Potential linear B-cells epitope change to a helix structure in the spike of Omicron 21L or BA.2 predicts increased SARS-CoV-2 antibodies evasion

Walid Al-Zyoud [a], Hazem Haddad [b,*]

[a] *Department of Biomedical Engineering, School of Applied Medical Sciences, German Jordanian University, Amman, 11180, Jordan*
[b] *Princess Haya Biotechnology Center, Jordan University of Science and Technology, Irbid, 22110, Jordan*

ARTICLE INFO

ABSTRACT

The world health organization has announced that SARS-CoV-2 Omicron variant (B.1.1.529), including the three versions; 21K (BA.1), 21L (BA.2) and 21M (BA.3) as a variant of concern (VOC) on November 2022. In this study, we used the specialized computational platforms to predict the stability and flexibility of the spike protein of Omicron. The aim of this study was to investigate the expected effect of Omicron spike mutations on its physiochemical properties. Findings of this study revealed 16 stabilizing mutations that might explain a newly gained environmental stability. We expect the new mutations to play a crucial role in changing the physiochemical properties of epitopes of the spike protein. The notable finding of SuerPose work was the potential linear B-cells epitope G252 → S255 that has been changed in the spike protein of the Omicron 21L to a helix structure which might confer an escape from human monoclonal antibodies.

1. Introduction

The researchers all over the world are concerned to understand in details the threat posed by the new variant of the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) that first appeared in Botswana and South Africa in November 2021, and called Omicron (B.1.1.529). Omicron was declared as a variant of concern (VOC) after many previous variants of concern dominated the COVID-19 confirmed cases worldwide, such as the variants Beta (B1.351) and Delta (B.1.617.2) variants. Previous reports have highlighted the elevated level of immune evasion of Beta variant from serum neutralizing antibodies (Wang et al., 2021; Chen et al., 2021). Whereas Delta variant had documented transmissibility, pathogenicity besides eroding the neutralizing effect by antibodies (Planas et al., 2021; Jhun et al., 2021a). What is unique about Omicron is that it has over 30 novel mutations in the spike protein, 15 of them are in the Receptor Binding Domain (RBD) if it is compared with prior variants (Cameroni et al., 2021) which might affect the physiochemical properties of the spike protein epitopes including hydrophilicity, surface flexibility, surface accessibility, and antigenicity. It is well known that the RBD is the primary target of antibodies binding (Piccoli et al., 2020; Greaney et al., 2021) which highlights the possibility of Omicron variant to express an antigenic shift and hence an escape from the antibodies. It was reported that the

peptide-based vaccine designs are safer because of its minimal allergic and toxic properties (Malonis et al., 2019; Poland et al., 2011). The vaccine developers of anti COVID-19 rarely take into consideration the epitopes of T-cells when studying the spike protein antigenicity. The aim of this study has focused on studying the predicted thermodynamics of the amino acid substitutions that happened in the spike protein of the Omicron variant. This may trigger the efforts of the vaccines developers to design resilient vaccines by mediating the human immune response by both the B and T cells to stop the spread of the SARS-CoV-2 virus all over the world. This also applies to the new design of monoclonal antibodies as a treatment for COVID-19. In this study, we used Normal Mode Analysis (NMA) via DynaMut as the main specialized computational web-server based platform to predict the stability and flexibility of the spike protein conformation of Omicron (Rodrigues et al., 2018). However, other computational methods can find the same predictions such as WEBnm@ v2.0, which is also a web-server based service with unique ability for comparative NMA on multiple protein structures (Tiwari et al., 2014).

---

2. Methods

2.1. Data retrieval

The spike protein sequences of Nextstrain clades 19A (WA1)(Wu et al., 2020), clade 20A (which includes the dominant mutation D614G) besides the clade 21K (Omicron; B.1.1.529) have been retrieved from the EpiCoV database of GISAID (https://www.gisaid.org/). The Omicron variant mutations included A67V, T95I, G142D, L212I,G339D, S371L,S373P,S375F,K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S,Q498R,N501Y,Y505H, T547K, D614G, H655Y,N679K, P681H,N764K,D796Y,N856K,Q954H, and L981F.

2.2. Normal mode analysis (NMA)

We studied the stability and flexibility dynamics prediction for and nearby the abovementioned 30 changes in an amino acid that represent the mutations of Omicron through a harmonic motion by using the Normal Mode Analysis (NMA) via DynaMut (Rodrigues et al., 2018). The-Tasser tool from the Zhang Lab (https://zhanglab.ccmb.med.umich.edu/COVID-19/)(Yang et al., 2014) was used to retrieve the wild type spike protein (QHD43416.pdb). Three versions of Omicron variants have been found in the EpiCoV database of GISAID and Nextstrain, as Omicron B.1.1.529 (BA.1 or clade 21K), Omicron B.1.1.529 (BA.3 or clade 21M) and Omicron B.1.1.529 (BA.2 or clade 21L).

2.3. Similarity matrix (DALI)

The DALI server (http://ekhidna2.biocenter.helsinki.fi/dali/) (Yang et al., 2014) was used to perform all-against-all protein databank (PDB) structure comparison for the wild type spike protein (clade 19A) versus the spike protein from variants of concern (DALI code: Nextstrain clade code): s001A: 20A, s002A: 19A, s003A: Beta, s004A: Iota, s005A: Delta, s006A: Gamma, s007A: Alpha, s008A: Epsilon, s009A: Lambda, s010A: Kappa, s011A: Eta, s012A: Theta, s013A: Omicron BA.1 or clade 21M, s014A: Omicron BA.3 or probable Omicron, s015A: Omicron BA.2 or clade 21L. The PDB files for the above-mentioned structures were built using the Phyre^2 a web portal for protein modeling, prediction and analysis at (http://www.sbg.bio.ic.ac.uk/phyre2) (Kelley et al., 2015).

2.4. Superimposition

The PDB files, which were built by Phyre^2, were used as input files for SuperPose web-server (version.1.0) to get 2D-difference distance matrix besides 3D-visualization based on the Template Modeling score (TM-score) retrieved by Zhang lab tool to study the structural similarity in the spike protein of each studied variant (Zhang and Skolnick, 2004; Xu and Zhang, 2010). It is worth acknowledging SuperPose v1.0 (2004) Rajarshi Maiti, Gary Van Domselaar, Haiyan Zhang, and David Wishart.

2.5. PSIPRED/(protein domain prediction)

The amino acid sequence of the spike protein of clade19A, clade20A, Omicron (BA.1 or clade 21M), Omicron (BA.3 or probable Omicron) and Omicron (BA.2 or clade 21L) were used as inputs for DomPred method as one of the prediction methods of the PSIRED workbench to compare the protein domain boundaries (Bryson et al., 2007; Buchan and Jones, 2019).

3. Results

3.1. Normal mode analysis (NMA) revealed 16 stabilizing mutations that gained environmental stability

The DynaMut results have revealed that the normal mode analysis (NMA) has predicted the thermodynamics of amino acid substitutions along the spike protein of Omicron. Out of 30 studied Omicron mutations, only two mutations, T478K and D614G, showed a tendency to increase the spike protein stability ($\Delta\Delta G$ DynaMut in kcal/mol) and flexibility ($\Delta\Delta S$ ENCoM in kcal.mol^{-1}.K^{-1}). Only four mutations, G142D,G446S,G496S, and T547K, showed a tendency to destabilize the protein with decreased molecule flexibility, the G142D has given a predicted lowest $\Delta\Delta S$ ENCoM value of -1.053 kcal.mol^{-1}.K^{-1} in decreasing the molecule flexibility. Out of the 30 studied Omicron mutations, 16 mutations predicted to have a tendency to increase the spike protein stability (e.g., A67V, T95I (giving the highest $\Delta\Delta S$ ENCoM value), G339D, S371L, S375F, N440K, S477N, T478K, Q493R, N501Y, D614G, H655Y, P681H, N856K, Q954H, and L981F). However, the other 14 mutations out of 30 predicted to show a tendency to destabilize the structure of the spike molecule (e.g., G142D, L212I, S373P, K417N, G446S, E484A, G496S, Q498R, Y505H (giving the lowest $\Delta\Delta S$ ENCoM value), T547K, N679K, N764K, D796Y and N969K). In terms of the change in the predicted flexibility, the NMA analysis revealed 18 mutations showed a tendency to decrease the spike protein flexibility (e.g., A67V, T95I, G142D, G339D, S371L, S375F, N440K, G446S, S477N, Q493R, G496S, N501Y, T547K, H655Y, P681H, N856K, Q954H, and L981F). Meanwhile, 12 mutations have shown a tendency to increase the spike molecule flexibility (e.g., L212I, S373P, K417N, T478K, E484A, Q498R, Y505H, D614G, N679K, N764K, D796Y and N969K, where E484A has given the highest increase in molecule flexibility with a $\Delta\Delta S$ ENCoM value of 0.762 kcal.mol^{-1}.K^{-1}). We have only selected the newly formed or missed interatomic interactions in the mutations that had a tendency to increase stability in the spike's structure protein to be tabulated in the last column of Table 1. Selected mutations, in Fig. 3, have a predicted tendency to increase the spike molecule stability, and hence transmissibility, with $\Delta\Delta G$ values greater than 0.5 kcal/mol except for the T478K, which has been added for its importance.

3.2. Structural similarity matrix (DALI) implies that Omicron 21L can be a new nucleus for new sub-clades

Fig. 1 shows the structural similarity matrix or heat map (Dali Z-scores) where the red color shows positive values; lighter hues for values near zero and white for negative values. The sharp drop in deviations at short distances might reach a space larger than 20 Å. The comparisons among the clades 21K, 21L and 21M versus the following previously reported 12 clades variants; Alpha or B.1.1.7 (20I/501Y.V1) or, Beta or B.1.351 (20H/501Y.V2), Gamma or P.1 (20J/501Y.V3, Delta or B.1.617.2 (21A/478K.V1), Epsilon or B.1.429 (21C/452R), Iota or B.1.526 (21F/253G.V1) showed high structural similarity neighborhood's near each other except for clade 19A, Lambda and Omicron clade 21L. The dendrogram results, as an average linkage clustering, have confirmed the structural similarity matrix (Dali Z-scores) shown in Fig. 2 a. The Z-score of 43.1 was the threshold of the structural similarity; where any score above the threshold is similar, and we judged any below the threshold dissimilar.

As a multidimensional scaling method, the correspondence analysis represents the positions of the data points with the most similar structural neighborhoods as clusters. In Fig. 2 b, the clade 19A (cluster A) was clustering distant in terms of linkage from other clusters; cluster A (19A clade), cluster B (Lambda), cluster C (clades of Omicron 21K, Eta, Epsilon, Gamma, Iota and Theta), cluster D (clade of Omicron 21L), cluster E (clades of 20A, Alpha, Beta, Delta, Kappa, and Omicron 21M). Lambda variant (cluster B) was also distant from the other clusters on the map. It was noticed that cluster C,D, and E were near neighbors. Interestingly, the Omicron clades 21M and 21K were adjacent neighbors to clades Alpha and Delta. In addition, Omicron 21L was an adjacent neighbor to clades Eta and Iota, with a relatively longer distance within its cluster, which might imply that 21L can be a new nucleus for new sub-clades.

Table 1

Normal Mode Analysis (NMA) by DynaMut for Omicron mutations.

#	Spike protein domains	Omicron mutations	ΔΔG DynaMut kcal/mol	ΔΔS ENCoM kcal.mol^{-1}.K^{-1}	If it has been found in Previous WHO or Nextstrain -named Variants	Interatomic interaction with surrounding Amino Acids for increasing flexibility only (−) Lost interatomic interaction (+) Gained interatomic interaction
1	NTD	A67V	1.794 (Stabilizing)	−0.539 (Decrease of molecule flexibility)	Eta & Iota	+I100, +G261
2	NTD	T95I	2.7 (Stabilizing)	−0.393 (Decrease of molecule flexibility)	Eta & Iota & Kappa	+Y266, +F186, +Y166
3	NTD	G142D	−0.191 (Destabilizing)	−1.053 (Decrease of molecule flexibility)	Kappa & Delta	
4	NTD	L212I	−0.386 (Destabilizing)	0.293 (Increase of molecule flexibility)	New	
5	RBD	G339D	0.017 (Stabilizing)	−0.086 (Decrease of molecule flexibility)	New	
6	RBD	S371L	0.554 (Stabilizing)	−0.057 (Decrease of molecule flexibility)	New	+S377, -S370,+V367
7	RBD	S373P	−0.314 (Destabilizing)	0.152 (Increase of molecule flexibility)	New	
8	RBD	S375F	0.65 (Stabilizing)	−0.084 (Decrease of molecule flexibility)	New	+F375 +N437
9	RBD	K417N	−0.287 (Destabilizing)	0.588 (Increase of molecule flexibility)	Beta, Delta & Gamma	
10	RBD	N440K	0.542 (Stabilizing)	−0.045 (Decrease of molecule flexibility)	New	-S438, +D442,+F486
11	RBD	G446S	−1.377 (Destabilizing)	−0.021 (Decrease of molecule flexibility)	New	
12	RBD	S477N	0.038 (Stabilizing)	−0.002 (Decrease of molecule flexibility)	Eta & Iota	
13	RBD	T478K	0.118 (Stabilizing)	0.012 (Increase of molecule flexibility)	Delta	
14	RBD	E484A	−0.728 (Destabilizing)	0.762 (Increase of molecule flexibility)	New	
15	RBD	Q493R	0.749 (Stabilizing)	−0.199 (Decrease of molecule flexibility)	New	-Y495,-Y489
16	RBD	G496S	−0.622 (Destabilizing)	−0.066 (Decrease of molecule flexibility)	New	
17	RBD	Q498R	−0.366 (Destabilizing)	0.26 (Increase of molecule flexibility)	New	
18	RBD	N501Y	0.502 (Stabilizing)	−0.211 (Decrease of molecule flexibility)	Alpha & Beta & Gamma & Theta	+Q498
19	RBD	Y505H	−0.937 (Destabilizing)	0.414 (Increase of molecule flexibility)	New	
20	Pre-Furin	T547K	−0.028 (Destabilizing)	−0.148 (Decrease of molecule flexibility)	New	
21	Pre-Furin	D614G	0.292 (Stabilizing)	0.103 (Increase of molecule flexibility)	Found in all	
22	Pre-Furin	H655Y	0.839 (Stabilizing)	−0.35 (Decrease of molecule flexibility)	Gamma	+A694,T696
23	Pre-Furin	N679K	−0.309 (Destabilizing)	0.053 (Increase of molecule flexibility)	New	
24	Pre-Furin	P681H	0.297 (Stabilizing)	−0.035 (Decrease of molecule flexibility)	Alpha & Theta	
25	Post-Furin	N764K	−0.066 (Destabilizing)	0.322 (Increase of molecule flexibility)	New	
26	Fusin Peptide	D796Y	−0.02 (Destabilizing)	0.139 (Increase of molecule flexibility)	New	
27	FP-HR1 linker (S2)	N856K	2.105 (Stabilizing)	−0.729 (Decrease of molecule flexibility)	New	+M470,V963
28	heptad repeat 1 (HR1)	Q954H	0.235 (Stabilizing)	−0.031 (Decrease of molecule flexibility)	New	
29	heptad repeat 1 (HR1)	N969K	−0.146 (Destabilizing)	0.31 (Increase of molecule flexibility)	New	
30	heptad repeat 1 (HR1)	L981F	0.937 (Stabilizing)	−0.662 (Decrease of molecule flexibility)	New	+S446,+L977,+I993,+E748,+R983, +L984, +A989

3.3. DomPred-boundaries for clades 20A, 21K, 21M, and 21L in PSI-BLAST alignment profile differed totally from clade 19A

Figs. 4–8 showed the protein domain boundary prediction using the Dom-Pred server for clades 19A, 20A,21K, 21L and 21M. For clade 19A, the putative domain boundaries in PSI-BLAST alignment profile by DPS were 6 domains; the domain boundary locations predicted DPS were at 163, 258, 436, 718, and 1113. For clades 20A, 21K, 21M, and 21L, the putative domain boundaries in PSI-BLAST alignment profile differed totally from clade 19A by DPS were 5 domains for each clade; the domain boundary locations predicted DPS were at (151, 269, 649, and 1113). (153, 265, 646, and 1110)‹ (151, 264, 600, and 1107), and (187, 268, 644, and 1113) respectively.
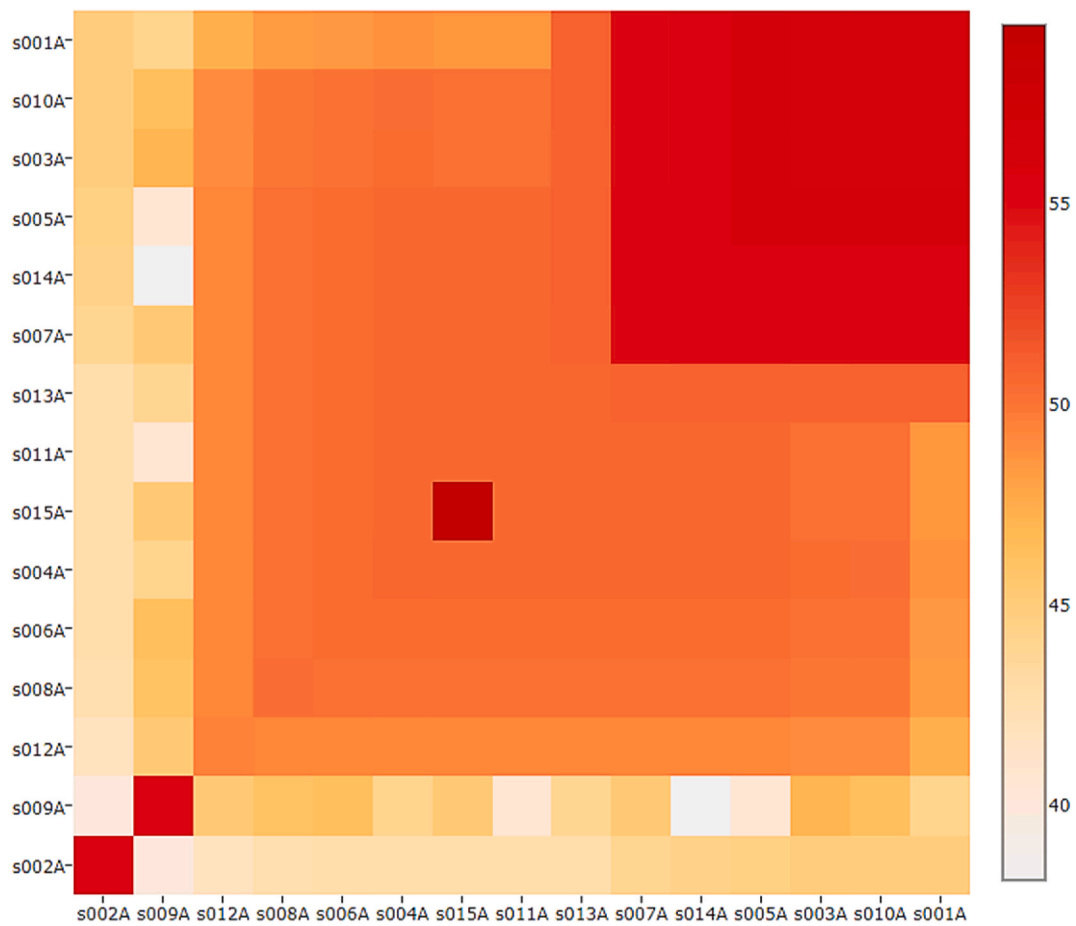
Fig. 1. Structural Similarity Matrix or Heat Map (Dali Z-scores): The color intensity shows the DALI-score, red for positive, lighter hues are approximately zero and white for negative values. The Sharp drop for deviations at short distances and damping of contributions from distances is longer than 20 Å.
* The codes on the x and y-axes of the matrix include s001A: 20A, s002A: 19A, s003A: Beta, s004A: Iota, s005A: Delta, s006A: Gamma, s007A: Alpha, s008A: Epsilon, s009A: Lambda, s010A: Kappa, s011A: Eta, s012A: Theta, s013A: Omicron BA.1 or clade 21K, s014A: Omicron BA.3 or clade 21M, s015A: Omicron BA.2 or clade 21L.
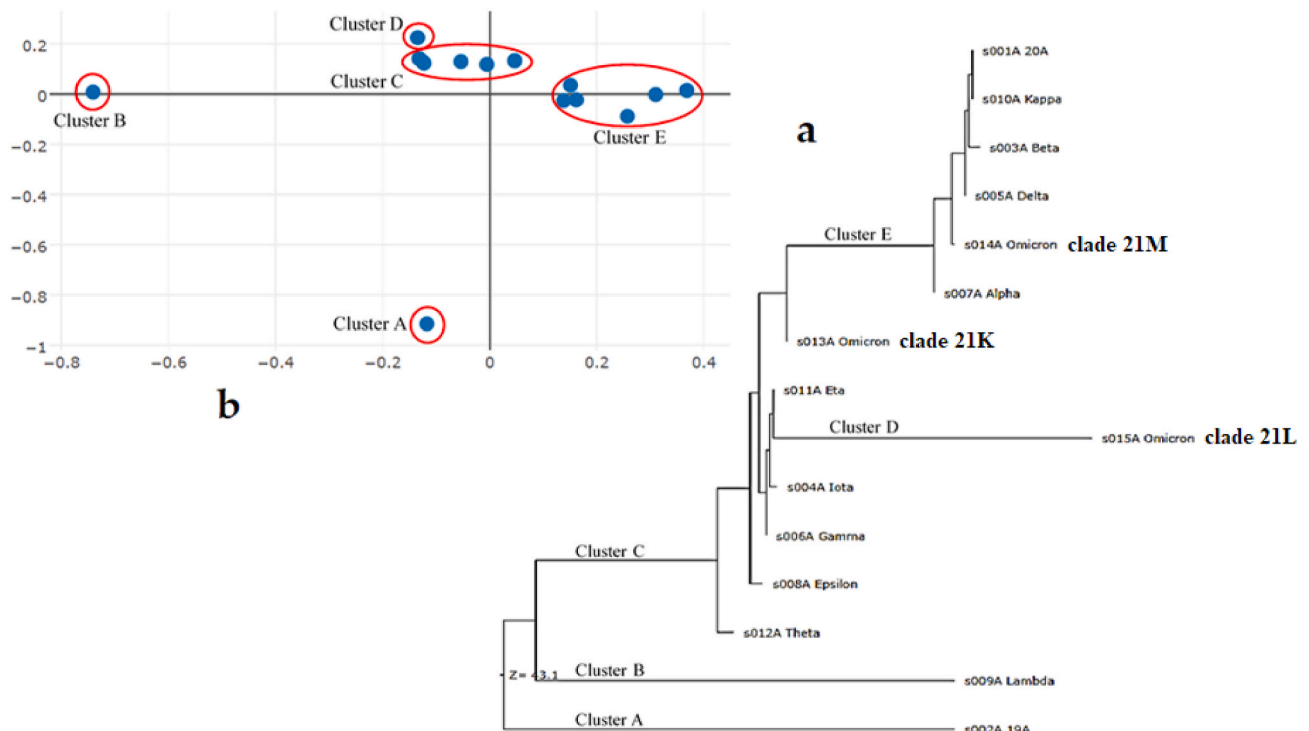
Fig. 2. a: Structural Similarity Dendrogram of 15 studied structures. Average linkage clustering of the structural similarity matrix derived from the dendrogram (Dali Z score = 43.1). **b**: Correspondence Analysis is a multidimensional scaling method. It positions data points with the most similar structural neighborhoods as clusters.
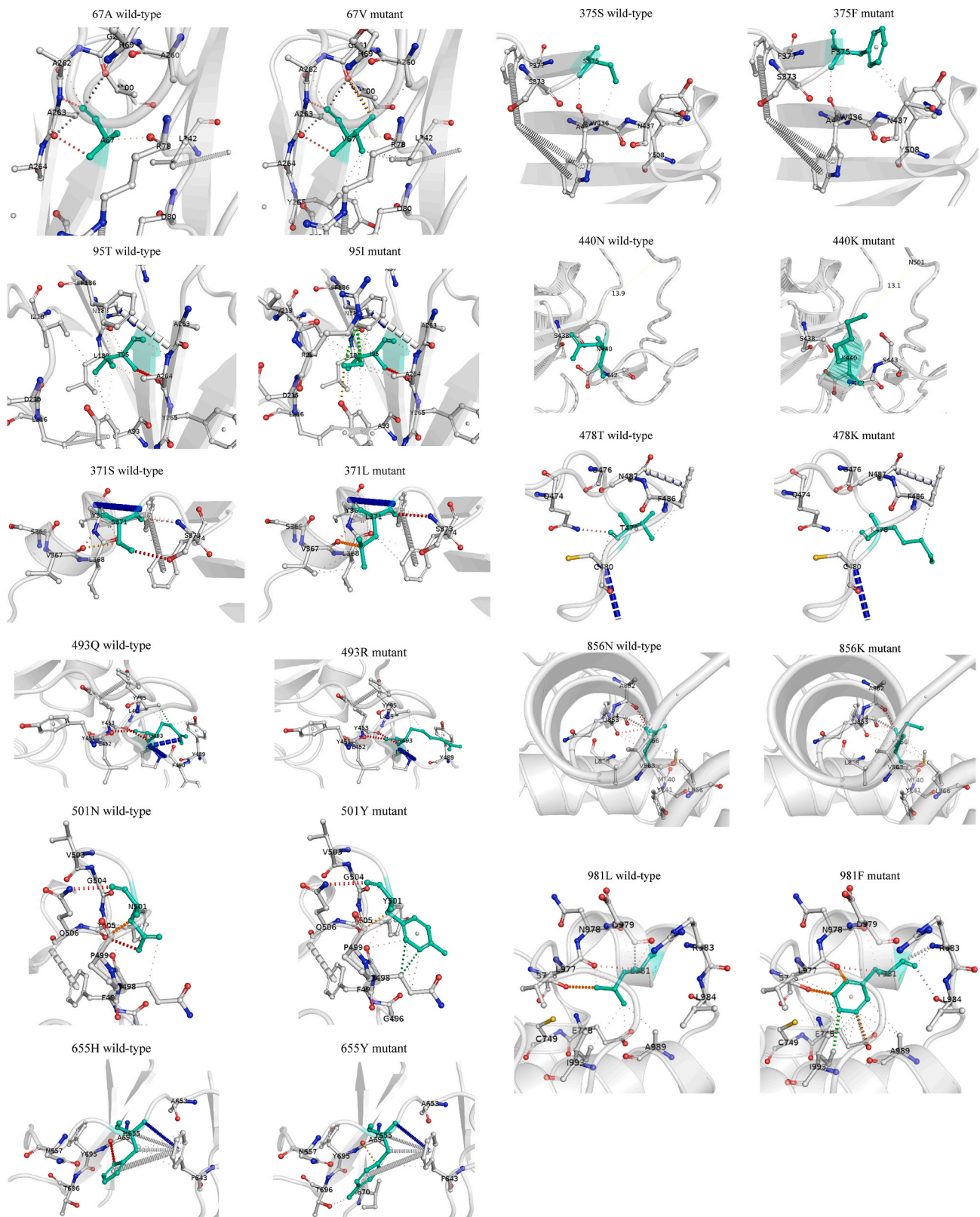
Fig. 3. The interatomic interactions as predicted by DynaMut only for the mutations showed a predicted tendency to increase the spike molecule stability; the wild-type and mutant residues are colored in light-green and represented as sticks alongside the surrounding residues involved in any interactions.
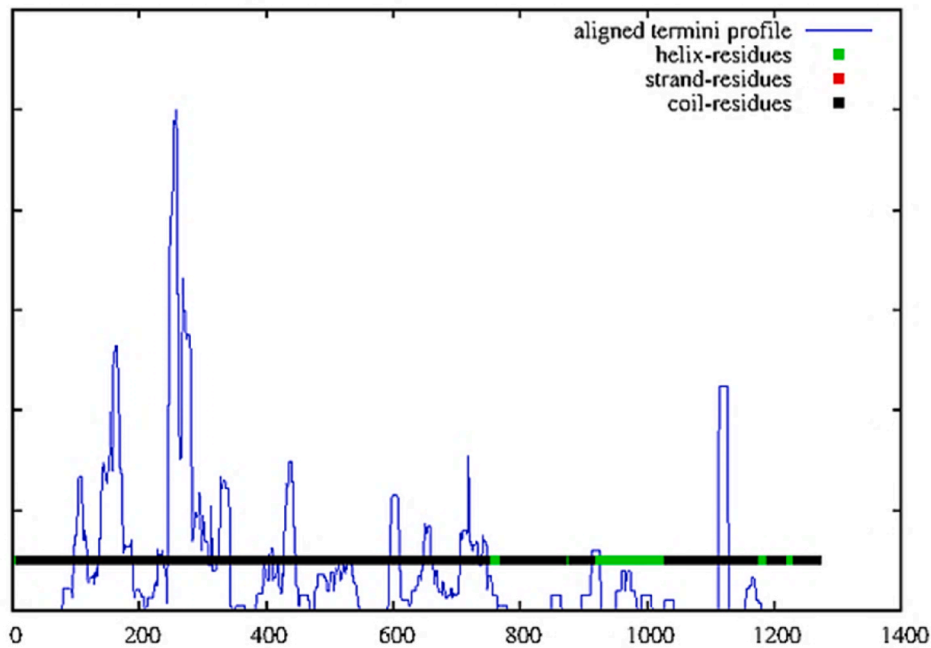
Fig. 4. Protein Domain Boundary Prediction Using the Dom-Pred Server for **clade 19A**. Putative domain boundaries in PSI-BLAST alignment profile: Number of predicted domains by DPS: **6**. Domain Boundary locations predicted DPS: 163, 258, 436, 718, and 1113.
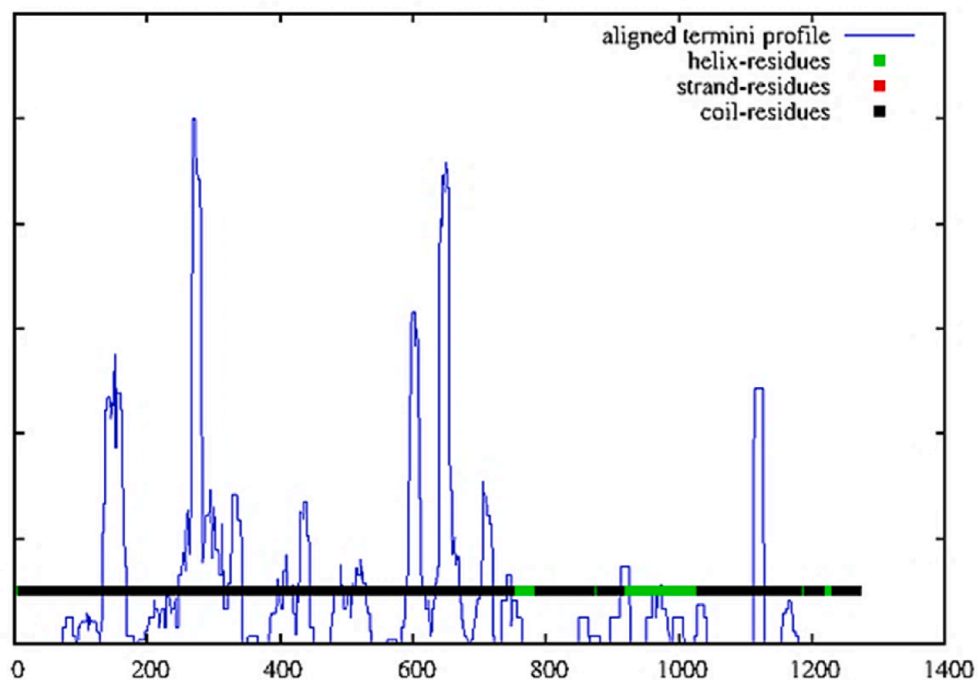

Fig. 5. Protein Domain Boundary Prediction Using the Dom-Pred Server for **clade 20A.** Putative domain boundaries in PSI-BLAST alignment profile: Number of predicted domains by DPS: **5**. Domain Boundary locations predicted DPS: 151, 269, 649, and 1113.

3.4. Superpose showed that Omicron clade 21L has changed the linear epitope of the amino acids G252 → S255 of the spike protein to a helix structure

In Template Modeling (TM), a 2D-difference distance matrices by SuperPose v.1.0 with 3D visualization by the TM-align for the spike protein variants were generated for the reference structures for clades 19A (wild type without D614G) and 20A (with D614G) both appeared in blue versus each structure of Omicron clades 21K, 21M, 21L (with

D614G and T478K) appeared in red, the two colors appeared overlaid each other to shows the structural change in angstrom (Å) unit. It is worth mentioning that the mutation T478K has the third highest frequency in the world with a total frequency with 1885726 as shown on the EpiCoV database of GISAID (https://www.gisaid.org/) and (https://users.math.msu.edu/users/weig/SARSCoV-2_Mutation_Tracker.html) as the data was last updated on January 10, 2022. The outcome of SuperPose is shown in Table 2. The summary of superimposition results include the structure of clade 19A versus Omicron clade 21K, 21M, and
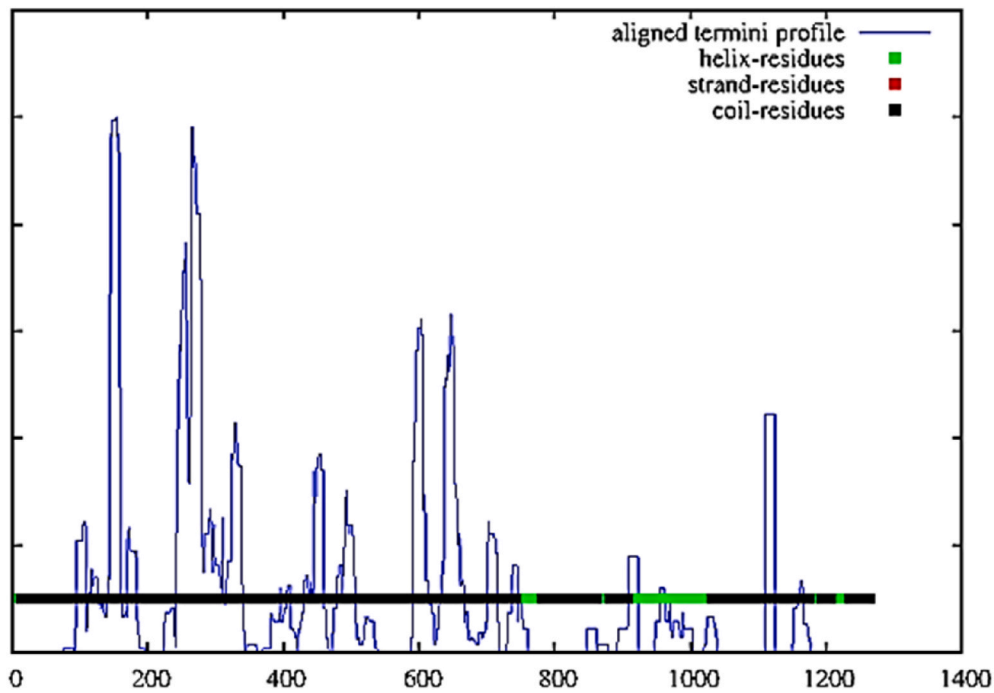
Fig. 6. Protein Domain Boundary Prediction Using the Dom-Pred Server for clade **21K**. Putative domain boundaries in PSI-BLAST alignment profile: Number of predicted domains by DPS: **5**. Domain Boundary locations predicted DPS: 153, 265, 646, and 1110.
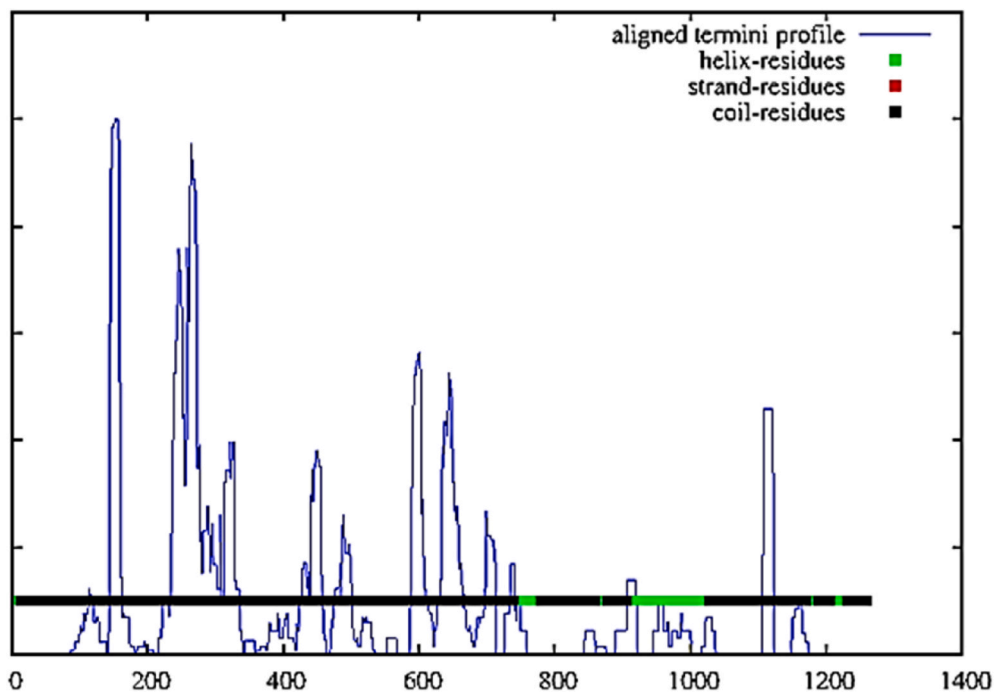

Fig. 7. Protein Domain Boundary Prediction Using the Dom-Pred Server for clade **21M**. Putative domain boundaries in PSI-BLAST alignment profile: Number of predicted domains by DPS: **5**. Domain Boundary locations predicted DPS: 151, 264, 600, and 1107.

21L have shown that the 3D-structural change was significant with a change range from 1.5 Å to 12 Å all over the Spike protein molecule. The superimposition results that include the structure of clades 19A and 20A versus Omicron clade 21K, 21M, and 21L have interestingly shown that the Omicron clade 21L has changed the linear epitope of the amino acids G252 → S255 of the spike protein to a helix structure, as shown in Fig. 9, with a change range reached to 12 Å which was not the case in Omicron clades 21K and 21M when compared with clade 19A and 20A.

4. Discussion

The spread Omicron (B.1.1.529) variant that first appeared on November 2021 in South Africa and Botswana is expected to be the most prevalent variant all over the world (Mannar et al., 2022). Actually, there are three Nextstrain (or PANGO Lineage) clades represent the Omicron variant now, which are 21K (BA.1), 21M (BA.3) and 21L (BA.2) from www.nextstrain.org. In this study, the thermodynamics attributes
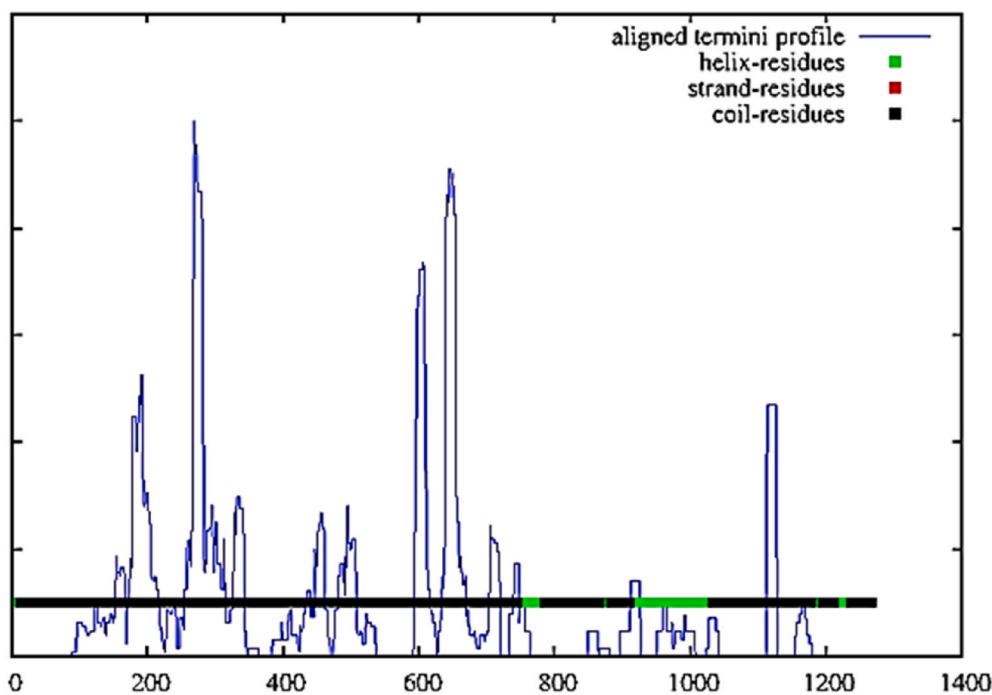
Fig. 8. Protein Domain Boundary Prediction Using the Dom-Pred Server for clade **21L**. Putative domain boundaries in PSI-BLAST alignment profile: Number of predicted domains by DPS: **5**. Domain Boundary locations predicted DPS: 187, 268, 644, and 1113.

of Omicron have been studied and compared against the wild type clade 19A (without substitution D614G) and 20A (with substitution D614G). We expect the new mutations of Omicron to play a crucial role in changing the physiochemical properties of the epitopes of the spike protein of SARS-CoV-2. The physiochemical properties of the epitopes include hydrophilicity, surface flexibility, surface accessibility, and antigenicity. Any changes in the protein physiochemical properties as an antigen usually lead to changes in the protein stability and antibody binding affinity. Usually, the changed spike protein stability is linked to changed transmissibility of the SARS-CoV-2 virus; meanwhile, a changed spike protein flexibility is linked with a changed affinity to bind antibodies. In this study, four platforms have been used to study the predicted effect of the new Omicron mutations, included Normal Mode Analysis (NMA) , DALI, DomPred Boundaries, and SuperPose. We have discussed the results of each platform in the following sections.
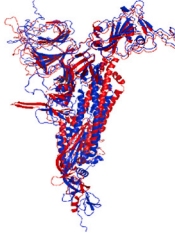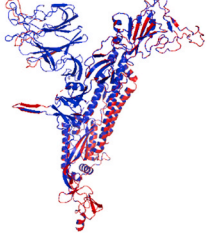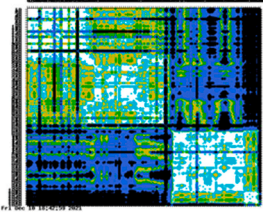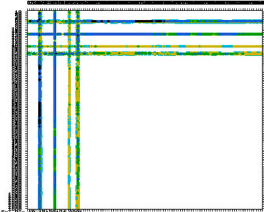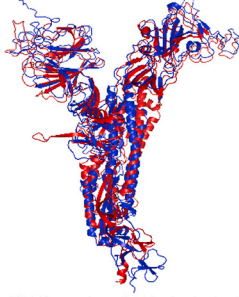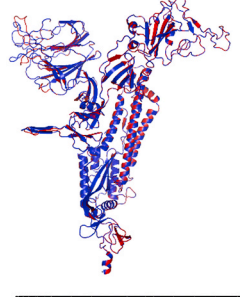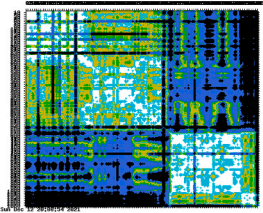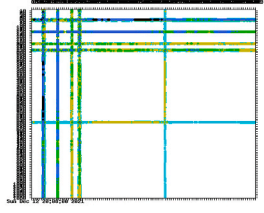
4.1. Normal mode analysis (NMA)

The results of the DynaMut have summarized in Table 1, where only two mutations, T478K and D614G, showed a predicted tendency to increase the stability and flexibility of the Omicron spike protein. The D614G mutation is a well-studied substitution that is a dominant in all clades appeared after the wild type clade 19A. However, few reports have highlighted the role of T478K mutation in the receptor-binding motif (RBM) in the spike protein. Some previous studies have reported that the mutation T478K, which has been appeared in Delta variant, has a unique role in increasing the virus infectivity and pathogenicity (Jhun et al., 2021b) which agrees with the findings of normal mode analysis in this study. Four mutations; G142D (appeared in Kappa and Delta Variants), G446S, G496S, and T547K (three new in the Omicron variant) have shown a predicted tendency to destabilize and decrease the Omicron spike protein flexibility. It was reported that the mutation G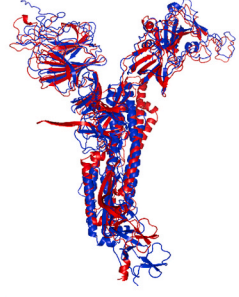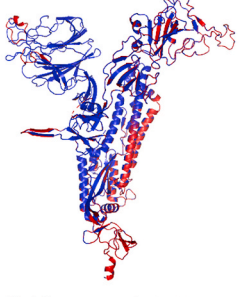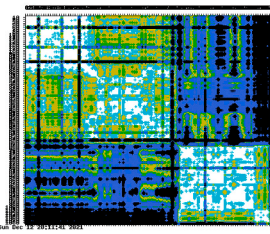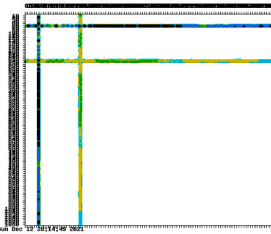142D with other mutations had an effect in improving the virus transmissibility, reducing the neutralization efficiency of receptor binding domain directed monoclonal antibodies and escaping from Human Leucocytes Antigen-24-induced cellular immunity, and reducing the efficacy of neutralization of convalescent sera and vaccines (Dubey

et al., 2021). The mutation G142D has been reported to play a role in SARS-CoV-2 transmission from human to the mink animal in the United States (Cai and Cai, 2021). No available data about G446 and T547K however, it was reported that G496S and Q493R, and Q498R have formed new salt bridges with residue D38 of the Angiotensin Converting Enzyme (ACE2) (Mannar et al., 2022). It is worth mentioning here that the Omicron mutation E484A has increased the spike protein flexibility in a higher value of $\Delta\Delta S$ ENCoM value, giving 0.762 kcal.mol^{-1}.K^{-1} when compared with E484Q with a value of $\Delta\Delta S$ ENCoM value giving 0.389 kcal.mol^{-1}.K^{-1} and E484K with a value of $\Delta\Delta S$ ENCoM value giving 0.490 kcal.mol^{-1}.K^{-1} (Al-Zyoud and Haddad, 2021a), the mutation in amino acid 484 was previously reported as a mutation confer antibody resistance (Liu et al., 2021a). Specifically, it was reported that the mutation Q493R has conferred a simultaneous resistance to the treatment by the monoclonal antibodies, namely *Bamlanivimab* and *Etesivimab* (Focosi et al., 2021). There is no room to discuss each mutation alone; however, we can say that a total number of 16 stabilizing mutations in competition with 14 destabilizing mutations implies the general tendency of stabilizing net effect in the spike protein molecule of the Omicron variant. For example, the group of mutations K417N, N440K, G446S and S477N have a net effect in increasing the spike molecule flexibility, which will affect the antibodies evasion because of the presence of K417N that make our findings agrees with a previous free energy study (Fratev, 2021). The NMA analysis showed a change in the interaction distance between 501N and 440N (wild-type) from 13.9 Å to 13.1 Å between 501N and 440K (mutant). The group of the neighboring mutations G339D, S371L, S373P, S375F showed a slight net effect to increase spike protein flexibility because of the presence of S373P, which had a predicted value of $\Delta\Delta S$ ENCoM giving 0.152 kcal.mol^{-1}.K^{-1}, there are no peer reviewed reports about this group of mutation at the time of writing. The group of mutations N501Y, Y505H and T547K had a predicted net effect of increasing the spike protein flexibility because of the presence of mutation Y505H with $\Delta\Delta S$ ENCoM value giving 0.414 kcal.mol^{-1}.K^{-1} that is a relatively high value that may affect the antibodies neutralization, and N501Y which has been reported as a substitution that enhances the transmission and infection of SARS-CoV-2 (Liu et al., 2021b). The findings of our NMA analysis

Table 2

SuperPose 2D-difference distance matrix with 3D-visualization by the TM-align for the spike protein from clades 19A, 20A, Omicron (21K, 21 M and 21L). The small box of horizontal plus marks is the color-coded superimposition marker in the angstrom unit (Å).

Compared Proteins	3D-Visualization of 19A blue colour	3D-Visualization of 20A blue color
Omicron BA.1 or 21K red color		
Outcome of SuperPose v 1.0		
Omicron BA.3 or 21M red color		
Outcome of SuperPose v 1.0		
Omicron BA.2 or 21L red color		
Outcome of SuperPose v 1.0		

have revealed an important mutation N764K that showed a similar effect of the above-mentioned mutation Y505H which agrees with a previous comparative computational study (Kumar et al., 2021). The last increasing flexibility mutation we want to highlight in this section is N969K with a $\Delta\Delta S$ ENCoM value of 0.31 kcal.mol^{-1}.K^{-1} in the heptad repeat 1 domain (HR1) which has a role in the fusion process between the viral and cellular membranes (Xia et al., 2020). At last, as there are 16 stabilizing mutations in the spike protein of the Omicron variant, this expected to explain the higher transmissibility over other variants, especially Delta, because of a new environmental stability feature.

The T478K was reported as a mutation in the functional Receptor Binding Domain (RBD) associated with the infectivity and pathogenicity of any variants, where is it found (Jhun et al., 2021c). The biological meaning of this is that the previously reported mutation D614G was the reason for giving the spike protein a tendency to increase transmissibility and alter BNT162b2 vaccine-elicited neutralization when compared with the wild type spike protein without D614G (Al-Zyoud and Haddad, 2021a; Plante et al., 2021; Isabel et al., 2020).

4.2. Similarity matrix (DALI)

The heat map analysis performed via DALI server showed that the three dimensional structure (3D-structure) of the spike protein in clades 20A, 19A, Beta, Iota, Delta, Gamma, Alpha, Epsilon, Lambda, Kappa, Eta, Theta, Omicron BA.1 or clade 21K, Omicron BA.3 or clade 21M, Omicron BA.2 or clade 21L differed from the wild type clade 19A because of mutation D614G in all of them. In addition, the structural similarity dendrogram, which is usually derived by average linkage clustering of the Structural Similarity Matrix (Dali Z score = 43.1) , and the correspondence analysis showed forming clusters A, B, C, D and E. The cluster E showed that Omicron 21M was a very adjacent neighbor between clades Alpha and Delta. The cluster C showed that the clades Omicron 21K was an adjacent neighbor to clades Alpha and Eta, where clusters E and C were close to each other. This implies that the Omicron variant has the most important mutations from clades Alpha, Delta and Eta. For example; the mutations A67V (found in Eta), T95I (found in Eta), G142E (found in Delta), S477N (found in Eta), T478K (found in Delta), N501Y (found in Alpha), D614G (Al-Zyoud and Haddad, 2021b) (found in every clade after the wild type), and P681H (found in Alpha). Cluster D showed a similarity of Omicron 21L between Eta and Iota, this variant has a new branch alone when compared with other Omicron variants and even with other previous clades, expected to have a biological effect on pathogenicity because of mutation H655Y that is present in clade Gamma and was reported to confer an escape from human

monoclonal antibodies when it has been studied in the domestic cats (Braun et al., 2021). This highlights the importance of studying the SARS-CoV-2 transmission from humans to animals and back to humans (Sharun et al., 2021). The mice were a reported as an animal model to study the respiratory dysfunction for SARS-CoV-2 (Gan et al., 2021). We have previously reported in details the similarity matrices for other clades and variants (Al-Zyoud and Haddad, 2021a).

4.3. DomPred-boundaries

The findings from protein domain boundary prediction analysis by the DomPred-bounderis has revealed that clade 19A has predicted domains by DPS of 6 domain boundary locations meanwhile the number of predicted domains by DPS in clade 20A and Omicron clades 21K, 21M and 21L was 5 for each of them. This implies a significant change at the spike protein domains boundaries and hence the final 3D-structure. Interestingly, the clade 21L showed a notable shift in the domain boundary locations by DPS if compared with the wild type clade 19A and even with other Omicron clades 21K and 21M. The results of DomPred-Boundaries highlight the role of the new mutations in Omicron in changing the locations of the proteases needed for spike protein to be cleaved before entering the human cells, which change the antigenicity of the spike protein. The findings of the DomPred-bounderies agree with the findings of SuperPose as discussed in the following section.

4.4. SuperPose

It was obvious from the analysis of SuperPose v 1.0 and TM-score that the 3D-structures of the three Omicron clades; 21K, 21M and 21L have completely changed because of the well-studied mutation D614G when compared to the wild type clade 19A. The Omicron 21L has the highest change at the level of the 3D-structure in Angstrom units as the change has reached 12 Å if compared with other Omicron clades 21K and 21M which reached to a range of change between 7 Å to 9 Å when compared with the clade 20A to exclude the effect of the mutation D614G. The remarkable finding of SuerPose work was the potential linear epitope G252 → S255 that has been changed in the spike protein of the Omicron clade 21L to a helix structure when compared with the wild type clade 19A and the subsequent clades 20A. 21K, and 21M as shown in Table 2 and Fig. 9. We expect this finding to play a role in the evasion of the antibodies neutralization because the linear epitope G252 → S255 is a part of a longer potential B-cell epitope sequence 242→263; this agrees with a previously reported study about subunit vaccine
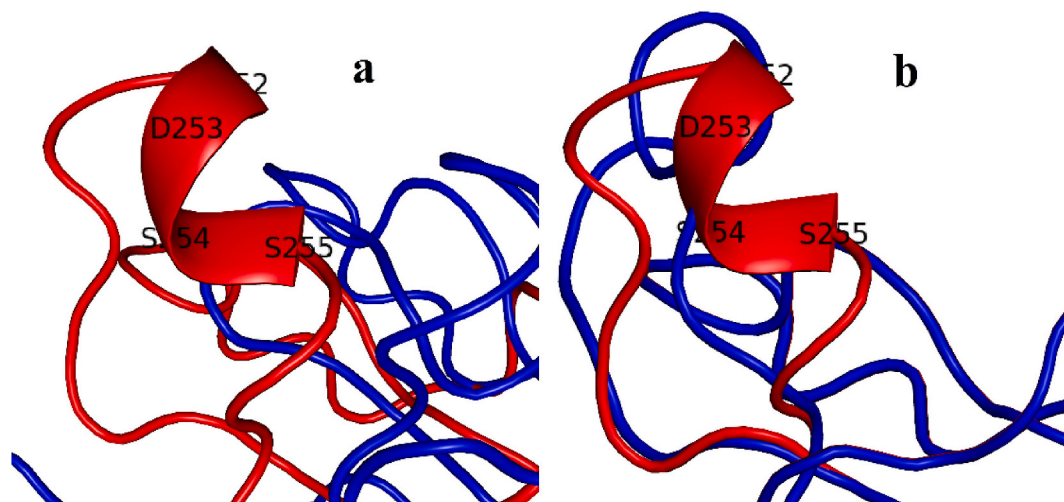

Fig. 9. The linear epitope **G252 → S255** has been changed in Omicron to the helix structure. **a)** The red structure represents Omicron 21L versus the blue structure represents clade 19A. **b)** The red structure represents Omicron 21L versus the blue structure represents clade 20A.

predicted design which has taken into consideration the physiochemical attributes of different B-cells and T-cells epitopes (Sarma et al., 2021).

5. Conclusion

The researchers all over the world are concerned with discerning the new SARS-CoV-2 Omicron variant. The thermodynamics prediction in this study revealed that there are 16 stabilizing mutations in the spike protein of the Omicron variant. This might explain the higher airborne transmissibility over other variants due to probably new tougher environmental stability feature. We expect the new Omicron mutations to play a crucial role in changing the physiochemical attributes of the B-cells epitopes in the spike protein of SARS-CoV-2. The most important finding of SuerPose work was the linear sequence G252 → S255 of the Omicron clade 21L which was reported as a part of the potential B-cells epitope in the spike protein that has been changed to a helix structure when compared with the wild type clade 19A and its successive clades 20A, 21K and 21M. We conclude, from our analysis of similarity matrix (DALI), that the similarity score of the Omicron variant 21L was between Eta and Iota, this variant has a new branch alone when compared with other Omicron variants and even other previous clades, this is expected to have a biological effect on its pathogenicity especially because of the presence of mutation H655Y which is also present in clade Gamma and was reported to confer an escape from human monoclonal antibodies when it has been previously studied in the domestic cats. The all findings in this study were based on predictions from specialized computational methods. As the computational predictions are universally verifiable claims, biological methods are always recommended.

CRediT authorship contribution statement

Walid Al-Zyoud: Conceptualization, Writing – original draft, Supervision, Validation, Writing – review & editing. **Hazem Haddad:** Conceptualization, Writing – original draft, Supervision, Validation, Writing – review & editing.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Al-Zyoud, W., Haddad, H., 2021a. Dynamics prediction of emerging notable spike protein mutations in SARS-CoV-2 implies a need for updated vaccines. Biochimie 191, 91–103. https://doi.org/10.1016/J.BIOCHI.2021.08.011.

Al-Zyoud, W., Haddad, H., 2021b. Mutational sensitivity of D614G in spike protein of SARS-CoV-2 in Jordan. Biochem. Biophys. Rep. 25, 100896 https://doi.org/10.1016/j.bbrep.2020.100896.

Braun, K.M., Moreno, G.K., Halfmann, P.J., Hodcroft, E.B., Baker, D.A., Boehm, E.C., Weiler, A.M., Haj, A.K., Hatta, M., Chiba, S., Maemura, T., Kawaoka, Y., Koelle, K., O'Connor, D.H., Friedrich, T.C., 2021. Transmission of SARS-CoV-2 in domestic cats imposes a narrow bottleneck. PLoS Pathog. 17 https://doi.org/10.1371/JOURNAL.PPAT.1009373.

Bryson, K., Cozzetto, D., Jones, D., 2007. Computer-assisted protein domain boundary prediction using the DomPred server. Curr. Protein Pept. Sci. 8, 181–188. https://doi.org/10.2174/138920307780363415.

Buchan, D.W.A., Jones, D.T., 2019. The PSIPRED protein analysis workbench: 20 years on. Nucleic Acids Res. 47, W402–W407. https://doi.org/10.1093/NAR/GKZ297.

Cai, H.Y., Cai, A., 2021. SARS-CoV2 spike protein gene variants with N501T and G142D mutation–dominated infections in mink in the United States. J. Vet. Diagn. Invest. 33, 939–942. https://doi.org/10.1177/10406387211023481.

Cameroni, E., Bowen, J.E., Rosen, L.E., Saliba, C., Zepeda, S.K., Culap, K., Pinto, D., VanBlargan, L.A., De Marco, A., di Iulio, J., Zatta, F., Kaiser, H., Noack, J., Farhat, N., Czudnochowski, N., Havenar-Daughton, C., Sprouse, K.R., Dillen, J.R., Powell, A.E., Chen, A., Maher, C., Yin, L., Sun, D., Soriaga, L., Bassi, J., Silacci-Fregni, C.,

Gustafsson, C., Franko, N.M., Logue, J., Iqbal, N.T., Mazzitelli, I., Geffner, J., Grifantini, R., Chu, H., Gori, A., Riva, A., Giannini, O., Ceschi, A., Ferrari, P., Cippà, P.E., Franzetti-Pellanda, A., Garzoni, C., Halfmann, P.J., Kawaoka, Y., Hebner, C., Purcell, L.A., Piccoli, L., Pizzuto, M.S., Walls, A.C., Diamond, M.S., Telenti, A., Virgin, H.W., Lanzavecchia, A., Snell, G., Veesler, D., Corti, D., 2021. Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. Nature. https://doi.org/10.1038/D41586-021-03825-4.

Chen, R.E., Zhang, X., Case, J.B., Winkler, E.S., Liu, Y., VanBlargan, L.A., Liu, J., Errico, J.M., Xie, X., Suryadevara, N., Gilchuk, P., Zost, S.J., Tahan, S., Droit, L., Turner, J.S., Kim, W., Schmitz, A.J., Thapa, M., Wang, D., Boon, A.C.M., Presti, R.M., O'Halloran, J.A., Kim, A.H.J., Deepak, P., Pinto, D., Fremont, D.H., Crowe, J.E., Corti, D., Virgin, H.W., Ellebedy, A.H., Shi, P.Y., Diamond, M.S., 2021. Resistance of SARS-CoV-2 variants to neutralization by monoclonal and serum-derived polyclonal antibodies, 2021 Nat. Med. 274, 717–726. https://doi.org/10.1038/s41591-021-01294-w, 27.

Dubey, A., Choudhary, S., Kumar, P., Tomar, S., 2021. Emerging SARS-CoV-2 variants: genetic variability and clinical implications, 2021 Curr. Microbiol. 791, 1–18. https://doi.org/10.1007/S00284-021-02724-1, 79.

Focosi, D., Novazzi, F., Genoni, A., Dentali, F., Gasperina, D.D., Baj, A., Maggi, F., 2021. Emergence of SARS-COV-2 spike protein escape mutation Q493R after treatment for COVID-19, emerg. Inf. Disp. 27, 2728. https://doi.org/10.3201/EID2710.211538.

Fratev, F., 2021. N501Y and K417N mutations in the spike protein of SARS-CoV-2 alter the interactions with both hACE2 and human-derived antibody: a free energy of perturbation retrospective study. J. Chem. Inf. Model. 61, 6079–6084. https://doi.org/10.1021/ACS.JCIM.1C01242/SUPPL_FILE/CI1C01242_SI_002.ZIP.

Gan, E.S., Syenina, A., Linster, M., Ng, B., Zhang, S.L., Watanabe, S., Rajarethinam, R., Tan, H.C., Smith, G.J.D., Ooi, E.E., 2021. A mouse model of lethal respiratory dysfunction for SARS-CoV-2 infection. Antivir. Res. 193, 105138 https://doi.org/10.1016/J.ANTIVIRAL.2021.105138.

Greaney, A.J., Loes, A.N., Crawford, K.H.D., Starr, T.N., Malone, K.D., Chu, H.Y., Bloom, J.D., 2021. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. Cell Host Microbe 29, 463–476. https://doi.org/10.1016/J.CHOM.2021.02.003 e6.

Isabel, S., Graña-Miraglia, L., Gutierrez, J.M., Bundalovic-Torma, C., Groves, H.E., Isabel, M.R., Eshaghi, A., Patel, S.N., Gubbay, J.B., Poutanen, T., Guttman, D.S., Poutanen, S.M., 2020. Evolutionary and structural analyses of SARS-CoV-2 D614G spike protein mutation now documented worldwide. Sci. Rep. 10, 14031 https://doi.org/10.1038/s41598-020-70827-z.

Jhun, H., Park, H.Y., Hisham, Y., Song, C.S., Kim, S., 2021a. SARS-CoV-2 delta (B.1.617.2) variant: a unique T478K mutation in receptor binding motif (RBM) of spike gene. Immune Net. 21 https://doi.org/10.4110/IN.2021.21.E32.

Jhun, H., Park, H.Y., Hisham, Y., Song, C.S., Kim, S., 2021b. SARS-CoV-2 delta (B.1.617.2) variant: a unique T478K mutation in receptor binding motif (RBM) of spike gene. Immune Net. 21 https://doi.org/10.4110/IN.2021.21.E32.

Jhun, H., Park, H.Y., Hisham, Y., Song, C.S., Kim, S., 2021c. SARS-CoV-2 delta (B.1.617.2) variant: a unique T478K mutation in receptor binding motif (RBM) of spike gene. Immune Net. 21 https://doi.org/10.4110/IN.2021.21.E32.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., Sternberg, M.J.E., 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. 10, 845–858. https://doi.org/10.1038/nprot.2015.053.

Kumar, S., Thambiraja, T.S., Karuppanan, K., Subramaniam, G., 2021. Omicron and Delta variant of SARS-CoV-2: a comparative computational study of spike protein. J. Med. Virol. https://doi.org/10.1002/JMV.27526.

Liu, L., Iketani, S., Guo, Y., Chan, J.F.-W., Wang, M., Liu, L., Luo, Y., Chu, H., Huang, Y., Nair, M.S., Yu, J., Chik, K.K.-H., Yuen, T.T.-T., Yoon, C., To, K.K.-W., Chen, H., Yin, M.T., Sobieszczyk, M.E., Huang, Y., Wang, H.H., Sheng, Z., Yuen, K.-Y., Ho, D. D., 2021a. Striking antibody evasion manifested by the Omicron variant of SARS-CoV-2. Nat. 2021 1–8. https://doi.org/10.1038/s41586-021-04388-0.

Liu, Y., Liu, J., Plante, K.S., Plante, J.A., Xie, X., Zhang, X., Ku, Z., An, Z., Scharton, D., Schindewolf, C., Widen, S.G., Menachery, V.D., Shi, P.Y., Weaver, S.C., 2021b. The N501Y spike substitution enhances SARS-CoV-2 infection and transmission. Nat. 2021 1–6. https://doi.org/10.1038/s41586-021-04245-0.

Malonis, R.J., Lai, J.R., Vergnolle, O., 2019. Peptide-based vaccines: current progress and future challenges. Chem. Rev. 120, 3210–3229. https://doi.org/10.1021/ACS.CHEMREV.9B00472.

Mannar, D., Saville, J.W., Zhu, X., Srivastava, S.S., Berezuk, A.M., Tuttle, K.S., Marquez, A.C., Sekirov, I., Subramaniam, S., 2022. SARS-CoV-2 Omicron variant: antibody evasion and cryo-EM structure of spike protein–ACE2 complex. Science. https://doi.org/10.1126/SCIENCE.ABN7760.

Piccoli, L., Park, Y.J., Tortorici, M.A., Czudnochowski, N., Walls, A.C., Beltramello, M., Silacci-Fregni, C., Pinto, D., Rosen, L.E., Bowen, J.E., Acton, O.J., Jaconi, S., Guarino, B., Minola, A., Zatta, F., Sprugasci, N., Bassi, J., Peter, A., De Marco, A., Nix, J.C., Mele, F., Jovic, S., Rodriguez, B.F., Gupta, S.V., Jin, F., Piumatti, G., Lo Presti, G., Pellanda, A.F., Biggiogero, M., Tarkowski, M., Pizzuto, M.S., Cameroni, E., Havenar-Daughton, C., Smithey, M., Hong, D., Lepori, V., Albanese, E., Ceschi, A., Bernasconi, E., Elzi, L., Ferrari, P., Garzoni, C., Riva, A., Snell, G., Sallusto, F., Fink, K., Virgin, H.W., Lanzavecchia, A., Corti, D., Veesler, D., 2020. Mapping neutralizing and immunodominant sites on the SARS-CoV-2 spike receptor-binding domain by structure-guided high-resolution serology. Cell 183, 1024–1042. https://doi.org/10.1016/J.CELL.2020.09.037 e21.

Planas, D., Veyer, D., Baidaliuk, A., Staropoli, I., Guivel-Benhassine, F., Rajah, M.M., Planchais, C., Porrot, F., Robillard, N., Puech, J., Prot, M., Gallais, F., Gantner, P., Velay, A., Le Guen, J., Kassis-Chikhani, N., Edriss, D., Belec, L., Seve, A., Courtellemont, L., Péré, H., Hocqueloux, L., Fafi-Kremer, S., Prazuck, T., Mouquet, H., Bruel, T., Simon-Lorière, E., Rey, F.A., Schwartz, O., 2021. Reduced

sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. Nat. 2021 596, 276–280. https://doi.org/10.1038/s41586-021-03777-9, 5967871.

Plante, J.A., Liu, Y., Liu, J., Xia, H., Johnson, B.A., Lokugamage, K.G., Zhang, X., Muruato, A.E., Zou, J., Fontes-Garfias, C.R., Mirchandani, D., Scharton, D., Bilello, J. P., Ku, Z., An, Z., Kalveram, B., Freiberg, A.N., Menachery, V.D., Xie, X., Plante, K.S., Weaver, S.C., Shi, P.Y., 2021. Spike mutation D614G alters SARS-CoV-2 fitness. Nature 592, 116–121. https://doi.org/10.1038/s41586-020-2895-3.

Poland, G.A., Ovsyannikova, I.G., Kennedy, R.B., Haralambieva, I.H., Jacobson, R.M., 2011. Vaccinomics and a new paradigm for the development of preventive vaccines against viral infections, omi. A J. Integr. Biol. 15, 625–636. https://doi.org/10.1089/OMI.2011.0032.

Rodrigues, C.H.M., Pires, D.E.V., Ascher, D.B., 2018. DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. Nucleic Acids Res. 46, W350–W355. https://doi.org/10.1093/nar/gky300.

Sarma, V.R., Olotu, F.A., Soliman, M.E.S., 2021. Integrative immunoinformatics paradigm for predicting potential B-cell and T-cell epitopes as viable candidates for subunit vaccine design against COVID-19 virulence. Biomed. J. 44, 447. https://doi.org/10.1016/J.BJ.2021.05.001.

Sharun, K., Dhama, K., Pawde, A.M., Gortázar, C., Tiwari, R., Bonilla-Aldana, D.K., Rodriguez-Morales, A.J., de la Fuente, J., Michalak, I., Attia, Y.A., 2021. SARS-CoV-2 in animals: potential for unknown reservoir hosts and public health implications. Vet. Q. 41, 181. https://doi.org/10.1080/01652176.2021.1921311.

Tiwari, S.P., Fuglebakk, E., Hollup, S.M., Skjærven, L., Cragnolini, T., Grindhaug, S.H., Tekle, K.M., Reuter, N., 2014. WEBnmat v2.0: web server and services for comparing protein flexibility. BMC Bioinf. 15, 1–12. https://doi.org/10.1186/S12859-014-0427-6/FIGURES/4.

Wang, P., Nair, M.S., Liu, L., Iketani, S., Luo, Y., Guo, Y., Wang, M., Yu, J., Zhang, B., Kwong, P.D., Graham, B.S., Mascola, J.R., Chang, J.Y., Yin, M.T., Sobieszczyk, M., Kyratsous, C.A., Shapiro, L., Sheng, Z., Huang, Y., Ho, D.D., 2021. Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7, | Nat. https://doi.org/10.1038/s41586-021-03398-2, 593.

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., Yuan, M.L., Zhang, Y.L., Dai, F.H., Liu, Y., Wang, Q.M., Zheng, J.J., Xu, L., Holmes, E.C., Zhang, Y.Z., 2020. A new coronavirus associated with human respiratory disease in China. Nature 579, 265–269. https://doi.org/10.1038/s41586-020-2008-3.

Xia, S., Zhu, Y., Liu, M., Lan, Q., Xu, W., Wu, Y., Ying, T., Liu, S., Shi, Z., Jiang, S., Lu, L., 2020. Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein. Cell. Mol. Immunol. 17, 765–767. https://doi.org/10.1038/s41423-020-0374-2.

Xu, J., Zhang, Y., 2010. How significant is a protein structure similarity with TM-score = 0.5? Bioinformatics 26, 889–895. https://doi.org/10.1093/bioinformatics/btq066.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang, Y., 2014. The I-TASSER suite: protein structure and function prediction. Nat. Methods 12, 7–8. https://doi.org/10.1038/nmeth.3213.

Zhang, Y., Skolnick, J., 2004. Scoring function for automated assessment of protein structure template quality. Proteins Struct. Funct. Genet. 57, 702–710. https://doi.org/10.1002/prot.20264.