Applied Chemical Engineering (2023) Volume 6 Issue 3
doi: 10.24294/ace.v6i3.2325

ORIGINAL RESEARCH ARTICLE

Omicron variants of SARS CoV-2 indicate how small molecules can interfere with spike glycoprotein trimerization

Neri Niccolai[1,2], Alfonso Trezza[1,*], Federico Marchini[1], Pietro Bongini[3], Monica Bianchini[3], Annalisa Santucci[1,4], Ottavia Spiga[1,4,*], Anna Visibelli[1]

[1] *Department of Biotechnology, Chemistry and Pharmacy, University of Siena, 53100 Siena, Italy*

[2] *Le Ricerche del BarLume Free Association, Ville di Corsano, Monteroni d'Arbia, 53014 Siena, Italy*

[3] *Department of Information Engineering and Mathematics, University of Siena, 53100 Siena, Italy*

[4] *Industry 4.0 Competence Center ARTES 4.0 Viale Rinaldo Piaggio, 56025 Pontedera, Pisa, Italy*

*** Corresponding authors:** Alfonso Trezza, alfonso.trezza2@unisi.it; Ottavia Spiga, ottavia.spiga@unisi.it

ABSTRACT

In our search for a possible achilles' heel of SARS-CoV-2, we explored the variability of 1,382,462 complete sequences of the viral spike glycoprotein, all the ones that we could retrieve from the NCBI SARS-CoV-2 databank as of 6 March 2023. Then, by using the Shannon entropy algorithm, we quantified the sequence variability of SARS-CoV-2 spike glycoprotein. With PDBePISA, we have performed a detailed analysis of protomer-protomer interfaces of the spike glycoprotein for two representative structures of different viral variants. The largest protomer-protomer contact patch that is present in the stem region of both structures is highly conserved. It is remarkable that the Asp796Tyr mutation, centered in this patch, is always present in all the Omicron variants. The structure of the SARS-CoV-2 Omicron spike glycoprotein trimer indicates that Tyr796 and Phe898 of the same protomer form a network of aromatic sidechains with Tyr707 of another protomer, yielding a strong constraint that stabilizes the spike glycoprotein quaternary assembly. We believe that the resulting structural stability of the viral trimer is among the key features for the successful proliferation of Omicron variants. This finding also supports the fact that disrupting this network of aromatic moieties with suitable small molecules would represent a powerful antiviral strategy.

Keywords: SARS-CoV-2; Omicron variants; sequence variability; spike glycoprotein; structure disassembly; interfering ligands

ARTICLE INFO

Received: 29 June 2023
Accepted: 14 August 2023
Available online: 12 October 2023

1. Introduction

Soon after the discovery that a new betacoronavirus named SARS-CoV-2 was responsible for the infections that have resulted in almost 7 million deaths since January 2020, the entire scientific community began to study ways to defeat this unexpected threat to human health. The initial major efforts for reducing viral infection were in the direction of developing efficient vaccines. However, once SARS-CoV-2 began to reveal all the mechanisms of its life cycle and several molecular structures of its protein repertoire were resolved, researchers moved also to the search for possible achilles' heels, to gain alternative strategies to vaccination that could reduce viral infectivity. Thus, researchers, with unprecedented, concerted efforts, proposed a large variety of different approaches to fight against COVID-19. They developed antibodies that could inhibit SARS-CoV-2 binding to the host cell receptor[1] and selected small molecules that could block the conformational change of the SARS-CoV-2 spike glycoprotein,

henceforth called SP, between its closed and fusion forms or that could interfere with the viral enzymatic repertoire[2]. Furthermore, the trimeric assembly of SP has been considered a target for antiviral therapies, once the quaternary structure has been experimentally disassembled with antibodies[3] and hindrance to SP trimerization has been predicted in the presence of suitable small molecules[4].

In the present report, we show how the huge amount of information now available for SARS-CoV-2 (https://www.ncbi.nlm.nih.gov/sars-cov-2/) can shed light on the overwhelming prevalence of its Omicron variants. The increased stability of the trimeric SP assembly, an important step for the viral life cycle[5], appears to be reached by the Omicron variants through additional constraints at the protomer-protomer interfaces (PPI). However, the successful evolution of these variants indicates also that interfering with SP trimer formation by using small molecules can be a successful alternative strategy for reducing viral proliferation.

2. Methods

We have used the NCBI SARS-CoV-2 databank (https://www.ncbi.nlm.nih.gov/sars-cov-2/) to retrieve an updated list of all the reported SARS-CoV-2 protein sequences to be aligned by using Clustal Omega with default parameters[6]. Sequence variability of the viral spike protein has been analyzed with the Shannon algorithm[7] whose variability index is expressed as a dimensionless parameter.

From the protein data bank, we have retrieved two quaternary structures of SARS-CoV-2 spike glycoproteins, PDB ID 6VSB[8] and 7WZ1[3], as the basis for the present study. PDBePISA analysis[9] yielded the complete list of residues located in close contact among the SP protomer interfaces of 6VSB and 7WZ1 structures. We have used Autodock/VinaXB[10] to calculate the binding affinity between one SP pocket and ligands contained in the DrugBank database v.5.1.5[11]. For structural data analysis and presentation, we have used the open-source software PyMOL v. 1.7.1.0 (the PyMOL molecular graphics system, version 1.7.1.0, Schrödinger, LLC).

3. Results

To explore the sequence variability of the SARS-CoV-2 spike protein we have considered the data freely accessible from NCBI among the SARS-CoV-2 resources. Thus, as of 6 March 2023, we could retrieve 1,382,462 complete SP sequences. For the alignment procedure, carried out with the Clustal Omega software[6], 148 SP sequences were removed, as they contained unassigned protein sequence positions, which could cause ambiguities in the present investigation. Due to the very large number of SP sequences in the used dataset, multiple alignments were performed in several subsets. Upon summation of all the obtained results, the SP sequence variability was quantified in terms of Shannon entropy, see **Figure 1**.

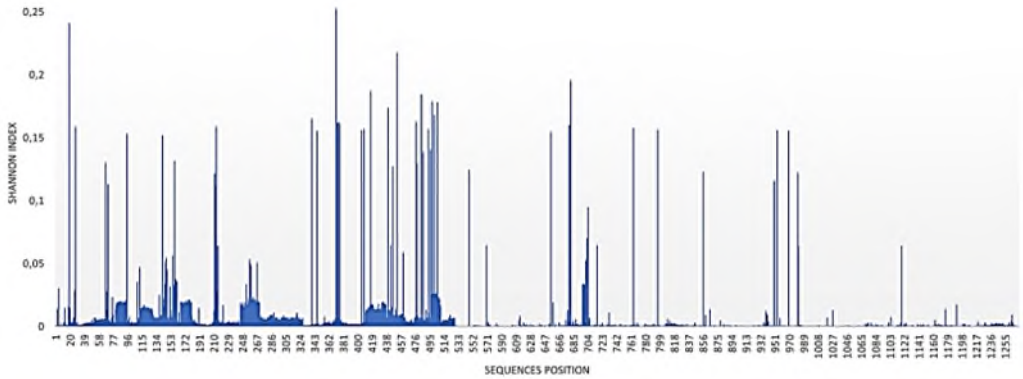

Figure 1. Position-specific Shannon entropies.

In **Figure 1**, the sequence variability of SARS-CoV-2 spike glycoprotein expressed in terms of Shannon entropy. The complete list of Shannon index values is given in Supplementary materials **Table S1**.

As already reviewed[12], extensive mutations occur along the entire SP sequence, particularly located in the SP S1 subunit spanning residues 1 to 685, being a major hindrance to permanent immunization through vaccinations[13]. As SARS-CoV-2, like all the other coronaviruses, requires a trimeric SP assembly to bind the host cell receptor, we performed a structure-driven exploration of the sequence variability that included protomer-protomer interfaces, henceforth named PPI. The analysis of PPI was based on two PDB entries, 6VSB[8] and 7WZ1[3], as representative structures for ancestral and Omicron variants, respectively, by using the PDBePISA tool[9], a service that is offered by the European bioinformatic institute at https://www.ebi.ac.uk/pdbe/pisa/. It should be noted that 7WZ1, in the file nomenclature, reports protomer B in the up conformation and protomer A and C in the down conformation.

As expected, we found that all the variability minima occur in protomer-protomer interfaces, identically distributed in the two investigated SP structures. A combined analysis of structural features with Shannon entropy indexes yields a tri-dimensional mapping of SP sequence variability, including all the Omicron-related mutations that have been recently listed[14]. Consistently with the latter findings, we have noticed that most of the amino acid replacements occur in the S1 subunit. The 7WZ1-driven structural inspection of all the reported Omicron mutations showed that only Asp796Tyr and Asn856Lys are positioned in the PPI region. The relevance of this finding is twofold, as it shows the efforts that SARS-CoV-2 puts to react against the human immunological defense by mutating outer residues and, at the same time, how it tries to reinforce the structural stability of its SP quaternary assembly with the inner ones, vide infra.

The PDBePISA survey on PPIs delineated the presence of many residues that are closer than 5A among the three SP protomers, which mostly are grouped in contact patches, acting as locks for the quaternary assembly. By restricting our analysis just to the interaction of protomer A and B, we have localized two of these contact patches, see **Figure 2**.

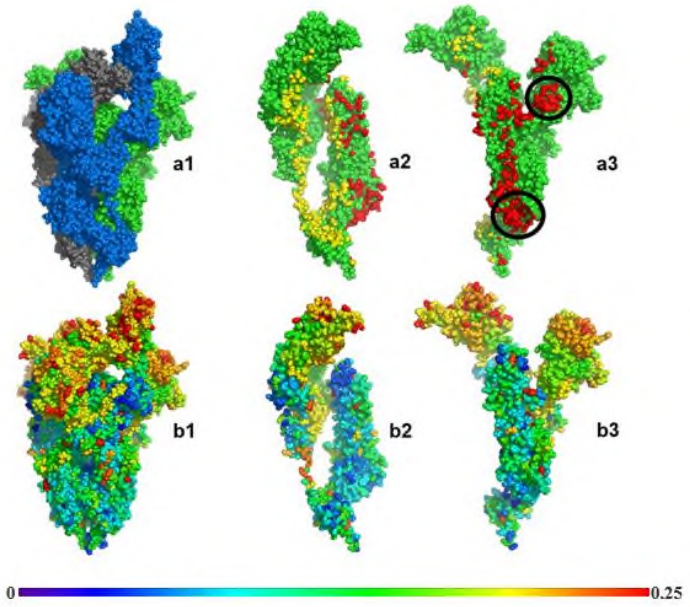

Figure 2. Spacefill representations of SARS-CoV-2 spike.

In **Figure 2**, spacefill representations of SARS-CoV-2 spike trimer from 7WZ1 PDB structure. In **Figure 2(a1)**, protomers A (green), B (blue) and C (grey) are shown; **Figure 2(a2,a3)** shows different orientations of protomer A with red and yellow atoms referring respectively to contacts closer than 5A from protomer B and C; **Figure 2(b1,b2,b3)** shows the same corresponding views where residues are colored according to their

Shannon indexes; in **Figure 2(a3)**, black circles highlight the main contact regions between protomer A and B. The color bar from blue to red (Shannon indexes of 0 and 0.25, respectively) represents the trend of residue Shannon indexes in their protein position.

In **Table S2** (Supplementary materials), we have reported the PDBePISA characterization of the contact patch in the A-B protomeric interface from the 7WZ1 structure (almost identical data were obtained for the 6VSB structure), where the druggable pocket that we have identified in a previous investigation is located[4]. **Figure 3A** shows the close approach of the Omicron variant Tyr796 residue of protomer A with the Tyr707 of protomer B to form a T-shaped structure with their tyrosyl side chains. An additional contact patch of protomer A, which involves residues in the region 38–49 that approach residues of the region 557–574 of protomer B, is described in **Figure 3B** (hydrogen bond network reported in Supplementary materials **Table S3**).

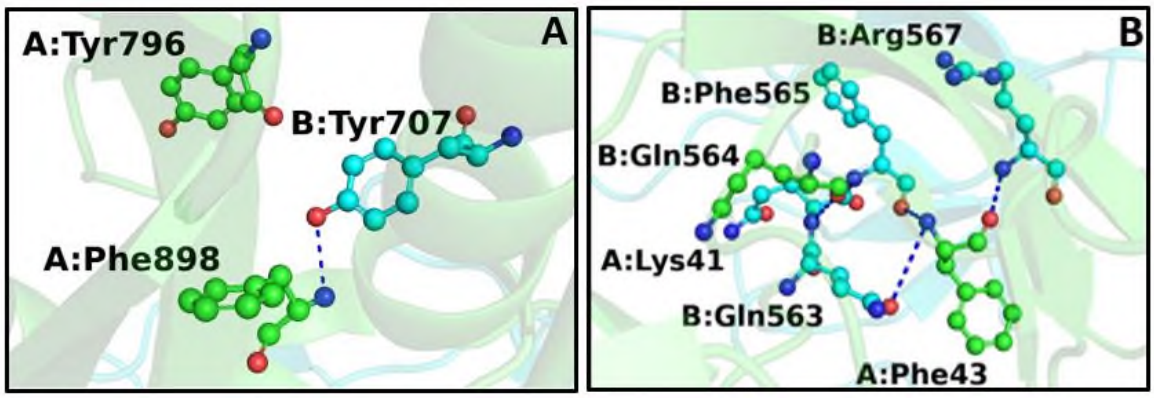

Figure 3. Protomer A and B contact patches.

In **Figure 3**, overview of contact patches between protomer A and B. In **Figure 3A**, the close contacts among the aromatic side chains of Tyr796 and Phe898 of protomer A and Tyr707 of protomer B. In **Figure 3B**, the close contacts among the side chains of Lys41 and Phe43 of protomer A and Gln563, Gln564, Phe565, and Arg567 of protomer B. A blue dotted line highlights the hydrogen bonding between residues.

Furthermore, we have considered 796–798, 883 and 896–898 sequence positions, as they define the pocket where the almost invariant Tyr707 can insert its aromatic side chain to stabilize the SP quaternary assembly (**Figure 4**).

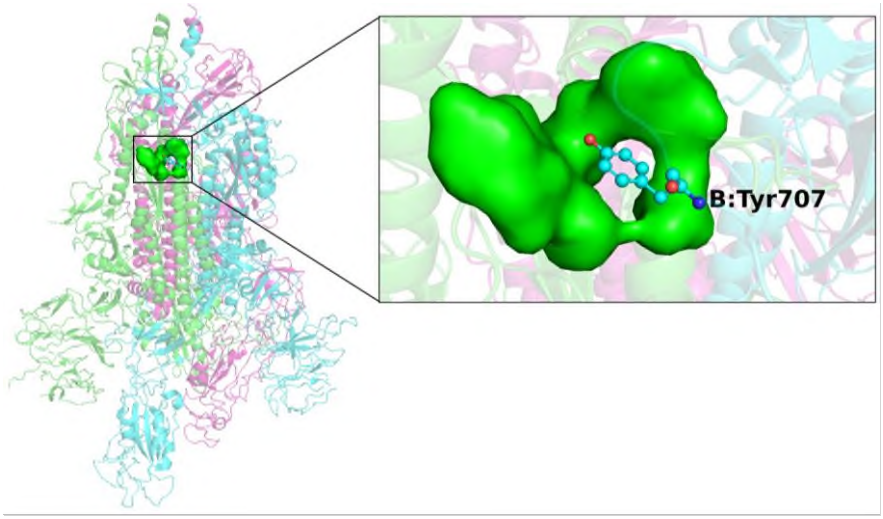

Figure 4. Glycoprotein pocket/Tyr707 potential interaction.

In **Figure 4**, potential interaction between glycoprotein pocket involved in stabilization of the SP

quaternary assembly (in green surface) and the almost invariant Tyr707 (cyan ball and sticks). Trimer SP is represented as green (chain A), cyan (chain B), and magenta (chain C) transparency cartoon. The enlargement is reported to clarify the potential interaction between the SP pocket and Tyr707 of the protomer B.

We have also monitored the variability at PPI position 856, a proposed locus of mutation typical of the Omicron variants[14]. There, the replacement of Asn with Lys yields an additional hydrogen bond between Lys856 NH3 of one protomer and the Thr572 OH of the protomer nearby.

The overall pattern of SP sequence mutations reveals that the Asp796Tyr represents a turning point in the SARS-CoV-2 evolution of the Omicron variants. Indeed, in the total 1,382,314 SP sequences of our refined dataset, Tyr796 is present 838,357 times and Asp796 is rather constantly present with 543,342 occurrences. Hence, we performed a separate count of SP sequences containing Tyr796, henceforth considered Omicron sequences, from all the other non-Omicron ones. Mutations in Omicron and non-Omicron sequences, probed in the above-mentioned sequence positions, are listed in **Table 1**, where all the amino acid replacements can be ascribed to single errors of the viral RNA-dependent RNA polymerase. Only in 10 cases, indeed, the reported Asp796Phe and Asp796Cys mutations of non-Omicron sequences require more than one nucleotide change. Consistently with the mutational profiles of Omicron variants[14], the PPI Asn856Lys mutation is present in 36% of the Omicron sequences, while it is only sporadic in the non-Omicron ones.

Table 1. With bold and normal characters are listed respectively the data obtained from sequences of Omicron and non-Omicron variants.

707		796		797		798	
Tyr	**785496**	**Tyr**	**838357**	**Phe**	**838291**	**Gly**	**837900**
Phe	**2**			**Cys**	**53**	**Asp**	**414**
				Leu	**10**	**Ser**	**29**
						Val	**5**
						Cys	**1**
Tyr	543976	Asp	543342	Phe	543918	Gly	543869
		His	230	Cys	6	Ser	44
		Asn	47	Leu	5	Ala	12
		Gly	33	Ile	1	Val	2
		Phe*	8			Asp	1
		Glu	3			Cys	1
		Cys*	2			Phe	1
883		896		897		898	
Thr	**837439**	**Ile**	**838305**	**Pro**	**838339**	**Phe**	**838336**
Ile	**817**	**Thr**	**22**			**Leu**	**1**
Ser	**9**	**Val**	**5**				
Met	**6**	**Leu**	**4**				
Ala	**1**	**Met**	**3**				
Thr	543927	Ile	544026	Pro	544049	Phe	544045
Ile	69	Val	9	Ser	1		
Ser	4	Leu	7				
		Thr	5				

Once identified the environment where Tyr707 residue inserts its aromatic side chain into the other contacting protomer, we used this information to run a docking simulation for selecting ligands that could interfere with SP trimerization. By using Autodock/VinaXB on the content of DrugBank 5.1.10[11], we selected the best 10 compounds based on their affinity on the target, with an affinity from −7.5 kcal/mol to −8.5 kcal/mol (**Table 2**), identifying paritaprevir as the best candidate for further experimental investigations. In **Table S4**, we have reported the molecular structures of these predicted ligands.

Figure 5 reports the sterical features of the interaction between paritaprevir and the above-described pocket. It is interesting to note that paritaprevir has been already proposed as an anti-SARS-CoV-2

experimental drug targeting the viral RNA-dependent RNA polymerase[15].

Table 2. Virtual screening results. Best 10 compounds based on their affinity against the target.

Best candidate-ligand ranking	Compound	ΔG (Kcal/mol)
1	Paritaprevir	−8.5
2	O-desmethyl midostaurin	−8
3	CGP52421	−8
4	Midostaurin	−7.9
5	Dutasteride	−7.9
6	Bendroflumethiazide	−7.7
7	Dihydroergotamine	−7.6
8	Ponatinib	−7.5
9	Lumacaftor	−7.5
10	Clofazimine	−7.5

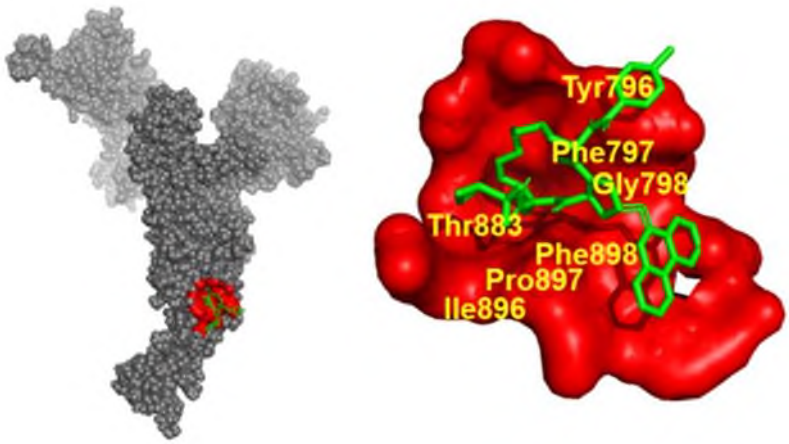

Figure 5. Surface representation of protomer A.

In **Figure 5**, surface representation of protomer A from the 7WZ1 PDB structure showing in red the druggable pocket predicted by a docking simulation. Structural details of the binding of paritaprevir, shown in green sticks, are given.

4. Discussion

The limiting step for betacoronavirus proliferation is the effectiveness of the SP quaternary assembly[5], which is driven by characteristic protomer-protomer interaction patterns. Hence, by using PDBePISA, we have explored PPIs of 6VSB and 7WZ1 PDB entries for a detailed investigation of their structural features and the evolution from non-Omicron to Omicron variants. In this way, we found almost identical major contact patches that are both characterized by low sequence variability, see **Figure 2**. As shown in **Table S1**, this SP moiety contains the pocket that we have already proposed as a possible binding site for small molecules with disassembling properties[4] likely caused by the presence of aromatic rings in their chemical structures (**Table S4**).

For monitoring the mutational pattern in conditions of low amino acid variability such as the one observed in PPIs, we focused our analysis on the residues contributing to the formation of the latter pocket, i.e., 796–798, 883 and 896–898, which always hosts the Tyr707 aromatic side chain of another protomer. Thus, the Asp796Tyr mutation appears as a landmark for distinguishing Omicron from non-Omicron variants. As shown

in **Figure 3**, the contact between protomers A and B is strongly constrained by the simultaneous involvement of Tyr707 in a hydrogen bond of its phenyl OH with Phe898 amide and a T-shaped interaction of its aromatic side chain with the one of Tyr796. It follows that the trimeric SP assembly is largely favored by the close-distance presence of two strong non-covalent interactions. Hence, the Asn856Lys mutation, very often found in Omicron variants, yields an additional hydrogen bond between different protomers, as above described, reinforcing the interaction among SP protomers. The im-proved protomer-protomer locking system that Omicron has evolved to stabilize the SP quaternary assembly through the Asp796Tyr mutation, is among the keys for its rapidly achieved predominance on all the other variants.

The SARS-CoV-2 strategy for neutralizing external assaults to its proliferation, arising from neutralizing antibodies and/or antiviral drugs, by introducing aromatic moieties in the 796-sequence position has been also observed for the viral evolution in chronic viral infections[16–18]. In that case, consistent with the data reported in **Table 1** his replaced the Asp796 residue in non-Omicron variant SP sequences for establishing an additional protomer-protomer interaction between the aromatic side chains of His796 and Tyr707.

The close correlation between the SP structural stability of Omicron variants and the predominance of their diffusion, suggests how SARS-CoV-2 has its achilles' heel just in the SP quaternary assembly. This finding indicates that the design of small molecules that can interfere with SP trimerization represents a practicable alternative approach to obtaining broad-spectrum antiviral drugs.

5. Conclusions

Interfering with the quaternary assembly of the SARS-CoV-2 spike glycoprotein as an alternative antiviral strategy was initially inferred[4] by the modes of protein trimerization that were observed for the similar betacoronavirus TGEV[5]. The large presence in vitro of unassembled protomers, indeed, indicated that the trimerization of the viral protein requires suitable conditions that can be chemically altered. The fact that antibodies can also disassemble the SARS-CoV-2 spike glycoprotein trimer[3] proved the remaining partial instability of Omicron SP, a very remarkable feature that can be exploited for antiviral therapies.

We have explored the distribution of druggable pockets just at the A-B interface, finding one possible candidate for binding a small molecule with antiviral activity. Interference with the SP quaternary assembly that SARS-CoV-2 needs to infect human cells has been predicted here by the docking simulation of paritaprevir with SP. The intrinsic instability of the SP trimer and the latter computational result should be the stimulus for experimental investigations to monitor the extent of the viral quaternary assembly during the viral protein expression in the presence of different compounds from suitable molecular libraries.

Interfering with the trimerization of SARS-CoV-2 spike glycoprotein seems to be a very promising strategy to block the infection not only of SARS-CoV-2 but also of other betacoronaviruses that share similarities in their protomer-protomer interactions. More in general, structure-driven exploration of the multimeric interfaces for viral proteins involved in the interaction with the host cell receptor to predict the existence of drug-gable sites, could represent a very suitable shortcut for the development of a new generation of antiviral drugs.

Supplementary materials

All other information not contained in this manuscript are available in supplementary materials.

Author contributions

Conceptualization, NN; methodology, NN, AT, AV and PB; validation, NN, AT and AV; formal analysis, NN, AT, AV, PB and FM; data curation, NN, AT, AV and PB; writing—original draft preparation, NN;

writing—review and editing, NN, AT and MB; visualization, NN, AT, AS, OS and MB; supervision, NN. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

Thanks are due to Riccio V for technical assistance and to Neri D and Temussi PA for helpful discussions.

Conflict of interest

The authors declare no conflict of interest.

References

1. Miljanovic D, Cirkovic A, Lazarevic I, et al. Clinical efficacy of anti-SARS-CoV-2 monoclonal antibodies in preventing hospitalisation and mortality among patients infected with Omicron variants: A systematic review and meta-analysis. *Reviews in Medical Virology* 2023; 33(4): e2439. doi: 10.1002/rmv.2439
2. Zhou Y, Wang H, Yang L, Wang Q. Progress on COVID-19 chemotherapeutics discovery and novel technology. *Molecules* 2022; 27(23): 8257. doi: 10.3390/molecules27238257
3. Zhan W, Tian X, Zhang X, et al. Structural study of SARS-CoV-2 antibodies identifies a broad-spectrum antibody that neutralizes the omicron variant by disassembling the spike trimer. *Journal of Virology* 2022; 96(16): e0048022. doi: 10.1128/jvi.00480-22
4. Bongini P, Trezza A, Bianchini M, et al. A possible strategy to fight COVID-19: Interfering with spike glycoprotein trimerization. *Biochemical and Biophysical Research Communications* 2020; 528(1): 35–38. doi: 10.1016/j.bbrc.2020.04.007
5. Delmas B, Laude H. Assembly of coronavirus spike protein into trimers and its role in epitope expression. *Journal of Virology* 1990; 64(11): 5367–5375. doi: 10.1128/JVI.64.11.5367-5375
6. Sievers F, Wilm A, Dineen D, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* 2011; 7: 539. doi: 10.1038/msb.2011.75
7. Bakan A, Dutta A, Mao W, et al. Evol and ProDy for bridging protein sequence evolution and structural dynamics. *Bioinformatics* 2014; 30(18): 2681–2683. doi: 10.1093/bioinformatics/btu336
8. Wrapp D, Wang N, Corbett KS, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* 2020; 367(6483): 1260–1263. doi: 10.1126/science.abb2507
9. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology* 2007; 372(3): 774–797. doi: 10.1016/j.jmb.2007.05.022
10. Koebel MR, Schmadeke G, Posner RG, Sirimulla S. AutoDock VinaXB: Implementation of XBSF, new empirical halogen bond scoring function, into AutoDock Vina. *Journal of Cheminformatics* 2016; 8(1): 27. doi: 10.1186/s13321-016-0139-1
11. Wishart DS, Feunang YD, Guo AC, et al. DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Research* 2018; 46(D1): D1074–D1082. doi: 10.1093/nar/gkx1037
12. Magazine N, Zhang T, Wu Y, et al. Mutations and evolution of the SARS-CoV-2 spike protein. *Viruses* 2022; 14(3): 640. doi: 10.3390/v14030640
13. Mistry P, Barmania F, Mellet J, et al. SARS-CoV-2 variants, vaccines, and host immunity. *Frontiers in Immunology* 2022; 12: 809244. doi: 10.3389/fimmu.2021.809244
14. Alkhatib M, Salpini R, Carioti L, et al. Update on SARS-CoV-2 omicron variant of concern and its peculiar mutational profile. *Microbiology Spectrum* 2022; 10(2): e0273221. doi: 10.1128/spectrum.02732-21
15. Gurung AB, Ali MA, Lee J, et al. The potential of Paritaprevir and Emetine as inhibitors of SARS-CoV-2 RdRp. *Saudi Journal of Biological Sciences* 2021; 28(2): 1426–1432. doi: 10.1016/j.sjbs.2020.11.078
16. Kemp SA, Collier DA, Datir RP, et al. Author correction: SARS-CoV-2 evolution during treatment of chronic infection. *Nature* 2022; 608(7922): E23. doi: 10.1038/s41586-022-05104-2
17. Araf Y, Akter F, Tang YD, et al. Omicron variant of SARS-CoV-2: Genomics, transmissibility, and responses to current COVID-19 vaccines. *Journal of Medical Virology* 2022; 94(5): 1825–1832. doi: 10.1002/jmv.27588
18. Farheen S, Araf Y, Tang YD, Zheng C. The Deltacron conundrum: Its origin and potential health risks. *Journal of Medical Virology* 2022; 94(11): 5096–5102. doi: 10.1002/jmv.27990