Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
In-silico screening to delineate novel antagonists to SARS-CoV-2 nucleocapsid protein

Mohd Fardeen Husain Shahanshah a, b, D. Anvitha a, Vandana Gupta a, *

a Ram Lal Anand College, University of Delhi, New Delhi, India
b Delhi Technological University, New Delhi, India

ARTICLE INFO

Keywords:
SARS-CoV-2
Nucleocapsid
FDA
Docking
COVID-19

ABSTRACT

Since its inception, SARS-CoV-2 has crossed all borders and continues rampaging around the globe, causing profound economic damage and heavy burden on the scientific community and the healthcare fraternity and facilities. With the emergence of new variants, the global pandemic has prolonged and raised concerns regarding the existing therapies. Most of the identified mutants have the potential to exacerbate the already existing crisis. In line with the urgent need for promising antivirals against the novel coronavirus, we conducted an in-silico drug docking study using SeeSAR and other bioinformatics tools and identified prospective molecules that target the nucleocapsid protein of SARS-CoV-2. The highly conserved N protein plays a crucial role in viral assembly and pathogenicity by interacting with the host ribosomal subunits and suppressing nonsense mediated decay (NMD) of viral mRNA by the host cell. In the current study, FDA approved drugs were docked into pockets created within the N protein including the crucial conserved residues and analyzed for their affinity. The docked compounds give us novel plausible models that can be inspected further and paves way for the development of potent therapeutics against SARS-CoV-2.

1. Introduction

As of May 5, 2022, more than 500 million cases of the novel coronavirus have been reported globally (Worldometer, 2022). Besides straining/crippling the healthcare fraternity and the scientific community, the pandemic has also disrupted the socio-economic balance. The profound debilitating impact of COVID-19 shall be remembered for generations to come (Shang et al., 2021). SARS-CoV-2 is a betacoronavirus that has a 30 knt, positive sense, single stranded RNA molecule as its genome. The RNA genome is surrounded by a nucleocapsid protein and a viral envelope which itself consists of three other structural proteins namely Membrane, Envelope and Spike (Wu et al., 2020; Wu A et al., 2020). Its extensive genome contains 14 open reading frames (ORFs) encoding 27 distinct proteins. This virus also shares homology with other notable viruses like SARS-CoV and MERS-CoV, both of which have led to outbreaks in the past. Transmission of the virus particle occurs primarily via respiratory aerosols but may also occur through fomites and human contact. Within the host, SARS-CoV-2 exhibits an array of signs and symptoms. Most common indications being fever, fatigue, cough, shortness of breath, loss of smell and taste. The mild infection may augment into critical complications like Acute Respiratory Distress Syndrome (ARDS) or severe pneumonia and enhance the probability of fatality particularly in patients with comorbid conditions. Furthermore, some individuals may not develop any symptoms, still acting as active carriers for the pathogen (Kaur and Gupta, 2020; Sharma et al., 2021; Giovanetti et al., 2021; Singh and Gupta, 2021).

With the successful development of efficacious vaccines, the progression of the pandemic seems to have slowed down but the advent of different variants of SARS-CoV-2 convey altogether a different reality. SARS-CoV-2 has acquired several adaptations while being transmitted from one region to another making it challenging to combat. There is a dissonance between the rate of mass immunization and rate of emergence of new viral variants. As long as the virus persists and continues to propagate within the population, the mutants will keep on emerging (Cascella et al., 2021; Wu A et al., 2020; Mistry et al., 2022; Thakur and Ratho, 2022). So far strategies to treat the coronaviral infection were limited to symptom management. Drugs such as remdesivir, ivermectin and hydroxychloroquine were suggested as potential antivirals; however, more controlled clinical trials are pertinent to undoubtedly prove their efficacy, which may be rendered ineffective in due course against...
the rapidly evolving variants.

Recently, the Medicines and Healthcare products Regulatory Agency (MHRA) authorized Molnupiravir, a ribonucleoside analog, for treating mild to moderate cases of COVID-19, making it the first specific anti-SARS-CoV-2 drug (Drugs, 2021; Aripaka, 2021). Nevertheless, there is still a lack of designated antiviral treatments against SARS-CoV-2. This necessitates the discovery of potent therapeutics against the novel coronavirus (Cascella et al., 2021; Kaur and Gupta, 2020; Sharma et al., 2020). New variants such as delta plus and omicron have the potential to prolong the pandemic and pose a significant threat to the world at large. The fight against the virus continues as it still remains a consequential health crisis (WHO, 2021).

With the wealth of literature and computational resources available, numerous in silico studies have been reported over such a brief span of time. Well known targets of SARS-CoV-2 such as Spike glycoprotein, RdRp (RNA dependent RNA Polymerase), PLpro, Mpro, Helicase and 3CLpro have been extensively researched due to their indispensable role in viral replication and pathogenesis (Peele et al., 2021). Lopinavir, umifenovir, hydroxychloroquine, ritonavir, ivermectin, favipiravir and remdesivir are some of the drug leads that have been repurposed against SARS-CoV-2. These antivirals and antiparasitic drugs specifically target the RdRp of the novel hCoV and obstruct the replication process of viral genome (Peele et al., 2021; Amin et al., 2021; Singh et al., 2020). Tocilizumab, an immunosuppressant used for rheumatoid arthritis has also shown antagonistic activity against SARS-CoV-2, although its safety and efficacy is yet to be ascertained (Amin et al., 2021; Singh et al., 2020). Natural compounds like tripterpenes, epigallocatechin gallate (EGCG), withaferin A and artemesunate have exhibited antiviral activity by interacting with the key viral proteases Mpro and PLpro. The binding energies from molecular docking studies ranged between –8.0 kcal/mol to –9.6 kcal/mol and were further validated by Molecular Dynamics (MD) simulations (Hisham Shady et al., 2020; Sharma and Deep, 2020). Commercially available drugs like temozolomide, remdesivir, baflofenycin A1 and colchicine displayed high binding affinity towards M protein of SARS-CoV-2 suggesting their potency as a COVID-19 therapeutic agent (Peele et al., 2021; Singh et al., 2020). Compounds like rapamycin, camostat, nafamostat, glycyrrhizic acid, theaflavin and curcumin were able to effectively inhibit the RNA binding domain of N protein of SARS-CoV-2 and displayed promising results in the in silico studies; however, more research is required to validate their efficacy and determine their modes of action (Tatar et al., 2021; Ray et al., 2020).

Nucleocapsid or N protein is one of the four structural proteins of SARS-CoV-2.Encoded by ORF-9, this 419 aa long, highly basic protein consists of two major functional and structural domains, an N-terminal RNA binding domain (NTD) and a C-terminal dimerization domain (CTD) (Zeng et al., 2020; Hasokszisz et al., 2020; Kang et al., 2020; Yoshimoto, 2020; Matsuo, 2021). Both these domains are copious in beta strands and are interspersed by three intrinsically disordered regions (IDRs), the N-arm, the central Ser/Arg (SR)-rich linker and the C-tail, that are characteristic to the protein (Zeng et al., 2020; Hasokszisz et al., 2020; Kang et al., 2020; Zhu et al., 2020). The IDRs are extremely dynamic and possess unique regions that impart transient helicity to protein, making it a suitable local binding interface for protein-RNA or protein-protein interactions (Cabuk et al., 2020; Zhu et al., 2020; Matsuo, 2021). The overall architecture of N protein of SARS-CoV-2 is highly conserved and exhibits great sequence identity to N proteins of other human coronaviruses such as SARS-CoV and MERS-CoV (Kang et al., 2020). Fig. 1 illustrates the structural organization of the N protein of SARS-CoV-2.

N protein of SARS-CoV-2 performs an array of functions, out of which packaging of the viral RNA genome and formation of a stable ribonucleoprotein (RNP) is central. Owing to its distinct surface chemistry, the positively charged regions of both NTD and CTD act as interfaces for RNA binding. Several studies suggest that the 3' end of the viral RNA genome associates with the N arm of the NTD. On the contrary, oligomerization of different regions within the protein are primarily mediated by the CTD (Hasokszisz et al., 2020; Kang et al., 2020). Although the two functional domains barely show any interaction with each other, presence of both is necessary for stabilizing the viral RNA (Cabuk et al., 2020). The serine and arginine rich central linker is abundant in phosphorylation sites and contributes to N protein’s multifarious activity (Prajapati et al., 2020; Zeng et al., 2020). The IDR downstream to CTD also plays a crucial role in establishing interactions with M protein (Prajapati et al., 2020). Binding of M protein to N protein-RNA complex stabilizes the internal core of virion (Malik, 2020). Besides fabrication of the RNP complex, N protein is also involved in viral assembly and budding. In conjunction with other structural and non structural proteins, N protein also aids in regulation of RNA replication and translation (Malik, 2020; Hasokszisz et al., 2020; Kang et al., 2020). Moreover, as indicated by some studies N protein also facilitates regulation of host-pathogen interactions by disrupting host cell metabolism and cell cycle progression (Prajapati et al., 2020; Yoshimoto, 2020; Kang et al., 2020). The structural protein can interact with host ribosomal subunits and prevent nonsense mediated decay (NMD) of the viral mRNA by host cell RNAases. It also obstructs the interferon signaling pathway and hampers host cells’ innate immune response (Siripilla et al., 2020; Gupta et al., 2020; Tufan et al., 2020; Wada et al., 2018). The multitude of vital functions performed by the N protein attests to its importance, making it an attractive target site for therapeutics.

This study is intended to provide a comprehensive overview on the potential of the N protein of SARS-CoV-2 as a therapeutic target, including an in depth analysis of the molecular docking patterns of a library of FDA approved drugs.

2. Methodology

In this study we performed an in-silico screening of a library of FDA approved drugs to delineate novel antagonists that target the N protein of SARS-CoV-2. Molecular docking was carried out against the crystal structures of the two essential domains of the N protein, the NTD (PDB ID: 7ACT) and the CTD (PDB ID: 7DE1) using SeeSAR which is an interactive, drug discovery platform that allows for virtual screening of
potential lead molecules against a particular target via sequential molecular docking. Developed by BiosolveIT, the softwares’ intuitive interface also supports ligand optimization by employing a multicentric approach that includes both intermolecular binding affinity and structural complementarity to maximize likelihood of success. The protein structures have been shown in Fig. 2.

Both SeeSAR and PyMOL were employed to construct and visualize an appropriate binding/docking pocket within the protein structures. For our docking pocket, we targeted crucial and conserved amino acid residues that reportedly interact with the viral RNA and aid in protein dimerization (Chen et al., 2007; Kang et al., 2020; Yang et al., 2021; Peng et al., 2020; Zhou et al., 2020). Some of the residues that were included within the NTD docking pocket were Thr54, Leu56, Arg88, Ala90, Lys102, Leu104, Tyr109, Tyr111, Glu171. CTD binding pocket included Lys256, Lys257, Pro258 Arg259, Lys261, Arg262, Lys266, Arg276 and Arg277. Both the docking pockets were heavily dominated by Arg and Lys which impart a positive charge to the protein surface, thereby allowing for an electrostatic interaction with the negatively charged RNA molecule (Fig. 3).

A library of 1615 FDA approved drugs was retrieved from ZINC15
NTD binding pocket. Meanwhile, in case of CTD docking, Mitoxandrone affinities were in the nanomolar range, suggesting more productive and PZC reported the best affinity (see Fig. 4). The predicted binding

### 3. Results

Multiply rounds of molecular docking-based virtual screening was done to obtain more precise data. The 2D plot of selected poses was prepared for the N-Terminal Domain, 1614 FDA approved drugs were docked onto a 25 aa long binding pocket. 6 poses for each molecule were generated. From this initial docking, 44 molecules were shortlisted based on their binding affinity, intermolecular and intramolecular clashes, torsion quality and LLE. These 44 molecules were then re-docked and 100 poses were generated for each. Ultimately 9 molecules made it to the final list, out of which only 2 reported binding affinities that lie in the nM range. Similarly for the C-Terminall Domain, 1614 FDA approved drugs were docked onto a 29 aa long binding pocket, 6 poses were generated for each molecule. 17 unique molecules made it to the first shortlist based on binding affinity, intermolecular and intramolecular clashes, torsion quality and LLE. These 17 molecules were then re-docked with 100 poses generated for each. Ultimately 9 molecules have been shortlisted in the concurrent list out of which only 3 reported binding affinities that lie in the nM range.

Dfo reported a Binding Affinity that spanned between 23 nm and 2400 nm. From the pose view it can be clearly seen that important amino acid residues like Lys102 and Leu104 are involved in ligand interaction (Fig. 5).

Mitoxandrone reported a binding affinity that ranged between 82 nm and 8208 nm. From the pose view it can be clearly seen that important amino acid residues like Lys361 and Pro258 are involved in ligand interaction (Fig. 6).

Based on the results obtained, we can hypothesize that the principal mode of action of both Dfo and Mitoxandrone will be to inhibit viral protein-protein/protein-RNA interactions. Though the predicted binding affinity in nanomolar range doesn’t seem to be very promising, lead optimization might insinuate stronger and more productive protein-ligand interactions that may be consequential. Once bound to the target, the proposed drugs may render the N protein inoperative by inducing conformational changes within it, eventually disrupting the

### 4. Discussions

For the N-Terminal Domain, 1614 FDA approved drugs were successfully docked onto a 29 amino acids binding pocket. 6 poses for each molecule were generated. From this initial docking, 44 molecules were shortlisted based on their binding affinity, intermolecular and intramolecular clashes, torsion quality and LLE. These 44 molecules were then re-docked and 100 poses were generated for each. Ultimately 9 molecules made it to the final list, out of which only 2 reported binding affinities that lie in the nM range. Similarly for the C-Terminal Domain, 1614 FDA approved drugs were docked onto a 25 aa long binding pocket, 6 poses were generated for each molecule. 17 unique molecules made it to the first shortlist based on binding affinity, intermolecular and intramolecular clashes, torsion quality and LLE. These 17 molecules were then re-docked with 100 poses generated for each molecule. Currently, 10 molecules have been shortlisted in the concurrent list out of which only 3 reported binding affinities that lie in the nM range.

Dfo reported a Binding Affinity that spanned between 23 nm and 2400 nm. From the pose view it can be clearly seen that important amino acid residues like Lys102 and Leu104 are involved in ligand interaction (Fig. 5).

Mitoxandrone reported a binding affinity that ranged between 82 nm and 8208 nm. From the pose view it can be clearly seen that important amino acid residues like Lys361 and Pro258 are involved in ligand interaction (Fig. 6).

Based on the results obtained, we can hypothesize that the principal mode of action of both Dfo and Mitoxandrone will be to inhibit viral protein-protein/protein-RNA interactions. Though the predicted binding affinity in nanomolar range doesn’t seem to be very promising, lead optimization might insinuate stronger and more productive protein-ligand interactions that may be consequential. Once bound to the target, the proposed drugs may render the N protein inoperative by inducing conformational changes within it, eventually disrupting the
viral assembly cascade and suppressing virion assembly. Dfo also exhibits binding with the CTD, which can further strengthen the binding of Dfo with N protein. The current indications of Dfo in treatment of iron overdose, potential application for Hypoxia, Diabetes Mellitus etc. which are the conditions of concern in severe COVID-19 patients (Singh et al., 2021), might have added advantage of management of these
conditions along with direct inhibition of viral assembly in COVID-19 patients. Mitoxandrone being an antineoplastic compound will have limited applicability in COVID-19 patients due its possible severe effects, but lead optimization to reduce its toxicity and enhance binding with N protein might be of help in the development of this compound as anti-SARS-CoV-2 molecule. Although, for the suggested lead compounds to be approved for clinical use, it is imperative to first warrant their potency and safety through multiple in vitro and in vivo studies.

5. Conclusion and future prospects

The damaging implications of COVID-19 have been suffered by nations all across the globe. Due to its persistent and erratic nature, a necessity for efficacious anti-SARS-CoV-2 therapeutics has emerged. The high rates of mutations of RNA viruses are correlated with enhanced immune escape, virulence and transmissibility. Hence, there is an imperative need for efficacious antivirals that can extricate SARS-CoV-2. This molecule will have added advantage of ameliorating the symptoms of COVID-19 such as increase in serum iron concentration, hypoxia and steroid induced diabetes and the pre-existing diabetes in COVID patients. Optimization of the binding site and the molecules may produce a more accurate and reliable result.

Funding information

This study is funded in part by star college scheme sanction letter no. BT/HRD/11/3/2020 provided by Department of Biotechnology, India.

CRediT authorship contribution statement

Mohd Fardeen Husain Shahanshah: Investigation, Data curation, Writing – original draft, Visualization. D. Anvitha: Investigation, Data curation, Visualization. Vandana Gupta: Conceptualization, Methodology, Resources, Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Prof. Vandana Gupta reports financial support was provided by Department of Biotechnology.

Acknowledgements

We thank the Department of Biotechnology, India, for the financial support provided under DBT Star College Scheme to the department of Microbiology, Ram Lal Anand College, required for the conduct of this study.

References

Amin, S.A., Banerjee, S., Ghosh, K., Gayen, S., Jha, T., 2021 Jan 1. 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 https://doi.org/10.1016/j.bmc.2020.115860. Epub 2020 Nov 6. PMID: 33191083; PMCID: PMC7647411.

Aripaka, P., 2021. Britain Approves Merck’s COVID-19 Pill in World First, 2021, November 5. https://www.reuters.com/business-healthcare-pharmaceuticals/britain-approves-mercks-oral-covid-19-pill-2021-11-04/. Retrieved November 6.

Cascella, M., Rajnik, M., Allegranzi, B., Dulebohn, S.C., Di Napoli, R., 2021 Jan. Features, evaluation, and treatment of coronavirus (COVID-19). 2021 Sep 2. In: StatPearls [Internet]. Treasure Island (FL). StatPearls Publishing. PMID: 32150360.

Chen, C.Y., Chang, C.K., Chang, Y.W., Sue, S.C., Bai, H.I., Riang, L., Hsiao, C.D., Huang, T.H., 2007 May 11. Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA. J. Mol. Biol. 368 (4), 1075–1086. https://doi.org/10.1016/j.jmb.2007.02.069. Epub 2007 Mar 2. PMID: 17379242; PMCID: PMC7094638.

Cubuk, J., Alston, J.J., Incicco, J.J., Singh, S., Stuchell-Brereton, M.D., Ward, M.D., Zimmerman, M.I., Vithani, N., Griffith, D., Wagoner, J.A., Bowman, G.R., Hall, K.B., Soranno, A., Holehouse, A.S., 2020 Dec. 2020.06.17.1358121. The SARS-CoV-2 Nucleocapsid Protein Is Dynamic, Disordered, and Phase Separates with RNA. bioRxiv [Preprint], 21. https://doi.org/10.1101/2020.06.17.1358121. Update in: Nat Commun. 2021 Mar 29;12(1):1936. PMID: 32587966; PMCID: PMC710622.

Dinesh, D.C., Chalupska, D., Silhan, J., Koutna, E., Nencka, R., Veverka, V., Boura, E., 2020 Dec 2. Structural basis of RNA recognition by the SARS-CoV-2 nucleocapsid protein phosphoprotein. PLoS Pathog. 16 (12), e1009100 https://doi.org/10.1371/journal. ppat.1009100 . PMID: 33264373; PMCID: PMC7735635.

Einhorn, M., Eick, S., Stützer, M., von Ehrenstein, G., Wank, M.P., 2021. SARS-CoV-2 Protease: Structure, Function, and Inhibition. Curr. Top. Microbiol. Immunol. 460, 1–86. https://doi.org/10.1007/978-3-030-76849-6_1. Epub 2020 Nov 12. PMID: 33248401; PMCID: PMC7647411.

Fig. 6. 2D plot -ligand (Mitoxandrone) and CTD binding site interaction.
