

Article

Comparative Analysis and Classification of SARS-CoV-2 Spike Protein Structures in PDB

Memoona Aslam ¹, M. Saqib Nawaz ², Philippe Fournier-Viger ² and Wenjin Li ^{1,*}

¹ Institute for Advanced Study, Shenzhen University, Shenzhen 518060, China

² College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China

* Correspondence: liwenjin@szu.edu.cn

Abstract: The Spike (S) protein of the SARS-CoV-2 virus that causes the COVID-19 disease is considered the most important target for vaccine, drug and therapeutic research as it attaches and binds to the ACE2 receptor of the host cells and allows the entry of this virus. Analysis and classification of newly determined S protein structures for SARS-CoV-2 are critical to properly understand their functional, evolutionary and architectural relatedness to already known protein structures. In this paper, first, the comparative analysis of SARS-CoV-2 S protein structures is performed. Through comparative analysis, the S protein structures in the PDB (protein data bank) database are compared and analyzed not only with each other but with the structures of other viruses for various parameters. Second, the S protein structures in PDB are classified into different variants, and the associated published literature is studied to investigate what kind of therapeutics (antibodies, T-cell receptors and small molecules) are used on the structures. This is the first study that classifies the S protein structures of the SARS-CoV-2 in PDB into various variants, and the obtained comparative analysis results could be beneficial to the research community, in general, and to crystallographers and health workers, in particular.

Keywords: SARS-CoV-2; DALI; PDB; variants; spike; antibodies/nanobodies



Citation: Aslam, M.; Nawaz, M.S.; Fournier-Viger, P.; Li, W. Comparative Analysis and Classification of SARS-CoV-2 Spike Protein Structures in PDB. *COVID* **2023**, *3*, 452–471. <https://doi.org/10.3390/covid3040034>

Academic Editor: Fernando Barroso

Received: 31 January 2023

Revised: 12 March 2023

Accepted: 27 March 2023

Published: 29 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2) virus [1] that causes the COVID-19 disease was initially identified in late 2019. This virus belongs to the family of single-strand RNA viruses. Their shape is spherical to pleomorphic and their length is between 80 and 160 nm. The genome of this virus contains 6 to 12 open reading frames (ORFs). The rest of the genome encodes four structural proteins (Envelope (E), Membrane (M), Nucleocapsid (N) and Spike (S)) and some non-structural proteins (Figure 1). ORF1b, the main reading frame, produces pp1a and pp1ab polypeptides that are cleaved by the PL^{pro} (papain-like protease) and 3CL^{pro} (3C-like cysteine protease). 3CL^{pro} is also called M^{pro} (main protease) [2–4]. Spike protein is a trimeric class I glycoprotein located on the surface of the viral envelope and contains the S1 and S2 subunits. S1 contains binding domains, such as the receptor-binding domain (RBD) and the N-terminal domain (NTD), and S2 contains a fusion peptide and two heptad repeats (HR) 1 and 2 domains that contribute to the virus's fusion [5].

SARS-CoV-2 can enter the cell membrane of the host when the S protein interacts with a receptor present on the host surface called “angiotensin-converting enzyme 2” (ACE2) [6–10]. Thus, the role of the S protein is fundamental in the pathogenesis, transmission and virulence of SARS-CoV-2. Compared to SARS-CoV, the S protein of SARS-CoV-2 binds to ACE2 with a higher affinity [11]. After binding, the entry of SARS-CoV-2 depends on the S protein priming process. This process is carried out by TMPRSS2 [8] (Figure 1), which is a type 2 serine protease present on the host cell surface. Blocking or preventing the binding of S proteins with ACE2 receptors is considered the first and most important

approach to block cell entry and stop the COVID-19 disease. The S protein is therefore considered the most important target for the COVID-19 vaccine, drug developments and antibody blocking.

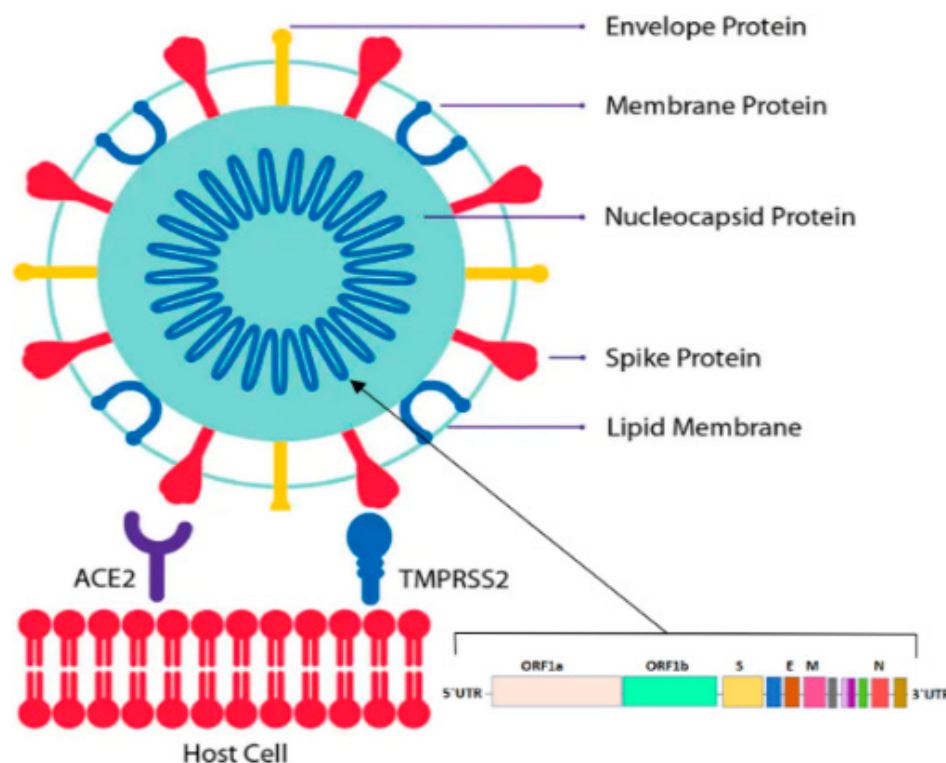


Figure 1. SARS-CoV-2 structure, binding and entry in the host. ACE2 and TMPRSS2 are the main determinants of entry. S protein uses ACE2 for binding and TMPRSS2 for its activation.

The S protein interaction with the ACE2 receptor can be found by using structural biology techniques, such as Cryo-EM (electron microscopy) and X-Ray crystallography. These techniques elucidate the 3D (three-dimensional) structures of proteins and their conformational changes. X-Ray crystallography provides high-resolution structural information about macromolecules. It is relative, cheap, simple and can produce a high atomic resolution. The Cryo-EM technique provides macromolecular structures at near-atomic resolution. Cryo-EM has certain advantages over X-Ray crystallography, such as avoiding the crystallization bottleneck and requiring lower protein amounts. Since the emergence of the SARS-CoV-2 virus, its protein structures are deposited at a fast speed in online databases. Some famous databases are protein data bank (PDB) [12] and Electron Microscopy Data Bank (EMDB) [13]. Until 5 November 2022, more than 195,000 protein structures are deposited in PDB. These online databases allow researchers to analyze any particular viral structure, its function and its molecular basis. These structures are used for (1) validating drug targets, (2) assessing the ability of the target drug, (3) characterizing the ligands, small molecules and other tool compounds that bind to the drug targets, (4) guiding the medicinal chemistry optimization of binding affinity and (5) overcoming the challenges in pre-clinical drug development.

Almost all proteins in general share structural similarities with each other due to the principles of chemistry and physics. These principles put a limitation on the number of ways a polypeptide chain in structures can be folded into a compact globule. Comparing and classifying new 3D protein structures can reveal biologically hidden and interesting similarities. These insights can contribute to a better understanding of not only the evolution but the structural architecture and protein's function, particularly for those that share little with well-characterized proteins in terms of sequence identity. These insights are important for drug discovery, identifying new folds for protein and for the proteome's

phylogenetic analysis. Considering the large number of S protein structures currently available in the PDB, we focus on comparative analysis and classification of the S protein structures of SARS-CoV-2. More specifically, the S protein structures of SARS-CoV-2 in PDB are analyzed and compared not only with each other but also with the protein structures of other viruses through structural alignment, similarity matrix, structures imposition and phylogeny. Moreover, the S protein structures of SARS-CoV-2 in PDB are classified into different variants. The associated published literature, in the last two and a half years, that offer structures for the S protein of SARS-CoV-2 is reviewed to study the therapeutics (antibodies, T-cell receptors and small molecules) used on the structures. We believe that obtained results by comparative analysis and structures classification could be beneficial to the research community, in general, and to crystallographers, biochemists and health workers, in particular.

2. Methods

DALI (Distance matrix alignment) server [14,15] is used for the comparison and analysis of SARS-CoV-2 S protein structures. DALI is famous for protein structures comparison and can be used to compare new structures with existing structures in PDB.

In order to maximize the DALI score, a set of one-to-one correspondences between the protein (sub)structures A and B are optimized:

$$DALI_{AB} = \sum_{i=1}^{LALI} \sum_{j=1}^{LALI} \begin{cases} \left(\theta - \frac{2|d_{ij}^A - d_{ij}^B|}{d_{ij}^A + d_{ij}^B} \right) e^{-\left(\frac{d_{ij}^A + d_{ij}^B}{2D} \right)^2}, & \text{if } i \neq j \\ \theta, & \text{if } i = j \end{cases} \quad (1)$$

where $LALI$ represents the number of residue pairs that are aligned, $D = 20 \text{ \AA}$, $\theta = 0.2$ and d_{ij}^A, d_{ij}^B represent the intra-molecular C α –C α distances in structures A and B, respectively. The expected Dali score (Equation (1)), for random pairwise comparison, increases with the number of amino acid residues in the proteins that are being compared.

Z-Score: The statistical significance of a pairwise comparison score is described by the DALI Z-score (Z_{AB}):

$$Z_{AB} = \frac{DALI_{AB} - m(L)}{\sigma(L)} \quad (2)$$

where $L = \sqrt{L_A L_B}$ is the average length of A and B. A large set of random pairs of structures is used to empirically derive the relation among m (mean score), L and σ (standard deviation). A polynomial fit gave the following approximation:

$$m(L) = \begin{cases} 7.95 + 0.71L - 0.000259L^2 - 0.00000192L^3, & \text{if } L \leq 400 \\ m(400) + L - 400, & \text{if } L > 400 \end{cases} \quad (3)$$

The empirical estimate for σ was $\sigma(L) = 0.5 \times m(L)$. Now, for every possible pair of domains, the Z-score is computed and the highest value is identified as the Z-score for a protein pair. Thus, Z-score reported by DALI is an optimal score for similarity that is defined as the sum of equivalent residue-wise C α –C α distances between two proteins. In other words, the 3D coordinates of proteins are used in the calculation of the Z-score to measure the degree of similarity among proteins. For two proteins, the large Z-score indicates more similarity that corresponds to the optimized set of residue equivalence discovered through the Monte Carlo optimization permutation of equivalent structural patterns. A Z-score < 2 is considered as spurious similarity and can be ignored.

Root Mean Square Deviation (RMSD): RMSD is a numerical measure of the difference between two structures and shows how well they align. For two protein structures A and B, RMSD can be calculated as follows:

$$RMSD_{AB} = \sqrt{\frac{1}{N} \sum_{i=1}^N d(a_i, b_i)^2} \quad (4)$$

where N is the number of C α atoms, and $d(a_i, b_i)$ distance is the difference in position of atom i in each molecule.

DALI supports three types of database searches: (1) PDB, (2) PDB25 and (3) AlphaFold (AF). DALI uses two strategies, pairwise and all against all, for structures comparison. The S protein structure of SARS-CoV-2 with PID 6VSB [11] (deposited to PDB on 10 February 2020) is used as the query structure. 6VSB is a 3D, 3.5 Å Cryo-EM-based SARS-CoV-2 S structure in the prefusion conformation. The main reason to select 6VSB as a query structure is that it is one of the earliest S protein structures deposited in PDB.

The COVID-19 coronavirus resource in the PDB website (rcsb.org) offers a quick link “www.rcsb.org/news/feature/5e74d55d2d410731e9944f52 (accessed on 19 November 2022)” to the submitted (1) Main proteases, (2) Spike proteins and RBDs and (3) Papain-like proteinases of SARS-CoV-2. More than 1000 and 500 S protein and M^{Pro} structures, respectively, of SARS-CoV-2 are present in the PDB. Each structure in PDB is associated with a paper link that deposited that structure to the PDB. For the S protein structures, the associated literature was collected and analyzed/read for the variants types and to see what kind of antiviral drugs were used on the S protein structures. The S protein structures and the published literature from PDB were collected in the time period 15 March 2022 to 20 September 2022.

The published literature related to the S protein structures of SARS-CoV-2 was also retrieved from the Web of Science (WoS) Core collection database. The keywords used for the articles collection were “Cryo-EM”, “SARS-CoV-2”, “COVID-19” and “Spike”. In total, WoS found 158 articles. After excluding the review articles, the total number of articles was 143. The temporal distribution of collected articles is listed in Table 1. We found that half of the S protein structures were deposited to PDB in 2021, while the other half in 2020 and 2022. The top five most cited papers [6,11,16–18] determined the S protein structures, respectively.

Table 1. Year-wise distribution of SARS-CoV-2 S protein structures publications.

Year	Documents (%)
2020	26 (18.2)
2021	75 (52.4)
2022	42 (29.4)

3. Results

The query structure 6VSB is compared with other protein structures in the PDB using DALI with four search subsets: PDB25, PDB50, PDB90 and PDB (or PDBall). DALI uses the subsets because the PDB database is highly redundant as the 3D structures of some proteins and their mutants are determined in various conditions. PDBall contains all structures in the PDB. Whereas PDB25, PDB50 and PDB90 are non-redundant subsets of PDB structures. In the PDB25 subset, those structures in PDB that are less than 25% identical in sequence are found. PDB50 and PDB90 search return PDB structures that are less than 50% and 90% identical in sequence. The three different sequence identity levels are derived from PDB using CD-Hit [19]. Thus, for 6VSB, DALI returned 46, 91, 178 and 2115 similar PDB structures in PDB25, PDB50, PDB90 and PDBall, respectively. Table 2 lists the most similar structures to 6VSB on the basis of the Z-score. The percent identity of aligned amino acids is also listed along with the Z-score. We also used BLAST [20] to see the first 100 similar structures to the query structure. The S protein structure of SARS-CoV-2 with PDB ID 7LQV has 100% identity while 6ZPO has 95% identity. Interestingly, all the 100 similar structures returned by BLAST belonged to the S protein structures of SARS-CoV-2.

Table 2. Similar SARS-CoV-2 S protein structures. Structures highlighted in purple are S protein structures in complex with antibodies. Bold PDB IDs (PIDs) represent those S protein structures that are present in multiple search strategies.

Parameters	PDB25	PDB50	PDB90
SARS-CoV-2 S protein	1 (7E9T (10.8, 78%))	3 (7N9E (42.9, 96%), 7QO9 (26.2, 73%), 7E9T (10.8, 78%))	48 (7AD1 (48.8, 97%), 7N9C (48.1, 99%), 7RA8 (48.1, 97%), 7SN3 (47.8, 97%), 7KE8 (47.6, 98%), 7LAA (47.4, 99%), 7TPH (47.3, 96%), 7Q9P (47.2, 96%), 7NS6 (47, 98%), 7C2L (46.6, 98%), 6XKL (46.6, 99%), 7LCN (46.1, 99%), 7TM0 (46, 97%), 7VX7 (45.8, 98%), 7THK (45.3, 98%), 7NDB (45.3, 99%), 7MTC (44.8, 100%), 7MTD (44.4, 99%), 7MTE (44.3, 100%), 7SBT (44.1, 94%), 7N0G (43.9, 98%), 7FCE (43.2, 99%), 7TOU (43.1, 99%), 7VXF (43, 99%), 7CN9 (42.9, 96%), 7N9E (42.8, 96%), 7P7A (42.7, 98%), 7WVO (42.5, 97%), 7TLA (42.5, 99%), 7R15 (42.2, 96%), 7FCD (41.3, 98%), 7WP9 (41.3, 97%), 7SG4 (41.3, 79%), 7TGE (41.2, 97%), 5X5B (40.1, 73%), 7WEV (39.6, 99%), 7N9T (39.5, 96%), 7L2D (38.3, 99%), 6ZGH (37.5, 99%), 7T9J (35.1, 98%), 7SBR (35.1, 98%), 7KRQ (33.9, 99%), 7TL1 (33.6, 96%), 6ZOZ (32.9, 99%), 7QO9 (26.2 73%), 7E7X (22.1, 88%), 7S0E (20.6, 92%), 7E9T (10.8, 78%))
S protein of other viruses	8	10	22
Other	37	78	108
Total	46	91	178

In the exhaustive PDB25 search results, only one structure (PID: 7E9T, Z-score of 10.8) belongs to the S protein of SARS-CoV-2 and eight structures are S protein structures of other viruses (listed in Table 3). 7E9T is the structure of the SARS-CoV-2 S protein in the post-fusion state that provided insights into the design of a viral entry inhibitor [21]. The remaining 37 are other structures that belong to other species and proteins/enzymes. Similarly, with PDB50/PDB90 search, 91/178 similar structures are found, out of which 10/22 structures belong to the S protein of the SARS-CoV-2 and 78/108 belong to the other. The majority of the structures in Tables 2 and 3 belong to the RBD region of S, with some belonging to the NTD region. The structures 7N9E, 7N9T and 7N9C belong to a group of nanobodies-bound RBDs of S protein structures that targets epitopes of three different classes [22]. Similarly, 7RA8 is the structure of SARS-CoV-2 S RBD antigenic site II in complex with a human monoclonal antibody (MAb) S2X259 [23]. In Table 3, the structures 6NB3, 6NB4 and 6NB6, 6NB7 are the S protein structures for MERS and SARS-CoV, respectively, with neutralizing antibodies (NAbs) LCA60 and S230 in different states, respectively [24]. Whereas 7AKJ is the SARS-CoV S protein structure with the Nab 47D11.

In PDB25 and PDB50, the percentage of the similar protein structures that are same to the S protein of SARS-CoV-2 is approximately 2.17% and 3.29%, respectively. Whereas in PDB90, the percentage is 27.5%. In the PDBall search, 2115 similar structures are found. Note that similar structures found in PDB25 are present in PDB50 and structures found in PDB50 are present in PDB90 and so on. From Table 3, we can see that the S protein structures of SARS-CoV-2 are similar to other coronaviruses (SARS-CoV, bat coronavirus and MERS-CoV, particularly) and to the viruses from other animals, such as pangolin, bat and mouse. Indeed, earlier genome sequencing also shows that SARS-CoV-2 is more similar (79%) to its predecessor (SARS-CoV-1), 50% to MERS-CoV and 96% to RaTG13

(bat coronavirus) [1,7]. In PDB50, the S structure of SARS-CoV-2 with PID 7QO9 is for the Omicron variant that was deposited to PDB within 1 month after the first emergence of the Omicron variant on 24 November 2021. The calculated RMSD value of 7QO9 against the wild-type S (6VSB) structure in [25] was 2.4 which includes particular modifications to the surfaces and local conformations involved in the antibody recognition. DALI's RMSD value for these structures is 2.5. In PDB90, the structures that belong to various variants are Alpha (7KE8, 7R15, 7KRQ), Beta (7RA8, 7Q9P, 7VXF, 7WEV, 7N9T, 7E7X), Gamma (7SBT), Delta (7TPH, 7TOU), Omicron (7TM0, 7THK, 7WVO, 7WP9, 7TGE, 7T9J, 7TL1, 7QO9) and Kappa (7SBR). The similar protein structures given in Table 3 for PDBall are those that are present in the first 1000 similar structures. For three search strategies (PDB25, PD50 and PDB90), the percent identity computed by DALI is different from BLAST. Overall, BLAST returned high similarity for structures compared to DALI.

Table 3. Similar S protein structures of other viruses. Structures in complex with antibodies are highlighted in purple.

Search Strategy	PID	Z-Score, % Identity	Blast % Identity	Spike Protein
PDB25	7CN8	44.8, 94	92.25	Pangolin (PCoV_GX)
	6CS2	35.7, 78	76.51	SARS-CoV
	6NZK	33.5, 31	35.82	Coronavirus OC43
	6NB4	32.7, 32	34.73	MERS-CoV
	5I08	27.4, 31	37.86	Prefusion human coronavirus HKU1
	6CV0	26.2, 25	35.69	Avian infectious bronchitis coronavirus
	6M15	25.8, 26	32.69	Rhinolophus bat coronavirus HKU2
	6JX7	17.8, 20	32.86	Feline infectious peritonitis (FIP) virus
PDB50	6NB3	34.2, 30	35.82	MERS-CoV
	6VV5	31.0, 28	33.89	Porcine epidemic diarrhea virus (PEDV)
	6B7N	30.5, 27	31.01	Prefusion porcine delta coronavirus
	5SZS	30.1, 28	29.86	Human coronavirus NL63 spike trimer
PDB90	7C2L	46.6, 98	99.75	Duvenhage virus phosphoprotein
	6NB7	40.7, 77	75.63	SARS-CoV
	7BBH	39.3, 92	89.27	Smuggled Guangdong pangolin
	6NB6	39.2, 33	75.63	SARS-CoV
	7AKJ	34.5, 79	73.15	SARS-CoV
	5WRG	32.6, 77	74.75	SARS-CoV
	6VSJ	30.6, 29	36.14	Mouse coronavirus
	6U7H	30.3, 27	55	Human coronavirus 229E
	6U7K	12.5, 12	59	PEDV
	6M3W	10.9, 71	89.82	SARS-CoV
PDBall	7CN4(C)	44, 95	94.48	Bat coronavirus RaTG13
	6CRW(C)	43.7, 78	76.61	Escherichia virus T4
	6ZGF(B)	43, 97	95.66	Bat coronavirus RaTG13
	6CRX(B)	43, 76	76.61	Escherichia virus T4
	6CS1(C)	42.4, 78	76.61	Escherichia virus T4
	6CRZ(B)	42.3, 76	76.61	Escherichia virus T4
	6CS0(B)	42.3, 76	76.61	Escherichia virus T4

We perform some more analysis by using the pairwise comparison in DALI to examine the similarity of the S protein structures of SARS-CoV-2 with the protein structures of other viruses. The 10 most similar structures (7CN8, 6CS2, 6NZK, 6NB4, 6JX7, 5I08, 6CV0, 6M15, 7E9T and 3AFK) alignment with the query structure in DALI is performed. More than 970 amino acids (AA) (Figure 2) and secondary structure elements (SSEs) (Figure 3) in 11 structures were aligned. Among 10 structures, only 7E9T is the S protein structure of SARS-CoV-2. Eight structures (listed in Table 2) belong to the S protein structure of other viruses, such as pangolin (7CN8), SARS-CoV (6CS2), human coronavirus OC43 (6NZK), MERS-CoV (6NB4), human coronavirus HKU1 (5I08), avian bronchitis coronavirus (6CV0),

Rhinolophus bat coronavirus HKU2 (6M15) and Feline infectious peritonitis (FIP) virus (6JX7). 3AFK is the crystal structure of *Agrocybe aegerita* lectin AAL. The structures at the top show the similar structures that share many frequent AAs at various positions. The most frequent AAs are Leucine (L), followed by Valine (V), Threonine (T), Alanine (A) and Serine (S), respectively. The structures at the bottom are not that much similar as they share less same AAs. Pangolin-CoV S structure (7CN8) is more similar to the query structure followed by 6CS2 and 6NZK.

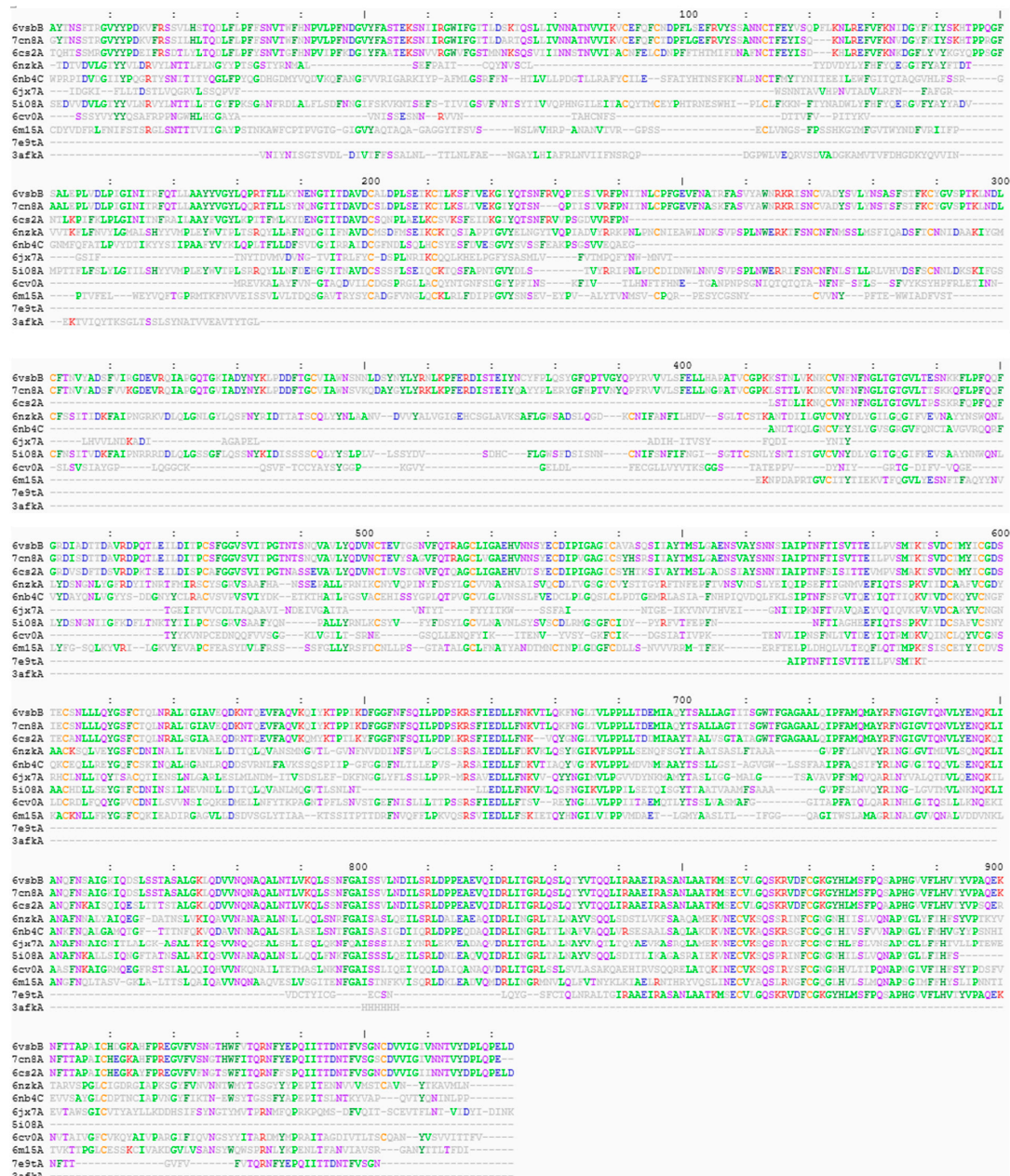


Figure 2. Structural alignment of AAs in 10 protein structures against the query structure. The most frequent AA type in each column is colored. For example, purple residues (S, T, N and Q) are uncharged, brown residues (R and K) are polar positively charged, blue residues (D and E) are polar negatively charged, dark green residues (F, Y, W and H) are non-polar except histidine (positive) but they are aromatic, light green residues (A, V, I and L) are non-polar residues and yellow residue (C) is a special case. The uppercase letters that are bold show conserved structurally equivalent positions with the query structure and non-color regions are variable regions.

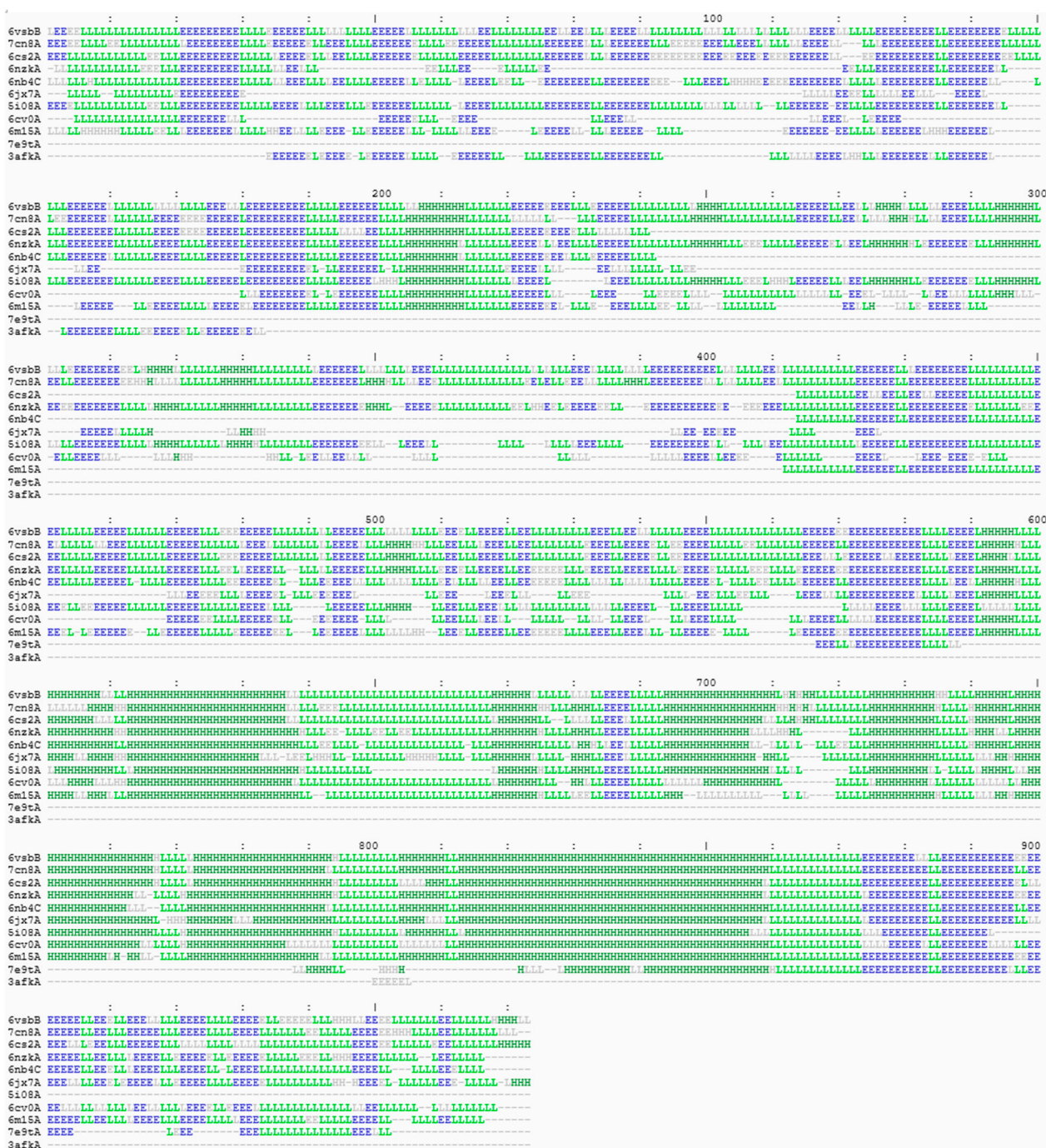


Figure 3. Alignment of SSE (assigned by the DSSP) in 10 protein structures against the query structure. Alpha helices (H) and coil (L) are represented with dark and light green color, while beta sheets (E) with blue color.

In the secondary structure, the two most common elements of a protein are α -helices and β -sheets (consist of β -strands). Traditionally, the secondary structures of a protein are characterized with three general states: (1) coil (C), (2) helix (H) and (3) strand (E). The Dictionary of Secondary Structure of Proteins (DSSP) [26] offers a finer classification of the secondary structures with eight states: (1) α -helix (H), (2) 3_{10} helix (G), (3) π -helix (I), (4) β -strand (E), (5) bridge (B), (6) turn (T), (7) bend (S) and (8) others (C). From Figure 3, the most frequent SSE are L, followed by H and E. Figure 4 also confirms that considered structure for superimposition with the query structure has more L, followed by H and E.

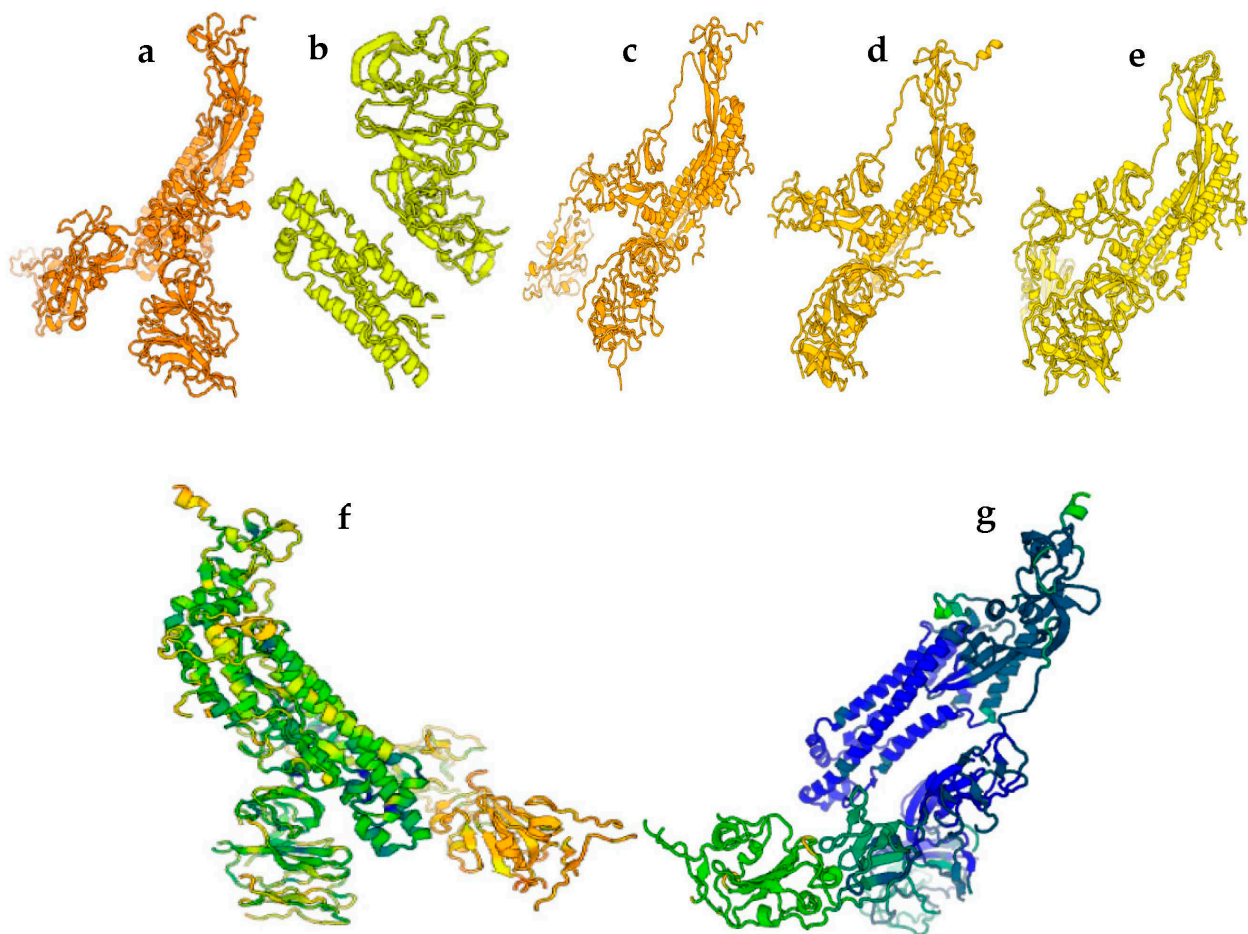


Figure 4. Ribbons structures of (a) 7CN8, (b) 7QO7, (c) 6CS2, (d) 6NZK, (e) 6JX7 and (f,g) sequence and structural conservations mapped onto the query structure. The color map for conservation mapping in DALI spans from blue (which shows the highest values) through green to red (which shows the lowest values). (f,g) Sequence and structural conservation of the five structures against the query structure, respectively. Blue color regions in (f,g) show high sequence and structural similarities in five structures. The green color in (f), which can be seen frequently, shows fewer similarities and the yellow color shows more dissimilarities. The green color in (g) represents strands of barrel that have reversed the direction in some structures. The A chains of all structures are used in comparison against the query structure.

Next, some protein structures from Tables 2 and 3 are compared against the query structure in the 3D structure view. The superimposition view provides sequence conservation and structure conservation. In the 3D superimposition view, the S protein structures for pangolin (Figure 4a), SARS-CoV-2 Omicron (Figure 4b), SARS-CoV (Figure 4c), human coronavirus OC43 (Figure 4d) and FIP virus (Figure 4e) are superimposed on the query structure (Figure 4f,g). DALI also computes traditional measures such as RMSD and LALI (the number of equivalent residues). In 3D superimposition, the RMSD value shows the average deviation in distance among the aligned C α atoms. RMSD value is 1.0 for those sequences that are half (50%) identical. Against the query structure, 6CS2 has the lowest RMSD value (1.9), followed by 7QO9 (2.5), 6NZK (3.4), 7CN8 (3.6) and 6JX7 (5.7). Whereas 7CN8 (6NZK and 6CS2) has 965 (876 and 796) aligned C- α traces (a reduced complexity representation to only show the α carbons of the AA), followed by 6JX7 (558) and 7QO9 (487). 6JX7 has the most equivalent AA residues (1245), followed by 6NZK (1175), 7CN8 (1125), 6CS2 (891) and 7QO9 (588). 7CN8 has more aligned AA (94%), followed by 6CS2 (78%), 7QO9 (73%), 6NZK (31%) and 6JX7 (29%). Thus, the FIP virus structure is most

dissimilar to the query structure. All the proteins in secondary structures have more coil (loop), followed by helix and strands.

The similarity matrix (Figure 5a) based on the Z-score of 6VSB with 61 other structures and dendrogram (Figure 5b) shows that 6VSB is closest to pangolin coronavirus structures (7CN8 and 7BBH) and occupies a dominant position in the fold cluster. In the literature, various studies suggested that bat and pangolin species are the most natural reservoir of SARS-CoV-2. Moreover, Zhang et al. [27] reported the structures for pangolin (PCoV_GX, PID: 7CN8) and bat (RaTG13, PID: 7CN4) S proteins. They found that in the overall structure, the S RBDs of PCoV_GX and RaTG13 are very similar to the RBD of SARS-CoV-2. The structural comparison of PCoV_GX, RaTG13 and SARS-CoV-2, their binding strength with ACE2 and their efficiency in facilitating pseudovirus cell entry suggest the cross-species transmission and evolution of SARS-CoV-2. Note that the bat RaTG13 (PID: 7CN4) structure is not included in the similarity matrix as DALI did not find this structure similarity with the query structure in the search subsets (PDB25, PDB50 and PDB90).

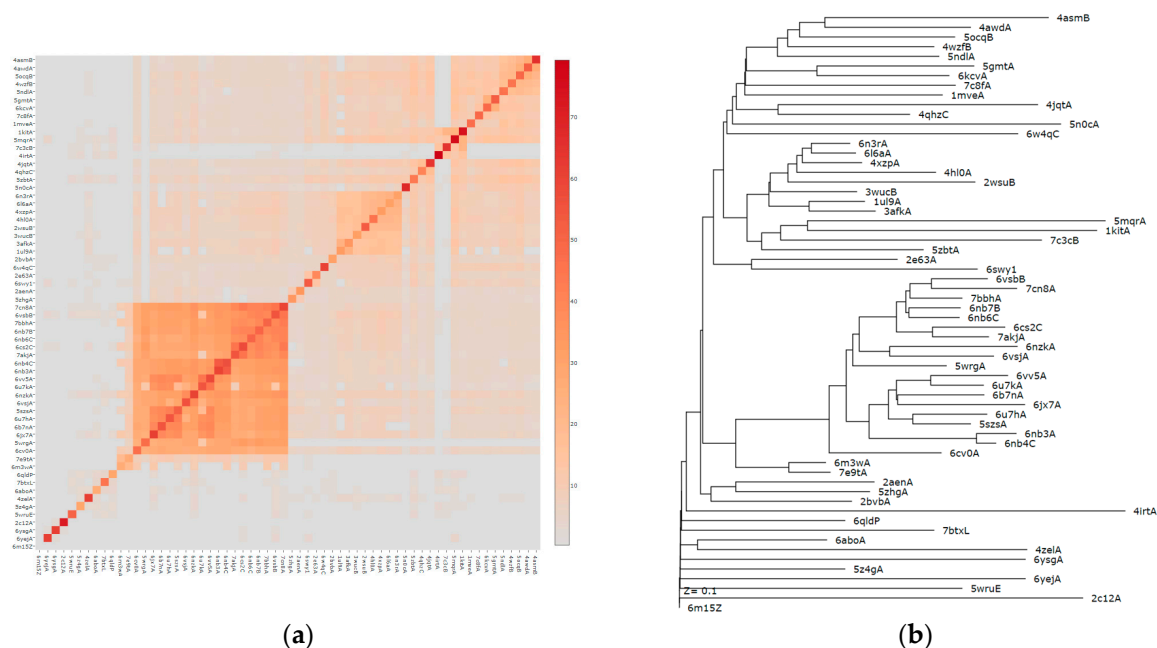


Figure 5. (a) Similarity matrix, based on Z-score. The query structure is closest to pangolin structures. (b) Structural similarity dendrogram. The presence of the query structure with the PCoV_GX S structure suggests that pangolin-CoV is the closest relative of SARS-CoV-2. The similarity matrix can be further divided into three main clusters. The 18 similar structures (PIDs: 7CN8, 7BBH, 6NB7, 6NB6, 6CS2, 7AKJ, 6NB4, 6NB3, 6VV5, 6U7K, 6NZK, 6VSJ, 5ZS, 6U7H, 6B7N, 6JX7, 5WRG and 6CV0) that form a cluster at the middle left are the S protein structures of other viruses (Table 3). Another small cluster, containing PIDs 6N3R, 6L6A, 4XZP, 4HLO, 2WSU, 3WUC, 3AFK and 1UL9, at the right middle mostly contains protein structures for Lectin and S-type Lectins (also known as Galectin) that are sugar-binding proteins of type *E. coli*. The third cluster (PIDs: 4ASM, 4AWD, 5OCQ, 4WZF, 5NDL, 5GMT, 6KCV, 7C8F, 1MVE, 1KIT and 5MQR) at the top right is for the structures of an enzyme that belongs to Lyase and Hydrolase. The phylogeny, which shows the evolutionary relationship, of the 62 proteins as a dendrogram is shown in (b). Labels are linked to structural summaries. The average linkage clustering of the 62 structures based on the DALI Z-score is used to generate the phylogeny. The dendrogram groups the most strongly similar structures quite well. Again, one can see that the 18 structures for the S protein of other viruses form a clade of closely related structures. Similarly, Lectin (Galectin) and Lyase (Hydrolase) form their own clades. Some structures such as the bottom four (6YSJ, 6YEJ, 7BTX and 4ZEL), with low Z-score values, are the outgroup as they are less related structures to 6VSB. Note that the branching order gets more or less arbitrary nearer the root.

The scatterplot (Figure 6) is for the relationship of RMSD with Z-score. As the structures are compared against 6VSB, the cluster for S protein structures is far from other two on the basis of high Z-scores. For the S protein structures, the outliers (6VV5, 6U7K, 7E9T and 6JX7) occurred due to low Z-scores, while for other two structure clusters, the outliers (2WSU and 5MQR, 5WRU) occurred due to high RMSD values. Some S protein structures such as 6U7H, 6B7N and 6NB7 have high RMSD values ≥ 5 . Rhinolophus bat coronavirus HKU2 (6M15) is not present as its Z-score and RMSD values against all structures were 0. The SARS-CoV-2 S protein structure (7E9T) is far from its respective clusters and close to the other two clusters because of the low Z-score. The two hydrolase structures (5MQR and 5WRU) at the top left have high RMSD values and low Z-score. Using the two similarity measures alone, particularly RMSD, cannot provide good clusters and comprehensive results for similarity. The correlation found between structures through two measures is consistent with the results obtained with the similarity matrix based on Z-score with small differences. In fact, the correlation results provide more detailed information, suggesting that it is better to use several similarity metrics than the use of each similarity measure alone. The pair of structures that are very close to each other but belong to different clusters in Figure 6 such as the pair 6VV5, 6W4Q, the pair 2WSU, 1KIT and the pair 3WUC, 1MVE have very low similarity identity against the query structure. BLAST returned percent identity of 33.89%, for 6VV5 and no significant similarity in sequence for 6W4Q, 2WSU, 1KIT, 3WUC and 1MVE against the query structure.

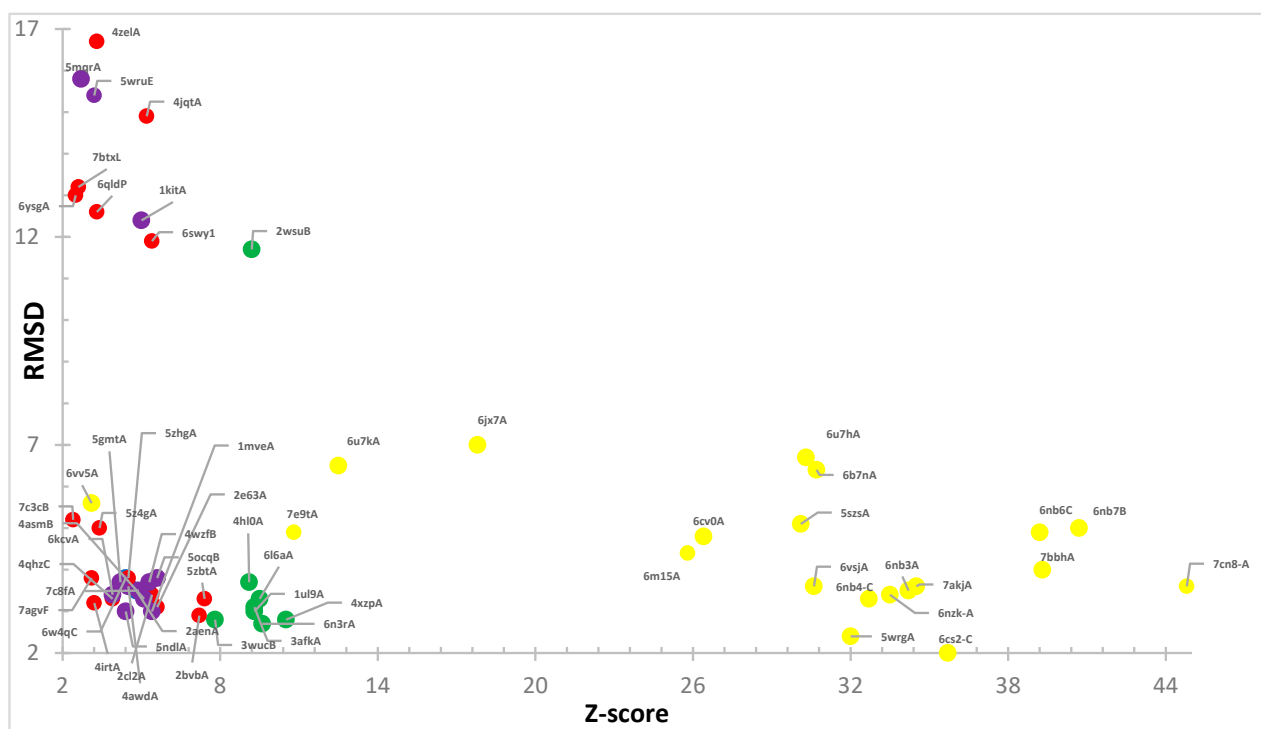


Figure 6. Relationship of two measures (Z-score and RMSD) for protein structures' similarity. The structures that belong to three clusters are colored (yellow for S protein structures, green for Lectin (Galectin) and purple for Lyase (Hydrolase)). Structures that do not belong to three clusters are red-colored. Lyase (Hydrolase) and Lectin (Galectin) are very close to each other on the basis of the correlation between RMSD and Z-score.

4. Structures Classification and Their Description

The S protein structures of SARS-CoV-2 belong to either the original Wuhan strains or to different variants in PDB. We classified them into different variants (Table 4). For VOCs, Omicron S protein structures are highest (28.63%), followed by Beta (24.41%), Delta (15.25%), Alpha (12.44%) and Gamma (7.51%). In PDB, 7.74% (0.5%) of S protein

structures belong to VOI Kappa (Epsilon and Brisdelta). For the S protein, approximately one-third of the published literature in PDB reported the structures to gain insight into this virus's functionalities and to properly understand how the evolution affects the virus transmissibility and immune evasion. The majority of this literature, 80%, focused on the S protein structure in SARS-CoV-2 variants. The remaining 20% studied the S protein structures in SARS-CoV-2 original strains. Omicron is the focus now as it is the only one left in the current VOCs list by the WHO. On the other hand, two-third of the published literature in PDB for S protein not only derived the structures but also investigated the therapeutic potential of antibodies, T-cell receptors (TCRs) and inhibitors for neutralizing SARS-CoV-2. For antiviral development, the S protein structures from the original strains are discussed more, followed by Beta, Omicron and Delta. The first structures for SARS-CoV-2 S protein were deposited to PDB in February 2020 by Chinese [6,17] and USA institutes [11,16].

Table 4. PDB IDs for SARS-CoV-2 variants. Red-colored PIDs represent those structures that are used in complex with antibodies.

Variant	PDB IDs	%
Alpha	7R0Z, 7R10, 7R13, 7R14, 7R15, 7R1A, 7SXX, 7SXY, 7SY1, 7SY2, 7SXR, 7FEM, 7FET, 7EDF, 7EDG, 7EDH, 7EDI, 7EDJ, 7EH5 , 7EKF, 7EAE, 7EB0, 7EB3, 7EB4, 7EB5, 7LWV, 7LWU, 7LWT, 7LX5, 7LWS, 7DX1, 7DX2, 7KRQ, 7KRR, 7KRS, 6XS6, 7KDI, 7KDH, 7KDL, 7KDJ, 7KDK, 7KDG, 7KEC, 7KEB, 7KEA, 7KE9, 7KE8, 7KE7, 7KE6, 7KE4, 7CM4, 7BNM, 7BNN	12.44
Beta	7S5P, 7S5Q, 7S5R, 7R11, 7R16, 7R17, 7WCR, 7WCZ, 7WD0, 7WD7, 7WD8, 7WD9, 7WDF , 7RNJ, 7VX1, 7VX4, 7VXD, 7VXF, 7VXK, 7VXM, 7WEV, 7Q9P, 7Q9M, 7Q9K, 7Q9J, 7Q9I, 7Q9G, 7Q9F, 7Q6E, 7Q0I, 7Q0H, 7Q0G, 7PS7, 7PS6, 7PS5, 7PS4, 7PS2, 7PS1, 7PS0, 7PRZ, 7PRY , 7SXT, 7SXV, 7SXW, 7SY3, 7SY4, 7FJS, 7FJO, 7FJN , 7V8C, 7V7Z, 7V80, 7EKG, 7V2A, 7V26, 7RAQ, 7V76, 7V77, 7E3L, 7E3K, 7LYK, 7LYL, 7LYN, 7LYO, 7LYP, 7LYQ, 7LYM, 7MY3, 7MY2, 7N1Y, 7N1X, 7N1W, 7N1V, 7N1U, 7N1T, 7N1Q, 7RA8, 7RAL, 7LWW, 7E7X, 7E7Y, 7E86, 7E88, 7E8C, 7E8F, 7CHO, 7CHP, 7CHS, 7EY0, 7EY4, 7EY5, 7EYA, 7EZV, 7ZFC, 7ZR7, 7ZR8, 7ZR9, 7ZRC	24.41
Gamma	7SXU, 7SY7, 8SY9, 7SBS, 7SBT, 7V79, 7V78, 7V7A, 7V81, 7V82, 7V83, 7V84, 7NXC, 7NX6, 7NX7, 7NX8, 7NX9, 7NXA, 7NXB , 7M8K, 7MJG, 7MJH, 7MJI, 7MIJ, 7MIK, 7MIL, 7MIJN, 7MIJM, 7MY3, 7MY2, 7EKC	7.51
Delta	7TOZ, 7TOY, 7TOX, 7TOV, 7TOU, 7TPL, 7TPH, 7TPF, 7TPE, 7TPC, 7TPA, 7TP9, 7TP8, 7TP7, 7TPF2, 7TP1, 7TP0, 7WBQ, 7SXS, 7SXZ, 7SY0, 7V7N, 7V7O, 7W9I, 7W9F, 7W9E, 7W9C, 7W9B, 7W99, 7W98, 7W97, 7W92 , 7V7P, 7V7Q, 7V7R, 7V7S, 7V88, 7V89, 7V8A, 7V8B, 7SBK, 7SBL, 7SBO, 7LQ7, 7Z0Y, 7Z0X, 7X7D, 7E9N, 7F46, 7E9O, 7E9P, 7E9Q, 7ENF, 7ENG, 7TEX, 7TEY, 7TEW, 7V7T, 7V7U, 7V7V, 7OR9, 7ORB, 7ORA, 7WWL, 7WWM	15.25
Omicron	7WVQ, 7WVP, 7WVO, 7WVN, 7U0N, 7U0D, 7WUH, 7TGY, 7TGW, 7TGX, 7TGE, 7WP9, 7WPA, 7WPB, 7WPC, 7WPD, 7WPE, 7WPF, 7WRV , 7QO7, 7SY5, 7SY6, 7QNW, 7WBL, 7WBP, 7T9J, 7T9K, 7T9L, 7TB4, 7R40, 7THK, 7TOW, 7TF8, 7TEI, 7TLD, 7TLC, 7TLA, 7TL9, 7TL1, 7THT, 7THE, 7TNW, 7TO4, 7SO9, 7SOA, 7UAP, 7UAQ, 7UAR , 7TM0, 7TLY, 7TN0, 7TLZ, 7WK5, 7WK4, 7WK6, 7WK9, 7WK2, 7WK3, 7WK8, 7WKA, 7WK9, 7QO9, 7UB0, 7UB5, 7UB6, 7RZQ, 7RZR, 7RZS, 7RZT, 7RZU, 7RZV, 7TIK, 7TB8, 7TC9, 7U., 7OD, 7TCA, 7TBE, 7TCC , 7TB4, 7WEE, 7WEE, 7WED, 7WEC, 7WEB, 7WEA, 7WE9, 7WE8, 7WE7, 7WLC, 7WHI, 7WHJ, 7WHK , 7WG6, 7WG7, 7WG8, 7WG9, 7WGB, 7WGC, 7WS9, 7WS8, 7WS8, 7WS7, 7WS6, 7WS5, 7WS4, 7WS3, 7WS2, 7WS1, 7WS0, 7WHH, 7ZF3, 7ZFB, 7ZFD, 7ZFE, 7ZFE, 7ZF4, 7ZF5, 7ZF6, 7ZF7, 7ZF8, 7ZF9, 7ZFA	28.63
Kappa	7VX5, 7VX9, 7VXA, 7VXB, 7VXC, 7VXE, 7VXI, 7V7D, 7V7E, 7V7F, 7V7G, 7V7H, 7V7I, 7V7J, 7V85, 7V86, 7V87, 7SBP, 7SBQ, 7SBR, 7SOB, 7SOC, 7SOD, 7SOE, 7SOF, 7TF5, 7TF4, 7TF3, 7TF2, 7TF1, 7TF0, 7TEZ	7.74
Mink	7R12, 7R18, 7R19, 7R1B, 7LWQ, 7LWP, 7LWO, 7LWN, 7LWM, 7LWL, 7LWK, 7LWJ, 7LWI	3
Epsilon	7N8H, 7N8I	0.50
Brisdelta	7ODL, 7OD3	0.50

4.1. Drug Development

S protein's vital role in viral infection makes it a target for antibody-blocking therapy, vaccine development and small molecule inhibitors against SARS-CoV-2. Table 5 summarizes the studies for S protein structures in complex with antibodies of SARS-CoV-2 variants.

Table 5. Summary of RBD (NTD)-directed NAb targeting the S protein of SARS-CoV-2 variants.

Ref	Domain	Antibodies	Variants
[28]	RBD	RBD-chAb15 and RBD-chAb 45	Alpha
[29]	RBD	S5D2, S5G2, S3H3	Alpha, Beta, Kappa Delta
[30]	RBD, NTD	Beta-6, -22, -24, -27, -29, -32, -38, -40, -43, -44, -47, -48, -49, -50, -53, -54, Covox-222, -45	Beta
[31]	RBD	T6	Wild-type, Beta, Gamma
[32]	RBD, NTD	N11, N9, BD-623, -508, -515,	Beta
[33]	RBD	P5A-1D2, P5A-3C8, P22A-1D1	Beta
[34]	RBD	BD-813, -744, -667, -771 -821, -804, -812, -836	Beta
[35]	RBD	222	Alpha, Beta, Gamma
[36]	RBD	8D3	Delta
[37]	RBD	222, 278, 45, 253 75,	Kappa and Delta
[38]	RBD	ZWD12, ZWC6	Alpha, Beta, Gamma, Kappa, Delta, Omicron
[39]	RBD	JMB2002	Omicron, Alpha, Beta, Gamma
[40]	NTD	C1520, C1791	Omicron
[40]	RBD	LY-CoV1404, S309, A19-46.1, S2E12, COV2-2196, B1-182.1, A23-58.1	Omicron
[41]	RBD	XGv265, -282, -289, -347,	VOC particularly Omicron
[42]	RBD	BN03	Omicron
[43]	RBD	Vir-S309	Beta and Omicron

4.1.1. Antibodies Based on S Protein of SARS-CoV-2 Variants

Yang et al. [28] examined the mutations' impact, particularly the N501Y mutation, on the Alpha variant by revealing the structural basis of two NAb (RBD-chAb15 and RBD-chAb 45) that bind to the RBD region of the S protein. Xu et al. [29] isolated nine MAbs from mice and investigated their neutralization potency against four variants. They identified three types of MAbs (S5D2, S5G2 and S3H3). S5D2 efficiently neutralized all variant pseudoviruses, but its potency against Delta decreased significantly. The S5G2 exhibited comparable neutralization towards the variants while S3H3 was able to neutralize, near equally, all variants. Liu et al. [30] extracted 674 MAbs from individuals infected with the Beta variant. Eighteen MAbs targeted the three mutations (E484K, K417N and N501Y) in Beta RBD. Liang et al. [31] developed and checked the efficacy of the mRNA-LNP vaccine against the Beta and Gamma variants as well as the wild type in the mouse. The protein structures were used to find the molecular basis for potent MAb T6 Fab. Cao et al. [32] reported the immune response of the plasma and NAb against Beta among individuals vaccinated with the inactivated vaccine (CoronaVac) or RBD-subunit vaccine (ZF2001) and those infected with the SARS-CoV-2. Nearly half of the identified anti-RBD NAb displayed neutralization reductions. Whereas the plasma from individuals of convalescents and CoronaVac vaccinees showed comparable neutralization reductions. For ZF2001, the

extended interval in the second and third doses produced improved neutralizing activity and more tolerance towards Beta compared to the conventional three-dose treatment. Zhang et al. [33] extracted 165 antibodies from eight SARS-CoV-2-infected patients. They find that potent NAbs have IGHV3-53/3-66 common in their heavy chain. NAb P5A-3C8 showed high defensive efficiency in a SARS-CoV-2-infected hamster. However, the K417N mutation in Beta RBD eliminated these antibodies' neutralizing activity.

Du et al. [34] characterized eight NAbs for the SARS-CoV-2 Beta variant. Five NAbs directly antagonized the binding. The remaining three neutralization mechanisms do not depend on the blocking of ACE2. Two NAbs (BD-812 and BD-836) showed high neutralizing potency against Beta. Dejnirattisai et al. [35] reported that compared to Beta, Gamma is noticeably less resistant to naturally occurring or vaccine-induced antibody responses. This suggested that mutations outside the RBD have an impact on neutralization. MAb 222 neutralized three VOCs. Wang et al. [36] found that the mutation (T478K) enhanced the Delta interaction with ACE2. MAb 8D3 was identified as a neutralizing cross-variant antibody. Moreover, the five tested MAbs, which target RBD, remained effective on Delta. Convalescent and vaccine sera were used by Liu et al. [37] to neutralize the two variants (B.1.617.1 and B.1.617.2). It was found that both variants' neutralization is reduced compared with the original strains. Moreover, the Gamma and Beta variants sera fail to neutralize the Delta variant which suggests that these variants are antigenically divergent. Chi et al. [38] developed various NAbs for Omicron and other VOCs elicited by the Ad5-nCoV vaccine and examined the responses of the human antibody after vaccination. The ZWD12 MAb showed potency and neutralization against six variants. ZWD12 and ZWC6 provided complete protection in the K18-hACE2 transgenic mouse model. Yin et al. [39] reported how the S protein of Omicron variant maintains the binding to the human ACE2 receptor and even gets stronger. The antibody JMB2002 effectively neutralized the four VOC, but not Delta.

Wang et al. [40] found six groups of NTD-directed NAbs from convalescent individuals and/or individuals vaccinated with the mRNA vaccine. Two NAbs C1520 and C1791 were able to recognize epitopes on the opposite faces of NTD. They found that vaccination increases the responses of anti-NTD against Omicron. Zhou et al. [41] identified RBD-directed NAbs to neutralize the Omicron variant. The sera from individuals who already received two or three doses of CoronaVac were examined [42] to see whether they could neutralize Omicron. They found that the third vaccine dose of CoronaVac significantly boosts the geometric mean neutralization antibody titer (16.5-fold) for Omicron. Moreover, a subset of MAbs, derived from memory B cells in individuals having three vaccine doses, neutralized all VOCs. Li et al. [43] identified two regions in the S RBD of Omicron recognized by NAbs. BN03, which is a bi-specific single-domain antibody, was produced that can be transported to the lung through inhalation. BN03 showed high neutralization efficiency in mice infected with SARS-CoV-2. Guo et al. [44] investigated how Omicron retains strong ACE2 interactions that help in evading certain NAbs and found one anti-RBD NAb (5105A) to validate how Omicron escapes the neutralization of the antibody. Nutalai et al. [45] investigated the three sub-lineages of Omicron. The RBD of BA.2 lineage has a high affinity for ACE2 compared to the lineages BA.1 and BA.1.1. On the other hand, the BA.2 lineage is more difficult to neutralize compared to the BA.1 lineage. Some mutations such as S371F (in BA.2 lineage) and R346K (in BA.1.1 lineage) reduced the neutralization effect of the Vir-S309 antibody. The structure and function analysis of potent MAbs obtained from Pfizer–BioNtech vaccinated individuals revealed two core clusters for these MAbs within the RBD. McCallum et al. [46] discussed the properties of the S protein of the Epsilon variant. In the signal peptide, this variant has one mutation (S13I). A signal peptide is a short peptide (generally 16–30 AAs long) located in the N-terminal of protein and contains information for protein secretion. One mutation (L452R) in RBD and one mutation (W152C) in NTD were also detected. These mutations reduced the neutralizing activity in RBD and NTD-directed MAbs. Table 5 summarizes the studies for S protein structures in complex with antibodies of SARS-CoV-2 variants.

Beside B cells, T cells are also critical against SARS-CoV-2 as they form a long-term memory response. Two studies [47,48] provided the structural information for T-cell receptors (TCRs). Nguyen et al. [47] determined three structures (PIDs: 7M8U, 7M8T and 7M8S) of the S-derived peptides presented by three frequently expressed HLA (human leukocyte antigen) molecules. Wu et al. [48] resolved the structures of TCR (YLQ7 (PID: 7N1D) and RLQ3 (PID: 7N1C), from patients infected with SARS-CoV-2 in complex with HLA-A2 and two epitopes of S protein (YLQ (PID: 7N1A) and RLQ (PID: 7N1B)). These TCRs elicited the T-cell responses of universal CD8+ in HLA-A2 patients. The public TCRs were found in various individuals which were unrelated to each other. On the other hand, private TCRs were distinct among individuals.

4.1.2. Small Protein Inhibitors

Small and stable proteins that can block the S binding to ACE2 were designed by Cao et al. [49]. The best mini-proteins bind with high affinity and prevented the infection of SARS-CoV-2 in mammalian Vero E6 cells. In mammalian cells, the mini-proteins do not need expression compared to antibodies. These proteins can be formulated directly to the nasal (or respiratory system) because of their small size and good stability. Transferring high concentrations of viral inhibitors into the respiratory system through the nose generally provides not only prophylactic protection but a therapeutic benefit to treat the early infection. Furthermore, the structural study [50] of inhibitors showed that the NTD region of the SARS-CoV-2 S binds biliverdin and bilirubin with nanomolar affinity. Particularly, biliverdin significantly reduced the reactivity of SARS-CoV-2 S. They reported the structure of SARS-CoV-2 S protein NTD with biliverdin (PID: 7B62). Potential drug targets include all of the enzymes involved in coronavirus replication. However, the focus in the antiviral studies for small-molecule inhibitors is on proteases, such as PL^{pro}, M^{pro} and RdRP (RNA-dependent RNA polymerase) [3,4,51–58].

5. Discussion

The S protein structures of SARS-CoV-2 in PDB are analyzed for similarity using DALI for various parameters. Generally, almost all proteins are structurally similar to other proteins. These similarities in protein structures arise from chemistry and physics principles. Finding the evolutionary relationships in protein structures can reveal surprising similarities. The comparative analysis of 3D S protein structures of SARS-CoV-2, obtained using Cryo-EM or X-Ray, reveals biologically interesting similarities that are hard to detect by only comparing the sequences. Protein structures are compared with sequence alignment, structures superimposition in 3D, similarity matrix and dendrogram. We believe that the obtained results may help in inferring the functional properties of SARS-CoV-2 S proteins. The correlation results of two similarity measures (Z-score and RMSD) showed that the usage of both measures provided better results compared to the ones when only one measure, i.e., RMSD, is employed. The earliest S protein structures of SARS-CoV-2 (PIDs: 6VSB, 6M17, 6VXX, 6VYB and 6M0J) were deposited to PDB in February 2020 [6,11,16,17]. These structures provided not only the basis but important details and information for the development of therapeutics that targeted the interaction and binding of the S protein with the ACE2 receptor. The reputed journals *Science*, *Cell*, *Nature Communication*, *Cell Reports* and *Nature* published the largest portion of studies for the S protein structures of SARS-CoV-2. Institutes from the USA (Harvard Medical School, Washington University and NIAID) published the highest number of papers, followed by China (CAS and Tsinghua University). Approximately 50% of studies related to the SARS-CoV-2 studies were published in 2021.

Besides comparing the S structures, we also classified them into different variants and found that the majority of the studies focused on the S of SARS-CoV-2 variants. More than 50% of variant structures belong to Omicron and Beta. The majority of the structures belong to the RBD region of S. Till now, almost all of the identified neutralizing antibodies/nanobodies for SARS-CoV-2 target two main regions (RBD or NTD) of the S

protein as both of them offer the two main neutralization targets for antibodies. Particularly, all MABs that have been approved or are undergoing clinical testing are aimed at the RBD regions of SARS-CoV-2, which targets a number of overlapping and non-overlapping epitopes [29].

Generally, the neutralization potency of RBD-targeted MABs correlates with its ACE2-blocking efficiency. On the other hand, the binding epitopes of NTD-directed neutralizing MABs overlap highly. This results in the formation of antigenic supersite. Other than NTD- and RBD-directed MABs, neutralizing MABs that bind to the S2 region of S were also reported. However, S2-directed MABs potency is very low. The studies that provide the structural information on NABs obtained from COVID-19 that bound to the SARS-CoV-2 RBD, NTD or spike trimer offer a highly thorough picture of the response of the B cell to SARS-CoV-2 virus. For SARS-CoV-2, the NABs may be short-lived. Moreover, they are not elicited in all patients. T cells also play an important role to combat this virus and form a long-term memory response. T cells tend to identify and recognize those parts in a virus that do not mutate rapidly. However, no comprehensive structural information is available on TCRs for the SARS-CoV-2 that bound to their peptide-MHC (pMHC) targets [49]. The main focus, till now, in immunity surveillance and vaccine development is on the role of NABs while less importance is given to understand the T cells' role on non-NABs that can offer protection through several mechanisms, such as Ab-dependent cellular cytotoxicity and opsonization.

Since late 2019, the SARS-CoV-2 virus has evolved a lot. All variants, particularly VOCs, contain multiple mutations in the S protein. Some changes were also found in the RBD, NTD and in the furin cleavage site between the S1 and S2 regions. Various variants of SARS-CoV-2 exhibit diverse mutation patterns in their S protein. These mutations gave this virus not only improved fitness but made it more infectious [59] and enhanced its ability for immune evasion. Because of this, SARS-CoV-2 shows increased resistance to antibody therapies and vaccines that were raised against the original strain [36,60] and against earlier VOCs. So far, the Omicron variant has undergone the most mutations: with more than 30 mutations in its S, 15 of them occur in the RDB. Omicron has the ability to nearly ablate the neutralizing effect of most FDA-approved antibody drugs, such as AZD1061, REGN10987, REGN10933, LY-CoV016 and Y-CoV555 [61]. The Omicron variant also reduces the effectiveness of approved vaccines, such as mRNA and inactivated virus vaccines [62]. Moreover, WHO Solidarity clinical trials [63] indicate that drugs, such as remdesivir, hydroxychloroquine and lopinavir, have minimal to no impact on COVID-19 patients who are hospitalized, as determined by overall mortality, the start of ventilation and length of stays at the hospital.

Thus, it is important, indeed critical, to continue the search for drugs and broadly NABs and define their epitopes for the development of long-term and broad-spectrum antiviral therapies for guiding the development of an effective vaccine for SARS-CoV-2. Several drug administration agencies have authorized and advised a third booster dose for all adults to battle the current revival of the epidemic since three doses can neutralize the Omicron with a 40-fold drop in viral titer, although two doses are less effective [64]. We believe that effective vaccine development, drug repurposing and other treatment therapies that can offer long-term immunity against the Omicron and all previous (and future) variants will be hot zones for future research directions. This is also evident from the deposition of more S protein structures in PDB that belong to Omicron. Moreover, recent works focused on finding antibodies for the Omicron S protein of SARS-CoV-2. However, more research is needed into the therapeutic implications of natural and vaccine-induced immunity in connection to the defense against infection and life-threatening illness. More contributions are also required on the role of the T cell to the host immune response (T-cell immunity). This may allow early, comprehensive and long-term as well as durable protection from SARS-CoV-2, particularly from VOCs and from emerging variants in the future.

6. Conclusions

Understanding the S protein structures of SARS-CoV-2 is critical to properly understand this virus and its main functionalities and for the development of effective drugs that block the S protein interaction and binding with the ACE2 receptor. In this work, we compare and analyze the S protein structures of SARS-CoV-2 that were deposited in the PDB. Moreover, the studies from PDB that deposited the S protein structures of SARS-CoV-2 were reviewed to classify the structures into variants and drug types. We found that most of the work used NAbs, particularly, MAbs, to neutralize the entry of the S protein in the host cell. The SARS-CoV-2 mortality rate is low but it is highly transmittable and new variants of concern, particularly the current VOC Omicron makes this virus more resistant against neutralizing antibodies that are already authorized and in use. VOCs carrying mutations in the binding domains, such as RBD and NTD, of S reduce the effectiveness of many NAbs and vaccines by evading neutralization. Thus, it is imperative to find and develop broadly neutralizing NAbs that can offer long-term and broad-spectrum antiviral therapies against VOCs, less susceptible to resistance and guide the design of an effective vaccine for SARS-CoV-2. Moreover, the T-cell immunity and its importance to control SARS-CoV-2 may have been underestimated thus far and needs more attention from academia.

For future work, an extension is to include the S protein structures of SARS-CoV-2 deposited to EMDB and to comparatively analyze the protein structures for the main protease (M^{pro}) of SARS-CoV-2. Moreover, performing a bibliometric analysis of the literature related to the SARS-CoV-2 S protein structures is also an interesting area. The literature can be retrieved from online scholarly databases such as Web of Science (WoS) and Scopus, and various software's such as VOSviewer can be used for bibliometric analysis. We are more interested in the future to check the effectiveness of known drug ligands on the S protein structures for SARS-CoV-2 variants using the molecular docking and molecular dynamics simulations. This will allow to investigate whether the performance of ligands on different variants and their lineages is the same or not.

Author Contributions: Conceptualization, M.A. and W.L.; methodology, M.A., M.S.N. and W.L.; software, M.A. and M.S.N.; validation, M.A. and M.S.N.; formal analysis, M.S.N. and W.L.; investigation, M.S.N., P.F.-V. and W.L.; data curation, M.A. and M.S.N.; writing—original draft preparation, M.A. and M.S.N.; writing—review and editing, M.A., P.F.-V. and W.L.; visualization, M.A.; supervision, W.L.; project administration, W.L.; funding acquisition, W.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Natural Science Foundation of Guangdong Province, China (Grant No. 2020A1515010984) and the Start-up Grant for Young Scientists (860-000002110384), Shenzhen University.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: The data used to support this study are included in this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wu, F.; Zhao, S.; Yu, B.; Chen, Y.-M.; Wang, W.; Song, Z.-G.; Hu, Y.; Tao, Z.-W.; Tian, J.-H.; Pei, Y.-Y.; et al. A new coronavirus associated with human respiratory disease in China. *Nature* **2020**, *579*, 265–529. [[CrossRef](#)] [[PubMed](#)]
2. Lata, S.; Akif, M. Comparative protein structure network analysis on 3CLpro from SARS-CoV-1 and SARS-CoV-2. *Proteins* **2021**, *89*, 1216–1225. [[CrossRef](#)] [[PubMed](#)]
3. Greasley, S.E.; Noell, S.; Plotnikova, O.; Ferre, R.; Liu, W.; Bolanos, B.; Fennell, K.; Nicki, J.; Craig, T.; Zhu, Y.; et al. Structural basis for the in vitro efficacy of nirmatrelvir against SARS-CoV-2 variants. *J. Biol. Chem.* **2022**, *298*, 101972. [[CrossRef](#)]
4. Fu, L.; Ye, F.; Feng, Y.; Yu, F.; Wang, Q.; Wu, Y.; Zhao, C.; Sun, H.; Huang, B.; Niu, P.; et al. Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease. *Nat. Commun.* **2020**, *11*, 4417. [[CrossRef](#)] [[PubMed](#)]

5. Nawaz, M.S.; Fournier-Viger, P.; He, Y. S-PDB: Analysis and classification of SARS-CoV-2 Spike protein structures. In Proceedings of the 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Las Vegas, NV, USA, 6–8 December 2022; pp. 2259–2265. [\[CrossRef\]](#)
6. Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020**, *581*, 215–220. [\[CrossRef\]](#)
7. Zhou, P.; Yang, X.-L.; Wang, X.-G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.-R.; Zhu, Y.; Li, B.; Huang, C.-L.; et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**, *579*, 270–273. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Kruger, N.; Herrler, T.; Erichsen, S.; Schiergens, T.S.; Herrler, G.; Wu, N.-H.; Nitsche, A.; et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **2020**, *181*, 281–292.e6. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Rawat, P.; Jemimah, S.; Ponnuswamy, P.K.; Gromiha, M.M. Why are ACE2 binding coronavirus strains SARS-CoV/SARS-CoV-2 wild and NL63 mild? *Proteins* **2021**, *89*, 389–398. [\[CrossRef\]](#)
10. Nawaz, M.S.; Fournier-Viger, P.; Shojaei, A.; Fujita, H. Using artificial intelligence techniques for COVID-19 genome analysis. *Appl. Intell.* **2021**, *51*, 3086–3103. [\[CrossRef\]](#)
11. Wrapp, D.; Wang, N.; Corbett, K.S.; Goldsmith, J.A.; Hsieh, C.-L.; Abiona, O.; Graham, B.S.; McLellan, J.S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020**, *367*, 1260–1263. [\[CrossRef\]](#)
12. Burley, S.K.; Berman, H.M.; Kleywegt, G.J.; Markley, J.L.; Nakamura, H.; Valenkar, S. Protein Data Bank (PDB): The Single Global Macromolecular Structure Archive. *Protein Crystallogr. Methods Mol. Biol.* **2017**, *1607*, 627–641. [\[CrossRef\]](#)
13. Lawson, C.L.; Baker, M.L.; Best, C.; Bi, C.; Dougherty, M.; Feng, P.; Ginkel, G.V.; Devkota, B.; Lagerstedt, I.; Ludtke, S.J.; et al. EMDDataBank.org: Unified data resource for CryoEM. *Nucleic Acids Res.* **2011**, *39* (Suppl. 1), D456–D464. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Holm, L. Dali server: Structural unification of protein families. *Nucleic Acids Res.* **2022**, *50*, W210–W215. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Holm, L. Using DALI for Protein Structure Comparison. In: Structural Bioinformatics. *Methods Mol. Biol.* **2020**, *2112*, 29–42. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Walls, C.; Park, Y.-J.; Tortorici, M.A.; Wall, A.; McGuier, A.T.; Veisler, D. Structure, function, and antigenicity of the SARS-CoV-2 Spike glycoprotein. *Cell* **2020**, *181*, 281–292.e6. [\[CrossRef\]](#)
17. Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **2020**, *367*, 1444–1448. [\[CrossRef\]](#)
18. Pinto, D.; Park, Y.-J.; Beltramello, M.; Walls, A.C.; Tortorici, M.A.; Bianchi, S.; Jaconi, S.; Culap, K.; Zatta, F.; Marco, A.D.; et al. Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **2020**, *583*, 290–295. [\[CrossRef\]](#)
19. Li, W.; Godzik, A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **2006**, *22*, 1658–1659. [\[CrossRef\]](#)
20. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [\[CrossRef\]](#)
21. Sun, D.; Sang, Z.; Kim, Y.J.; Xiang, Y.; Cohen, T.; Belford, A.K.; Huet, A.; Conway, J.F.; Sun, J.; Taylor, D.J.; et al. Potent neutralizing nanobodies resist convergent circulating variants of SARS-CoV-2 by targeting diverse and conserved epitopes. *Nat. Commun.* **2021**, *12*, 4676. [\[CrossRef\]](#)
22. Tai, L.; Zhu, G.; Yang, M.; Cao, L.; Xing, X.; Yin, G.; Chan, C.; Qin, C.; Rao, C.; Wang, X.; et al. Nanometer-resolution in situ structure of the SARS-CoV-2 postfusion spike protein. *Proc. Natl. Acad. Sci. USA* **2020**, *118*, e2112703118. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Tortorici, M.A.; Czudnochowski, N.; Starr, T.N.; Marzi, R.; Walls, A.C.; Zatta, F.; Bowen, J.E.; Jaconi, S.; Di'Iulio, D.; Wang, A.; et al. Broad sarbecovirus neutralization by a human monoclonal antibody. *Nature* **2021**, *597*, 103–108. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Walls, A.C.; Xiong, X.; Park, Y.-J.; Tortorici, M.A.; Snijder, J.; Quispre, J.; Cameoni, E.; Gopal, R.; Dai, M.; Lanzavecchia, A.; et al. Unexpected receptor functional mimicry elucidates activation of coronavirus fusion. *Cell* **2020**, *183*, 1732. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Ni, D.; Lau, K.; Turelli, P.; Raclot, C.; Beckert, B.; Nazarov, S.; Pojer, F.; Myasnikov, A.; Stahlberg, H.; Trono, D. Structural analysis of the Spike of the Omicron SARS-COV-2 variant by cryo-EM and implications for immune evasion. *BioRxiv* **2021**. [\[CrossRef\]](#)
26. Kabsch, W.; Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577–2637. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Zhang, S.; Qiao, S.; Yu, J.; Zeng, J.; Shan, S.; Tian, L.; Lan, J.; Zhang, L.; Wang, W. Bat and pangolin coronavirus spike glycoprotein structures provide insights into SARS-CoV-2 evolution. *Nat. Commun.* **2021**, *12*, 1607. [\[CrossRef\]](#)
28. Yang, T.J.; Yu, P.-Y.; Chang, Y.-C.; Liang, K.-H.; Tso, H.-C.; Ho, M.-R.; Chen, W.-Y.; Lin, H.-T.; Wu, H.-C.; Hsu, S.-T.D. Effect of SARS-CoV-2 B.1.1.7 mutations on spike protein structure and function. *Nat. Struct. Mol. Biol.* **2021**, *28*, 731–739. [\[CrossRef\]](#)
29. Xu, S.; Wang, Y.; Wang, Y.; Zhang, C.; Hong, Q.; Gu, C.; Xu, R.; Wang, T.; Yang, Y.; Zang, J.; et al. Mapping cross-variant neutralizing sites on the SARS-CoV-2 spike protein. *Emerg. Microbes. Infect.* **2022**, *11*, 351–367. [\[CrossRef\]](#)
30. Liu, C.; Zhou, D.; Nutalai, R.; Duyvesteyn, H.M.E.; Tuekprakhon, A.; Ginn, H.M.; Dejnirattisai, W.; Supasa, P.; Mentzer, A.J.; Wang, B.; et al. The antibody response to SARS-CoV-2 Beta underscores the antigenic distance to other variants. *Cell Host Microbe* **2021**, *30*, 53–68.e12. [\[CrossRef\]](#)
31. Liang, Q.; Wang, Y.; Zhang, S.; Sun, J.; Sun, W.; Li, J.; Liu, Y.; Li, M.; Cheng, L.; Jiang, Y.; et al. RBD trimer mRNA vaccine elicits broad and protective immune responses against SARS-CoV-2 variants. *iScience* **2022**, *25*, 104043. [\[CrossRef\]](#)
32. Cao, Y.; Yisimayi, A.; Bai, Y.; Huang, W.; Li, X.; Zhang, Z.; Yuan, T.; An, R.; Wang, J.; Ziao, T.; et al. Humoral immune response to circulating SARS-CoV-2 variants elicited by inactivated and RBD-subunit vaccines. *Cell Res.* **2021**, *31*, 732–741. [\[CrossRef\]](#)

33. Zhang, Q.; Ju, B.; Ge, J.; Chan, J.F.-W.; Cheng, L.; Wang, R.; Huang, W.; Fang, M.; Chen, R.; Zhou, B.; et al. Potent and protective IGHV3-53/3-66 public antibodies and their shared escape mutant on the spike of SARS-CoV-2. *Nat. Commun.* **2021**, *12*, 4210. [\[CrossRef\]](#)
34. Du, S.; Liu, P.; Zhang, Z.; Xiao, T.; Yasimayi, A.; Huang, W.; Wang, Y.; Cao, Y.; Xie, X.S.; Xiao, J. Structures of SARS-CoV-2 B.1.351 neutralizing antibodies provide insights into cocktail design against concerning variants. *Cell Res.* **2021**, *31*, 1130–1133. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Dejnirattisai, W.; Zhou, D.; Supasa, P.; Liu, C.; Mentzer, A.J.; Ginn, H.M.; Zhao, Y.; Duyvesteyn, H.M.E.; Tuekprakhon, A.; Nutalai, R.; et al. Antibody evasion by the P.1 strain of SARS-CoV-2. *Cell* **2021**, *184*, 2939–2954.e9. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Wang, Y.C.; Liu, C.; Zhang, C.; Wang, Y.; Hong, Q.; Xu, S.; Li, Z.; Yang, Y.; Huang, Z.; Cong, Y. Structural basis for SARS-CoV-2 Delta variant recognition of ACE2 receptor and broadly neutralizing antibodies. *Nat. Commun.* **2022**, *13*, 871. [\[CrossRef\]](#)
37. Liu, C.; Ginn, H.M.; Dejnirattisai, W.; Supasa, P.; Wang, B.; Tuekprakhon, A.; Nutalai, R.; Zhou, D.; Mentzer, A.J.; Zhao, Y.; et al. Reduced neutralization of SARS-CoV-2 B.1.617 by vaccine and convalescent serum. *Cell* **2021**, *184*, 4220–4236.e13. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Chi, X.; Guo, Y.; Zhang, G.; Sun, H.; Zhang, J.; Li, M.; Chen, Z.; Han, J.; Zhang, Y.; Zhang, X.; et al. Broadly neutralizing antibodies against Omicron-included SARS-CoV-2 variants induced by vaccination. *Signal Transduct. Target Ther.* **2022**, *7*, 139. [\[CrossRef\]](#)
39. Yin, W.; Xu, Y.; Xu, P.; Cao, X.; Wu, C.; Gu, C.; He, X.; Wang, X.; Huang, S.; Yuan, Q.; et al. Structures of the Omicron spike trimer with ACE2 and an anti-Omicron antibody. *Science* **2022**, *375*, 1048–1053. [\[CrossRef\]](#)
40. Wang, Z.; Muecksh, F.; Cho, A.; Gaebler, C.; Hoffmann, H.-H.; Ramos, V.; Zong, S.; Cipolla, M.; Johnson, B.; Schmidt, F.; et al. Analysis of memory B cells identifies conserved neutralizing epitopes on the N-terminal domain of variant SARS-Cov-2 spike proteins. *Immunity* **2022**, *55*, 998–1012.e8. [\[CrossRef\]](#)
41. Zhou, T.; Wang, L.; Misasi, J.; Pegu, A.; Zhang, Y.; Harris, D.R.; Olia, A.S.; Talana, C.A.; Yang, E.S.; Chen, M.; et al. Structural basis for potent antibody neutralization of SARS-CoV-2 variants including B.1.1.529. *Science* **2022**, *376*, eabn8897. [\[CrossRef\]](#)
42. Wang, K.; Jia, Z.; Bao, L.; Wang, L.; Cao, L.; Chi, H.; Hu, Y.; Li, Q.; Zhou, Y.; Zhu, Q.; et al. Memory B cell repertoire from triple vaccinees against diverse SARS-CoV-2 variants. *Nature* **2022**, *603*, 919–925. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Li, C.; Zhan, W.; Yang, Z.; Tu, C.; Hu, G.; Zhang, X.; Song, W.; Du, S.; Zhu, Y.; Huang, K.; et al. Broad neutralization of SARS-CoV-2 variants by an inhalable bispecific single-domain antibody. *Cell* **2022**, *185*, 1389–1401. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Guo, H.; Gao, Y.; Li, T.; Li, T.; Lu, Y.; Zheng, L.; Liu, Y.; Yang, T.; Luo, F.; Song, S.; et al. Structures of Omicron spike complexes and implications for neutralizing antibody development. *Cell Rep.* **2022**, *39*, 110770. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Nutalai, R.; Zhou, D.; Tuekprakhon, A.; Ginn, H.M.; Supasa, P.; Liu, C.; Huo, J.; Mentzer, A.J.; Duyvesteyn, H.M.E.; Djokaite-Guraliuc, A.; et al. Potent cross-reactive antibodies following Omicron breakthrough in vaccinees. *Cell* **2022**, *185*, 2116–2131. [\[CrossRef\]](#) [\[PubMed\]](#)
46. McCallum, M.; Bassi, J.; Marco, A.D.; Chen, A.; Walls, A.C.; Di'Iulio, J.; Tortorici, M.A.; Navarro, M.-J.; Sillaci-Frengi, C.; Saliba, C.; et al. SARS-CoV-2 immune evasion by the B.1.427/B.1.429 variant of concern. *Science* **2021**, *373*, 648–654. [\[CrossRef\]](#)
47. Nguyen, A.T.; Szeto, C.; Jayasinghe, D.; Lobos, C.A.; Halim, H.; Chatzileontiadou, D.S.M.; Grant, E.J.; Gras, S. SARS-CoV-2 Spike-Derived peptides presented by HLA molecules. *Biophysica* **2021**, *1*, 194–203. [\[CrossRef\]](#)
48. Wu, D.; Kolesnikov, A.; Yin, R.; Guest, J.D.; Gowathaman, R.; Shmelev, A.; Serdyuk, Y.; Dianov, D.V.; Efimov, G.A.; Pierce, B.G.; et al. Structural assessment of HLA-A2-restricted SARS-CoV-2 spike epitopes recognized by public and private T-cell receptors. *Nat. Commun.* **2022**, *3*, 19. [\[CrossRef\]](#)
49. Cao, L.; Greshnik, I.; Coventry, B.; Case, J.B.; Miller, L.; Kozodoy, L.; Chen, R.E.; Garter, L.; Walls, A.C.; Park, Y.-J.; et al. De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* **2020**, *370*, 426–431. [\[CrossRef\]](#)
50. Rosa, A.; Pye, V.E.; Graham, C.; Muir, L.; Seow, J.; Ng, K.W.; Cook, N.J.; Rees-Spear, C.; Parker, E.; Santos, M.S.D.; et al. SARS-CoV-2 can recruit a heme metabolite to evade antibody immunity. *Sci. Adv.* **2021**, *7*, eabg7607. [\[CrossRef\]](#)
51. Ma, C.; Xia, Z.; Sacco, M.D.; Hu, Y.; Townsend, J.A.; Meng, X.; Choza, J.; Tan, H.; Jang, J.; Gongora, M.V.; et al. Discovery of Di- and Trihaloacetamides as covalent SARS-CoV-2 main protease inhibitors with high target specificity. *J. Am. Chem. Soc.* **2021**, *143*, 20697–20709. [\[CrossRef\]](#)
52. Quan, B.X.; Shuai, H.; Xia, A.-J.; Hou, Y.; Zeng, R.; Liu, X.-L.; Lin, G.-F.; Qiao, J.-X.; Li, W.-P.; Wang, F.-L.; et al. An orally available Mpro inhibitor is effective against wild-type SARS-CoV-2 and variants including Omicron. *Nat. Microbio.* **2022**, *7*, 716–725. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Yin, W.; Mao, C.; Luan, X.; Shen, D.-D.; Shen, Q.; Su, H.; Wang, X.; Zhou, F.; Zhao, W.; Gao, M.; et al. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by Remdesivir. *Science* **2020**, *368*, 1499–1504. [\[CrossRef\]](#)
54. Peng, Q.; Peng, R.; Yuan, B.; Wang, M.; Zhao, J.; Fu, L.; Qi, J.; Shi, Y. Structural Basis of SARS-CoV-2 Polymerase Inhibition by Favipiravir. *Innovation* **2021**, *2*, 100080. [\[CrossRef\]](#)
55. Neydenova, K.; Muir, K.W.; Wu, L.-F.; Zhang, Z.; Coscia, F.; Peet, M.J.; Castro-Hartmann, P.; Qian, P.; Sader, K.; Dent, K. Structure of the SARS-CoV-2 RNA-dependent RNA polymerase in the presence of favipiravir-RTP. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2021946118. [\[CrossRef\]](#)
56. Narayanan, A.; Narwal, M.; Majowicz, S.A.; Varricchio, C.; Toner, S.A.; Ballatore, C.; Brancale, A.; Murakami, K.; Jose, J. Identification of SARS-CoV-2 inhibitors targeting Mpro and PLpro using in-cell-protease assay. *Commun. Biol.* **2022**, *5*, 169. [\[CrossRef\]](#)

57. Rosas-Lemus, M.; Minasov, G.; Shuvalova, L.; Inniss, N.L.; Kiryukhina, O.; Brunzelle, J.; Satchell, K.J.F. High-resolution structures of the SARS-CoV-2 2'-O-methyltransferase reveal strategies for structure-based inhibitor design. *Sci. Signal.* **2020**, *3*, 202. [[CrossRef](#)]
58. Ren, P.X.; Shang, W.-J.; Yin, W.-C.; Ge, H.; Wang, L.; Zhang, X.-L.; Li, B.-Q.; Li, H.-L.; Xu, Y.-C.; Xu, E.H.; et al. A multi-targeting drug design strategy for identifying potent anti-SARS-CoV-2 inhibitors. *Acta Pharmacol. Sin.* **2022**, *43*, 483–493. [[CrossRef](#)] [[PubMed](#)]
59. Wrobel, A.G.; Benton, D.J.; Roustan, C.; Borg, A.; Hussain, S.; Martin, S.R.; Rosenthal, P.B.; Skehel, J.L.; Gamblin, S.J. Evolution of the SARS-CoV-2 spike protein in the human host. *Nat. Commun.* **2022**, *13*, 1178. [[CrossRef](#)] [[PubMed](#)]
60. Browner, P.J.M.; Caniels, T.G.; van der Straten, K.; Snitselaar, J.L.; Aldon, Y.; Bangaru, S.; Torres, J.L.; Okba, N.M.A.; Claireaux, M.; Kerster, G.; et al. Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* **2020**, *369*, 643–650. [[CrossRef](#)]
61. Cao, Y.; Wang, J.; Jian, F.; Xiao, T.; Song, W.; Yisimayi, A.; Huang, W.; Li, Q.; Wang, P.; An, R.; et al. Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature* **2022**, *602*, 657–663. [[CrossRef](#)]
62. Malik, J.A.; Ahmed, S.; Mir, A.; Shinde, M.; Bender, O.; Alshammari, F.; Ansari, M.; Anwar, S. The SARS-CoV-2 mutations versus vaccine effectiveness: New opportunities to new challenges. *J. Infect. Public Health* **2022**, *15*, 228–240. [[CrossRef](#)] [[PubMed](#)]
63. WHO Solidarity Trial Consortium. Repurposed antiviral drugs for COVID-19—Interim WHO Solidarity Trial Results. *N. Engl. J. Med.* **2021**, *384*, 497–511. [[CrossRef](#)] [[PubMed](#)]
64. Garcia-Beltran, W.F.; Denis, K.J.S.; Hoelzemer, A.; Lam, E.C.; Nitido, A.D.; Sheehan, M.L.; Berrios, C.; Ofoman, O.; Chang, C.C.; Hauser, B.M.; et al. mRNA-based COVID-19 vaccine boosters induce neutralizing immunity against SARS-CoV-2 Omicron variant. *Cell* **2022**, *185*, 457–466. [[CrossRef](#)] [[PubMed](#)]

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