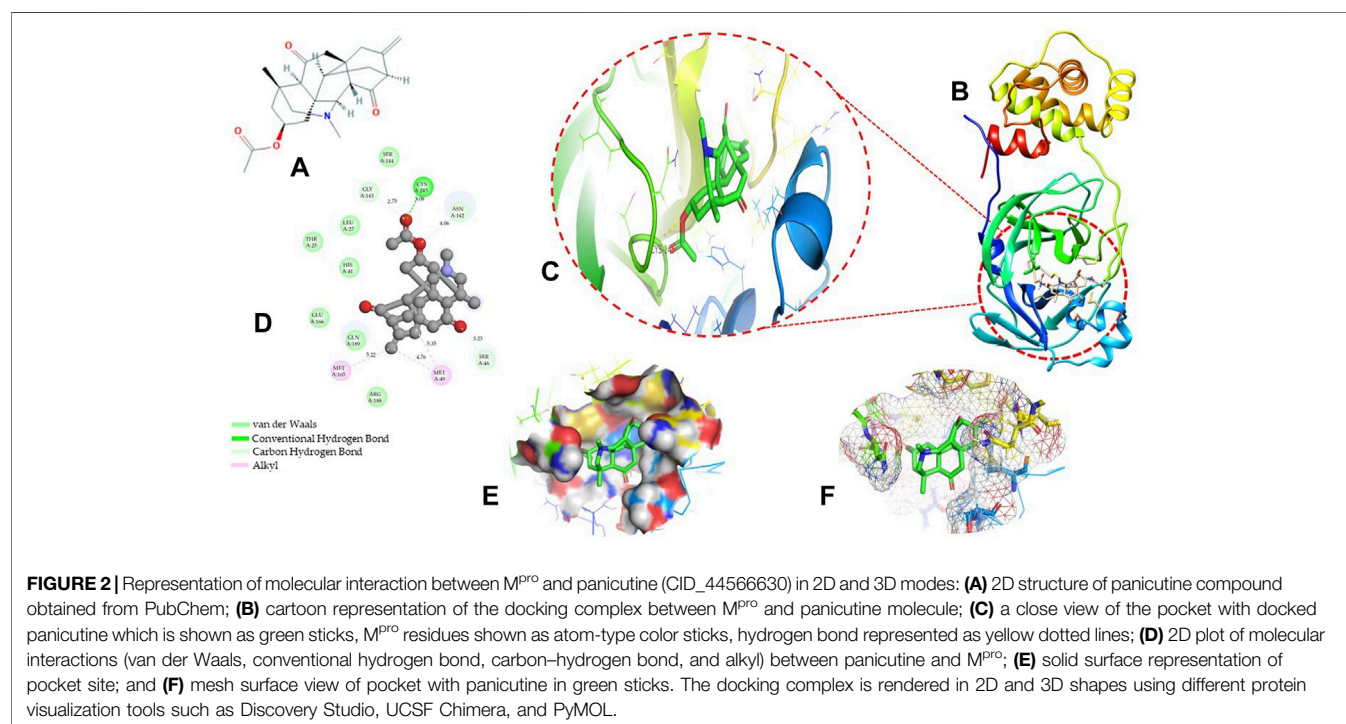
Binding Free Energy Calculations

After the MD simulations of the protein–ligand complexes were completed, the molecular mechanics/Poisson–Boltzmann surface area (MM/PBSA) binding free energy was calculated for all four docking complexes using the g_mmpbsa script program developed by Kumari et al. (2014). The g_mmpbsa module was embedded in the GROMACS package to integrate high-throughput MD simulation with calculations of binding free energy (Kumari et al., 2014; Aldeghi et al., 2017). The major components of energy, namely, binding energy (kJ/mol), van der Waal energy (ΔE_{vdW}), electrical energy (ΔE_{elec}), polar solvation energy (ΔG_{polar}), and solvent-accessible surface area (SASA) were calculated using the MM/PBSA method through the MD trajectories as described in previous reports (Kumar et al., 2020; Kumar et al., 2021; Mishra et al., 2021; Pathak et al., 2021).

In general, the following equation can be used to calculate the MM/PBSA method-based binding free energy of docking complexes:

TABLE 1 | Details of molecular docking of Jadwar-derived alkaloids against the main protease of COVID-19.

Compound name	PubChem ID	Chemical structure		Binding energy (kcal/mol)	Molecular interactions
		2D	3D		
Panicutine	CID_44566630			-7.4	Hydrogen bond: CYS145 (3.08 Å) Carbon-hydrogen bond: SER46 (3.23 Å), ASN142 (4.06 Å), GLY143 (2.75 Å) Alkyl: MET49 (4.76, 4.77, 5.35 Å), MET165 (5.22 Å)
Vilmorriane	CID_44566629			-7.0	Hydrogen bond: CYS145 (5.01 Å), GLN189 (4.77 Å) Carbon-hydrogen bond: HIS41 (5.05) Alkyl: MET165 (5.63 Å)
Denudatine	CID_161515			-6.0	Hydrogen bond: HIS41 (5.26 Å) Carbon-hydrogen bond: GLN189 (4.35 Å) Alkyl: MET49 (4.92 Å), CYS145 (3.71 Å), MET165 (4.59 Å)
Condelphine	CID_441720			-5.9	Hydrogen bond: CYS145 (5.00 Å), HIS164 (6.57 Å), GLU166 (4.36 Å, 4.74 Å) Carbon-hydrogen bond: THR26 (4.74 Å), SER46 (3.56 Å), MET165 (3.76 Å) Alkyl: CYS44 (3.83 Å), MET49 (4.58 Å)



$$\Delta G_{MMPBSA} = \langle G_{complex} - G_{protein} - G_{ligand} \rangle_{complex}$$

where $G_{protein}$ and G_{ligand} denote the total free energies of the isolated protein and ligand in the solvent, and $G_{complex}$ depicts the total free energy of the docking complex, respectively.

RESULTS AND DISCUSSION

Molecular Docking

Molecular docking has become one of the most important molecular modeling methods to study the mechanism of

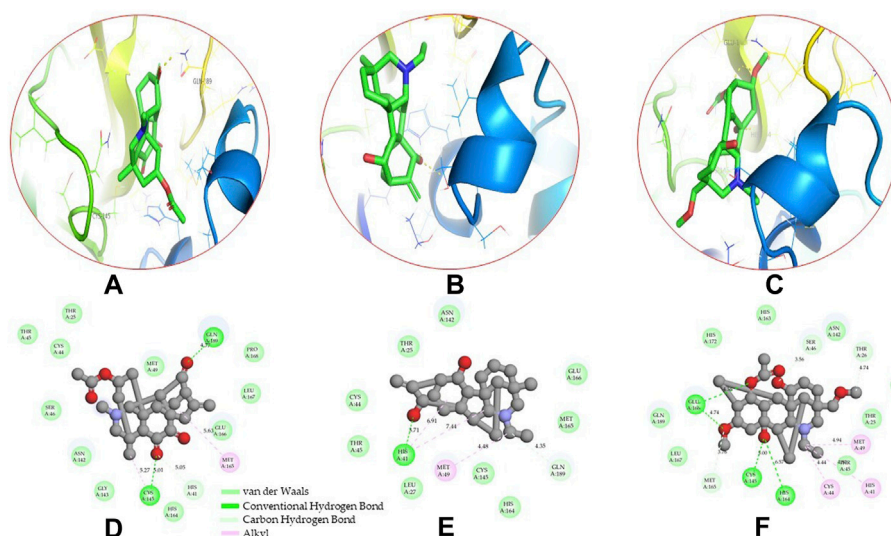


FIGURE 3 | 3D and 2D representations of the docking complexes between vilmorrianone, denudatine, condelphine, and M^{pro} of COVID-19; **(A)** 3D binding mode of vilmorrianone compound with M^{pro} active site; **(B)** 3D binding mode of denudatine compound with M^{pro} active site; **(C)** 3D binding mode of condelphine compound with M^{pro} active site. Molecules are represented as green sticks, M^{pro} residues depicted as atom-type color sticks, while yellow dotted lines show the hydrogen bonds between natural alkaloids and M^{pro} of COVID-19; **(D)** 2D representation of M^{pro} and vilmorrianone complex; **(E)** 2D representation of M^{pro} and denudatine complex; and **(F)** 2D representation of M^{pro} and condelphine complex. Molecular interactions such as van der Waals, conventional hydrogen bond, carbon–hydrogen bond, and alkyl are indicated by different colors. 3D and 2D representations of docking complexes are rendered with the help of PyMOL and Discovery Studio programs, respectively.

interaction between enzymes and ligands and investigate how receptors and ligands fit together in a significant manner (Meng et al., 2011; Jee et al., 2017; Salmaso and Moro 2018). This revolutionary method has been utilized in the process of computer-aided drug design (CADD) to discover potential drug candidates against different diseases (Wadood et al., 2013; Yu and MacKerell 2016; Salmaso and Moro 2018). There is a critical need for effective drugs against COVID-19 since treatments are still limited. The screening of drug candidates against important targets of SARS-CoV-2 is a valid approach to solve this dilemma. In the present study, to investigate the inhibition potency and gain insight into the possible mechanisms of Jadwar-derived natural alkaloids: panicutine, vilmorrianone, denudatine, and condelphine, molecular docking was performed on M^{pro} of COVID-19. These compounds were docked with the M^{pro} binding pocket. These bioactive molecules have been found to have a binding energy of -7.4 , -7.0 , -6.0 , and -5.9 kcal/mol, respectively. Out of these four docked compounds, the panicutine molecule ranked as the top interacting molecule with M^{pro} based on the calculated higher negative binding energy. Details of the chemical structures (2D and 3D), PubChem IDs, binding energy scores, hydrogen bonding, carbon–hydrogen (C–H) bonds, and alkyl interactions between the M^{pro} of COVID-19 and the four natural alkaloids are presented in **Table 1**. Two different programs, Discovery Studio and PyMOL, have been utilized to visualize the docking interactions in the form of 2D and 3D plots. As evident from **Figure 2**, the panicutine formed the hydrogen bond with Cys145 (3.08 Å) residue. Three residues namely, Ser46 (3.23 Å), Asn142 (4.06 Å), and Gly143 (2.75 Å), demonstrated the C–H bonds.

Panicutine was found to have two alkyl bonds with Met49 (4.76, 4.77, 5.35 Å) and Met165 (5.22 Å) residues. Seven residues such as Thr25, Leu27, His41, Ser144, Glu166, Arg188, and Gln189 manifest van der Waals (VdW) interactions. In the vilmorrianone compound, two residues Cys145 (5.01 Å) and Gln189 (4.77 Å) exhibit hydrogen bond interactions. The residues His41 (5.05) and Met165 (5.63 Å) showed C–H bonds and alkyl interactions, respectively. Eleven residues, namely, Thr25, Cys44, Thr45, Ser46, Met49, Asn142, Gly143, His164, Glu166, Leu167, and Pro168, demonstrated VdW interactions (**Figures 3A,D**). As shown in **Figures 3B and E**, residue His41 (5.26 Å) formed a single hydrogen bond with the denudatine molecule. Residue Gln189 (4.35 Å) showed C–H bonds. Three residues, namely, Met49 (4.92 Å), Cys145 (3.71 Å), and Met165 (4.59 Å), manifest alkyl bonds. Seven residues, namely, Leu27, Thr25, Cys44, Thr45, Asn142, His164, and Glu166, interacted with the denudatine compound *via* VdW. In the case of condelphine, three residues, namely, Cys145 (5.00 Å), His164 (6.57 Å), and Glu166 (4.36 Å, 4.74 Å), formed hydrogen bonds. Three residues Thr26 (4.74 Å), Ser46 (3.56 Å), and Met165 (3.76 Å) interacted through C–H bonds. Two residues Cys44 (3.83 Å) and Met49 (4.58 Å) formed alkyl bonds. Condelphine also manifests VdW interaction with nine residues, namely, Thr25, Thr45, His41, Asn142, Gly143, His163, Leu167, His172, and Gln189 (**Figures 3C and F**). We also predicted the contact maps of interactions between Jadwar-derived alkaloids and M^{pro} using the PDBsum web server to study the relative positions of the ligand molecules to the binding site of Mpro. As evident from **Supplementary Figure S1**, the predicted contact interaction maps show the hydrogen bond distances between 2.72 and 3.32 Å, while the nonbonded

interaction shows a higher range of distances between the binding site and ligand molecules. The binding site residues which are not involved in forming the hydrogen bond with ligands also show a higher range of distances. Previously, different *in silico* and *in vitro* studies reported a similar trend of tightly fitting inhibitors in the binding pocket of SARS-CoV-2 M^{Pro}, which confirms our study (Park et al., 2015; Aanouz et al., 2020; Bello et al., 2020; Chikhale et al., 2020; Krupanidhi et al., 2020; Kumar et al., 2020; Muhammad et al., 2020; Tripathi et al., 2020; Mitra et al., 2021; Varadharajan et al., 2021; Ali et al., 2022; Linani et al., 2022).

The interest to develop effective antiviral therapy against COVID-19 from the natural products of medicinal plants has increased globally. In several studies, phytochemicals were found to exhibit promising inhibitory effects against the M^{Pro} of COVID-19 (Benarba and Pandiella 2020; Huang et al., 2020; Verma et al., 2020; Ahmad et al., 2021). Kumar et al. (2020) demonstrated that three natural bioactive molecules, namely, ursolic acid, carvacrol, and oleanolic acid, were able to inhibit M^{Pro} protein in a significant manner. The binding affinity of these natural metabolites was found to be −5.9, −4.0, and −6.0 kcal/mol, respectively. In a recent study, Mishra et al. (2021) investigated the inhibition potencies of a set of natural compounds from medicinal plants as promising inhibitors against SARS-CoV-2 M^{Pro}. Based on the integrated molecular docking and modeling analysis, four phytochemicals, namely, amentoflavone, guggulsterone, puerarin, and piperine, were reported as top-ranked molecules. In a recent follow-up study, Kumar et al. (2021) explored sesame-derived natural compounds as antiviral molecules against M^{Pro}. The virtual screening of an in-house library which contains 36 natural sesame-derived compounds exhibited four bioactive molecules, namely, sesamin, sesaminol, sesamol, and sesamolol, as the top interacting compounds to the M^{Pro} of COVID-19. Out of these four molecules, the sesamin compound showed a higher negative binding energy of −6.7 kcal/mol and formed three strong hydrogen bonds with Asn151 (5.46 Å), Ser158 (4.38 Å), and Arg298 (6.05 Å) residues. The interaction mechanisms of natural compounds against M^{Pro} of COVID-19 reported in these studies are consistent with the docking results of the present study.

To date, several studies have been performed to investigate the inhibition potential of alkaloid compounds against different drug targets of SARS-CoV-2, namely, spike glycoprotein, nucleocapsid, angiotensin-converting enzyme 2 (ACE2), RdRp, and 3CL^{Pro} (Garg and Roy 2020; El-Demerdash et al., 2021; Ghosh et al., 2021; Ismail et al., 2021; Majnooni et al., 2021). A docking study reported thalimonine and sophaline D as potential inhibitors against the M^{Pro} of COVID-19 with binding energies of −8.39 and −8.79 kcal/mol, respectively (Garg and Roy 2020). In another study, a *Justicia adhatoda*-derived alkaloid compound, namely, anisotine, showed interaction with two catalytic residues (His41 and Cys145) of the M^{Pro} with a binding score of −7.9 kcal/mol (Ghosh et al., 2021). Of note, two alkaloids (quinoline and quinazoline) have been previously shown to be effective against three-drug targets of COVID-19, namely, M^{Pro}, spike glycoprotein, and ACE2 (Ismail et al., 2021). In a recent study, El-Demerdash et al. (2021) performed the virtual screening approach to screen a library of alkaloids to identify the promising inhibitors of multidrug targets for SARS-CoV-19. Based

TABLE 2 | Pharmacokinetics evaluation of natural alkaloids derived from Jadwar (*D. denudatum*).

Drug likeness	MW (g/mol)	Consensus log Po/W (range ≤5)	No. of H-bond acceptors (range ≤10)	No. of H-bond donors (range ≤5)	Molar refractivity (range 40–130)	Lipinski	Veber	Bioavailability score (range 0.4–0.6)	Synthetic accessibility (range >6)	TPSA (Å ²) (range >100)	No of rotatable bonds (range 1–10)	Solubility (mg/ml)
Alkaloids	(range ≤500 g)											
Panicutine (CID_44566630)	383.48	2.32	5	0	107.23	Yes	Yes	0.55	5.78	63.68	2	7.14e-02 (Soluble)
Vilmorrianone (CID_44566629)	397.46	1.72	6	0	107.43	Yes	Yes	0.55	5.82	80.75	2	9.82e-02 (Soluble)
Denudatine (CID_161515)	343.50	2.85	3	2	103.06	Yes	Yes	0.55	5.91	43.70	1	4.05e-01 (Soluble)
Condephine (CID_441720)	449.58	1.71	7	2	121.71	Yes	Yes	0.55	6.17	88.46	6	2.05e+00 (Soluble)

on the docking affinity, pentacyclic alkaloids, crambescidin and crambescidin, have been proposed as the top interacting molecules for five drug targets of COVID-19, namely, the M^{Pro}, spike glycoprotein, nucleocapsid phosphoprotein, membrane glycoprotein, and a nonstructural protein (nsp10). Our docking results may support previous reports on the inhibition potential of natural alkaloids against the M^{Pro} of COVID-19.

Pharmacokinetics Evaluation

The promising docking results enabled us to explore the ADMET properties of compounds, therefore, pharmacokinetic characteristics of four Jadwar-derived natural alkaloids (panicutine, vilmorrianone, denudatine, and condelphine) were evaluated prior to conducting the MD analysis using automated SwissADME server. The calculated pharmacokinetic properties of these compounds are shown in **Table 2**. Panicutine, vilmorrianone, denudatine, and condelphine have the following molecular weights, respectively: 383.48, 397.46, 343.50, and 449.58 g/mol; these four natural alkaloids have a molecular weight ≤ 500 g/mol. The molecular weight characteristics of these molecules suggested that they can easily be transported, diffused, and absorbed in the body in a significant manner (Lipinski 2004). The LogP values of panicutine, vilmorrianone, denudatine, and condelphine compounds were found to be 2.32, 1.72, 2.85, and 1.71, respectively, which meet the essential conditions of the Lipinski's rule of five. The calculated number of hydrogen bond donors of these four molecules was less than five which is in accordance with ADME as the number of H-bond donors must be ≤ 5 . The pharmacokinetics analysis suggested that panicutine, vilmorrianone, denudatine, and condelphine alkaloids represent the following values of the topological polar surface (TPSA): 63.68, 80.75, 43.70, and 88.46 Å². The lower TPSA values indicate the acceptable range of results and were found to be consistent with previous reports which showed the natural products as promising inhibitors of SARS-CoV-2 M^{Pro} (Kumar et al., 2020; Kumar et al., 2021; Mishra et al., 2021). The Jadwar-derived natural alkaloids proposed in the present study also meet the essential criteria of Veber's rule which defines the oral bioavailability of drug-like molecules. Panicutine, vilmorrianone, denudatine, and condelphine have molar refractivity values of 107.23, 107.43, 103.06, and 121.71, respectively; these alkaloids also present the synthetic accessibility (SA) scores of 5.78, 5.82, 5.91, and 6.17, respectively. In the drug designing process, SA has been considered as one of the essential parameters (Ertl and Schuffenhauer 2009). The calculated SA score of these molecules was found to be < 10 , which conforms with previous reports and reveals that these alkaloids can be synthesized easily (El-Demerdash et al., 2021; Ghosh et al., 2021; Kumar et al., 2021). Altogether, the pharmacokinetics evaluation suggested that these Jadwar-derived natural alkaloids harbor favorable drug-likeness properties and could be considered as therapeutic agents.

Docking Complexes Showed Stability Throughout the Molecular Dynamics Simulations on 200 ns

Over the past years, MD simulations-based methods have expanded dramatically in the field of structural biology and drug discovery to design novel therapeutics against contagious

diseases (Gajula et al., 2016; Hollingsworth and Dror, 2018). These revolutionary methods provide the ability to assess the stability and behavior of biological macromolecules and their molecular interactions with ligand molecules at very fine temporal resolution. In the present study, MD simulations were conducted for 200 ns using the docked conformation of M^{Pro}-panicutine, M^{Pro}-vilmorrianone, M^{Pro}-denudatine, and M^{Pro}-condelphine complexes to evaluate the stability and investigate the molecular interactions at the atomic level. The dynamic behavior of the simulated systems was analyzed using different functions, namely, RMSD, RMSF, hydrogen bond, Rg, and SASA.

Root Mean Square Deviations

Calculating the RMSD plot is a well-established method to investigate the stability of docking complexes. All the ligand and backbone RMSDs were graphically studied to check the stability of the docking complexes. From the RMSD plot shown in **Figure 4A**, it can be observed that the M^{Pro} backbone exhibited a constant range of stability throughout the simulation with a range between ~ 0.1 and ~ 0.43 nm. The average RMSD values for the M^{Pro} complexes with panicutine, vilmorrianone, denudatine, and condelphine were 0.25, 0.17, 0.19, and 0.21 nm, respectively (**Table 3**). The M^{Pro} Apo (pink) which has been considered as the control showed an average RMSD value of 0.2 nm. Along with the control, most of the docking complexes demonstrated small fluctuations between 20 and 80 ns. After 80 ns, the docking complexes showed stability up to 180 ns. As shown in **Figure 4A**, panicutine (green) showed the largest fluctuations between 20 and 100 ns; however, after 100 ns, panicutine (green) reflected stability on the 200-ns time scale around ~ 0.3 nm. Condelphine (black) presented as the second most fluctuated; this molecule showed two fluctuations between 10–80 and 180–200 ns. Two complexes, vilmorrianone (blue) and denudatine (red), also exhibited small fluctuations between 185 and 200 ns; however, no conformational changes were noted in the receptor protein structure upon the binding of ligand molecules. As compared with the control, all four ligands presented similar patterns of stability and average RMSD with small conformational changes. As expected, the calculated ligand RMSD plot also showed the constant range of target molecules' stability with small fluctuations over time. The average ligand RMSD values of panicutine, vilmorrianone, denudatine, and condelphine were 0.08, 0.07, 0.05, and 0.07, respectively. As shown in **Figure 4B**, the measured ligand plot represents the stability of ligands with small fluctuations in condelphine (black), panicutine (green), and vilmorrianone (blue). Condelphine (black) showed small fluctuations at the initial point (5–40 ns) on 0.9 nm. At the starting point, panicutine (green) also depicted small fluctuations up to 10 ns. After 10 ns, the vilmorrianone (blue) showed two fluctuations, the first between 10 and 20 ns on ~ 0.9 nm and the second between 175 and 185 ns on ~ 0.5 nm. Denudatine (red) reflected the straight line without notable fluctuations throughout the simulations on 200 ns. Therefore, the binding of denudatine made the complex more stable. Based on the analysis of the protein backbone and ligand RMSD plots, it

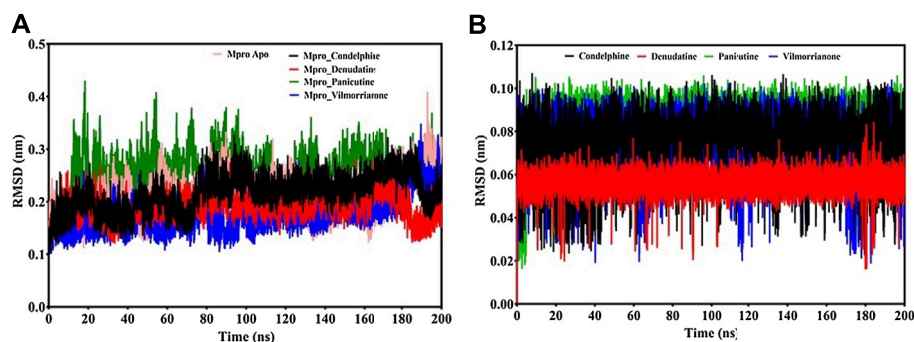


FIGURE 4 | RMSD analysis: **(A)** backbone RMSD plot of docking complexes and **(B)** ligand RMSD plot of docking complexes [control (pink), M^{pro}-condelphine (black), M^{pro}-denudatine (red), M^{pro}-panicutine (green), and M^{pro}-vilmorrianone (blue)].

TABLE 3 | Calculated values of different components of MD simulations such as protein backbone RMSD, ligand RMSD, RMSF, Rg, SASA, and H-bonds.

Components	Apo	Condelfine	Denudatine	Vilmorrianone	Panicutine
RMSD protein (nm)	0.2	0.21	0.19	0.17	0.25
RMSD ligand (nm)	-	0.07	0.05	0.07	0.08
RMSF (nm)	0.19	0.1	0.1	0.11	0.14
Rg (nm)	2.24	2.2	2.23	2.23	2.25
SASA (nm ²)	153.41	147.57	150.46	150.22	153.96
H-bonds (#)	-	2	1	1	1

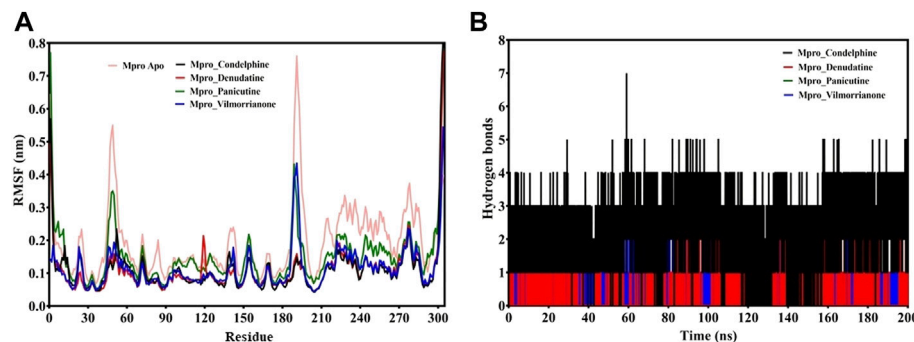


FIGURE 5 | **(A)** RMSF plot of docking complexes and **(B)** distribution of hydrogen bonds [control (pink), M^{pro}-condelphine (black), M^{pro}-denudatine (red), M^{pro}-panicutine (green), and M^{pro}-vilmorrianone (blue)].

can be concluded that the measured RMSDs demonstrated minimally and the docking complexes were stable with significant potential to compare with structures available in the structure repositories.

Root Mean Square Fluctuations

The RMSF plot analysis was conducted with the primary goal to assess the mobility of residues upon binding of the ligand molecules. Per the general phenomena of the RMSF analysis, a high fluctuations score depicts more flexibility and unstable bonds, while a low value represents the correct structure regions in the docking complexes (Martínez 2015; Gajula et al., 2016). The RMSFs of the alpha carbon atoms of all simulated systems were investigated in the present study and

are shown in **Figure 5A**. All five simulated systems, namely, the control, M^{pro}-panicutine, M^{pro}-vilmorrianone, M^{pro}-denudatine, and M^{pro}-condelphine, demonstrated a close pattern of fluctuations throughout the simulation on a 200-ns time scale. The average RMSF values of the control, M^{pro}-panicutine, M^{pro}-vilmorrianone, M^{pro}-denudatine, and M^{pro}-condelphine docking complexes were 0.19, 0.14, 0.11, 0.1, and 0.1 nm, respectively. These values clearly reflect that all docking complexes show relatively less conformation fluctuations than the control system. It is observed from the RMSF plot on **Figure 5A** that two complexes, M^{pro}-panicutine (green) and M^{pro}-vilmorrianone (blue) show the highest peak between 180 and 200 residues on 0.4 nm. The fewer fluctuations noted in the protein-ligand complexes support the docking

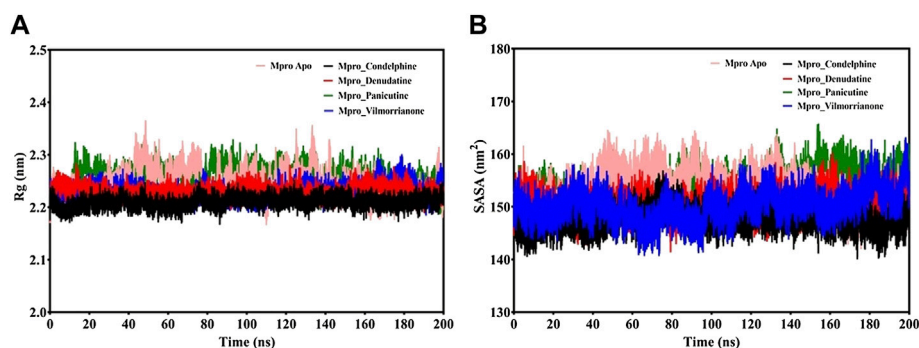


FIGURE 6 | (A) Rg plot of docking complexes and **(B)** SASA plot of docking complexes [control (pink), M^{Pro}-condelphine (black), M^{Pro}-denudatine (red), M^{Pro}-panicutine (green), and M^{Pro}-vilmorrianone (blue)].

findings and reveal that the M^{Pro} significantly interacts with panicutine, vilmorrianone, denudatine, and condelphine compounds.

Hydrogen Bond Analysis

To determine the strength of binding of panicutine, vilmorrianone, denudatine, and condelphine to the target M^{Pro}, the number of intermolecular hydrogen bonds was calculated by utilizing the MD trajectories. Panicutine (green), vilmorrianone (blue), and denudatine (red) formed one hydrogen bond with M^{Pro}, while condelphine (black) manifested two hydrogen bonds with the target receptor throughout the MD simulation on the scale of 200 ns. These results were the confirmation of the hydrogen bond interactions predicted using the molecular docking approach. **Figure 5B** represents the distribution of hydrogen bonds. The ligand molecules had a constant range of hydrogen bonds between one and two during the whole simulation. The results of the hydrogen bond analysis revealed that the intermolecular hydrogen bonds were stable, and the natural alkaloids considered in the present study could maintain a strong molecular interaction with the active site of the M^{Pro} in a significant manner.

Radius of Gyration and Solvent Accessible Surface Area Analysis

MD trajectories corresponding to the four docking complexes (M^{Pro}-panicutine, M^{Pro}-vilmorrianone, M^{Pro}-denudatine, and M^{Pro}-condelphine) were further evaluated with the help of integrated Rg and SASA analyses. The Rg plot analysis was calculated to extract the compactness of the simulated systems with the time scale. As evident from **Figure 6A**, the Rg values of the control and docking complexes are noted between ~2.2 and ~2.3 nm during the simulation on 200 ns. The average Rg values of the control, M^{Pro}-panicutine, M^{Pro}-vilmorrianone, M^{Pro}-denudatine, and M^{Pro}-condelphine docking complexes are 2.24, 2.25, 2.23, 2.23, and 2.2 nm, respectively. These Rg values demonstrate that all the protein–ligand complexes except M^{Pro}-panicutine (green) showed relatively less value than the control. The calculated Rg values for these alkaloid

compounds are in consent with previous findings which have reported that the medicinal plants–derived bioactive molecules are potential inhibitors against M^{Pro} of COVID-19 (Kumar et al., 2020; Mishra et al., 2021). In the present study, the calculated Rg values confirm the stability of every docking complex and reflect that the binding of the natural alkaloids does not induce structural changes throughout the simulation on a 200-ns time scale.

The analysis of the SASA plot is an important step toward the investigation of the receptor exposed to solvents during the MD simulations on different nanoseconds. Per the general rule, a higher SASA value indicates the expansion of protein volume. During MD simulations, low fluctuations for the docking complexes are always considered good. The calculated SASA plot for the control and docking complexes of the natural alkaloids and M^{Pro} is given in **Figure 6B**. The SASA values are noted between 145 and 160 nm². The average Rg values of the control, M^{Pro}-panicutine, M^{Pro}-vilmorrianone, M^{Pro}-denudatine, and M^{Pro}-condelphine docking complexes are 153.41, 153.96, 150.22, 150.46, and 147.57 nm², respectively. All the complexes showed the SASA values as less than those of the control system except for the M^{Pro}-panicutine complex. The outcome of the SASA analysis suggests the stability of the docking complexes and also indicates that binding of panicutine, vilmorrianone, denudatine, and condelphine does not affect the protein folding.

Calculation of Binding Free Energy

To achieve accurate binding free energy estimation of the protein ligand complexes (M^{Pro}-panicutine, M^{Pro}-vilmorrianone, M^{Pro}-denudatine, and M^{Pro}-condelphine), we employed the MmPbSaStat.py python script embedded in the g_mmpbsa module. The MM/PBSA method is a widely used accurate method to calculate the ligand-binding affinities (Hou et al., 2011; Genheden and Ryde 2015). The scores of the calculated binding free energy (van der Waal energy, electrical energy, polar solvation energy, and SASA) are provided in **Table 4**. As reported in previous reports, the final binding energy of the protein–ligand complex is represented by the cumulative sum of the different energies (van der Walls, electrostatic, polar solvation, and SASA) (Jee et al., 2017; Pathak et al.,

TABLE 4 | Calculated free binding energy of docking complexes of M^{Pro} and natural alkaloids (panicutine, vilmorrianone, denudatine, and condelphine).

Complex	Binding energy (kJ/mol)	van der Waal energy (ΔE_{vdW}) (kJ/mol)	Electrical energy (ΔE_{elec}) (kJ/mol)	Polar solvation energy (ΔG_{polar}) (kJ/mol)	SASA energy (kJ/mol)
Panicutine	-140.758 \pm 23.707	-168.672 \pm 19.599	-29.311 \pm 13.245	71.015 \pm 10.117	-13.790 \pm 1.331
Vilmorrianone	-147.091 \pm 10.059	-192.801 \pm 7.972	-14.844 \pm 7.307	78.543 \pm 9.421	-14.988 \pm 0.868
Denudatine	-138 \pm 13.873	-160.823 \pm 11.989	-34.509 \pm 11.582	70.585 \pm 9.442	-13.811 \pm 0.802
Condelphine	-127.939000 \pm 13.931	-155.616 \pm 12.975	-61.506 \pm 8.677	104.370 \pm 9.388	-15.187 \pm 1.198

2021). In the present study, all types of energies such as van der Waals, electrostatic, polar solvation, and SASA contributed to molecular interactions between the alkaloids and M^{Pro} of COVID-19 in a significant manner. The calculated binding free energy of the Jadwar-derived alkaloids is as follows: panicutine, -140.758 \pm 23.707 kJ/mol; vilmorrianone, -147.091 \pm 10.059 kJ/mol; denudatine, -138 \pm 13.873 kJ/mol; and condelphine, -127.939000 \pm 13.931 kJ/mol. As a general fact, more negative values of the free binding energy depicted a stronger molecular interaction and increased affinity between the receptor protein and ligand molecules. Vilmorrianone (-147.091 \pm 10.059 kJ/mol) possesses the maximum negative binding energy when compared with the other natural alkaloids considered in the present study, while panicutine (-140.758 \pm 23.707 kJ/mol) exhibited the second least binding energy based on the MM/PBSA method of estimation. The complex of M^{Pro}-vilmorrianone showed lower binding free energy because of its stable interactions with binding site residues of M^{Pro} at the atomic level. As evident from the MD simulation results, vilmorrianone also showed RMSD behavior in an acceptable range and displayed stability with small fluctuations throughout the MD simulation with an average RMSD value of 0.07 nm. Taken together, these natural alkaloids with maximum negative energy support the concept of design and validate the CADD approach; they also demonstrate and assure the inhibition potential of Jadwar-derived alkaloids against the M^{Pro} of COVID-19.

CONCLUSION

The M^{Pro} of SARS-CoV-2 is a well-validated drug target due to its principal role in viral replication. The screening of phytochemicals against this important drug target has become a promising strategy in the design of potential drug candidates using the CADD approach. In the present study, the inhibitory potential of four natural Jadwar-derived alkaloids, namely, panicutine, vilmorrianone, denudatine, and condelphine, that targeted M^{Pro} was investigated using the integrated molecular docking and modeling methods. Based on the docking results, we demonstrated that all four bioactive molecules significantly bind and stably interact with the active site of M^{Pro}. Furthermore, MD simulations analysis was performed over 200 ns to evaluate the binding position and structural stability of the docking complexes using different components from the MD trajectories. The calculation of the binding free energy supported the MD simulation in a significant manner and confirmed their stability at the atomic level. Furthermore, Lipinski's rule of five and ADME properties-based validation of these natural compounds suggested positive drug-likeness properties, which is an essential step

toward demonstrating drug safety. The inhibition potential of these Jadwar-derived natural alkaloids against M^{Pro} can also be validated in the wet-lab setting with the aid of cell culture and small animal experiments. Based on the molecular modeling investigations, the current study suggests that panicutine, vilmorrianone, denudatine, and condelphine have the potential to inhibit the M^{Pro} of COVID-19 and, in the future, may be candidates for anti-viral therapy.

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

AK and DK contributed to the conception and design of the study. AK drafted the manuscript. MS contributed to data analysis. DK supervised the study. CR and DK edited the final manuscript. All authors gave their final approval and agreed to be accountable for all aspects of the work.

FUNDING

This work was supported by grants from the Canadian Institutes of Health Research, Genome Canada/Atlantic Genome, Research Nova Scotia, Dalhousie Medical Research Foundation, and the Li-Ka Shing Foundation.

ACKNOWLEDGMENTS

The authors are thankful to Sanjay Kumar, the CSIR-Central Drug Research Institute (CDRI), Lucknow, India, for technical support in the molecular dynamics simulations. Nikki Kelvin (*The Journal of Infection in Developing Countries*) provided editing assistance throughout the preparation of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2022.898874/full#supplementary-material>

REFERENCES

- Aanouz, I., Belhassan, A., El-Khatibi, K., Lakhli, T., El-Idrissi, M., and Bouachrine, M. (2020). Moroccan Medicinal Plants as Inhibitors against SARS-CoV-2 Main Protease: Computational Investigations. *J. Biomol. Struct. Dyn.* 39, 2971–2979. doi:10.1080/07391102.2020.1758790
- Abdelli, I., Hassani, F., Bekkel Briki, S., and Ghalem, S. (2020). In Silico study the Inhibition of Angiotensin Converting Enzyme 2 Receptor of COVID-19 by Ammoides Verticillata Components Harvested from Western Algeria. *J. Biomol. Struct. Dyn.* 39, 1–14. doi:10.1080/07391102.2020.1763199
- Abid, M., Gosh, A. K., and Ali Khan, N. (2017). In Vivo Psychopharmacological Investigation of Delphinium Denudatum and Amaranthus Spinosa Extracts on Wistar Rats. *Bcn* 8, 503–512. doi:10.29252/NIRP.BCN.8.6.503
- Abraham, M. J., and Gready, J. E. (2011). Optimization of Parameters for Molecular Dynamics Simulation Using Smooth Particle-Mesh Ewald in GROMACS 4.5. *J. Comput. Chem.* 32, 2031–2040. doi:10.1002/jcc.21773
- Ahmad, H., Ahmad, S., Khan, E., Shahzad, A., Ali, M., Tahir, M. N., et al. (2017). Isolation, crystal Structure Determination and Cholinesterase Inhibitory Potential of Isotalatizidine Hydrate from Delphinium Denudatum. *Pharm. Biol.* 55, 680–686. doi:10.1080/13880209.2016.1240207
- Ahmad, S., Zahiruddin, S., Parveen, B., Basist, P., Parveen, A., Parveen, R., et al. (2021). Indian Medicinal Plants and Formulations and Their Potential against COVID-19—Preclinical and Clinical Research. *Front. Pharmacol.* 11, 2470. doi:10.3389/fphar.2020.578970/BIBTEX
- Aldeghi, M., Bodkin, M. J., Knapp, S., and Biggin, P. C. (2017). Statistical Analysis on the Performance of Molecular Mechanics Poisson-Boltzmann Surface Area versus Absolute Binding Free Energy Calculations: Bromodomains as a Case Study. *J. Chem. Inf. Model.* 57, 2203–2221. doi:10.1021/acs.jcim.7b00347
- Ali, S., Alam, M., Khatoun, F., Fatima, U., Elsbali, A. M., Adnan, M., et al. (2022). Natural Products Can Be Used in Therapeutic Management of COVID-19: Probable Mechanistic Insights. *Biomed. Pharmacother.* 147, 112658. doi:10.1016/j.biopha.2022.112658
- Alsobaie, S. (2021). Understanding the Molecular Biology of SARS-CoV-2 and the COVID-19 Pandemic: A Review. *Idr* 14, 2259–2268. doi:10.2147/IDR.S306441
- Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., and Garry, R. F. (2020/2020). The Proximal Origin of SARS-CoV-2. *Nat. Med.* 26, 450–452. doi:10.1038/s41591-020-0820-9
- Arden, K. E., Nissen, M. D., Sloots, T. P., and Mackay, I. M. (2005). New Human Coronavirus, HCoV-NL63, Associated with Severe Lower Respiratory Tract Disease in Australia. *J. Med. Virol.* 75, 455–462. doi:10.1002/jmv.20288
- Atta-ur-Rahman, and Choudhary, M. I. (1999). Recent Studies on Bioactive Natural Products. *Pure Appl. Chem.* 71, 1079–1081. doi:10.1351/PAC199971061079/MACHINEREADABLECITATION/RIS
- Atta-ur-Rahman, Nasreen, A., Akhtar, F., Shekhani, M. S., Clardy, J., Parvez, M., et al. (1997). Antifungal Diterpenoid Alkaloids from Delphinium Denudatum. *J. Nat. Prod.* 60, 472–474. doi:10.1021/NP960663N
- Bello, M., Martínez-Muñoz, A., and Balbuena-Rebolledo, I. (2020). Identification of Saquinavir as a Potent Inhibitor of Dimeric SARS-CoV2 Main Protease through MM/GBSA. *J. Mol. Model.* 26, 340. doi:10.1007/s00894-020-04600-4
- Benarba, B., and Pandiella, A. (2020). Medicinal Plants as Sources of Active Molecules against COVID-19. *Front. Pharmacol.* 11, 1189. doi:10.3389/fphar.2020.01189/BIBTEX
- Berman, H. M. (2000). The Protein Data Bank. *Nucleic Acids Res.* 28, 235–242. doi:10.1093/nar/28.1.235
- Burley, S. K., Berman, H. M., Bhikadiya, C., Bi, C., Chen, L., Di Costanzo, L., et al. (2018). RCSB Protein Data Bank: Biological Macromolecular Structures Enabling Research and Education in Fundamental Biology, Biomedicine, Biotechnology and Energy. *Nucleic Acids Res.* 47, D464–D474. doi:10.1093/nar/ky1004
- Chen, T., Wu, D., Chen, H., Yan, W., Yang, D., Chen, G., et al. (2020). Clinical Characteristics of 113 Deceased Patients with Coronavirus Disease 2019: Retrospective Study. *BMJ* 368, m1295. doi:10.1136/bmj.m1295
- Cheng, V. C. C., Lau, S. K. P., Woo, P. C. Y., and Yuen, K. Y. (2007). Severe Acute Respiratory Syndrome Coronavirus as an Agent of Emerging and Reemerging Infection. *Clin. Microbiol. Rev.* 20, 660–694. doi:10.1128/cmr.00023-07
- Chikhale, R. V., Gurav, S. S., Patil, R. B., Sinha, S. K., Prasad, S. K., Shakya, A., et al. (2020). Sars-cov-2 Host Entry and Replication Inhibitors from Indian Ginseng: an In-Silico Approach. *J. Biomol. Struct. Dyn.*, 1–12. doi:10.1080/07391102.2020.1778539
- Chowdhury, P. (2021). In Silico investigation of Phytoconstituents from Indian Medicinal Herb 'Tinospora Cordifolia (Giloy)' against SARS-CoV-2 (COVID-19) by Molecular Dynamics Approach. *J. Biomol. Struct. Dyn.* 39, 6792–6809. doi:10.1080/07391102.2020.1803968/SUPPL_FILE/TBSD_A_1803968_SM4930.PDF
- 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, 42717. doi:10.1038/srep42717
- Das, S., Sarmah, S., Lyndem, S., and Singha Roy, A. (2020). An Investigation into the Identification of Potential Inhibitors of SARS-CoV-2 Main Protease Using Molecular Docking Study. *J. Biomol. Struct. Dyn.*, 1–11. doi:10.1080/07391102.2020.1763201
- de Leuw, P., and Stephan, C. (2017). Protease Inhibitors for the Treatment of Hepatitis C Virus Infection. *GMS Infect. Dis.* 5, Doc08. doi:10.3205/ID000034
- de Wit, E., van Doremalen, N., Falzarano, D., and Munster, V. J. (2016). SARS and MERS: Recent Insights into Emerging Coronaviruses. *Nat. Rev. Microbiol.* 14, 523–534. doi:10.1038/nrmicro.2016.81
- El-Demerdash, A., Metwaly, A. M., Hassan, A., Abd El-Aziz, T. M., Elkaeed, E. B., Eissa, I. H., et al. (2021/2021). Comprehensive Virtual Screening of the Antiviral Potentialities of Marine Polycyclic Guanidine Alkaloids against SARS-CoV-2 (COVID-19). *Biomolecules* 11, 460. doi:10.3390/Biom11030460
- Ertl, P., and Schuffenhauer, A. (2009). Estimation of Synthetic Accessibility Score of Drug-like Molecules Based on Molecular Complexity and Fragment Contributions. *J. Cheminform* 1, 8. doi:10.1186/1758-2946-1-8
- Fehr, A. R., and Perlman, S. (2015). Coronaviruses: An Overview of Their Replication and Pathogenesis. *Coronaviruses* 1282, 1–23. doi:10.1007/978-1-4939-2438-7_1
- Fuhrmann, J., Rurainski, A., Lenhof, H.-P., and Neumann, D. (2010). A New Lamarckian Genetic Algorithm for Flexible Ligand-Receptor Docking. *J. Comput. Chem.* 31, 1911. doi:10.1002/jcc.21478
- Gajula, M., Kumar, A., and Ijaq, J. (2016). Protocol for Molecular Dynamics Simulations of Proteins. *BIO-PROTOCOL* 6, e2051. doi:10.21769/bioprotoc.2051
- Garg, S., and Roy, A. (2020). In Silico analysis of Selected Alkaloids against Main Protease (Mpro) of SARS-CoV-2. *Chemico-Biological Interactions* 332, 109309. doi:10.1016/j.cbi.2020.109309
- Genheden, S., and Ryde, U. (2015). The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities. *Expert Opin. Drug Discov.* 10, 449–461. doi:10.1517/17460441.2015.1032936
- Ghosh, R., Chakraborty, A., Biswas, A., and Chowdhuri, S. (2021). Identification of Alkaloids from Justicia Adhatoda as Potent SARS CoV-2 Main Protease Inhibitors: An In Silico Perspective. *J. Mol. Struct.* 1229, 129489. doi:10.1016/j.molstruc.2020.129489
- Helmy, Y. A., Fawzy, M., Elswad, A., Sobieh, A., Kenney, S. P., and Shehata, A. A. (2020). The COVID-19 Pandemic: A Comprehensive Review of Taxonomy, Genetics, Epidemiology, Diagnosis, Treatment, and Control. *Jcm* 9, 1225. doi:10.3390/jcm9041225
- Hollingsworth, S. A., and Dror, R. O. (2018). Molecular Dynamics Simulation for All. *Neuron* 99, 1129–1143. doi:10.1016/j.neuron.2018.08.011
- Hou, T., Wang, J., Li, Y., and Wang, W. (2011). Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J. Chem. Inf. Model.* 51, 69–82. doi:10.1021/CI100275A/SUPPL_FILE/CI100275A_SI_001.PDF
- Huang, J., Tao, G., Liu, J., Cai, J., Huang, Z., and Chen, J. X. (2020). Current Prevention of COVID-19: Natural Products and Herbal Medicine. *Front. Pharmacol.* 11, 588508. doi:10.3389/fphar.2020.588508/BIBTEX
- Huang, J., and Mackerell, A. D. (2013). CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* 34, 2135–2145. doi:10.1002/jcc.23354
- Hulswit, R. J. G., de Haan, C. A. M., and Bosch, B.-J. (2016). Coronavirus Spike Protein and Tropism Changes. *Adv. Virus Res.* 96, 29–57. doi:10.1016/bs.aivir.2016.08.004
- Ismail, E. M. O. A., Shantier, S. W., Mohammed, M. S., Musa, H. H., Osman, W., and Mothana, R. A. (2021). Quinoline and Quinazoline Alkaloids against Covid-19: An In Silico Multitarget Approach. *J. Chem.* 2021, 1–11. doi:10.1155/2021/3613268

- Jain, R., Siddiqui, M. Z., Waseem, M., Raghav, R., Jabin, A., Jhanjee, S., et al. (2021). *Therapeutic Usefulness of Delphinium Denudatum (Jadwar): An Update* 97 *International Journal of Herbal Medicine*, 97–100. Available at: www.florajournal.com (Accessed February 10, 2022).
- Jamwal, S., Gautam, A., Elsworth, J., Kumar, M., Chawla, R., and Kumar, P. (2020). An Updated Insight into the Molecular Pathogenesis, Secondary Complications and Potential Therapeutics of COVID-19 Pandemic. *Life Sci.* 257, 118105. doi:10.1016/j.lfs.2020.118105
- Jee, B., Kumar, S., Yadav, R., Singh, Y., Kumar, A., and Sharma, N. (2017). Ursolic Acid and Carvacrol May Be Potential Inhibitors of Dormancy Protein Small Heat Shock protein16.3 of Mycobacterium tuberculosis. *J. Biomol. Struct. Dyn.* 36, 3434–3443. doi:10.1080/07391102.2017.1389305
- Jin, Z., Zhao, Y., Sun, Y., Zhang, B., Wang, H., Wu, Y., et al. (2020). Structural Basis for the Inhibition of SARS-CoV-2 Main Protease by Antineoplastic Drug Carmofur. *Nat. Struct. Mol. Biol.* 27, 529–532. doi:10.1038/s41594-020-0440-6
- Joshi, R. S., Jagdale, S. S., Bansode, S. B., Shankar, S. S., Tellis, M. B., Pandya, V. K., et al. (2020). Discovery of Potential Multi-Target-Directed Ligands by Targeting Host-specific SARS-CoV-2 Structurally Conserved Main Protease. *J. Biomol. Struct. Dyn.*, 1–16. doi:10.1080/07391102.2020.1760137
- Khan, S. A., Zia, K., Ashraf, S., Uddin, R., and Ul-Haq, Z. (2020). Identification of Chymotrypsin-like Protease Inhibitors of SARS-CoV-2 via Integrated Computational Approach. *J. Biomol. Struct. Dyn.* 39, 2607–2616. doi:10.1080/07391102.2020.1751298
- Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., et al. (2020). PubChem in 2021: New Data Content and Improved Web Interfaces. *Nucleic Acids Res.* 49, D1388–D1395. doi:10.1093/nar/gkaa971
- Krupanidhi, S., Abraham Peele, K., Venkateswarulu, T. C., Ayyagari, V. S., Nazneen Bobby, M., John Babu, D., et al. (2020). Screening of Phytochemical Compounds of Tinospora Cordifolia for Their Inhibitory Activity on SARS-CoV-2: an In Silico Study. *J. Biomol. Struct. Dyn.* 1, 1–5. doi:10.1080/07391102.2020.1787226
- Kumar, A., Choudhir, G., Shukla, S. K., Sharma, M., Tyagi, P., Bhushan, A., et al. (2020). Identification of Phytochemical Inhibitors against Main Protease of COVID-19 Using Molecular Modeling Approaches. *J. Biomol. Struct. Dyn.* 39, 3760–3770. doi:10.1080/07391102.2020.1772112
- Kumar, A., Kumar, R., Sharma, M., Kumar, U., Gajula, M., and Singh, K. (2018). Uttarakhand Medicinal Plants Database (UMPDB): A Platform for Exploring Genomic, Chemical, and Traditional Knowledge. *Data* 3, 7. doi:10.3390/data3010007
- Kumar, A., Mishra, D. C., Angadi, U. B., Yadav, R., Rai, A., and Kumar, D. (2021). Inhibition Potencies of Phytochemicals Derived from Sesame against SARS-CoV-2 Main Protease: A Molecular Docking and Simulation Study. *Front. Chem.* 9, 773. doi:10.3389/FCHEM.2021.744376/BIBTEX
- Kumari, R., Kumar, R., and Lynn, A. (2014). g_mmpbsa-A GROMACS Tool for High-Throughput MM-PBSA Calculations. *J. Chem. Inf. Model.* 54, 1951–1962. doi:10.1021/ci500020m
- Kutzner, C., Páll, S., Fechner, M., Esztermann, A., Groot, B. L., and Grubmüller, H. (2019). More Bang for Your Buck: Improved Use of GPU Nodes for GROMACS 2018. *J. Comput. Chem.* 40, 2418–2431. doi:10.1002/jcc.26011
- Lee, N., Allen Chan, K. C., Hui, D. S., Ng, E. K. O., Wu, A., Chiu, R. W. K., et al. (2004). Effects of Early Corticosteroid Treatment on Plasma SARS-Associated Coronavirus RNA Concentrations in Adult Patients. *J. Clin. Virol.* 31, 304–309. doi:10.1016/j.jcv.2004.07.006
- Linani, A., Benarous, K., Bou-Salah, L., Yousfi, M., and Goumri-Said, S. (2022). Exploring Structural Mechanism of COVID-19 Treatment with Glutathione as a Potential Peptide Inhibitor to the Main Protease: Molecular Dynamics Simulation and MM/PBSA Free Energy Calculations Study. *Int. J. Pept. Res. Ther.* 28, 1–16. doi:10.1007/S10989-022-10365-6/TABLES/2
- Lipinski, C. A. (2004). Lead- and Drug-like Compounds: the Rule-Of-Five Revolution. *Drug Discov. Today Tech.* 1, 337–341. doi:10.1016/j.ddtec.2004.11.007
- Majnooni, M. B., Fakhri, S., Bahrami, G., Naseri, M., Farzaei, M. H., and Echeverria, J. (2021). Alkaloids as Potential Phytochemicals against SARS-CoV-2: Approaches to the Associated Pivotal Mechanisms. *Evid Based. Complement. Altern. Med.* 2021, 1–21. doi:10.1155/2021/6632623
- Martinez, L. (2015). Automatic Identification of Mobile and Rigid Substructures in Molecular Dynamics Simulations and Fractional Structural Fluctuation Analysis. *PLoS ONE* 10, e0119264. doi:10.1371/JOURNAL.PONE.0119264
- Meng, X.-Y., Zhang, H.-X., Mezei, M., and Cui, M. (2011). Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Cad* 7, 146–157. doi:10.2174/157340911795677602
- Mengist, H. M., Dilnessa, T., and Jin, T. (2021). Structural Basis of Potential Inhibitors Targeting SARS-CoV-2 Main Protease. *Front. Chem.* 9, 7. doi:10.3389/FCHEM.2021.622898/BIBTEX
- Mishra, A., Pathak, Y., Kumar, A., and Mishra, S. (2021). Natural Compounds as Potential Inhibitors of SARS-CoV-2 Main Protease: An In-Silico Study. *Asian Pac. J. Trop. Biomed.* 11, 155. doi:10.4103/2221-1691.310202
- Mitra, D., Verma, D., Mahakur, B., Kamboj, A., Srivastava, R., Gupta, S., et al. (2021). Molecular Docking and Simulation Studies of Natural Compounds of Vitex Negundo L. Against Papain-like Protease (PLpro) of SARS CoV-2 (Coronavirus) to Conquer the Pandemic Situation in the World. *J. Biomol. Struct. Dyn.*, 1–22. doi:10.1080/07391102.2021.1873185
- Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., et al. (2009). AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* 30, 2785–2791. doi:10.1002/jcc.21256
- Muhammad, I. A., Muangchoo, K., Muhammad, A., Ajingi, Y. u. S., Muhammad, I. Y., Umar, I. D., et al. (2020). A Computational Study to Identify Potential Inhibitors of SARS-CoV-2 Main Protease (Mpro) from Eucalyptus Active Compounds. *Computation* 8, 79. doi:10.3390/computation8030079
- Mukherjee, S., Dasgupta, S., Adhikary, T., Adhikari, U., and Panja, S. S. (2021). Structural Insight to Hydroxychloroquine-3c-like Proteinase Complexation from SARS-CoV-2: Inhibitor Modelling Study through Molecular Docking and MD-simulation Study. *J. Biomol. Struct. Dyn.* 39, 7322–7334. doi:10.1080/07391102.2020.1804458/SUPPL_FILE/TBSD_A_1804458_SM5671.DOCX
- Nizami, Q., and Jafri, M. A. (2006). Unani Drug, Jadwar (Delphinium Denudatum Wall.) - a Review. *Indian J. Traditional Knowledge* 5, 463–467. Available at: <https://eurekamag.com/research/012/897/012897519.php> (Accessed February 10, 2022).
- O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., and Hutchison, G. R. (2011). Open Babel: An Open Chemical Toolbox. *J. Cheminform* 3, 33. doi:10.1186/1758-2946-3-33
- Ou, T., Mou, H., Zhang, L., Ojha, A., Choe, H., and Farzan, M. (2021). Hydroxychloroquine-mediated Inhibition of SARS-CoV-2 Entry Is Attenuated by TMPRSS2. *Plos Pathog.* 17, e1009212. doi:10.1371/journal.ppat.1009212
- Pak, A., Adegboye, O. A., Adekunle, A. I., Rahman, K. M., McBryde, E. S., and Eisen, D. P. (2020). Economic Consequences of the COVID-19 Outbreak: the Need for Epidemic Preparedness. *Front. Public Health* 8, 241. doi:10.3389/FPUH.2020.00241/BIBTEX
- Pal, M., Berhanu, G., Desalegn, C., and Kandi, V. (2020). Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2): An Update. *Cureus* 12, e7423. doi:10.7759/CUREUS.7423
- Park, J.-Y., Ko, J.-A., Kim, D. W., Kim, Y. M., Kwon, H.-J., Jeong, H. J., et al. (2015). Chalcones Isolated from Angelica Keiskei Inhibit Cysteine Proteases of SARS-CoV. *J. Enzyme Inhib. Med. Chem.* 31, 23–30. doi:10.3109/14756366.2014.1003215
- Patel, C. N., Kumar, S. P., Pandya, H. A., and Rawal, R. M. (2021). Identification of Potential Inhibitors of Coronavirus Hemagglutinin-Esterase Using Molecular Docking, Molecular Dynamics Simulation and Binding Free Energy Calculation. *Mol. Divers.* 25, 421–433. doi:10.1007/S11030-020-10135-W/FIGURES/7
- Pathak, Y., Mishra, A., Choudhir, G., Kumar, A., and Tripathi, V. (2021). Rifampicin and Letemovir as Potential Repurposed Drug Candidate for COVID-19 Treatment: Insights from an In-Silico Study. *Pharmacol. Rep.* 73, 926–938. doi:10.1007/s43440-021-00228-0
- Payne, S. (2017). Family Coronaviridae. *Fam. Coronaviridae Viruses*, 149–158. doi:10.1016/B978-0-12-803109-4.00017-9
- Pelletier, S. W. (1996). *Alkaloids: Chemical and Biological Perspectives*. New York, NY: Elsevier.
- Rahman, S., Ali Khan, R., and Kumar, A. (2002). Experimental Study of the Morphine De-addiction Properties of Delphinium Denudatum Wall. *BMC Complement. Altern. Med.* 2, 6. doi:10.1186/1472-6882-2-6
- Raj, V., Park, J. G., Cho, K.-H., Choi, P., Kim, T., Ham, J., et al. (2021). Assessment of Antiviral Potencies of Cannabinoids against SARS-CoV-2 Using

- Computational and *In Vitro* Approaches. *Int. J. Biol. Macromolecules* 168, 474–485. doi:10.1016/j.ijbiomac.2020.12.020
- Rangsinth, P., Sillapachaiyaporn, C., Nilkhet, S., Tencomnao, T., Ung, A. T., and Chuchawankul, S. (2021). Mushroom-derived Bioactive Compounds Potentially Serve as the Inhibitors of SARS-CoV-2 Main Protease: An *In Silico* Approach. *J. Traditional Complement. Med.* 11, 158–172. doi:10.1016/J.JTCME.2020.12.002
- Rastogi, M., Pandey, N., Shukla, A., and Singh, S. K. (2020). SARS Coronavirus 2: from Genome to Infectome. *Respir. Res.* 21, 1–15. doi:10.1186/S12931-020-01581-Z
- Raza, M., Shaheen, F., Choudhary, M. I., Rahman, A.-u., Sombati, S., Suria, A., et al. (2003). Anticonvulsant Effect of FS-1 Subfraction Isolated from Roots of *Delphinium Denudatum* on Hippocampal Pyramidal Neurons. *Phytother. Res.* 17, 38–43. doi:10.1002/PTR.1072
- Salmaso, V., and Moro, S. (2018). Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Front. Pharmacol.* 9, 923. doi:10.3389/FPHAR.2018.00923/BIBTEX
- Shamsi, A., Mohammad, T., Anwar, S., Amani, S., Khan, M. S., Husain, F. M., et al. (2021). Potential Drug Targets of SARS-CoV-2: From Genomics to Therapeutics. *Int. J. Biol. Macromolecules* 177, 1–9. doi:10.1016/j.ijbiomac.2021.02.071
- Sharma, M., Prasher, P., Mehta, M., Zacconi, F. C., Singh, Y., Kapoor, D. N., et al. (2020). Probing 3CL Protease: Rationally Designed Chemical Moieties for COVID-19. *Drug Dev. Res.* 81, 911–918. doi:10.1002/DDR.21724
- Shi, Y., Yu, X., Zhao, H., Wang, H., Zhao, R., and Sheng, J. (2020). Host Susceptibility to Severe COVID-19 and Establishment of a Host Risk Score: Findings of 487 Cases outside Wuhan. *Crit. Care* 24, 108. doi:10.1186/s13054-020-2833-7
- Singh, K. P., Kumar, A., and Kumar, U. (2018). *Medicinal Plants of Uttarakhand*. New Delhi: Astral International.
- Singh, N., and Chopra, K. L. (2011). Diterpene Alkaloids. Isolation and Study of Two New Alkaloids. *J. Pharm. Pharmacol.* 14, 288–293. doi:10.1111/J.2042-7158.1962.TB11096.X
- Sousa Da Silva, A. W., and Vranken, W. F. (2012). ACPYPE - AnteChamber PYthon Parser interface. *BMC Res. Notes* 5, 367–368. doi:10.1186/1756-0500-5-367/FIGURES/3
- Su, S., Wong, G., Shi, W., Liu, J., Lai, A. C. K., Zhou, J., et al. (2016). Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. *Trends Microbiol.* 24, 490–502. doi:10.1016/j.tim.2016.03.003
- Tang, W., Hu, T., Hu, B., Jin, C., Wang, G., Xie, C., et al. (2020). Prevalence and Correlates of PTSD and Depressive Symptoms One Month after the Outbreak of the COVID-19 Epidemic in a Sample of home-quarantined Chinese university Students. *J. Affective Disord.* 274, 1–7. doi:10.1016/j.jad.2020.05.009
- Teli, D. M., Shah, M. B., and Chhabria, M. T. (2021). *In Silico* Screening of Natural Compounds as Potential Inhibitors of SARS-CoV-2 Main Protease and Spike RBD: Targets for COVID-19. *Front. Mol. Biosciences* 7, 429. doi:10.3389/FMOLB.2020.599079/BIBTEX
- Tripathi, M. K., Singh, P., Sharma, S., Singh, T. P., Ethayathulla, A. S., and Kaur, P. (2020). Identification of Bioactive Molecule from *Withania Somnifera* (Ashwagandha) as SARS-CoV-2 Main Protease Inhibitor. *J. Biomol. Struct. Dyn.*, 1–14. doi:10.1080/07391102.2020.1790425
- Ullrich, S., and Nitsche, C. (2020). The SARS-CoV-2 Main Protease as Drug Target. *Bioorg. Med. Chem. Lett.* 30, 127377. doi:10.1016/J.BMCL.2020.127377
- Varadharajan, V., Arumugam, G. S., and Shanmugam, S. (2021). Isatin-based Virtual High Throughput Screening, Molecular Docking, DFT, QM/MM, MD and MM-PBSA Study of Novel Inhibitors of SARS-CoV-2 Main Protease. *J. Biomol. Struct. Dyn.*, 1–16. doi:10.1080/07391102.2021.1904003
- 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, 2615–2623. doi:10.1021/jm020017n
- Verma, S., Twilley, D., Esmear, T., Oosthuizen, C. B., Reid, A. M., Nel, M., et al. (2020). Anti-SARS-CoV Natural Products with the Potential to Inhibit SARS-CoV-2 (COVID-19). *Front. Pharmacol.* 11, 561334. doi:10.3389/FPHAR.2020.561334/BIBTEX
- Wadood, A., Ahmed, N., Shah, L., Ahmad, A., Hassan, H., and Shams, S. (2013). *In-silico* Drug Design: An Approach Which Revolutionarised the Drug Discovery Process. *OA Drug Des. Deliv.* 1, 3. doi:10.13172/2054-4057-1-1-1119
- Wang, L., Wang, Y., Ye, D., and Liu, Q. (2020a). Review of the 2019 Novel Coronavirus (SARS-CoV-2) Based on Current Evidence. *Int. J. Antimicrob. Agents* 55, 105948. doi:10.1016/j.ijantimicag.2020.105948
- Wang, W., Xu, Y., Gao, R., Lu, R., Han, K., Wu, G., et al. (2020b). Detection of SARS-CoV-2 in Different Types of Clinical Specimens. *JAMA* 323, 1843–1844. doi:10.1001/jama.2020.3786
- Wang, Y., Lv, Z., and Chu, Y. (2015). HIV Protease Inhibitors: a Review of Molecular Selectivity and Toxicity. *Hiv* 7, 95. doi:10.2147/HIV.S79956
- Wilder-Smith, A., and Osman, S. (2020). Public Health Emergencies of International Concern: a Historic Overview. *J. Trav. Med.* 27, 19. doi:10.1093/JTM/TAAA227
- Wu, A., Peng, Y., Huang, B., Ding, X., Wang, X., Niu, P., et al. (2020). Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host & Microbe* 27, 325–328. doi:10.1016/j.chom.2020.02.001
- Yadav, R., Imran, M., Dhamija, P., Chaurasia, D. K., and Handu, S. (2020). Virtual Screening, ADMET Prediction and Dynamics Simulation of Potential Compounds Targeting the Main Protease of SARS-CoV-2. *J. Biomol. Struct. Dyn.* 39, 6617–6632. doi:10.1080/07391102.2020.1796812
- Yu, S., and Yu, M. (2020). Severe Acute Respiratory Syndrome Coronavirus 2-Induced Neurological Complications. *Front. Cel Dev. Biol.* 8, 605972. doi:10.3389/fcell.2020.605972
- Yu, W., and MacKerell, A. D. (2016). “Computer-Aided Drug Design Methods,” in *Methods in Molecular Biology* (New York: Springer), 85–106. doi:10.1007/978-1-4939-6634-9_5
- Yuan, M., Wu, N. C., Zhu, X., Lee, C.-C. D., So, R. T. Y., Lv, H., et al. (2020). A Highly Conserved Cryptic Epitope in the Receptor Binding Domains of SARS-CoV-2 and SARS-CoV. *Science* 368, 630–633. doi:10.1126/science.abb7269
- Zafar, S., Ahmad, M. A., and Siddiqui, T. A. (2001). Protective Role of *Delphinium Denudatum* (Jadwar) against Morphine Induced Tolerance and Dependence in Mice. *J. ethnopharmacology* 78, 95–98. doi:10.1016/S0378-8741(01)00317-8
- Zhang, S.-f., Tuo, J.-l., Huang, X.-b., Zhu, X., Zhang, D.-m., Zhou, K., et al. (2018). Epidemiology Characteristics of Human Coronaviruses in Patients with Respiratory Infection Symptoms and Phylogenetic Analysis of HCoV-OC43 during 2010-2015 in Guangzhou. *PLOS ONE* 13, e0191789. doi:10.1371/JOURNAL.PONE.0191789

Conflict of Interest: DK is a Canada Research Chair in Translational Vaccinology and Inflammation.

The remaining authors declare 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.

Copyright © 2022 Kumar, Sharma, Richardson and Kelvin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.



Evaluation of Clove Phytochemicals as Potential Antiviral Drug Candidates Targeting SARS-CoV-2 Main Protease: Computational Docking, Molecular Dynamics Simulation, and Pharmacokinetic Profiling

OPEN ACCESS

Edited by:

Arvind Ramanathan,
Argonne National Laboratory (DOE),
United States

Reviewed by:

Debsindhu Bhowmik,
Oak Ridge National Laboratory (DOE),
United States
Kemal Yelekci,
Kadir Has University, Turkey

*Correspondence:

Balamuralikrishnan
Balasubramanian
bala.m.k@sejong.ac.kr
Wen-Chao Liu
liuwc@gdau.edu.cn

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Biological Modeling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 12 April 2022

Accepted: 19 May 2022

Published: 28 June 2022

Citation:

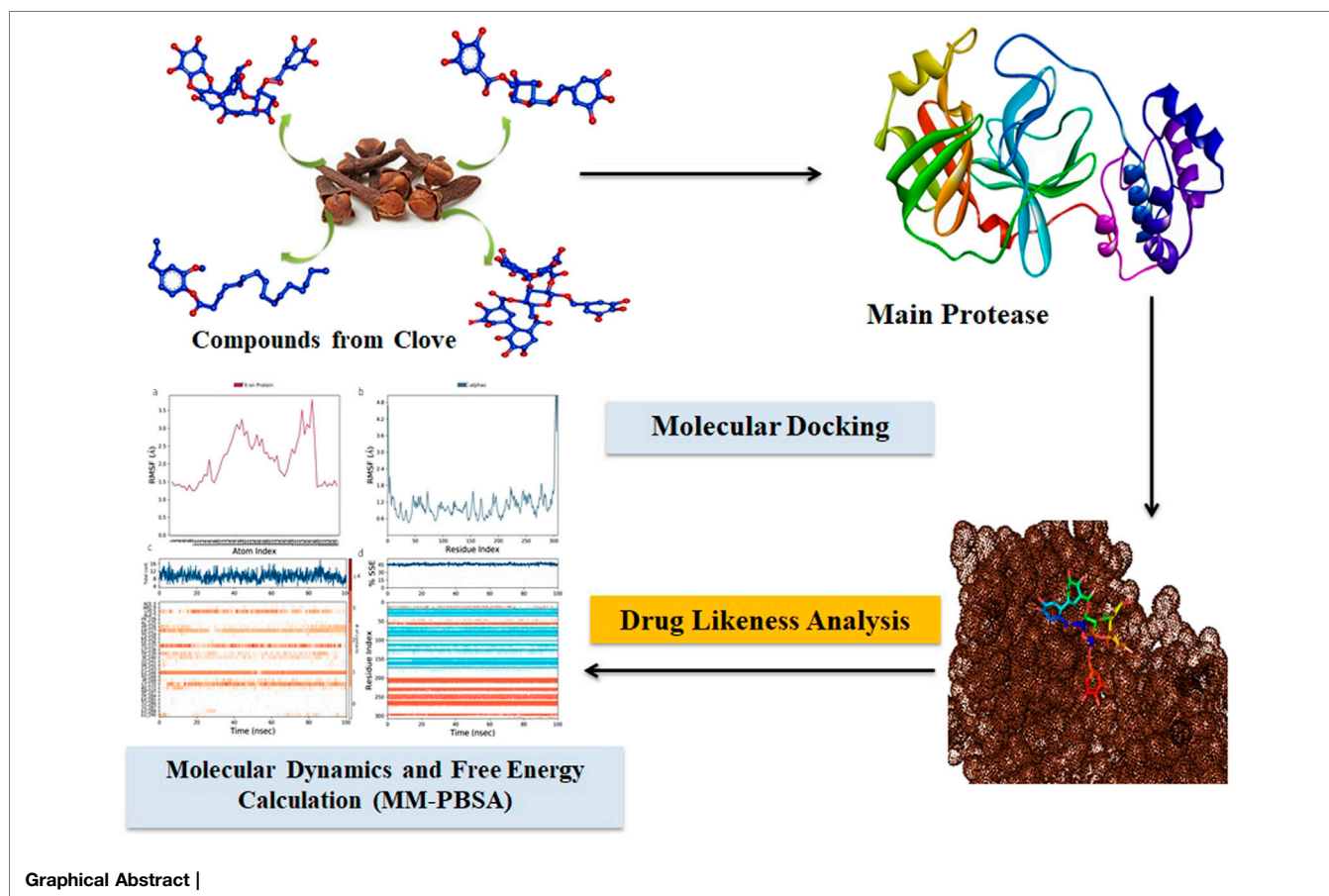
Chandra Manivannan A, Malaisamy A,
Eswaran M, Meyyazhagan A,
Arumugam VA, Rengasamy KRR,
Balasubramanian B and
Liu W-C (2022) Evaluation of Clove
Phytochemicals as Potential Antiviral
Drug Candidates Targeting SARS-
CoV-2 Main Protease: Computational
Docking, Molecular Dynamics
Simulation, and
Pharmacokinetic Profiling.
Front. Mol. Biosci. 9:918101.
doi: 10.3389/fmolb.2022.918101

Arun Chandra Manivannan¹, Arunkumar Malaisamy^{2†}, Murugesw Eswaran^{3†},
Arun Meyyazhagan⁴, Vijaya Anand Arumugam⁵, Kannan R. R. Rengasamy⁶,
Balamuralikrishnan Balasubramanian^{7*} and Wen-Chao Liu^{8*}

¹Department of Microbiology, Alagappa University, Karaikudi, India, ²Integrative Biology Division, International Centre for Genetic Engineering and Biotechnology, New Delhi, India, ³Department of Bioinformatics, Bharathiar University, Coimbatore, India, ⁴Department of Life Sciences, CHRIST (Deemed to be University), Bengaluru, India, ⁵Department of Human Genetics and Molecular Biology, Bharathiar University, Coimbatore, India, ⁶Department of Pharmacology, Centre for Transdisciplinary Research, Saveetha Dental College, Saveetha Institute of Medical and Technical Sciences (SIMATS), Chennai, India, ⁷Department of Food Science and Biotechnology, College of Life Science, Sejong University, Seoul, South Korea, ⁸Department of Animal Science, College of Coastal Agricultural Sciences, Guangdong Ocean University, Zhanjiang, China

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus can cause a sudden respiratory disease spreading with a high mortality rate arising with unknown mechanisms. Still, there is no proper treatment available to overcome the disease, which urges the research community and pharmaceutical industries to screen a novel therapeutic intervention to combat the current pandemic. This current study exploits the natural phytochemicals obtained from clove, a traditional natural therapeutic that comprises important bioactive compounds used for targeting the main protease of SARS-CoV-2. As a result, inhibition of viral replication effectively procures by targeting the main protease, which is responsible for the viral replication inside the host. Pharmacokinetic studies were evaluated for the property of drug likeness. A total of 53 bioactives were subjected to the study, and four among them, namely, eugenie, syzyginin B, eugenol, and casuarictin, showed potential binding properties against the target SARS-CoV-2 main protease. The resultant best bioactive was compared with the commercially available standard drugs. Furthermore, validation of respective compounds with a comprehensive molecular dynamics simulation was performed using Schrödinger software. To further validate the bioactive phytochemicals and delimit the screening process of potential drugs against coronavirus disease 2019, *in vitro* and *in vivo* clinical studies are needed to prove their efficacy.

Keywords: clove, casuarictin, phytochemicals, SARS-CoV-2, main protease



INTRODUCTION

A sudden outbreak of respiratory illness with unknown etiology arose in Wuhan, China, and was later diagnosed as a novel coronavirus (nCoV) with a novel variant called severe acute respiratory syndrome (SARS)-CoV-2. Coronaviruses (CoVs) are a family of enveloped RNA viruses comprising seven human coronaviruses (HCoVs) causing human and animal infection (Biswaranjan, 2022). These are zoonotic obligate intracellular organisms and primarily infect respiratory and associated regions, and this novel virus spreads through air transmission when an infected person coughs or sneezes (Mittal et al., 2020; Majdi et al., 2022). HCoV-229E, HCoV-NL63, HCoV-OC43, and HCoV-HKU1 are usually seasonal, causing mild respiratory diseases that are best known for causing common cold, while other variants like CoV229E and OC43 can provoke pneumonia. Since 3 decades of the virus domination from the first spread of severe strains of middle east respiratory syndrome (MERS)-CoV, SARS-CoV-1, and the current SARS-CoV-2, the lesser known virus is stuck on the headlines for its high pathogenicity and high proliferation rate (Liu et al., 2020; Majdi et al., 2022). The viral infection prevails in one's body with a highly specific recognition between the virus spike proteins through HCoV-specific receptors. Meanwhile, the

main protease (M^{Pro}) is the vital enzyme for processing viral polyproteins produced after being translated from RNA. This M^{Pro} is structurally present as a homodimer, which is made up of two promoters of three domains, namely I, II, and III, whose amino acid ranges are 8–101, 102–184, and 201–303, respectively, and a long loop (185–200) that connects domains II and III (Kumar et al., 2021; Mengist et al., 2021). Every possible sector has unleashed its potency over research and development to targeting the SARS-CoV-2 virus to wind this raging pandemic (Marcelino et al., 2022). Traditional medicinal practices based on herbs and their extracts are gaining momentum. Their formulations are widely given with supplementary allopathic treatment for the earlier recovery (Ang et al., 2020; Mhya et al., 2021; Vanshylla et al., 2021; Ashande et al., 2022; Nair et al., 2022). *Syzygium aromaticum*, a native species of the Maluku Islands in Indonesia, traditionally found its importance as a flavoring additive for food. However, clove finds its use in ayurvedic and Chinese medicine (Cortés-Rojas et al., 2014). These aromatic flower buds are from a tree species of the Myrtaceae family that inhabits tropical climates. The current status of coronavirus disease 2019 (COVID-19) infection on 5 March 2022 crossed 440,807,756 cases with 5,978,096 deaths worldwide, and India's status was 42,951,556 cases with 514,589 deaths.

This research evaluates the antiviral efficacy of clove-based phytochemicals by computationally using the autodock protocol by targeting the M^{Pro} of the SARS-CoV-2 virus (KuchiBhotla et al., 2021; Marcelino et al., 2022). Moreover, the computational world has witnessed a recent trend of large natural compound data retrieval to target various structures of the SARS-CoV-2 virus to define a proper therapeutic. The extensive pathophysiological mechanisms underlying viral infections as well as the related adverse effects of the currently available conventional medicines necessitate the development of a novel drug discovery process. With the conventional methods, finding a new drug is very challenging, and also, it is the costliest method ever. *In silico* methods were carried out to facilitate the virtual screening of the best drug candidate to overcome the challenging drawbacks. Pharmacological studies were conducted to analyze the bioavailability and dynamics inside the host by profiling the ADMET properties of phytochemicals for the suitable drug candidate through extending absorption, distribution, metabolism, excretion, and toxicity (ADMET) (Kar et al., 2021). The molecular dynamics simulation studies were conducted to dive deep into their extent and efficiency by sustaining the administered system over a defined time of 100 ns of total simulation. Investigation of clove-extracted compounds against a target of the M^{Pro} of the SARS-CoV-2 virus leads to effective therapeutic drugs. Further validation of all our compounds was compared with the commercial antiviral drugs to combat COVID-19.

MATERIALS AND METHODS

Protein Preparation

The M^{Pro} acts as a target of treatment for various viral infectious agents, including SARS CoV-2, MERS-CoV, noroviruses, enteroviruses, and rhinoviruses. The M^{Pro} is a disparate protein homing in the infected individual, and as it catalyzes most maturation cleavage events, this proteolytic compound is an explicit target for effective lead screening (Luan et al., 2020). The X-ray crystal structure of M^{Pro} (PDB ID-6LU7) was retrieved from the RCSB Protein Data Bank Database. The structure was subjected to protein preparation using autodock tools, a graphical user interface program, which was exploited for the preparation, running, and analyzing the docking simulations. Water molecules, ligand groups (inhibitor), and other nonspecific molecules were removed, polar hydrogen was added with the merging nonpolar hydrogens, and partial charges were assigned (Arunkumar et al., 2021). The grids are placed in the region that possesses the nature of an active site since grids direct the ligand toward the binding site (Forli et al., 2016).

Ligand Preparation

Data on active phytochemicals present in clove were acquired from the curated databases of Indian Medicinal Plants, Phytochemistry, and Therapeutics (IMPPAT) (Mohanraj et al., 2018). These subsequent structures were retrieved from the PubChem repository, and the related structures were retried from the Zinc Database in the output format

of the structure data file. These were converted to the PDB format using Open Babel software; energy minimization was carried out using pyrxtool applying the molecular mechanics force field and optimized for further exploitation of the ligand. The complete dataset of phytochemical names and their IDs used in this study is provided in **Supplementary Table S1**.

Molecular Docking

The molecular docking was performed in autodock tools with an extensive suite of python molecular viewers. First, the site-specific docking was carried out with the aid of autodock 4.2; during docking, the protein was placed as a rigid molecule and the ligand was flexible (Trott and Olson., 2010). The studies were carried out using the Lamarckian genetic algorithm with the genetic algorithm parameters comprising 2.5×10^6 energy appraisals and a maximum number of 2.7×10^4 generations with a mutation rate of 0.02 with a crossover rate of 0.8. Pseudo Solis and Wets parameters for local search were performed and introduced 300 iterations. Finally, 50 independent runs for each compound were placed, with the grid dimension of $76 \times 76 \times 76$ and with a spacing of 0.375 \AA (Seeliger and de Groot., 2010).

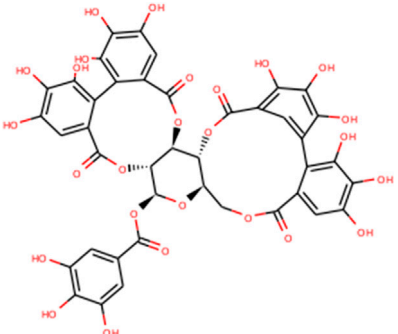
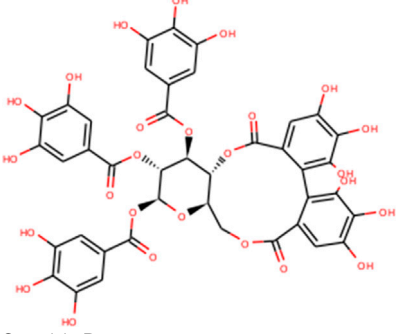
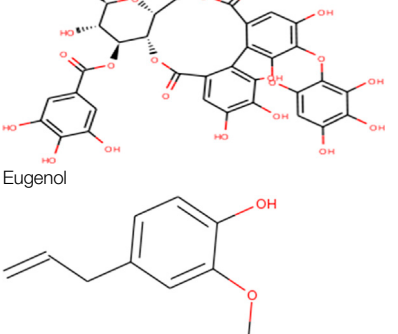
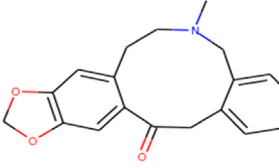

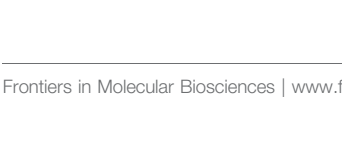
Pharmacokinetics Evaluation

Target prediction studies compute the probable macromolecular target site of the screened small molecules; this methodology aids in tracing the bioactivity, side effects, and off-targets. In addition, the ADMET analysis divulges the pharmacokinetics that a ligand must boat to establish its function in the administered body (Arunkumar et al., 2022). The top-ranked compound was evaluated for the ADMET analysis using the Qikrop module on Schrödinger's Maestro platform (Schrödinger Release 2021-2: QikProp, Schrödinger, LLC, New York, NY, 2021).

Molecular Dynamics Simulation

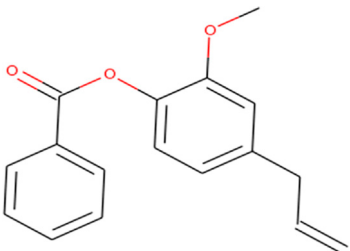
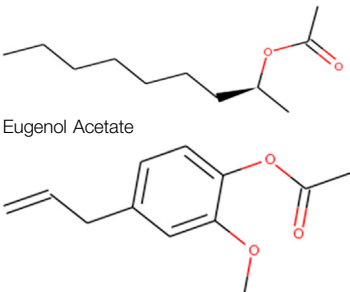
Because molecules are dynamic in nature, studying their motions at the molecular and atomistic levels is critical to comprehending the crucial physicochemical processes. In all other computational applications, molecular dynamics simulation stands alone as the essential computational technique for capturing the dynamic events of scientific interest. Based on the molecular interaction and binding score of the small molecule against the target molecule, the top-ranked complex molecules were selected for the molecular dynamics simulation studies. First, the complex molecule was preprocessed using the protein preparation wizard module; then the structure was refined by optimizing the hydrogen bond and applying the force field OPLS3e for energy minimization. OPLS3e improves the accuracy of small-molecule conformational propensities, solvation, and protein–ligand binding performance benchmarks (Roos et al., 2019). Furthermore, the complex molecule was solvated using a system builder module to a hydration model (TIP3P) in the 3D orthorhombic box with a buffer distance of 10 \AA . Finally, the whole system is designated for the simulation time of 100 ns with 1,000 frame trajectory points under a default NPT ensemble of

TABLE 1 | Interaction between the amino acid residue of COVID -19 protease and ligands at receptor sites.

Ligands	Interacting Species and Bond Distance		
	Hydrogen Binding Interaction	Hydrophobic Interaction	Electrostatic Interaction
Casuarictin, (CHEMBL1076705) 	Thr199 (3.02), Asp197 (3.79), Arg131 (2.76), Lys137 (2.46), and Leu287 (3.53)	Leu287 (5.30), Leu272 (5.12), and Tyr239 (5.20)	—
Eugenilin 	Ala285 (3.38), Leu287 (3.08), Lys236 (2.67), Asn238(2.57), Lys137 (2.85), Thr199 (3.07), and Arg131 (2.98)	Tyr237 (5.04)	Lys236 (3.98)
Syzyginin B 	Lys137 (2.70), Arg131 (2.10), Asp289(281), Glu228 (3.38), Leu287 (3.33), and Thr199 (2.53)	Leu287 (4.89)	—
Eugenol 	Phe294(3.79), Gln110 (2.29), and Ala129 (1.96)	-	—
Protopine 	Leu271 (3.46) and Leu287 (2.51)	Leu287 (4.66)	—
Eugenyl Benzoate 	Cys145 (3.24), His163 (3.55), Ser144 (3.13), and Phe140(3.62)	Met165 (5.09), Cys145 (4.53), and His163 (4.68)	—

(Continued on following page)

TABLE 1 | (Continued) Interaction between the amino acid residue of COVID -19 main protease and ligands at receptor sites.

Ligands	Interacting Species and Bond Distance		
	Hydrogen Binding Interaction	Hydrophobic Interaction	Electrostatic Interaction
 1-Methyloctyl Acetate	Cys145 (2.34), His163 (2.61, Phe140(3.35), and His163 (2.14)	—	Glu166 (4.78)
 Eugenol Acetate	Gln110 (2.18), Asn151(2.25), and Ser158(3.55)	—	—

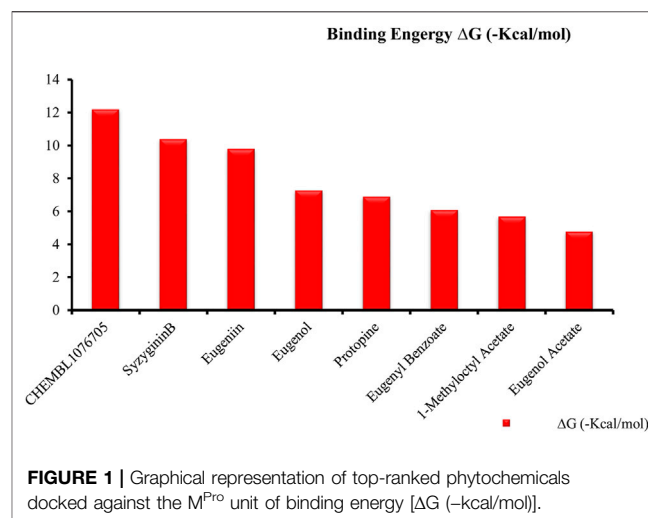
constant pressure, temperature, and atom number. They were performed using the Desmond module on Schrödinger's Maestro platform (Schrödinger Release 2021-2: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2021. Maestro-Desmond Interoperability Tools, Schrödinger, New York, NY, 2021).

RESULTS AND DISCUSSION

Despite fast-tracking the research of COVID cure, no potential lead molecules that can effectively break the viral proliferation chain within an individual are identified (Anju et al., 2020). Moreover, many studies are performed on phytochemicals from medicinal plants for their efficacy against the current COVID-19 disease. SARS-CoV-2 viruses are also reported to get disseminated into various body organs and contaminate the environment in more than one route (Xu et al., 2016; Deng-hai Zhang et al., 2020; Garg et al., 2020; Biswaranjan, 2022). The discovery of the SARS-CoV-2 M^{Pro} has opened the door for an effective approach to drug discovery that can be enabled *via* a virtual combinational mode employing computational tools (Guo et al., 2021; Marcelino et al., 2022).

Molecular Docking and Interactions

About 53 phytochemicals extracted from clove were subjected to molecular docking experiments against the main protease; 27 compounds were native to clove, and 23 are chemicals structurally similar to a few native compounds retrieved from



the ZINC database. Around 60–90% of clove phytochemicals are eugenol, eugenyl acetate, caryophyllene, and acetugenol (Xu et al., 2016). The US Food and Drug Administration (FDA) categorized the clove essential oil as generally recognized as safe, and the World Health Organization has drafted the daily intake quantity for cloves as 2.5 mg/kg of an individual (Sink et al., 2007; Kulkarni et al., 2020). The M^{Pro} is responsible for proper viral replication in SARS CoV-2. Hence, any potential leads can effectively inhibit viral replication inside the host system (Mothay and Ramesh, 2020; Narkhede et al., 2020). Thus, any compound manifesting the disarming of the M^{Pro} can be taken

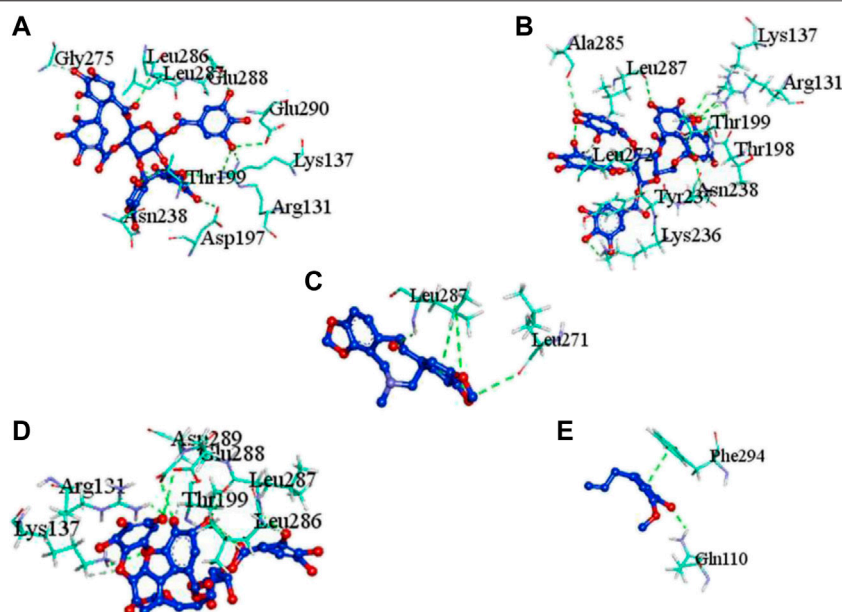


FIGURE 2 | Molecular docking interaction of target proteins with clove's phytochemicals of (A) casuarictin, (B) eugenini, (C) protopine, (D) syzyginin B, and (E) eugenol.

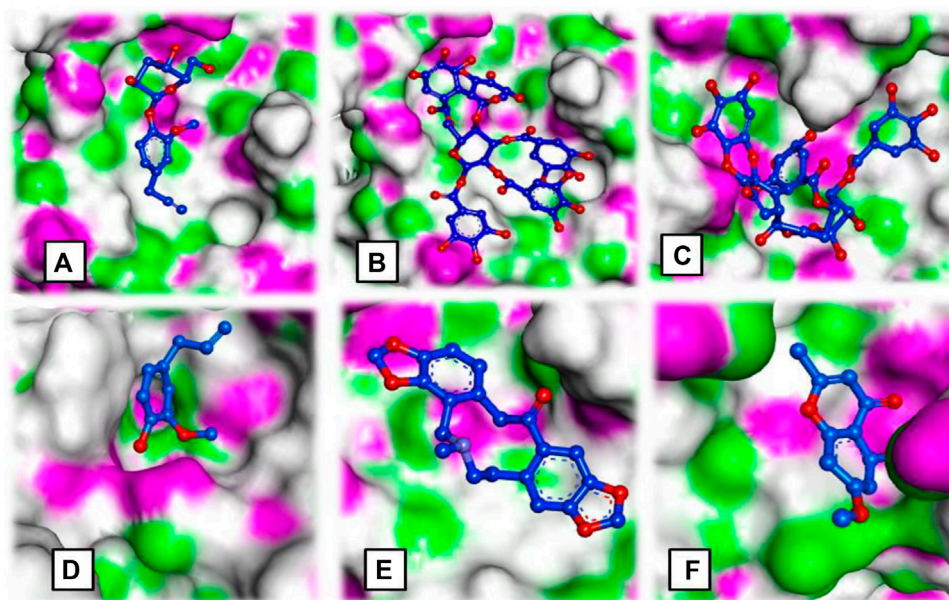


FIGURE 3 | Ligand occupancy in an active site of a target protein complex molecule of (A) syringin, (B) eugenini, (C) syzyginin, (D) eugenol acetate, (E) protopine, and (F) 1-methyloctyl acetate.

for further clinical studies. Still, various leading laboratories in dry and wet labs worldwide are thriving hard to ace the race to screen for highly efficient drugs for the SARS-CoV-2 virus that could effectively treat any variant of the same.

Recently, large quantities of natural compounds are being exploited to act against the deadly virus. A total of 53 compounds were subjected to robust docking using

autodock Vina tools against the main viral protease; four compounds, namely, casuarictin, eugenini, syzyginin B, and eugenol, are identified to bind with the M^{Pro} with the numerically lowest binding energies (most negative) such as -12.2 kcal/mol, -9.8 kcal/mol, -10.4 kcal/mol, and -7.3 kcal/mol, respectively. Since the binding energy score was negative for almost all the compounds, those exhibiting values more

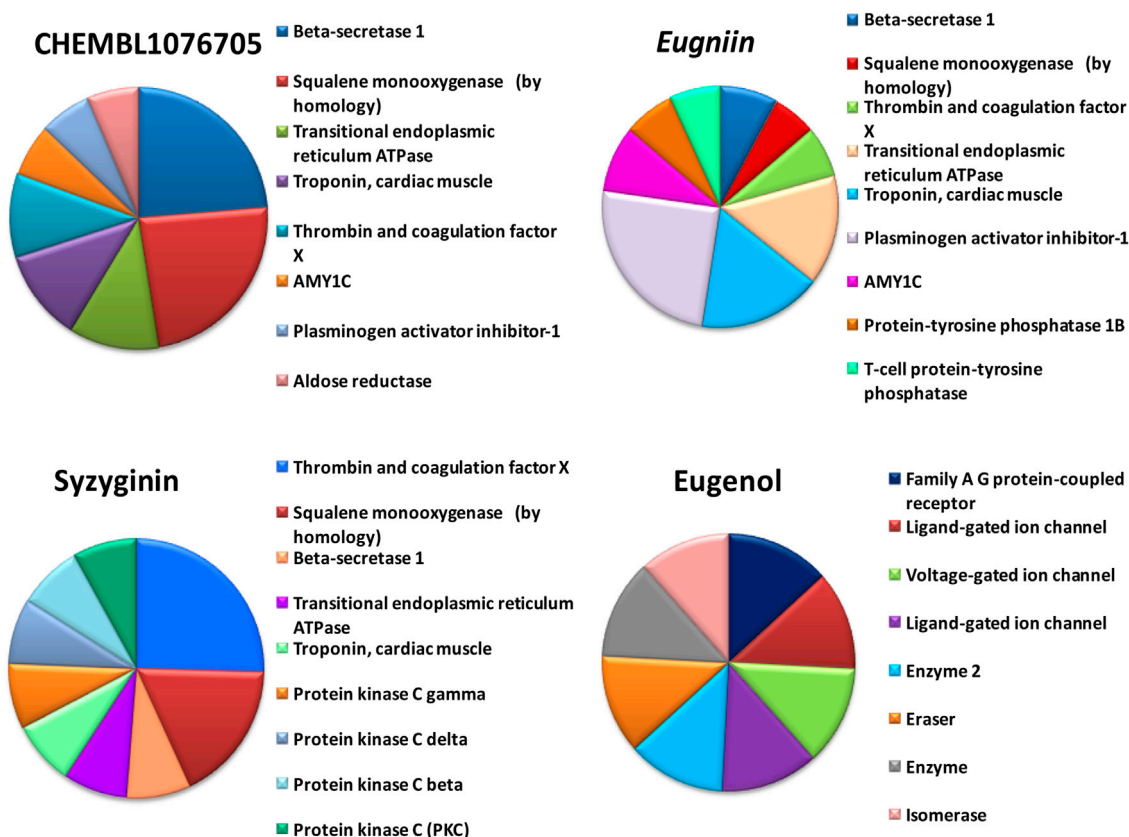


FIGURE 4 | Off-target prediction of selected compounds including beta secretase, squalene mono-oxygenase, thrombin, and coagulation factor.

than -7 were censored, and the remaining were subjected to further *in silico* modeling (Figure 1).

The complex molecules (protein and ligand) were interrogated for their characters of post docking analysis using the Biovia Discovery studio tool. Hydrogen interaction was the predominant force for bond formation and spatial arrangement of the ligands within interacting pockets. In contrast, hydrophobic and electrostatic interactions are facilitated by a hydrogen bond. Casuarictin demonstrates the highest binding energy with an alliance of five hydrogen bonds with amino acids Thr199 (3.02 Å), Asp197 (3.79 Å), Arg131 (2.76 Å), Lys137 (2.46 Å), and Leu287 (3.53 Å) and three hydrophobic interactions with amino acids Leu287 (5.30 Å), Leu272 (5.12 Å), and Tyr239 (5.20 Å).

Next to this, syzyginin B establishes a hydrogen bond with six amino acids, Lys137 (2.70 Å), Arg131 (2.10 Å), Asp289 (2.81 Å), Glu228 (3.38 Å), Leu287 (3.33 Å), and Thr199 (2.53 Å), and one amino acid forging the hydrophobic interaction at Leu287 (4.89 Å). Finally, eugenin with seven hydrogen bonds at locations Ala285 (3.38 Å), Leu287 (3.08 Å), Lys236 (2.67 Å), Asn238 (2.57 Å), Lys137 (2.85 Å), Thr199 (3.07 Å), and Arg131 (2.98 Å) fashions a binding energy within the range of -9.8 kcal/mol. Table 1 describes the types of interaction various ligands are experiencing. These bonds are responsible for arresting the ligand efficiently within the active site of the protein. Figures 2, 3, 4

render the delineation of various bonds formed in screened compounds. The remaining 53 complex molecule interactions are listed in Supplementary Table S1, and for the remaining phytochemicals, interactions of their 3D structure with 2D interactions are represented in Supplementary Figures S1–19.

Our top-ranked compound of eugenin (-9.8 kcal/mol), syzyginin B (-10.4 kcal/mol), eugenol (-7.3 kcal/mol), and casuarictin (-12.2 kcal/mol) has shown higher binding affinity than the commercially available drugs such as remdesivir (-6 kcal/mol), liponavir (-4 kcal/mol), tamiflu (-5.72 kcal/mol), plaquenil (-5.59 kcal/mol), and mycophenolic (-6.02 kcal/mol) (Arunkumar et al., 2021; Khater and Nassar, 2021). Eugenol is extracted from cloves, is abundantly present in the clove oil, is used for antispasmodic treatment, and acts as a carminative to treat gastrointestinal disorders. They are explored in bioactivities of antioxidant, anti-inflammatory, antiviral, insect-repellent, antimicrobial, and antiparasitic properties with various other related infections (Xu et al., 2016). The additional properties of cloves are strong, pungent, spicy odor and pungent combined aromatic taste, and cloves play a role in industrial application in perfumes, soaps, histological cleaning agents, and anesthetic fishes (Taylor and Roberts 1999). Eugenol exploited against spike glycoprotein for the treatment of SARS-CoV-2 has been reported computationally, with the attempted clinical phase in the official Siddha formulation of Kabasura

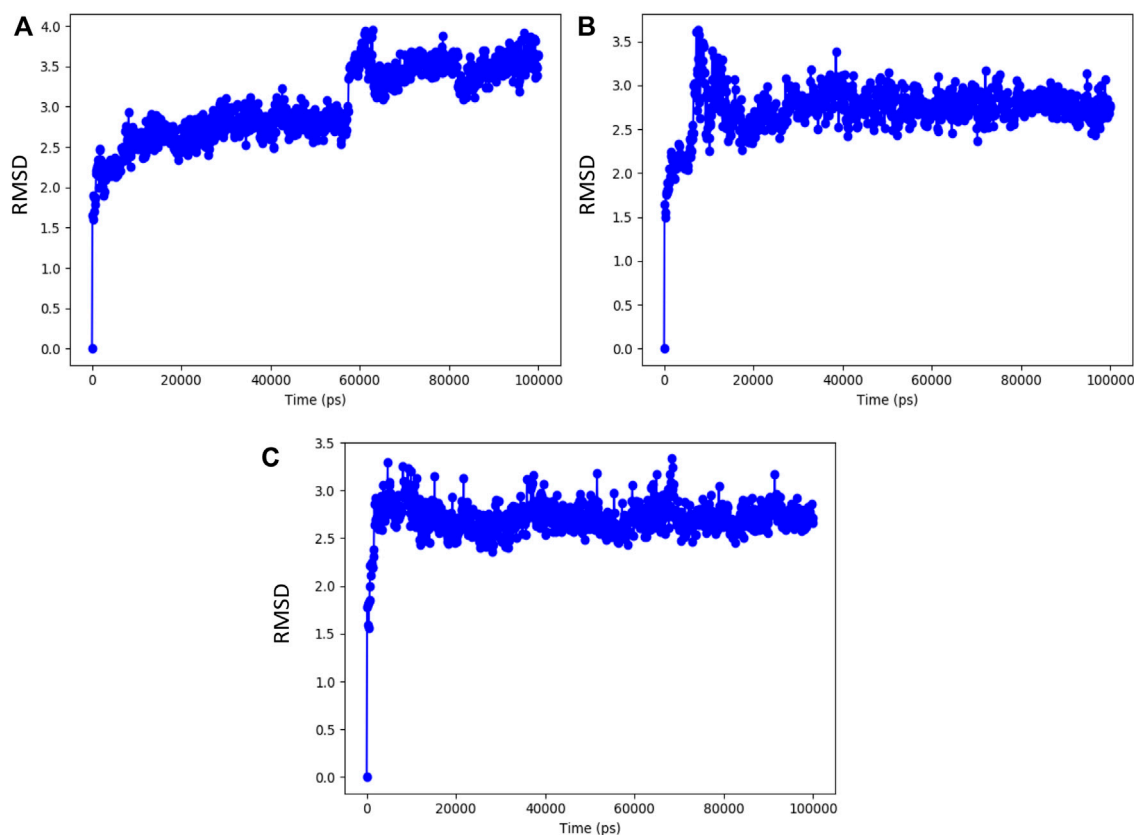


FIGURE 5 | RMSD value of the complex molecule of the main protease with (A) syzyginin B, (B) eugenin, and (C) eugenol.

Kudineer (Kiran et al., 2022). Moreover, the crisis on physical health due to continuous steroid supplementation within individuals infected with the SARS-CoV-2 virus stresses the need for alternative medicine.

Prediction of Probable Off-Target Activity

By tracing the probable drug reaction within a host and its assumed interaction, it might undergo intruding the host metabolism, which is provided in hit compound target prediction; this paves the way for the preparedness of any drug that can elicit any adverse reaction for patients. Additionally, it provides the researcher with a putative thought to remodel or restrict the further analysis of screened drug-like molecules (Drwal et al., 2014; Daina et al., 2019). For example, eugenol was predicted to be a class 4 toxic substance. The predicted LD50 is 1.930 mg/kg; it does not elicit carcinogenicity, hepatotoxicity, immunotoxicity, mutagenicity, and cytotoxicity and does not interfere with the signaling and stress response pathways. On the other hand, syzyginin B, casuarictin, and eugenin are classified as class 5 toxic substances and possess an LD50 value of 2.260 mg/kg. This ternion exhibits mild reactivity and less than the recommended level of reactivity toward phosphoprotein (tumor repressor) p53. Still, the concerning part is that it may

result in immunotoxicity under unregulated administration. At the same time, the last couple may interfere with the mitochondrial membrane protein and aryl hydrocarbon receptor and may initiate reactions adding up alarm for carcinogenicity, albeit with negligible probability. These results were obtained by combining two toxicity prediction web tools, Swiss Target prediction tools (Gupta et al., 2013). Casuarictin, syzyginin B, and eugenin are identified to have common off-target interactions whose extent of integration either overtaking the active antagonist property or least significant can be identified only upon wet-lab studies (Mohamed et al., 2021). The common targets of these three compounds include beta secretase, squalene mono-oxygenase, thrombin and coagulation factor, troponin, and cardiac muscles with varying proportions (Figure 4).

Molecular Dynamics Simulation

The resulting top-ranked docking complex molecule was considered for performing the molecular dynamics simulation for further validation. In these simulation studies, the protein interaction with the ligand molecules was studied throughout the total simulation time of 100 ns with 1,000 projection points (Frames). The macromolecules and ligand causing interactions throughout the simulation time are called contacts, classified

TABLE 2 | ADME/T properties of various compounds that exhibited a high binding efficiency.

Title (and Range)	Syzyginin (ZINC230067171)	Syzyginin B	Eugeniin	Casuarictin	Protopine	Eugenol
#amine (0–1)	0	0	0	0	1	0
#amidine (0)	0	0	0	0	0	0
#acid (0–1)	0	0	0	0	0	0
#amide (0–1)	0	0	0	0	0	0
CNS (-2 inactive and +2 active)	-2	-2	-2	-2	2	0
mol MW (130.0–725.0)	756.54	756.54	938.672	936.657	353.374	164.204
Dipole (1.0–12.5)	3.342	7.316	4.159	10.409	2.299	1.859
SASA (300.0–1,000.0)	915.82	924.397	1,186.696	1,133.803	508.221	401.412
FOSA (0.0–750.0)	105.614	97.55	66.353	76.301	310.695	204.373
FISA (7.0–330.0)	612.263	621.364	835.169	810.688	26.505	51.309
PISA (0.0–450.0)	197.944	205.484	285.174	246.814	171.021	145.731
WPSA (0.0–175.0)	0	0	0	0	0	0
Volume (500.0–2000.0)	1786.437	1800.373	2,304.514	2,217.198	973.528	641.616
donorHB (0.0–6.0)	12	12	15	14	0	1
accptHB (2.0–20.0)	19.6	19.6	22.95	21.95	7	1.5
dip2/V (0.0–0.13)	0.006252	0.029726	0.007507	0.048864	0.005431	0.005386
ACxDN.5/SA (0.0–0.05)	0.074137	0.073449	0.074901	0.072437	0	0.003737
Glob (0.75–0.95)	0.777,464	0.774,251	0.711,015	0.725,266	0.934,713	0.89624
QPpolrz (13.0–70.0)	58.24	58.87	75.312	74.02	34.44	18.347
QPlogPC16 (4.0–18.0)	23.298	23.568	31.713	29.985	9.054	5.793
QPlogPo/w (-2.0–6.5)	-2.911	-2.905	-3.469	-3.237	1.727	2.666
QPlogS (-6.5–0.5)	-3.622	-3.716	-4.733	-5.061	-0.963	-2.35
QPPCaco (<25 poor, >500 great)	0.015	0.013	0	0	1,385.006	3,231.023
QPlogBB (-3.0–1.2)	-7.033	-7.191	-11.716	-10.736	0.76	-0.103
QPPMDCK (<25 poor, >500 great)	0.003	0.003	0	0	778.255	1757.464
QPlogKp (-8.0 to -1.0)	-10.762	-10.903	-13.891	-13.959	-3.468	-1.568
IP(eV) (7.9–10.5)	8.542	8.383	9.017	8.897	8.923	8.729
EA (eV) (-0.9–1.7)	0.881	0.658	0.634	0.737	0.406	-0.232
QPlogKhsa (-1.5–1.5)	-1.035	-1.025	-1.246	-0.984	-0.428	-0.113
Human Oral Absorption	1	1	1	1	3	3
Percent Human Oral Absorption (>80% is high, <25% is poor)	0	0	0	0	93.285	100
RuleOfFive (maximum is 4)	3	3	3	3	0	0
RuleOfThree (maximum is 3)	2	2	2	2	0	0

based on hydrogen bonds, hydrophobic interactions, ionic bonds, and water bridges. The molecular dynamics simulation output was investigated with a root-mean-square deviation (RMSD) value around 3 Å distance, representing the stability of the complex molecule. Syzyginin B starts the stability at around 60 ns in the first phase with the deviation of around 1 Å distance and remains stable around 2.5 Å. After that, it deviates from the cavity site of 1 Å distance, which holds the stability in 3.5 Å up to a total simulation time of 100 ns. On the other hand, eugeniin deviates in the initial phase up to 3.5 Å and stabilizes with minimal deviation, which remains stable from 20 to 100 ns around 2.8 Å RMSD. Finally, eugenol started the initial phase around 2.9 Å and remains stable across 100 ns of the total simulation time within 2.8 Å. Moreover, all the three complex molecules show better results from the molecular dynamics simulation studies, depicted by the graph in **Figure 5**.

The protein–ligand contact of clove phytochemical eugeniin showed an interaction of discontinuous contacts in the active site amino acid of LYS5, GLY170, SER139, and PHE140 and continuous contacts with GLU166, LYS137, and GLN27 across the total simulation time. On the other hand, the remaining two complex molecules of syzyginin B and eugenol showed the major discontinuous contacts (**Supplementary Figures S20–22**).

Pharmacokinetics

Computational ADMET screening can reduce the cost of high capital-consuming wet-lab trials that may end up in failure on many occasions (Hage-Melim et al., 2020; Tongqiang Zhang, et al., 2020). In the current study, not all the selected molecules were following the optimal limit of ADMET properties. With the octanal/water partition coefficient, aqueous solubility, and brain/blood partition coefficient, all compounds align within an acceptable range of -2.0–6.5, -6.5–0.5, and -3.0–2.1, respectively (Gleeson et al., 2009). A major deviation was exhibited by casuarictin (-13.959) and eugeniin (-13.959) for skin permeability, where the optimum range is between -8.0 and -1.0. Eugenol exhibited a minimum metabolic reaction limited to 3, followed by protopine with five reactions. At the same time, the next two molecules deviate from an acceptable band of 1–8, where both eugeniin and casuarictin exhibit 15 reaction exceptions.

All molecules are efficient enough to bind with human serum albumin within the accepted range of -1.5 to 1.5. Eugenol and protopine are highly capable of oral absorption and hold a good van der Waals surface area of polar nitrogen and oxygen atoms and carbonyl carbon atoms. Eugeniin and casuarictin initially exhibited high binding energy (mostly negative). However, they face the most number of violations, albeit their ability to infuse toxicity within an individual is low, which suggests that its

efficacy can be taken for the next level of analysis as hit compounds (Liao and Nicklaus, 2009; Zhong et al., 2013) (Table 2).

The ability of candidates to get accommodated within the traditional laws of drug-likeness and physiochemical property limits are major initial screening processes, aiming to screen out unfit candidates (Mohamed et al., 2021). For example, Lipinski's rule of five (RO5) is a major factor for drug-likeness that aids in identifying the potential compounds from a pool of drug-like molecules which must have strong gastrointestinal absorption, high oral bioavailability, and descent membrane permeability, with their $\log p \leq 5$; $MW \leq 500$ Da, $HBDs \leq 5$, and $HBA \leq 10$ (Gleeson et al., 2009). Natural compounds that have already been identified for treatment purposes are reported to violate RO5. On the other hand, due to the significant efficacy of natural compounds, those from marine-based and terrestrial resource-based compounds are accepted, although they have been identified as violating RO5 (Zhong et al., 2013; Mohanraj et al., 2018). The new framework proposed by the FDA over relaxed and diluted norms for FDA approval of drugs supports the idea to migrate the sample for further intensive trials so as to find the therapeutic scope beyond RO5.

Violation of log P is still not a concern for major cancer drugs (DeGoey et al., 2018). Casuarictin violates Veber's rule for polar surface area (TPSA), resulting in a range $\leq 140 \text{ \AA}$ by holding a value of 444.18. The existing drugs that are prescribed for being effective against COVID-19 virus-like lopinavir, ritonavir, and remdesivir also violate Lipinski's RO5, where lopinavir disobeys with $MW > 500$, many rotatable bonds being >10 , ritonavir and remdesivir at features like $MW > 500$, many rotatable bonds > 10 , and $TPSA > 140 \text{ \AA}$ (Kar et al., 2021). Despite few notable fluctuations expressed by these selected compounds, there are updated relaxed norms by FDA over exploiting potential drugs that still disobey drug-likeness properties and the existing record of supplementing drugs of a high molecular weight, disobeying RO5 (DeGoey et al., 2018).

CONCLUSION

Using biological sources to find alternative and successful drug candidates could be a long-term strategy for improving the COVID-19 drug discovery process. Cloves are a rich source of bioactive chemicals such as eugenol, which have been shown to have antiviral and immunostimulatory activities. In the current computational approach, clove phytochemicals of a total of 53 compounds were investigated for the molecular docking experiment against the M^{Pro} of SARS-CoV-2. Interestingly, among them, eugenin, syzygininB, eugenol, and casuarictin have shown possible antagonist properties against the M^{Pro} with

significant binding energies. Furthermore, the top-ranked phytochemicals were validated with the molecular dynamics simulation and revealed three compounds, namely, syzyginin B, eugenin, and eugenol, as strongly interacting compounds that got stabilized with the least deviation from the site of interaction over the observed total simulation time. Moreover, the phytochemicals were assessed for their pharmacokinetic properties, shown to be druggable with no significant violation of any ADMET profiling parameters. As a result, these cloves' phytochemicals may be viable candidates against SARS-CoV-2. Eugenol is one of the formulations of official siddha as Kabasura Kudineer for the treatment of COVID-19. However, extensive research is required to determine its efficacy as an antiviral drug, particularly *in vitro* trials against SARS-CoV-2. Finally, the innovative findings of this study could have a significant impact on the advancement of COVID-19 antiviral drug interventions in the near future (Trott and Olson, 2010; Lin et al., 2020; Khanna et al., 2021).

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 authors.

AUTHOR CONTRIBUTIONS

This research article is written by the collaboration between the authors. Conceptualization, ME, BB, and W-CL; writing the original manuscript, AC, AM; methodology, data curation, and formal analysis, AC, AM, ME, and AM; organizing the working groups, BB, ME; writing review/editing, VA, KR, W-CL; interpretation and review/revision, ME, AM, BB. All Authors revised and approved the final article.

ACKNOWLEDGMENTS

The authors thank the International Centre for Genetic Engineering and Biotechnology (ICGEB), New Delhi, for Molecular dynamic simulation and ADMET studies using Schrodinger software.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2022.918101/full#supplementary-material>

REFERENCES

- Ang, L., Lee, H. W., Kim, A., and Lee, M. S. (2020). Herbal Medicine for the Management of COVID-19 during the Medical Observation Period: a Review of Guidelines. *Integr. Med. Res.* 9 (3), 100465. doi:10.1016/j.imr.2020.100465

- Anju, M., Sreeush, M. G., Valsala, V., Smitha, B. R., Hamza, F., Bharathi, G., et al. (2020). Understanding the Role of Nutrient Limitation on Plankton Biomass over Arabian Sea via 1-D Coupled Biogeochemical Model and bio-Argo Observations. *J. Geophys. Res. Oceans* 125 (6), e2019JC015502. doi:10.1029/2019jc015502
- Arunkumar, M., Gunaseelan, S., Kubendran Aravind, M., Mohankumar, V., Anupam, P., Hari Krishnan, M., et al. (2021). Marine Algal Antagonists

- Targeting 3CL Protease and Spike Glycoprotein of SARS-CoV-2: a Computational Approach for Anti-COVID-19 Drug Discovery. *J. Biomol. Struct. Dyn.*, 1–28. doi:10.1080/07391102.2021.1921032
- Arunkumar, M., Mahalakshmi, M., Ashokkumar, V., Aravind, M. K., Gunaseelan, S., Mohankumar, V., et al. (2022). Evaluation of Seaweed Sulfated Polysaccharides as Natural Antagonists Targeting *Salmonella typhi* OmpF: Molecular Docking and Pharmacokinetic Profiling. *Beni-Suef Univ. J. Basic Appl. Sci.* 11 (1), 8. doi:10.1186/s43088-021-00192-x
- Ashande, C., Masunda, A., Ngbolua, K. T., N., Kilembe, J. T., Matondo, A., Liyongo Clément, I., et al. (2022). Glucose Oxidase as a Model Enzyme for Antidiabetic Activity Evaluation of Medicinal Plants: *In Vitro* and *In Silico* Evidence. *Nat. Resour. Hum. Health* 2 (2), 265–273. doi:10.53365/nrfhh/144779
- Biswaranjan, P. (2022). Human Health Care against COVID-19 via Environmental Management. *Nat. Resour. Hum. Health* 2 (2), 142–149. doi:10.53365/nrfhh/143833
- Cortés-Rojas, D. F., de Souza, C. R. F., and Oliveira, W. P. (2014). Clove (*Syzygium Aromaticum*): a Precious Spice. *Asian Pac. J. Trop. Biomed.* 4 (2), 90–96. doi:10.1016/S2221-1691(14)60215-X
- Daina, A., Michielin, O., and Zoete, V. (2019). SwissTargetPrediction: Updated Data and New Features for Efficient Prediction of Protein Targets of Small Molecules. *Nucleic Acids Res.* 47 (W1), W357–W364. doi:10.1093/nar/gkz382
- DeGoey, D. A., Chen, H.-J., Cox, P. B., and Wendt, M. D. (2018). Beyond the Rule of 5: Lessons Learned from AbbVie's Drugs and Compound Collection. *J. Med. Chem.* 61 (7), 2636–2651. doi:10.1021/acs.jmedchem.7b00717
- Deng-hai Zhang, D.-h., Wu, K.-l., Zhang, X., Deng, S.-q., and Peng, B. (2020). *In Silico* screening of Chinese Herbal Medicines with the Potential to Directly Inhibit 2019 Novel Coronavirus. *J. Integr. Med.* 18 (2), 152–158. doi:10.1016/j.joim.2020.02.005
- Drwal, M. N., Banerjee, P., Dunkel, M., Wettig, M. R., and Preissner, R. (2014). ProTox: a Web Server for the *In Silico* Prediction of Rodent Oral Toxicity. *Nucleic Acids Res.* 42 (W1), W53–W58. doi:10.1093/nar/gku401
- Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., and Olson, A. J. (2016). Computational Protein-Ligand Docking and Virtual Drug Screening with the AutoDock Suite. *Nat. Protoc.* 11 (5), 905–919. doi:10.1038/nprot.2016.051
- Garg, S., Garg, M., Prabhakar, N., Malhotra, P., and Agarwal, R. (2020). Unraveling the Mystery of Covid-19 Cytokine Storm: From Skin to Organ Systems. *Dermatol. Ther.* 33 (6), 13859. doi:10.1111/dth.13859
- Gleeson, P., Bravi, G., Modi, S., and Lowe, D. (2009). ADMET Rules of Thumb II: A Comparison of the Effects of Common Substituents on a Range of ADMET Parameters. *Bioorg. Med. Chem.* 17 (16), 5906–5919. doi:10.1016/j.bmc.2009.07.002
- Guo, S., Xie, H., Lei, Y., Liu, B., Zhang, L., Xu, Y., et al. (2021). Discovery of Novel Inhibitors against Main Protease (Mpro) of SARS-CoV-2 via Virtual Screening and Biochemical Evaluation. *Bioorg. Chem.* 110, 104767. doi:10.1016/j.bioorg.2021.104767
- Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., and Raghava, G. P. S. (2013). *In Silico* Approach for Predicting Toxicity of Peptides and Proteins. *PLoS ONE* 8 (9), e73957. doi:10.1371/journal.pone.0073957
- Hage-Melim, L. I. D. S., Federico, L. B., de Oliveira, N. K. S., Francisco, V. C. C., Correia, L. C., de Lima, H. B., et al. (2020). Virtual Screening, ADME/Tox Predictions and the Drug Repurposing Concept for Future Use of Old Drugs against the COVID-19. *Life Sci.* 256, 117963. doi:10.1016/j.lfs.2020.117963
- Kar, P., Sharma, N. R., Singh, B., Sen, A., and Roy, A. (2021). Natural Compounds from *Clerodendrum* Spp. As Possible Therapeutic Candidates Against SARS-CoV-2: An *In Silico* Investigation. *J. Biomol. Struct. Dyn.* 39 (13), 4774–4785. doi:10.1080/07391102.2020.1780947
- Khanna, K., Kohli, S. K., Kaur, R., Bhardwaj, A., Bhardwaj, V., Ohri, P., et al. (2021). Herbal Immune-Boosters: Substantial Warriors of Pandemic Covid-19 Battle. *Phytomedicine* 85, 153361. doi:10.1016/j.phymed.2020.153361
- Khater, I., and Nassar, A. (2021). *In Silico* Molecular Docking Analysis for Repurposing Approved Antiviral Drugs Against SARS-CoV-2 Main Protease. *Biochem. Biophys. Rep.* 27, 101032.
- Kiran, G., Karthik, L., Shree Devi, M. S., Sathiyarajeswaran, P., Kanakavalli, K., Kumar, K. M., et al. (2022). *In Silico* computational Screening of Kabasura Kudineer - Official Siddha Formulation and JACOM against SARS-CoV-2 Spike Protein. *J. Ayurveda Integr. Med.* 13 (1), 100324. doi:10.1016/j.jaim.2020.05.009
- Kuchi Bhotla, H., Balasubramanian, B., Arumugam, V. A., Pushparaj, K., Easwaran, M., Baskaran, R., et al. (2021). Insinuating Cocktailed Components in Biocompatible-Nanoparticles Could Act as an Impressive Neo-Adjuvant Strategy to Combat COVID-19. *Nat. Resour. Hum. Health* 1 (1), 3–7. doi:10.53365/nrfhh/140607
- Kulkarni, S. A., Nagarajan, S. K., Ramesh, V., Palaniyandi, V., Selvam, S. P., and Madhavan, T. (2020). Computational Evaluation of Major Components from Plant Essential Oils as Potent Inhibitors of SARS-CoV-2 Spike Protein. *J. Mol. Struct.* 1221, 128823. doi:10.1016/j.molstruc.2020.128823
- Kumar, A., Choudhir, G., Shukla, S. K., Sharma, M., Tyagi, P., Bhushan, A., et al. (2021). Identification of Phytochemical Inhibitors against Main Protease of COVID-19 Using Molecular Modeling Approaches. *J. Biomol. Struct. Dyn.* 39 (10), 3760–3770. doi:10.1080/07391102.2020.1772112
- Liao, C., and Nicklaus, M. C. (2009). Comparison of Nine Programs Predicting pKa Values of Pharmaceutical Substances. *J. Chem. Inf. Model.* 49 (12), 2801–2812. doi:10.1021/ci900289x
- Lin, Y., Hu, Z., Alias, H., and Wong, L. P. (2020). Knowledge, Attitudes, Impact, and Anxiety Regarding COVID-19 Infection Among the Public in China. *Front. Public Health* 8, 236. doi:10.3389/fpubh.2020.00236
- Liu, Y.-C., Kuo, R.-L., and Shih, S.-R. (2020). COVID-19: The First Documented Coronavirus Pandemic in History. *Biomed. J.* 43 (4), 328–333. doi:10.1016/j.bj.2020.04.007
- Luan, B., Huynh, T., Cheng, X., Lan, G., and Wang, H.-R. (2020). Targeting Proteases for Treating COVID-19. *J. Proteome Res.* 19 (11), 4316–4326. doi:10.1021/acs.jproteome.0c00430
- Majdi, H., ZarKalai, F., Yeddes, W., and Saidani, M. (2022). Phenolic Compounds as Antiviral Agents: An *In-Silico* Investigation against Essential Proteins of SARS-CoV-2. *Nat. Resour. Hum. Health* 2 (1), 62–78. doi:10.53365/nrfhh/143085
- Marcelino, R. C., de Araújo, L. P., Bueno de Moraes Borba, J., and da Silveira, N. J. F. (2022). Molecular Docking Study Involving Bioactive Natural Compounds against SARS-CoV-2 Proteins. *Nat. Resour. Hum. Health.* doi:10.53365/nrfhh/147375
- Mengist, H. M., Dilnessa, T., and Jin, T. (2021). Structural Basis of Potential Inhibitors Targeting SARS-CoV-2 Main Protease. *Front. Chem.* 9, 622898. doi:10.3389/fchem.2021.622898
- Mhya, D., Nuhu, A., and Mankilik, M. (2021). *In-silico* Discovery of Antidiabetic Drug Potential of Balanites Aegyptiaca Leaf's Phenolic Compounds. *Nat. Resour. Hum. Health* 1 (2), 91–97. doi:10.53365/nrfhh/142375
- Mittal, A., Manjunath, K., Ranjan, R. K., Kaushik, S., Kumar, S., and Verma, V. (2020). COVID-19 Pandemic: Insights into Structure, Function, and hACE2 Receptor Recognition by SARS-CoV-2. *PLoS Pathog.* 16 (8), e1008762. doi:10.1371/journal.ppat.1008762
- Mohamed, H. S., El-Serwy, W. S., and El-Serwy, W. S. (2021). Synthesis, Molecular Docking, *In Silico* ADME Predictions, and Toxicity Studies of N-Substituted-5-(4-Chloroquinolin-2-Yl)-1,3,4-Thiadiazol-2-Amine Derivatives as COVID-19 Inhibitors. *Russ. J. Bioorg. Chem.* 47 (1), 158–165. doi:10.1134/S1068162021010155
- Mohanraj, K., Karthikeyan, B. S., Vivek-Ananth, R. P., Chand, R. P. B., Aparna, S. R., Mangalapandi, P., et al. (2018). IMPPAT: A Curated Database of Indian Medicinal Plants, Phytochemistry and Therapeutics. *Sci. Rep.* 8 (1), 4329. doi:10.1038/s41598-018-22631-z
- Mothay, D., and Ramesh, K. v. (2020). Binding Site Analysis of Potential Protease Inhibitors of COVID-19 Using AutoDock. *VirusDis.* 31 (2), 194–199. doi:10.1007/s13337-020-00585-z
- Nair, P. P., Nair, P. G., and Pratap Shankar, K. M. (2022). An Ayurvedic Personalized Prophylactic Protocol in COVID-19. *J. Ayurveda Integr. Med.* 13 (1), 100351. doi:10.1016/j.jaim.2020.08.004
- Narkhede, R. R., Pise, A. v., Cheke, R. S., and Shinde, S. D. (2020). Recognition of Natural Products as Potential Inhibitors of COVID-19 Main Protease (Mpro): *In-Silico* Evidences. *Nat. Prod. Bioprospect.* 10 (5), 297–306. doi:10.1007/s13659-020-00253-1
- Roos, K., Wu, C., Damm, W., Reboul, M., Stevenson, J. M., Lu, C., et al. (2019). OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *J. Chem. Theory Comput.* 15 (3), 1863–1874. doi:10.1021/acs.jctc.8b01026
- Seeliger, D., and de Groot, B. L. (2010). Ligand Docking and Binding Site Analysis with PyMOL and Autodock/Vina. *J. Comput. Aided Mol. Des.* 24 (5), 417–422. doi:10.1007/s10822-010-9352-6
- Sink, T. D., Strange, R. J., and Sawyers, R. E. (2007). Clove Oil Used at Lower Concentrations Is Less Effective Than MS-222 at Reducing Cortisol Stress Responses in Anesthetized Rainbow Trout. *North Am. J. Fish. Manag.* 27 (1), 156–161. doi:10.1577/m05-103.1

- Taylor, P. W., and Roberts, S. D. (1999). Clove Oil: An Alternative Anaesthetic for Aquaculture. *North Am. J. Aquac.* 61 (2), 150–155. doi:10.1577/1548-8454(1999)061<0150:coaaaf>2.0.co;2
- Tongqiang Zhang, T., Cui, X., Zhao, X., Wang, J., Zheng, J., Zheng, G., et al. (2020). Detectable SARS-CoV-2 Viral RNA in Feces of Three Children during Recovery Period of COVID-19 Pneumonia. *J. Med. Virol.* 92 (7), 909–914. doi:10.1002/jmv.25795
- Trott, O., and Olson, A. J. (2010). AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* 31 (2), 455–461. doi:10.1002/jcc.21334
- Vanshylla, K., Di Cristanziano, V., Kleipass, F., Dewald, F., Schommers, P., Gieselmann, L., et al. (2021). Kinetics and Correlates of the Neutralizing Antibody Response to SARS-CoV-2 Infection in Humans. *Cell host microbe* 29 (6), 917–929. doi:10.1016/j.chom.2021.04.015
- Xu, J.-G., Liu, T., Hu, Q.-P., and Cao, X.-M. (2016). Chemical Composition, Antibacterial Properties and Mechanism of Action of Essential Oil from Clove Buds against *Staphylococcus aureus*. *Molecules* 21 (9), 1194. doi:10.3390/molecules21091194
- Zhong, H., Mashinson, V., Woolman, T., and Zha, M. (2013). Understanding the Molecular Properties and Metabolism of Top Prescribed Drugs. *Curr. Top. Med. Chem.* 13 (11), 1290–1307. doi:10.2174/15680266113139990034
- Conflict of Interest:** The authors declare 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.

Copyright © 2022 Chandra Manivannan, Malaisamy, Eswaran, Meyyazhagan, Arumugam, Rengasamy, Balasubramanian and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read
for greatest visibility
and readership



FAST PUBLICATION

Around 90 days
from submission
to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,
and constructive
peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers
acknowledged by name
on published articles

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: frontiersin.org/about/contact



REPRODUCIBILITY OF RESEARCH

Support open data
and methods to enhance
research reproducibility



DIGITAL PUBLISHING

Articles designed
for optimal readership
across devices



FOLLOW US

@frontiersin



IMPACT METRICS

Advanced article metrics
track visibility across
digital media



EXTENSIVE PROMOTION

Marketing
and promotion
of impactful research



LOOP RESEARCH NETWORK

Our network
increases your
article's readership