Advances In Multidisciplinary & Scientific Research
A Multidisciplinary & Interdisciplinary Journal
Volume 10, No 4, December 2024 Series

Generating Novel Small Molecule Drugs for Selected SARS-CoV-2 Proteins: The Medgnosis GenAI Approach

[1]*Obi, E. D., [1]Yentumi, J.A, [1]Mbatuegwu, D., [2]Ayobami, F. & [1]Obi, T.
[1]Autogon Inc. 3002 Falls at Fairdale, 77057, USA
[2]Department of Neurosurgery, University of Texas, Houston, Texas, USA.
E-mail*:* eobi@autogon.ai, joshua@autogon.ai, david@autogon.ai, fidelixayobami@gmail.com, oladimejioyinkansola36@gmail.com;
Phone: [1]*+1 832 925 1036, [1]+233 54 100 6410, [1]+234 816 873 2219, [2]+1 281 202 7601, +1 832 940 3505;

ABSTRACT

The Coronavirus disease 2019 (COVID-19) pandemic caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has resulted in an unprecedented public health crisis. Owing to the novelty of the virus and its high mutability, there are currently limited SARS-CoV-2-specific treatments available despite extensive research that has been performed on the virus. Therefore, rapid development of effective therapies, especially small molecules against SARS-CoV-2, is urgently needed. In this study, we targeted three core structural proteins of the virus, spike glycoprotein, replicase polyprotein 1a, and replicase polyprotein 1ab, which have been implicated in advancing viral infection. Using our GenAI system, we generated forty-seven novel small molecules targeting these three proteins. We further tested, *in silico*, how these molecules compared to the existing drugs, Paxlovid and Remdesivir. Our data showed that several of the predicted molecules outperform existing drugs with higher binding affinities. Our system presents a much quicker and inexpensive method to develop small-molecule drugs against specific protein targets, representing an important advancement in the fight against novel viruses and their resulting diseases.

Keywords*:* SARS-CoV-2, Small molecules, Spike protein, GenAI, Drug Discovery

Aims Research Journal Reference Format:
Obi, E. D., Yentumi, J.A, Mbatuegwu, D., Ayobami, F. & Obi, T. (2024). Generating Novel Small Molecule Drugs for Selected SARS-CoV-2 Proteins: The Medgnosis GenAI Approach. Advances in Multidisciplinary and Scientific Research Journal Vol. 10. No. 4. Pp 7-18. www.isteams.net/aimsjournal. dx.doi.org/10.22624/AIMS/V10N4P1

1. BACKGROUND

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), which causes the viral disease COVID-19, exploded globally in December 2019, rapidly spreading through the human population, infecting close to 148.5 million people and resulting in up to 3.1 million deaths, as reported by the World Health Organization.[1] Coronaviruses obtain their characteristic names based on the spherical nature of their outer membranes. They are enveloped viruses with a single-stranded, non-segmented RNA genome.[2,3] The genome consists of four structural proteins: spike (S), envelope (E), membrane (M), and nucleocapsid (N) (Fig. 1B).

These four structural proteins are interspersed with 15 non-structural proteins, namely, NSP1, NSP2, NSP3, NSP4, NSP5, NSP6, NSP7, NSP8, NSP9, NSP10, NSP12, NSP13, NSP14, NSP15, and NSP16, together with eight accessory proteins, namely, 3a, 3b, 6, 7a, 7b, 8b, 9b, and ORF14 (Fig. 1A).[3] These proteins perform several functions that contribute to viral replication and assembly.

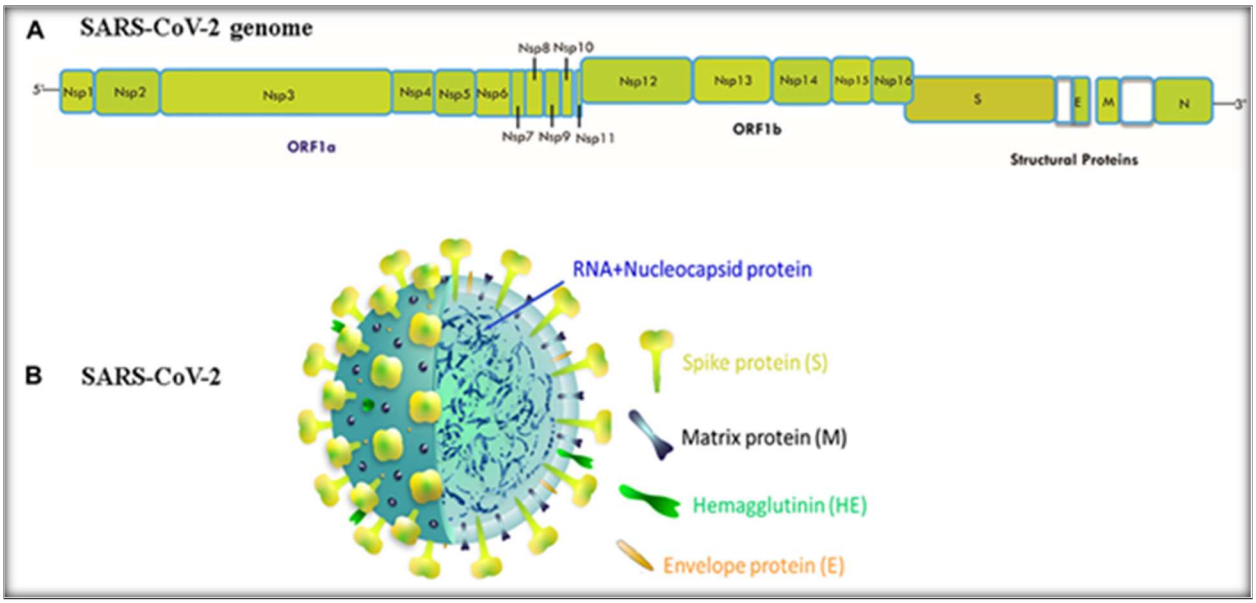


Fig. 1A: Genome organization of SARS-CoV-2 depicting an arrangement of genes in the 5' to 3' direction. 1B: SARS-CoV-2 structure depicting its main structural proteins.[4]

Viral infection begins with the attachment of the receptor-binding domain of the spike (S) protein to angiotensin conversion enzyme II (ACE II), which is found on several cell surfaces. The viral RNA genome is then translated into polyproteins pp1a and pp1ab, which undergo further cleavage by the viral proteases NSP3-PLpro and NSP5-Mpro) to generate further genomic and subgenomic RNAs.[2–4] Structural proteins from the subgenomic mRNAs are then inserted into the endoplasmic reticulum (ER) of the infected cells, undergo further cleavage, and result in newly synthesized SARS-CoV-2 particles that emerge from the ER and are passed to extracellular vesicles for viral release from the cell.[3,5,6]

Current treatment methods for COVID-19 include symptomatic treatment, supportive care, and repurposed drugs, all of which are geared toward inhibiting cellular viral entry.[7] Nazerian et al (2021)[8] report using human recombinant soluble ACE II (hrsACE II) that neutralizes the virus in the serum. Small-molecule drugs targeting the receptor-binding domain (RBD) of the spike protein can also block viral entry, although this target is often poorly conserved across the coronavirus family.[8,9] Examples include remdesivir, nirmatrelvir, and ritonavir (Fig. 2).

Drug name	Class	Mechanism of action	Adverse effect
Chloroquine and Hydroxychloroquine	Antiparasitic	Inhibition of host cell receptor glycosylation to block viral entry, acidification of the endosomal and proteolytic process.	High doses can lead to respiratory arrest, cardiac arrest, and hypokalemia.
Lopinavir/Ritonavir	Antiviral	Inhibition of 3CL protease. Inhibition of viral replication.	Risks for pediatric patients
Remdesivir	Antiviral	RNA-dependent RNA polymerase inhibitor. Block viral replication	There is no information on whether overdosing can cause any adverse effects
Heparin	Anticoagulant and anti-inflammatory	Heparin binds to the RBD of the SARS-CoV-2 protein S, inhibiting viral infection	Platelet count usually decreases to between days 5 and 12
Tocilizumab	Monoclonal antibody	IL-6 inhibiting receptor. Cytokine storm reduction blockade.	Overdose-neutropenia
Anakinra	Immune Response Modulator	Monoclonal antibody that acts against the IL-1 receptor	Rheumatoid arthritis (incidence > 10%)
Baricitinib	Immune Response Modulator	Antiviral activity Inhibitor of clathrin-mediated endocytosis Janus kinases 1 and 2 (JAK1/2 inhibitor)	Multiple adverse reactions
Camostat Mesilate	Antiviral	TMPRSS2 inhibitor that prevents replication Viral Blocks viral mutation	Rash, pruritus, nausea, abnormal values from laboratory tests and diarrhea
Molnupiravir	Antiviral	It works by inducing mutagenesis in viral RNA, causing the newly formed RNA strand chain to terminate	Mild adverse effects
Paxlovid	Antiviral	It inhibits viral replication at a stage known as proteolysis, which occurs before viral RNA replication	Absent

Fig. 2: Examples of medicines used in moderating the symptoms and effects of COVID-19 disease.[3]

The COVID-19 pandemic has had an unprecedented global impact, highlighting the urgent need for rapid and effective therapeutic interventions against emerging infectious diseases. Traditional drug discovery methods are often too slow to keep pace with rapidly evolving pathogens like SARS-CoV-2. Current treatment options for COVID-19 are limited and come with significant drawbacks. For instance, drugs like Paxlovid and Remdesivir have shown varying degrees of efficacy and may not be readily accessible to all populations. Additionally, their effectiveness can be compromised by the virus's high mutation rate, leading to the emergence of resistant strains.

There is therefore a critical need to address this gap by enabling the swift generation of novel small molecules targeting specific protein sequences of the virus. Developing such targeted therapies can overcome the limitations of existing treatments by providing more effective, accessible, and adaptable solutions. This approach not only holds promise for combating COVID-19 but also sets a precedent for responding more efficiently to future pandemics caused by rapidly evolving pathogens. In this paper, a total of 1200 small molecules were generated against the three disease protein targets of SARS-CoV-2 – spike glycoprotein, replicase 1a, and replicase 1ab to find out anti-SARSCoV-2 molecules effective against these protein targets.

2. METHODOLOGY

2.1 Protein Structure Retrieval and Molecule Generation

All protein sequences and crystal structure files in pdb format were retrieved from the Protein Data Bank (PDB) (https://www.rcsb.org). The PDB IDs for each of the proteins were 6LVN, 5R7Y, and 1Q2W for the spike glycoprotein, replicase 1ab, and replicase 1a respectively. The target protein sequence was then passed to our GenAI system for molecule generation.

2.2 Molecular Docking

Using the PDB ID of the proteins of interest, the crystal structure was retrieved, and the SMILE representations of the molecules generated by our system were then passed to our double docking system for molecule docking to begin. The first docking system uses the DiffDock® docking tool (https://github.com/gcorso/DiffDock) which finds binding pockets using deep learning techniques. Using twenty (20) inference steps with a ligand conformation size of ten (10) samples per protein-ligand complex, initial molecular docking was performed. The results from the DiffDock molecular docking process were then fed into the GNINA® docking protocol (https://github.com/gnina/gnina) which also uses deep-learning convolutional networks to find the best binding pockets for each docking pose received from DiffDock. Gnina has an autobox_ligand feature which we used to determine the binding center location of the protein binding site. Using a measured DiffDock confidence score threshold of -1.5 and GNINA minimized affinity score of ≤ -5.0 kcal/mol together with other custom metrics, the best molecules generated were selected as hit molecules.

2.3 ADMETox Prediction

For each molecule generated, we used an industry-based standard ADMET predictive machine learning model, ADMET-Ai® (https://github.com/swansonk14/admet_ai) to generate ADMET descriptors, such as molecular weight of the molecule, solubility (logP), hydrogen bond acceptor and donor values, quantitative estimation of drug-likeness (QED), bioavailability, and Lipinski score amongst many others.

2.4 Report Generation

A comprehensive hypertext markup language (HTML) report detailing the performance and molecule characteristics of each of the molecules was produced at the end of the molecule generation cycle.

3. RESULTS

We used our GenAI system to generate strong small-molecule candidates that exhibit better and stronger **binding affinity** against selected disease proteins of SARS-CoV-2. The model generated thousand two hundred (1200) potential small-molecule drug candidates (400 molecules for each protein target) against the spike protein (Fig. 4A), replicase 1a (Fig. 4B), and replicase 1ab (Fig. 4C) viral proteins of the SARS-CoV-2 virus.

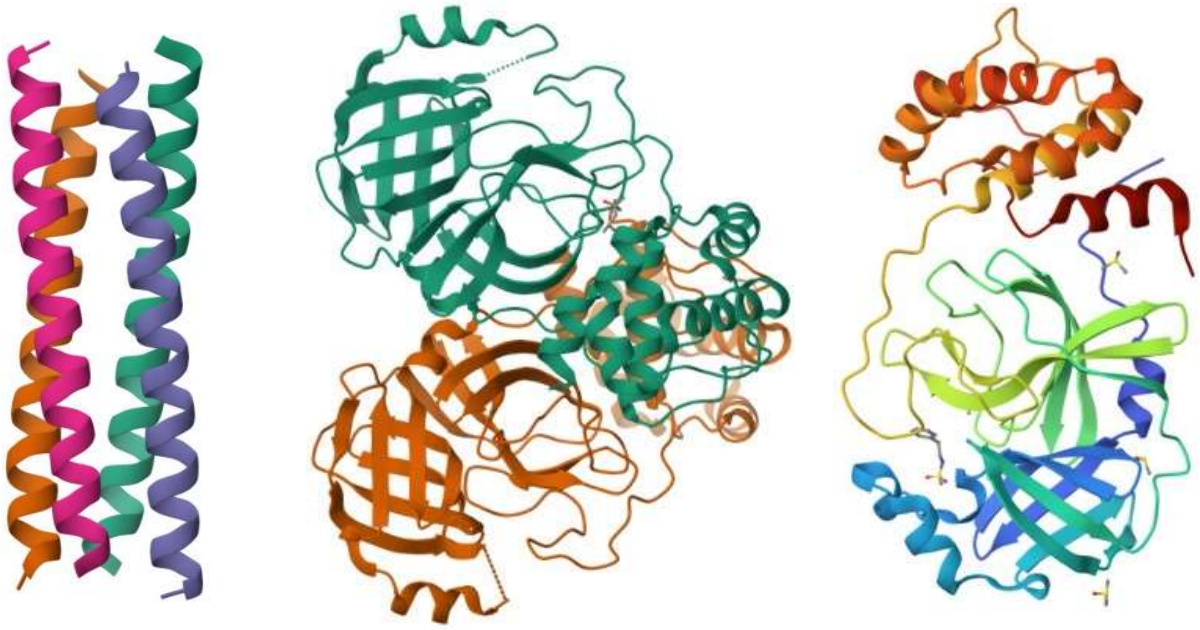
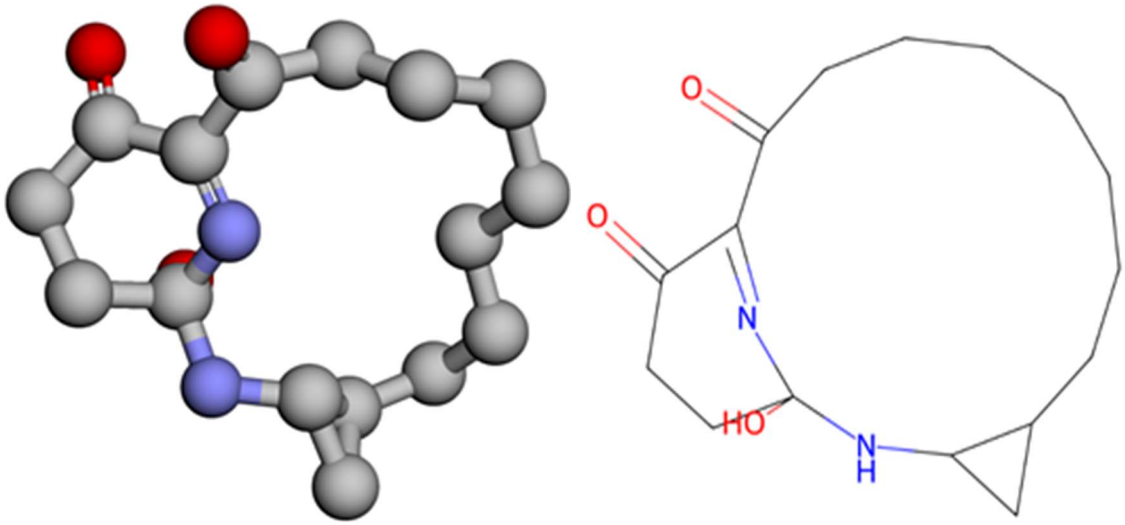


Figure 4: Asymmetric 2D models of A: Spike Protein B: Replicase Polyprotein 1a and C: Replicase Polyprotein 1ab (Retrieved from the Protein Data Bank)

Comparing our generated molecules with existing drugs (Paxlovid and Remdesivir), our molecules demonstrated potentially better inhibitory function, as shown by the higher binding affinities shown in the reports below (Table 3). Interestingly, when provided with the full protein sequence of the Spike protein, which has no FDA-approved small-molecule drug, our hit molecule selection criteria showed one molecule having a strong binding affinity as the hit molecule for this protein target, out of the 400 molecules generated (Table 1 and 2). Fig. 5A and B depict the structure of this hit molecule and Fig. 6 shows the molecule in a binding pocket of the spike glycoprotein. This hit molecule demonstrated favorable drug-like properties, as suggested by its compliance with Lipinski's rule, a QED score above 0.7, and an acceptable logP value. With a high BBB Martin score, this compound may be promising for CNS-targeted therapies. Furthermore, a high bioavailability score is a significant advantage of oral administration. These characteristics were consistent across all the hit molecules generated for the three target proteins.



A B

Fig. 5A: Ball-and-stick representation of hit molecules generated against the SARS-CoV-2 spike protein. B: 2D representation of hit molecule generated against SARS-CoV-2 spike protein

Table 1: Docking Results of Top Medgnosis GenAI Hit Molecule for SARS-CoV-2 Spike Protein

Hit Molecule	MW	Docking Confidence	GNINA Minimized Affinity (Binding Score)	SA Score	Similarity Score
Hit 1	292.379	-1.46	-5.0601	5.3065	0

Table 2: Pharmacokinetic Properties of Top Medgnosis GenAI Hit Molecule for SARS-CoV-2 Spike Protein

Hit Molecule	logP	Lipinski	QED	AMES	BBB Martins	Bioavailability Ma	Carcinogens Lagunin	ClinTox	HBA	HBD	TPSA
Hit 1	1.7279	4	0.7124	0.3024	0.9481	0.8791	0.0897	0.3518	5	2	78.76

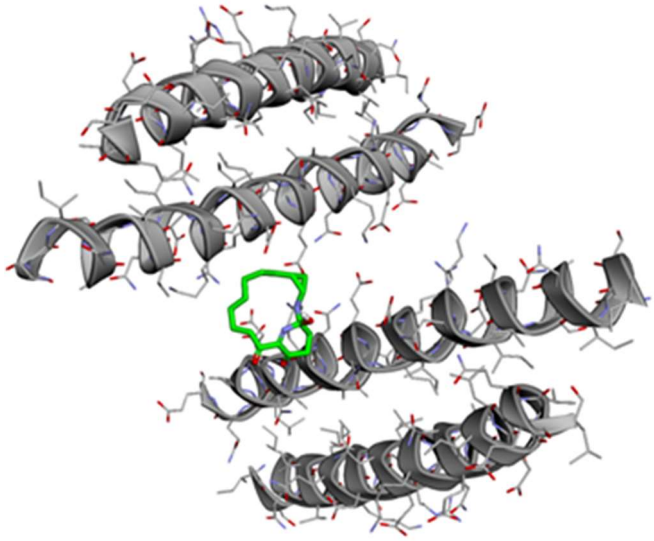


Fig. 6: Hit molecule 1 in a binding pocket of SARS-CoV-2 Spike Protein

Table 3: Comparison of Docking Results Between Current FDA-Approved Drugs and Top Five Medgnosis GenAI Molecules for SARS-CoV-2 Replicase 1a Protein

Hit Molecule	MW	DiffDock Confidence	GNINA Minimized Affinity (Binding Score)	Synthesis Accessibility Score	Similarity Score
Paxlovid	499.534	-2.05	-5.86869	4.5762	1.0
Hit 1	292.427	-1.54	-7.51241	6.1688	0.0
Hit 2	331.293	-0.75	-7.42374	6.9442	0.0
Hit 3	400.48	-1.12	-7.06427	7.1690	0.1068
Hit 4	276.253	-0.98	-6.29543	6.0944	0.0806
Hit 5	306.394	-1.5	-6.14947	6.7826	0.0488

Table 4: Pharmacokinetic Properties Comparison between Current FDA-Approved Drugs and Top Five Medgnosis GenAI Molecules for SARS-CoV-2 Replicase 1a Protein

Hit Molecule	logP	Lipinski	QED	AMES	BBB Martins	Bioavailability Ma	Carcinogens Lagunin	ClinTox	HBA	HBD	TPSA
Paxlovid	1.09718	4	0.5037	0.3477	0.8643	0.5797	0.1177	0.3045	5	3	131.4
Hit 1	1.6382	4	0.5444	0.9736	0.8336	0.9454	0.7638	0.5014	5	3	76.54
Hit 2	-3.34232	2	0.2452	0.9984	0.2704	0.9844	0.5296	0.4343	14	6	160.28
Hit 3	-1.0967	3	0.1150	0.8937	0.3168	0.8094	0.2566	0.3891	11	4	118.04
Hit 4	-1.29352	3	0.1885	0.9982	0.5279	0.9689	0.6370	0.1610	11	3	108.92
Hit 5	-0.80354	4	0.4353	0.9084	0.2704	0.9421	0.4352	0.3908	8	5	94.1

Table 5: Comparison of Docking Results Between Current FDA-Approved Drugs and Top Five Medgnosis GenAI Molecules for SARS-CoV-2 Replicase 1ab Protein

Hit Molecule	MW	DiffDock Confidence	GNINA Minimized Affinity (Binding Score)	Synthesis Accessibility Score	Similarity Score
Remdesivir	499.534	-2.23	-6.40293	4.5763	1.0
Paxlovid	602.585	-2.11	-6.03841	4.8153	1.0
Hit 1	344.465	-0.68	-7.31727	7.4939	0.0
Hit 2	356.327	-0.89	-7.22897	6.0819	0.0
Hit 3	375.438	-0.99	-7.03432	5.0479	0.0
Hit 4	494.583	-1.2	-6.80267	6.4957	0.0
Hit 5	624.726	-1.18	-6.44186	6.6155	0.0

Table 6: Pharmacokinetic Properties Comparison between Current FDA-Approved Drugs and Top Five Medgnosis GenAI Molecules for SARS-CoV-2 Replicase 1ab Protein

Hit Molecule	logP	Lipinski	QED	AMES	BBB Martins	Bioavailability Ma	Carcinogens Lagunin	ClinTox	HBA	HBD	TPSA
Remdesivir	1.0972	4	0.5037	0.3477	0.8643	0.5797	0.1177	0.3045	5	3	131.4
Paxlovid	2.3122	2	0.1640	0.2176	0.4953	0.6716	0.0754	0.6902	13	4	203.55
Hit 1	-1.5114	3	0.2076	0.9545	0.2162	0.8880	0.2587	0.4148	9	7	121.86
Hit 2	-1.4984	3	0.2600	0.8776	0.1721	0.8516	0.2295	0.1690	11	5	129.79
Hit 3	-3.5911	2	0.1768	0.9946	0.3840	0.9033	0.8506	0.5391	12	8	174.65
Hit 4	-5.6551	2	0.0976	0.9810	0.0595	0.7570	0.7160	0.5110	17	12	208.4
Hit 5	-3.2775	1	0.0648	0.8783	0.0527	0.7881	0.3546	0.2962	21	12	238.3

4. DISCUSSION

The novel coronavirus pandemic has had a significant impact on global health care. The healthcare industry has achieved significant results in the development of small-molecule drugs and vaccines that contribute to curtailing the spread and casualty rate of the virus; however, more work needs to be done in the area of developing even more effective medications against viral proteins that exacerbate the infection. To this end, artificial intelligence can significantly improve the rate of discovery of new drugs for such targets.

Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) are critical properties for evaluating the suitability of potential compounds as drug molecules.[10] Experimentally evaluating these properties to appropriately select and optimize lead molecules can be a costly and labor-intensive endeavor. Several machine learning-based tools, such as ADMETlab,[11] admetSAR,[12,] and Pharmaco Kinetics Knowledge Base (PKKB),[13] have been developed to accurately determine these properties computationally, providing a baseline from which *in vivo* studies can proceed.

Our GenAI pipeline employs a machine learning-based ADMET property prediction tool called ADMETAi®. This is a fast and accurate ADMET property prediction tool that uses a graph neural network to make predictions.[14] In determining the physicochemical properties of the molecule, the results suggested that the hit molecule generated against the spike protein had an optimal molecular weight that satisfied the Lipinski rule of five, which states that oral drugs should have a molecular weight of less than 500 Da for optimal bioavailability. At 292.379 Da, this compound was well within the desirable range, potentially facilitating good oral absorption and blood-brain barrier penetration.

Additionally, this molecule had an octanol-water partition coefficient (logP) value of approximately 1.7, demonstrating a good balance between aqueous solubility and membrane permeability. The logP coefficient measures the lipophilicity of a molecule, with a value range generally accepted to be less than 5. Approximately ninety (90%) percent of the drugs developed for human consumption are absorbed by the small intestine;[15] thus, drug permeability across this membrane is one of the most important pharmacokinetic properties of a good molecule. The Topological Polar Surface Area (TPSA), measured in Å^2, determines the likelihood of a potential molecule crossing the intestinal barrier. Values <140 Å^2 are associated with good intestinal absorption, whereas values <90 Å^2 are favorable for BBB penetration. At 78.76 Å^2, our lead compound is predicted to have good absorption properties and potential CNS penetration.[16]

Drug-likeness is crucial for evaluating compounds for drug development, as it helps identify potential failures early, improving success rates and reducing costs.[17] Using the Quantitative Estimation of Drug-likeness (QED) score, one can immediately determine whether a molecule could be a drug. Ranging from 0 to 1, with higher values indicating greater drug-likeness, the lead molecule obtained a value of 0.7124, suggesting that this compound has good overall drug-like properties. The Ames test uses bacterial strains such as *Salmonella typhimurium* to assess the mutagenic properties of chemicals by monitoring reverse mutations. This is a preliminary screening tool for evaluating the carcinogenic potential of chemicals or drug-like molecules.[18] Lower values indicate a lower mutagenic potential.

A score of 0.3024 combined with a low carcinogenicity and moderate clinical toxicity score of 0.0897 and 0.3518, respectively, suggests a relatively low mutagenic and carcinogenic risk, together with a low risk of toxicity in a clinical setting of the lead compound. This molecular profile presents a promising hit compound with well-balanced physicochemical properties and favorable predicted characteristics. The compound demonstrated good drug-likeness scores, low toxicity risks, and promising bioavailability. Its moderate synthetic accessibility, which measures the ease of synthesizing a compound, suggests a balance between novelty and the feasibility of production.

5. CONCLUSION

The successful application of our GenAI system to SARS-CoV-2 underscores its potential to revolutionize how we develop small-molecule drugs against novel viruses. We aim to contribute to the global effort in managing the current pandemic and to pave the way for more agile responses to future infectious disease outbreaks. Our approach represents a significant advancement in the fight against rapidly emerging pathogens, offering hope for more effective and timely therapeutic intervention upon validation via in vitro and in vivo studies.

REFERENCES

1. Nguyen, T. T., Abdelrazek, M., Nguyen, D. T., et al. (2022). Origin of novel coronavirus causing COVID-19: A computational biology study using artificial intelligence. *Machine Learning Applications*, 9, 100328. https://doi.org/10.1016/j.mlwa.2022.100328
2. Sheely, J. J. (2023). A review of Covid-19 detailed study of the history, life cycle, diagnosis, and prevention of coronavirus. *Asian Journal of Research in Infectious Diseases*, 12, 50-61. https://doi.org/10.9734/ajrid/2023/v12i2241

3. Da Silva Torres, M. K., Bichara, C. D. A., de Almeida, M. de N. S., et al. (2022). The complexity of SARS-CoV-2 infection and the COVID-19 pandemic. *Frontiers in Microbiology*, 13, 789882. https://doi.org/10.3389/fmicb.2022.789882
4. Zeyaullah, M., AlShahrani, A. M., Muzammil, K., et al. (2021). COVID-19 and SARS-CoV-2 variants: Current challenges and health concern. Frontiers in Genetics, 12, 693916. https://doi.org/10.3389/fgene.2021.693916
5. Tseng, Y., Chang, C., Wang, S., Huang, K., & Wang, C. (2013). Identifying SARS-CoV membrane protein amino acid residues linked to virus-like particle assembly. *PLoS ONE*, 8(4). https://doi.org/10.1371/journal.pone.0064013
6. Ulasli, M., Verheije, M. H., de Haan, C. A. M., & Reggiori, F. (2010). Qualitative and quantitative ultrastructural analysis of the membrane rearrangements induced by coronavirus. *Cellular Microbiology*, 12(6), 844-861. https://doi.org/10.1111/j.1462-5822.2010.01437.x
7. Li, X., Mi, Z., Liu, Z., & Rong, P. (2024). SARS-CoV-2: Pathogenesis, therapeutics, variants, and vaccines. *Frontiers in Microbiology*, 15, 1334152. https://doi.org/10.3389/fmicb.2024.1334152
8. Nazerian, Y., Vakili, K., Ebrahimi, A., & Niknejad, H. (2021). Developing cytokine storm-sensitive therapeutic strategy in COVID-19 using 8P9R chimeric peptide and soluble ACE2. *Frontiers in Cell and Developmental Biology*, 9, 717587. https://doi.org/10.3389/fcell.2021.717587
9. Shang, J., Wan, Y., Luo, C., et al. (2020). Cell entry mechanisms of SARS-CoV-2. *Proceedings of the National Academy of Sciences of the United States of America*, 117(20), 11727-11734. https://doi.org/10.1073/pnas.2003138117
10. Li, X., Tang, L., Li, Z., Qiu, D., Yang, Z., & Li, B. (2023). Prediction of ADMET properties of anti-breast cancer compounds using three machine learning algorithms. *Molecules*, 28(5), 52326. https://doi.org/10.3390/molecules28052326
11. Dong, J., Wang, N. N., Yao, Z. J., et al. (2018). ADMETlab: A platform for systematic ADMET evaluation based on a comprehensively collected ADMET database. *Journal of Cheminformatics*, 10, 29.
12. Cheng, F., Li, W., Zhou, Y., et al. (2012). admetSAR: A comprehensive source and free tool for assessment of chemical ADMET properties. *Journal of Chemical Information and Modeling*, 52(12), 3099-3105.
13. Cao, D., Wang, J., Zhou, R., et al. (2012). ADMET evaluation in drug discovery. 11. PharmacoKinetics Knowledge Base (PKKB): A comprehensive database of pharmacokinetic and toxic properties for drugs. *Journal of Chemical Information and Modeling*, 52(5), 1132-1137.
14. Swanson, K., Walther, P., Leitz, J., et al. (n.d.). ADMET-AI: A machine learning ADMET platform for evaluation of large-scale chemical libraries. https://doi.org/10.5281/zen
15. Azman, M., Sabri, A. H., Anjani, Q. K., Mustaffa, M. F., & Hamid, K. A. (2022). Intestinal absorption study: Challenges and absorption enhancement strategies in improving oral drug delivery. *Pharmaceuticals*, 15(8), 975. https://doi.org/10.3390/ph15080975
16. Sun, D., Gao, W., Hu, H., & Zhou, S. (2022). Why 90% of clinical drug development fails and how to improve it?. *Acta pharmaceutica Sinica. B*, 12(7), 3049–3062. https://doi.org/10.1016/j.apsb.2022.02.002
17. Li, B., Wang, Z., Liu, Z., et al. (2024). DrugMetric: Quantitative drug-likeness scoring based on chemical space distance. *Briefings in Bioinformatics*, 25(4), 321. https://doi.org/10.1093/bib/bbae321

18. Uesawa, Y. (2024). Progress in predicting Ames test outcomes from chemical structures: An in-depth re-evaluation of models from the 1st and 2nd Ames/QSAR international challenge projects. *International Journal of Molecular Sciences*, 25(3), 31373. https://doi.org/10.3390/ijms25031373