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

Biochimie 191 (2021) 91–103

Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

Research paper

Dynamics prediction of emerging notable spike protein mutations in SARS-CoV-2 implies a need for updated vaccines

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 Centre, Jordan University of Science and Technology, Irbid, 22110, Jordan

ARTICLE INFO

Article history:
Received 31 May 2021
Received in revised form
26 August 2021
Accepted 27 August 2021
Available online 8 September 2021

Keywords:
Vaccines
Spike protein
SARS-CoV-2
COVID-19
D614G
Thermodynamics
Superimposition

ABSTRACT

The spike protein of SARS-CoV-2 plays a crucial role in binding with the human cell surface, which causes its pathogenicity. This study aimed to predict molecular dynamics change of emerging variants in the spike protein. In this study, several structural biology tools, such as SuperPose, were utilized to study spike protein structures' thermodynamics, superimposition, and the spike protein disulphide bonds. This questions the current vaccines efficacies that were based on the Nextstrain clade 19A that first documented in Wuhan and lacks any variants. The prediction results of this study have exhibited the stabilizing role of the globally dominant variant, the D614G; clade 20A, and other variants in addition to their role in increasing the flexibility of the spike protein of the virus. The SuperPose findings have revealed a conformational change impact of D614G in allowing the polybasic Furin cleavage site (^{682}RRAR↓S^{686}) to be closer to the receptor-binding domain (RBD) and hence more exposed to cleavage. The presence of D614G in any clade or subclade, such as 20A, B.1.1.7 (20I/501Y.V1) or Alpha, B.1.351 (20H/501Y.V2) or Beta, P.1 (20J/501Y.V3) or Gamma, B.1.617.2 (21A/478K.V1) or Delta, has increased its stability and flexibility and unified the superimposition among all clades which might impact the virus ability to escape the antibodies neutralization by changing the antigenicity drift of the protein three-dimensional (3D) structure from the wild type clade 19A; this is in agreement with previous study. In conclusion, a new design for the current vaccines to include at least the mutation D614G is immediately needed.

1. Introduction

The current pandemic of COVID-19 has allowed access to high throughput genomic sequencing data for SARS-CoV-2. Genetic variation is a difference in the DNA or RNA sequence among the same species' organism members [1]. There are many globally emerging and circulating genetic variants of SARS-CoV-2 from the beginning of the pandemic [2]. The Centers for Disease Control and Prevention (CDC) have classified such viral genetic variants in the USA into variants of interest, variants of concern, and variants of high consequences. The variants of interest are changes in the viral genetic sequence that might affect the virus's ability to escape the immune system. The variants of concern are expected to be more associated with increased hospitalization or deaths. A piece of solid evidence associates the variants of high consequences with an impact on Medical Countermeasures (MCM), such as failure of diagnostics [3]. In July 2020, Rambaut et al. presented a sound and dynamic nomenclature system to trace the expanding phylogenetic diversity in clusters of variants along the genome sequence of the SARS-CoV-2 virus; called PANGO lineages [4]. Linage spread is associated with an epidemiological event in designated geographical zones. PANGO linages are now in a global report that resulted from investigating haplotypes of novel coronavirus spreading mutations. PANGO reports, which are available online, currently include five linages; B.1.1.7 or known as UK linage (N501Y, P681H and other mutations) [5], B.1.351 or known as South Africa linage (501Y.v2) [6], P.1 or known as Brazil linage (E484K, N501Y and K417T) [6], A.23.1 linage (F157L, V367F, Q613H and P681R) [6] and B.1.525 linage (E484K, Q677H, F888L) [7]. One of the most well-studied emerging and globally distributed mutations in the spike protein of the SARS-CoV-2 virus is D614G which now occupies the first rank in terms of the total mutation frequency (frequency = 1349735 counts, the last update on 08.08.2021).

* Corresponding author.
E-mail addresses: walid.alzyoud@gju.edu.jo (W. Al-Zyoud), hazem_haddad1981@just.edu.jo (H. Haddad).

https://doi.org/10.1016/j.biochi.2021.08.011
0300-9084/

Previous reports have hypothesized that the mutation D614G impacts the viral fitness and infectivity [8–10]. In addition, it has been reported that D614G had increased viral titers and viral replication in the upper respiratory airway in the human epithelial cells [11]. It is worth mentioning that the pandemic has triggered international efforts to build up vast and significant genome evolution within real-time tracking databases such as Nextstrain. Nextstrain has its unique naming and grouping system for SARS-CoV-2 variants into what is called clades; Nextstrain relies on Python scripts to maintain pathogen genomic sequences reposited initially in open source databases with an ability to perform robust phylodynamic data curation, analysis and visualization with a continuous update to bind the scientific and public health potentials [12]. Nextrain, the database gets updated every two months. Recently the World Health Organization (WHO) started an initiative to name the variants of concern and variants of interest by using the simple Greek alphabet e.g., Alpha, Beta, Gamma, etc. to ease its understanding by non-scientists.

Since its discovery in Wuhan in December 2019 (named Wuhan-Hu-1 or WA1), SARS-CoV-2, the wild type clade, has been called 19A, without any amino acid changes in the spike protein and 19B; without D614G but other changes in the spike protein. The clade 19A has evolved into several co-circulating variants presently the major clades from 2020 until now are: 20A: basal pandemic lineage bearing spike protein mutation D614G that is globally distributed, 20B: derived from 20A, 20C: derived from 20A, 20D: derived from 20B, 20E: derived from 20A, 20F: derived from 20B, 20G: derived from 20C, 20H/501Y.V2: derived from 20C, 20I/501Y.V1: derived from 20B. The standard and WHO names of all SARS-CoV-2 clades are summarized in Table 1. In this study, we studied variants of interest or concern in terms of mutational sensitivity utilizing Phyre2 [13] and prediction of stability changes in the spike protein of SARS-CoV-2 upon mutations using DynaMut by Normal Mode Analysis (NMA) [14] considering the clade 20A, which bears the D614G the new global dominant variant, as a new reference sequence instead of the clade 19A due to the fact that all other clades and subclades have been derived from the clade 20A.

2. Methods

2.1. Data retrieval

We have made a subset of 28 complete 1273 spike amino acid sequences; most of them represent the variants of interest or variants of concerns with total frequency >1000 and found as single genomic mutations on the SARS-CoV-2 mutation tracker of genome wide analysis from the EpiCoV database of GISAID (https://www.gisaid.org/) and (https://users.math.msu.edu/users/weig/SARS-CoV-2_Mutation_Tracker.html). The chosen sequences of single mutations included the following amino acid substitutions: L5F, L18F, D80A, S98F, A222V, A262S, P272L, K417 N, N439K, L452R, Y453F, S477 N, E484K, E484Q, N501T, N501Y, E583D, D614G, Q675H, Q675P, Q677H, Q677P, P681H, P681R, A701V, D1163Y, G1167V, V1176F in addition to the wild type spike protein (NCBI Reference Sequence: YP_009724390.1) [15]. The genetic variants' sequences were identified and aligned against the sequence of the wild type clade (19A or WA1). We used the sequence of the single amino acid variants mentioned above for Normal Mode Analysis (NMA) and Multi AgEnt STability pRedictiOn (MAESTRO) because these two tools can work on only single mutations at a time. We used Nextstrain clades for Similarity Matrix and Superimposition because these tools work on the overall two dimensional (2D) and three dimensional structures (3D) with multiple mutations usually found in one clade.

2.2. Normal mode analysis (NMA)

The stability of the above-mentioned single amino acid variants was analyzed by Normal Mode Analysis (NMA) using DynaMut. This user-friendly web server introduces the dynamics components to the mutation analysis that significantly eased the study of the predicted and approximate dynamics of the conformation around targeted amino acid substitutions through harmonic motion [14]. For NMA analysis, the spike protein QHD43416.pdb, which I-Tasser from the wild type already predicted, was used (https://zhanglab.ccmb.med.umich.edu/COVID-19/) [16]. The stability analysis included the Site-Directed Mutator (SDM) and DUET servers. The impact of amino acid changes in any given protein is predicted based on the positive or negative changes in Gibbs free energy ($\Delta\Delta G^{\circ}$) between unfolded versus folded states in that protein [14]. It is well known that a protein $\Delta\Delta G^{\circ}$can represent its stability as a value. As its name implies, DUET integrates two prediction approaches; the first is the SDM approach which relies on the 3D structure of the target protein in predicting its stability and the second is the mCSM approach with prediction that has to yield a consensus prediction with SDM [12].

2.3. Multi AgEnt STability pRedictiOn (MAESTRO)

The wild type spike protein, QHD43416.pdb, was exposed to potential disulphide bonds scan and evaluation by implementing a Multi AgEnt STability pRedictiOn (MAESTRO), a web tool to predict changes in unfolding free energy upon point mutation based on a multi-agent machine learning system [17]. Applying MAESTRO, all amino acid pairs with a beta-carbon to beta-carbon (C^{β}-C^{β}) distance within 5 Å are considered potential binding partners. All possible amino acid pairs are subsequently mutated to cysteine if they were not originally cysteine, and then the cysteines get rated by a final disulphide score called Sss. The score values of the Sss sort the higher negative score as the best.

Table 1
Standard Naming System of SARS-CoV-2 variants.

WHO label	Additional amino acid changes monitored*	Earliest documented samples	Variants Standard Naming System
Alpha	+S.484K +S.452R	United Kingdom, Sep-2020	B.1.1.7 (20I/501Y.V1) or Alpha
Beta	+S.L18F	South Africa,May-2020	B.1.351 (20H/501Y.V2) or Beta
Gamma	+S.681H	Brazil, Nov-2020	P.1 (20J/501Y.V3) or Gamma
Delta	+S.417 N	India, Oct-2020	B.1.617.2 (21A/478K.V1) or Delta
Epsilon	+S.W152C	California, USA, 20-May-2020	B.1.429 (21C/452R) or Epsilon
Eta		Multiple countries, Dec-2020	B.1.525 (21D/484K.V3) or Eta
Theta		Japan on March 12, 2021	P.3 (21E) or Theta
Iota		USA, Nov-2020	B.1.526 (21F/253G.V1) or Iota
Kappa		India, Oct-2020	B.1.617.1 (21B/452R.V3) or Kappa
Lambda		Peru, Dec-2020	C.37 (21G/452Q.V1) or Lambda

Table 2
Normal Mode Analysis (NMA) by DynaMut.

#	SARS-CoV-2 Spike Variants of interest	Spike Protein Region	ΔΔG DynaMut *kcal/mol*	$\Delta\Delta S_{Vib}$ ENCoM $kcal.mol^{-1}.K^{-1}$	Global Percentage from GSAID August 8, 2021	Interatomic interaction with surrounding Amino Acids for increasing flexibility only (−) Lost interatomic interaction (+) Gained interatomic interaction
1	L5F	SP	−0.105 (Destabilizing)	−0.258 (Decrease of molecule flexibility)	3.49	
2	L18F	NTD	0.985 (Stabilizing)	−0.641 (Decrease of molecule flexibility)	5.94	
3	D80A	NTD	0.692 (Stabilizing)	0.199 (Increase of molecule flexibility)	1.19	-W64,-F69,-Y26,-P82,-H66,+L242
4	S98F	NTD	0.788 (Stabilizing)	−0.348 (Decrease of molecule flexibility)	1.40	
5	A222V	NTD	2.134 (Stabilizing)	−0.547 (Decrease of molecule flexibility)	8.13	
6	A262S	NTD	0.828 (Stabilizing)	−0.408 Decrease of molecule flexibility)	0.54	
7	P272L	NTD	1.297 (Stabilizing)	−0.467 (Decrease of molecule flexibility)	0.38	
8	K417 N	RBD	−0.287 (Destabilizing)	0.588 (Increase of molecule flexibility)	1.29	
9	N439K	RBD	2.132 (Stabilizing)	−0.703 (Decrease of molecule flexibility)	1.25	
10	L452R	RBD	0.227 (Stabilizing)	−0.014 (Decrease of molecule flexibility)	17.73	
11	Y453F	RBD	−0.087 (Destabilizing)	−0.131 (Decrease of molecule flexibility)	0.07	
12	S477 N	RBD	0.038 (Stabilizing)	−0.002 (Decrease of molecule flexibility)	2.51	
13	E484K	RBD	−0.187 (Destabilizing)	0.490 (Increase of molecule flexibility)	6.14	-Y489,-F490 +F486
14	E484Q	RBD	−0.488 (Destabilizing)	0.389 (Increase of molecule flexibility)	0.28	-Y489,-F490
15	N501T	RBD	0.701 (Stabilizing)	−0.131 (Decrease of molecule flexibility)	0.15	
16	N501Y	RBD	0.502 (Stabilizing)	−0.211 (Decrease of molecule flexibility)	43.91	
17	E583D	S1	−0.365 (Destabilizing)	0.215 (Increase of molecule flexibility)	0.32	-L533,K535
18	D614G	PRE-FURIN	0.292 (Stabilizing)	0.103 (Increase of molecule flexibility)	98.05	-R646
19	Q675H	PRE-FURIN	−0.621 (Destabilizing)	0.332 (Increase of molecule flexibility)	0.79	+S691,-Y660
20	Q675P	PRE-FURIN	−0.125 (Destabilizing)	0.254 (Increase of molecule flexibility)	0.00	+S691,+S673,+Q677,+Q690,-Y660
21	Q677H	PRE-FURIN	2.661 (Stabilizing)	−0.760 (Decrease of molecule flexibility)	1.37	
22	Q677P	PRE-FURIN	1.053 (Stabilizing)	−0.176 (Decrease of molecule flexibility)	0.18	
23	P681H	PRE-FURIN	0.297 (Stabilizing)	−0.035(Decrease of molecule flexibility)	42.47	
24	P681R	PRE-FURIN	0.788 (Stabilizing)	−0.066 (Decrease of molecule flexibility)	15.37	
25	A701V	POST- FURIN	0.192 (Stabilizing)	0.033 (Increase of molecule flexibility)	2.50	-N703
26	D1163Y	HR2	0.365 (Stabilizing)	−0.168 (Decrease of molecule flexibility)	0.18	
27	G1167V	HR2	−0.127 (Destabilizing)	0.076 (Increase of molecule flexibility)	0.13	-I1169,+D1163, -D1165
28	V1176F	HR2	−0.252 (Destabilizing)	−0.178 (Decrease of molecule flexibility)	2.88	

2.4. Similarity matrix

We have applied DALI server, which performs an all-against-all structure comparison approach at (http://ekhidna2.biocenter.helsinki.fi/dali/) [18]. A set of 39 PDB structures of the wild type and the mutant clades of Nextstrain, listed in Figs. 2 and 3, were built by phyre2 at (http://www.sbg.bio.ic.ac.uk/phyre2) [13], the GISAID accessions of the chosen clades in this study; Alpha, Beta, Gamma, Delta, Epsilon, Eta Theta, Iota, Kappa, and Lambda are listed in Table S in the appendix. The DALI's program outputs a root-mean-square deviation (RMSD) and superimposed coordinates because this is customary and, in many cases, provides an informative visualization of superimposed chain traces. The all-against-all structural comparison usually performs a few rounds of transitive alignment (involving triplets to improve the score of the weakest link) followed by refinement as long as the sum of Z-scores over the matrix increases. From the similarity matrix, a dendrogram was derived using average linkage clustering. DALI is closest to the consensus observation of average over independent agents giving the best approximation of structural truth equivalences based on explicit 3D superimposition; the score depends on the radial distance from one focal point; this gives 2D alignment extra robustness compared to 3D alignment.

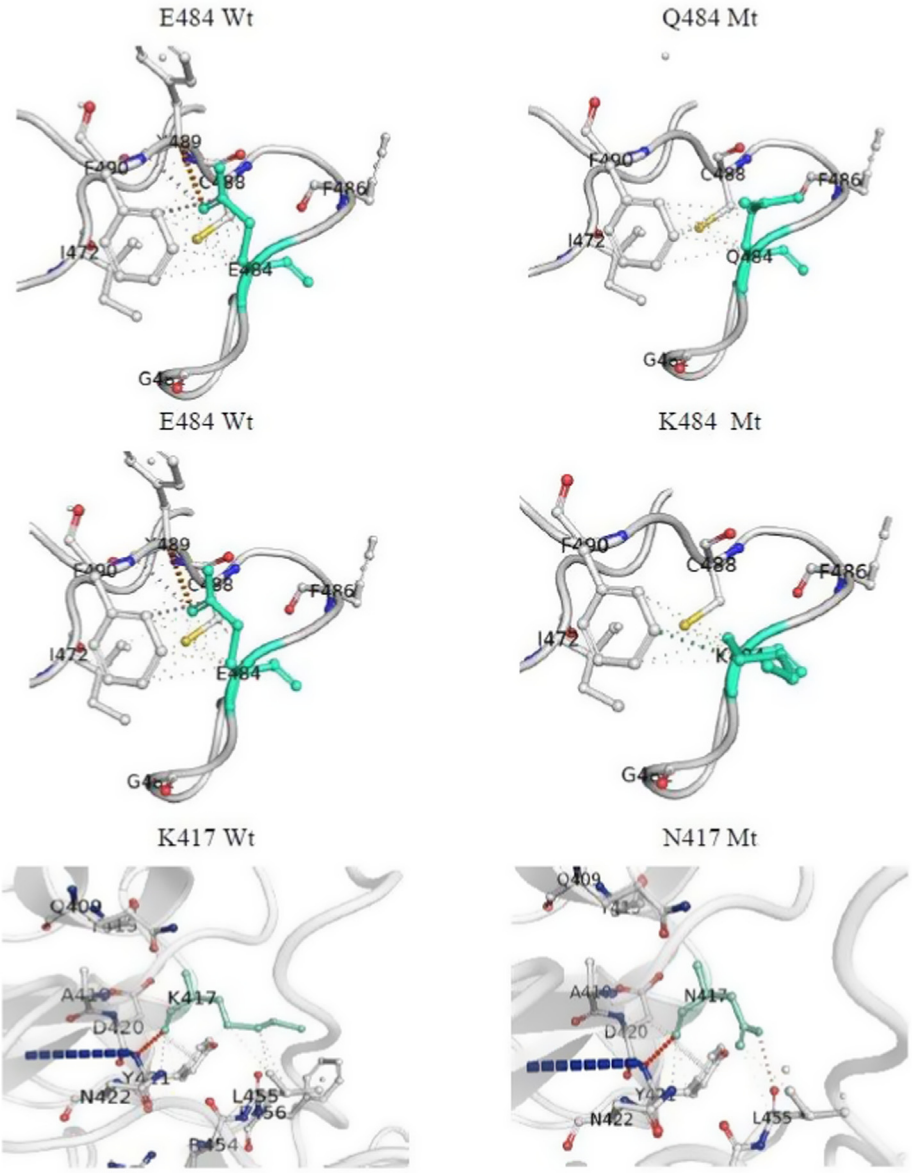


Fig. 1. DynaMut prediction of interactomic interactions for only increasing the molecule flexibility, wild-type and mutant residues are coloured in light-green and represented as sticks alongside the surrounding residues involved in any interactions.

2.5. Superimposition

SuperPose, the webserver version 1.0, was operated to calculate the spike protein superposition by comparing the protein data bank (PDB) files of the wild type version of the protein versus the mutants which phyre2 has built, listed in Fig. 4 and Tables 4 and 5 (we acknowledge SuperPose v1.0 (2004) Rajarshi Maiti, Gary Van Domselaar, Haiyan Zhang, and David Wishart). The Template Modeling score (TM-score) by Zhang lab was employed to compare the structural similarity between any two spike proteins [19,20].

3. Results

The results of normal mode analysis (NMA) by DynaMut shown in Table 2 for the predicted thermodynamic effect of studied variants on the stability and flexibility of the spike protein of SARS-CoV-2 has revealed that out of 28 studied variants, 10 variants increased the protein flexibility (e.g., E484K, D080A, K417 N, A701V, E484Q, D614G, Q675P, E583D, G1167V and Q675H) meanwhile 18 variants decreased the protein flexibility (e.g., A262S, Q677P, N501T, Q677H, L452R, V1176F, N501Y, P681R, L5F, L18F, S98F, A222V, P272L, N439K, Y483F, S477 N, D1163Y and P681H). On the other hand, only 10 variants have destabilized the spike protein (e.g., E484K, K417 N, V1176F, E484Q, Q675P, G1167V, E553D, L5F, Y453F and Q675H) meanwhile, 18 variants (e.g., A262S, D080A, Q677P, N501T, A701V, Q677H, L452R, N501Y, D614G, P681R, L18F, S98F, A222V, P272L, N439K, S477 N, D1163Y and P681H) showed stabilizing tendency. Interestingly, the results of the NMA analysis revealed that some variants are stabilizing but with increased protein flexibility (e.g., D080A, A701V and D614G). However, some other variants have destabilized the protein with increased flexibility (e.g., E484K, K417 N, E484Q, Q675P, G1167V, E583D and

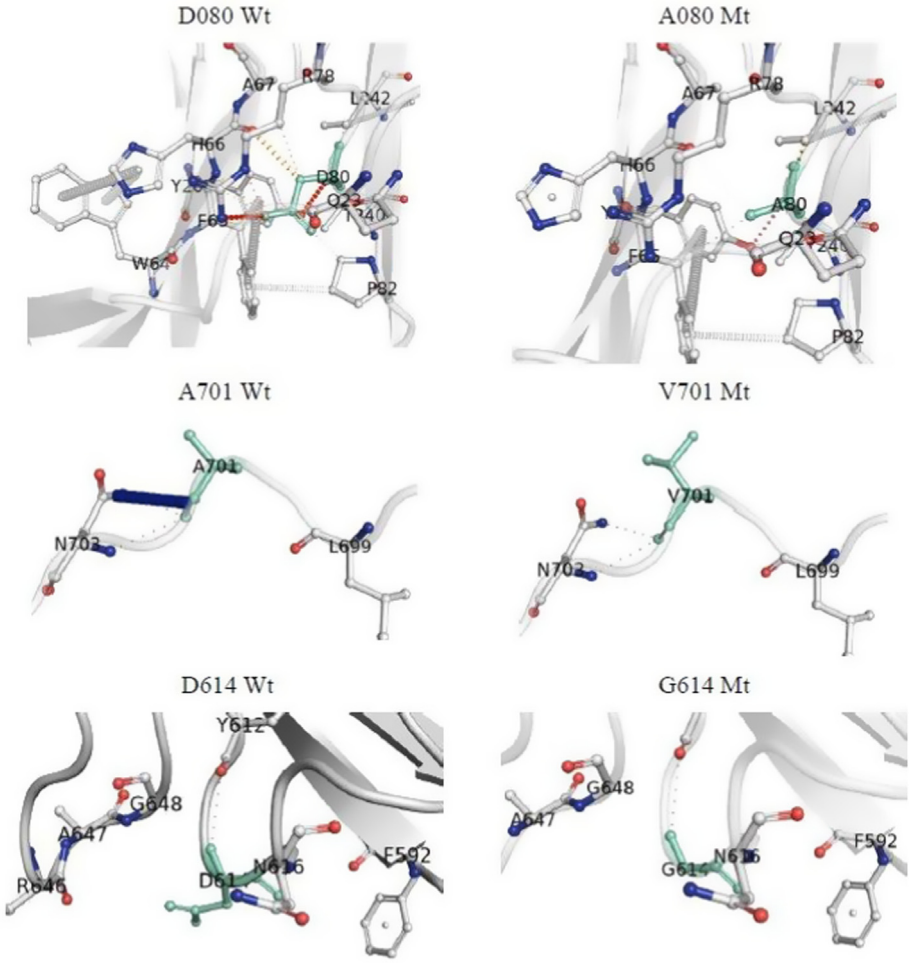


Fig. 1. (*continued*).

Q675H). In more detail, the variants that increased protein flexibility showed a notable change in the interatomic molecular interactions with surrounding amino acids, as shown in Fig. 1, which offers a comparison of the Ensemble NMA of wild type versus single variants. The wild type and variant sequences were extracted and aligned from their respective 3D structures. For example, E484K has lost interatomic interactions with Y489 and F490 but gained interatomic interactions with F486, as shown in Table 2. Discussing E484K has led us to compare it with E484Q, which has lost interatomic interactions with Y489 and F490 without gaining interatomic interactions with F486, which might explain the higher negative (destabilizing) value of DynaMut ΔΔG equal to −0.187 kcal/mol and -0.488 kcal/mol respectively. It is important that Q675H had the highest negative (destabilizing) value of DynaMut ΔΔG with -0.621 kcal/mol, and Q677H had the highest (stabilizing) value of DynaMut ΔΔG with 2.661 kcal/mol.

Table 3 shows the disulphide score (Sss) prediction by another tool, MAESTRO, along the reference sequence of the wild type. The prediction shows that cysteine 743 (C743) with cysteine 749 (C749) appeared as the best two partners to form a disulphide bond, giving the highest negative Sss value of -4.610, at pH = 7. Any amino acid substitution in these two cysteines would significantly affect the protein folding verified by the prediction tool; Phyre2 gives high mutational sensitivity (Appendix, Figure S). We have noticed that some variants of interest are located within the locations of the two disulphide bond partners. For example, E484K or E484Q is situated in the amino sequence C_{480}-NGV**E**GFN-C_{488}. The disulphide partners with an Sss score of -4.352 where the amino acids acid substitution in that location is expected to increase the spike protein thermodynamic flexibility as shown in the DynaMut results in Table 2. Other examples include the sites of N439K, Y453F, S477 N, N501Y or N501T, and L452R; the variants located in the amino sequence between C391with C525 the disulphide partners with Sss score of −2.653 where the amino acid substitution in these two locations are expected to decrease the spike protein thermodynamic flexibility as shown in the DynaMut results in Table 2.

Fig. 2 shows the structural similarity matrix (Dali Z-scores) where the colour dimension shows red for positive; lighter hues are near zero and white for negative values. The sharp drop of deviations at short distances and damping contributions from distances might reach a space larger than 20 Å. The comparison among the clade 20A and its derived subclades variants and the following 10 clades variants; B.1.1.7 (20I/501Y.V1) or Alpha, B.1.351 (20H/501Y.V2) or Beta, P.1 (20J/501Y.V3) or Gamma, B.1.617.2 (21A/478K.V1) or Delta, B.1.429 (21C/452R) or Epsilon, B.1.526 (21F/253G.V1) or Iota, in the presence of D614G in all of them, in addition to clade 19A, showed high structural similarity neighbourhood's near each other except for clade 19A and 20A/S.Q675H . This has been confirmed by the dendrogram results derived by average linkage clustering of the structural similarity matrix (Dali Z-scores) shown in Fig. 3a. The Z-score of 43.7 was the threshold to judge the structural similarity; any scores above the threshold were

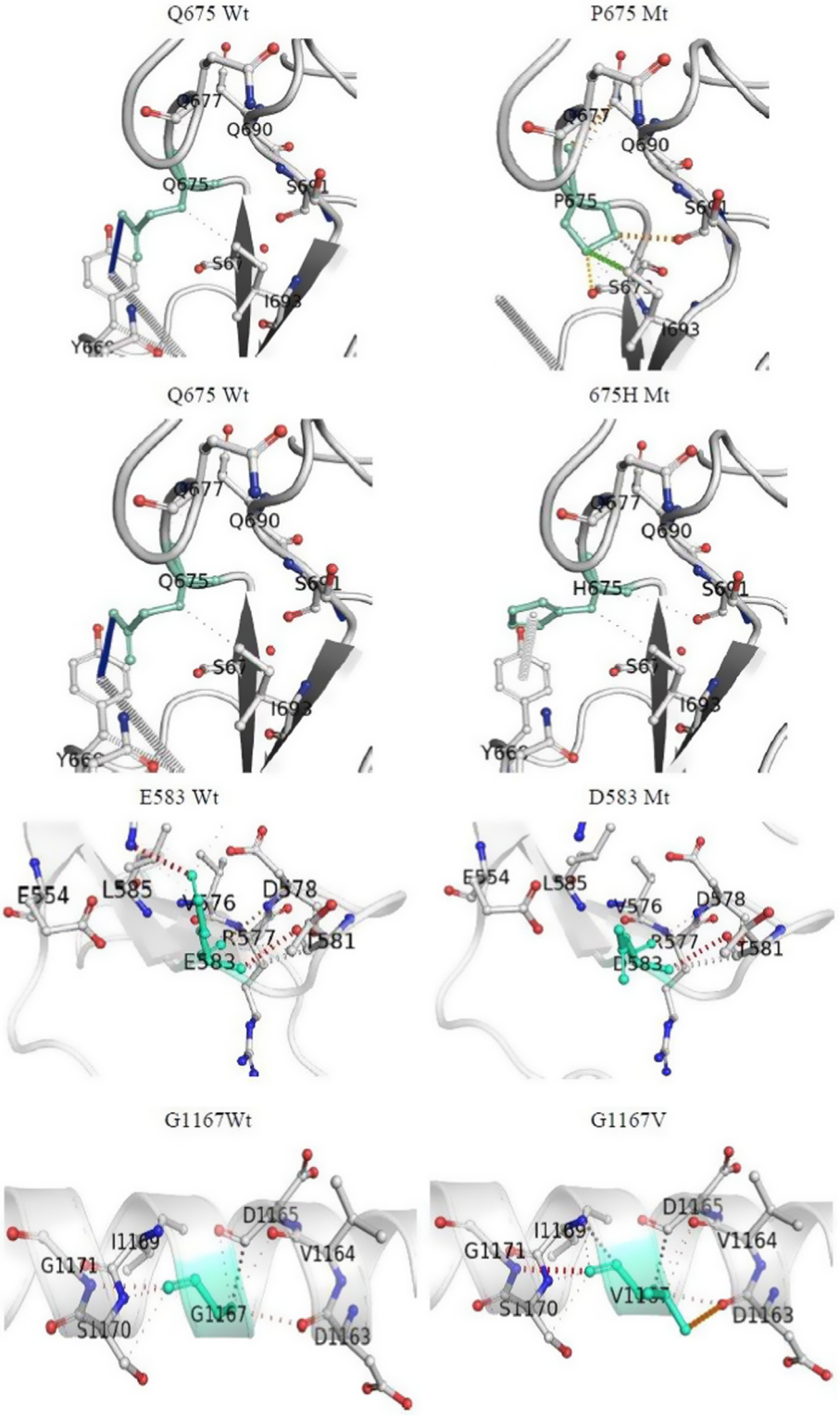


Fig. 1. (*continued*).

considered similar, and any score below the threshold was deemed dissimilar. Correspondence analysis represents a multidimensional scaling method; this analysis positions the data points with the most similar structural neighbourhoods in clusters. In Fig. 3b, the clades 19A and 20A/S.Q675H were clustering (cluster A) far in terms of linkage from clades 20A and its subclades clusters (cluster B, C and D). This had led us to consider the clade 20A with D614G to be as a new reference structure instead of the clade 19A when we carried out superimposition structure analysis with Template Modeling score (TM-score) since the global percentage of the D614G is 98% as shown in Table 2 (data on 08-08-2021).

In the Template Modeling (TM), the reference structure appears

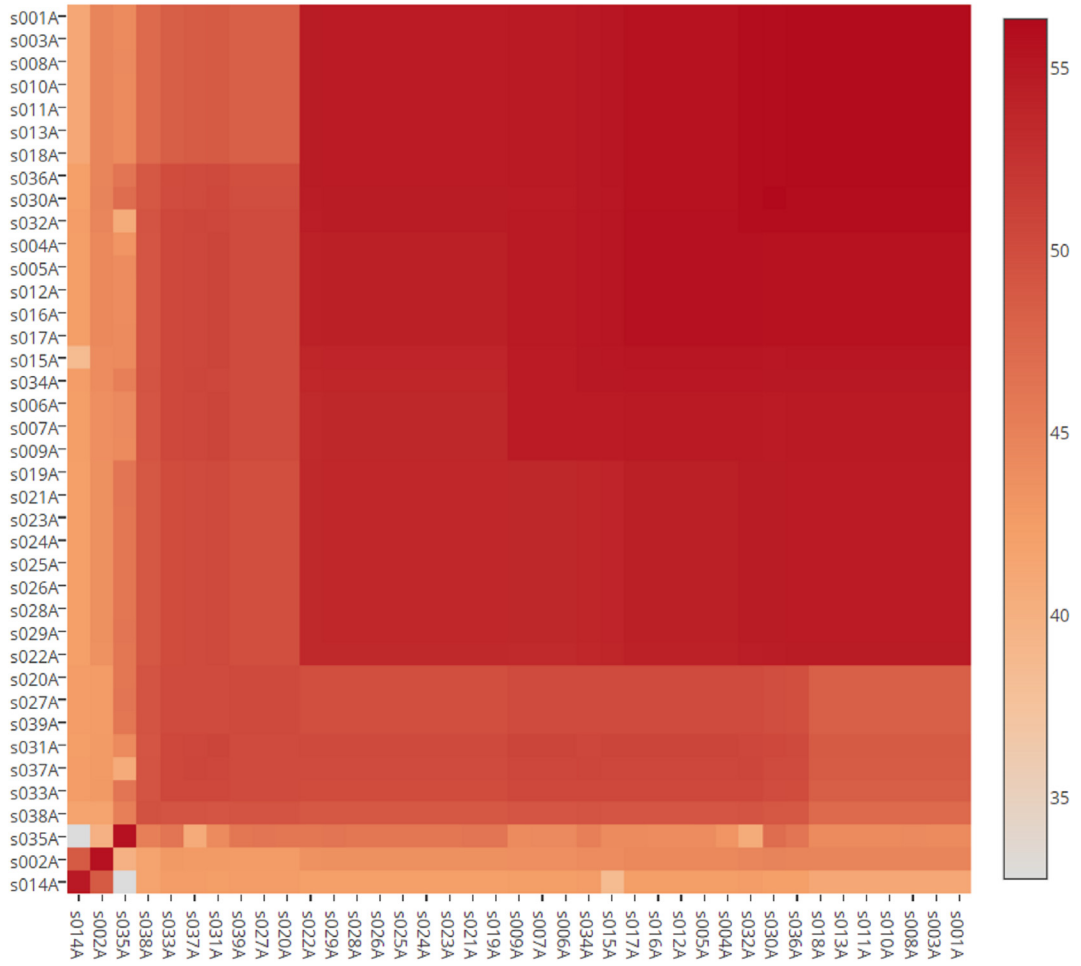


Fig. 2. Heat map or structural similarity matrix (Dali Z-scores): The colour dimension shows the DALI-score, red for positive, lighter hues are near zero and white for negative values. Sharp drop for deviations at short distances and damping of contributions from distances longer than 20 Å.
* The codes on the x and y axes of the matrix include: s001A: 20A, s002A: 19A, s003A: 20A/S.A262S, s004A: 20A/S.A701V, s005A: 20A/S.D80A, s006A: 20A/S.E484K, s007A: 20A/S.E484Q, s008A: 20A/S.K417 N, s009A: 20A/S.L452R, s010A: 20A/S.N501T, s011A: 20A/S.N501Y, s012A: 20A/S.P681H, s013A: 20A/S.P681R, s014A: 20A/S.Q675H, s015A: 20A/S.Q675P, s016A: 20A/S.Q677H, s017A: 20A/S.Q677P, s018A: 20A/S.V1176F, s019A: 20A/S.A222V, s020A: 20A/S.L18F, s021A: 20A/S.S477 N, s022A: 20A/S.N439K, s023A: 20A/S.S98F, s024A: 20A/S.L5F, s025A: 20A/S.P272L, s026A: 20A/S.D1163Y, s027A: 20A/S.E583D, s028A: 20A/S.G1167V, s029A: 20A/S.Y453F, s030A: Beta, s031A: Iota, s032A: Delta, s033A: Gamma, s034A: Alpha, s035A: Epsilon, s036A: Lambda, s037A: Kappa, s038A: Eta, s039A: Theta.

in blue, and the other structures to be compared with appearing in red, so when two colours appeared, this indicates a structural change in the angstrom (Å) unit. The outcome of SuperPose v 1.0 with 3D visualization for the spike protein D614G variant of the clade 20A versus the clade 19A is shown in Fig. 4 and Tables 4 and 5 The summary of our superimposition results includes that the following we-named first group of variants; 20A/S.A262S, 20A/S.L452R, 20A/S.N501T, 20A/S.N501Y, 20A/S.P681H, 20A/S.P681R, 20A/S.V1176F, 20A/S.N439K, 20A/S.S98F, 20A/S.L5F, 20A/S.P272L, 20A/S.D1163Y, 20A/S.N439K, 20A/S.S98F, 20A/S.L5F, 20A/S.P272L, 20A/S.D1163Y and 20A/S.G1167V have no structural superimposition changes at all, which means red structures were only apparent; this is in agreement with DynaMut results for variants stabilizing the spike protein structure. The second group of variants with enormous superimposition changes with more than 12 Å for amino acids along the spike protein sequence included 20A versus 19A on the one hand and the 20A/S.Q675H versus 20A on the other hand. The DynaMut results showed that D614G and the Q675P had decreased the spike protein stability. The third group of variants with located superimposition change up to 12 Å for many structural regions included the following variants; 20A/S.L18F and B.1.429 (21C/452R at the Furin cleavage, which both hidden a Furin region. The 20A/S.701 and B.1.526 (21F/253G.V1) or Iota have changed a Furin region to Alpha helix structure. The B.1.525 (21D/484K.V3) or Eta changed an RBD region from a linker structure to Alpha helix. The C.37 (21G/452Q.V1) or Lambda has changed an RBD region to Beta sheet. The B.1.351 (20H/501Y.V2) or Beta has changed an RBD region to Beta sheet. The fourth group included the superimposition with up to 9 Å; in this group where the B.1.617.1 (21B/452R.V3) or Kappa and the B.1.617.2 (21A/478K.V1) or Delta both have changed an RBD region and a Furin region to Alpha helix. B.1.1.7 (20I/501Y.V1) or Alpha has changed an RBD region to an Alpha helix and Beta sheet. The 20A/S.677H has changed the orientation of the Beta sheet of the Furin region as shown in Table 5.

The fifth group of variants with located superimposition change, up to 5 Å in the RBD, included the following variants: the 20A/S.D80A, 20A/S.A222V, 20A/S.K417 N, 20A/S.477 N, A20/S E484K, A20/S E484Q, 20A/S.Y453F and 20A/S.E583D have changed a Furin region to an Alpha helix structure. The A20/S.Q677P has changed a Furin region to a small Beta sheet. The P.1 (20J/501Y.V3) or Gamma and the P.3 (21E) or Theta have changed an RBD region from a linker structure to an Alpha helix.

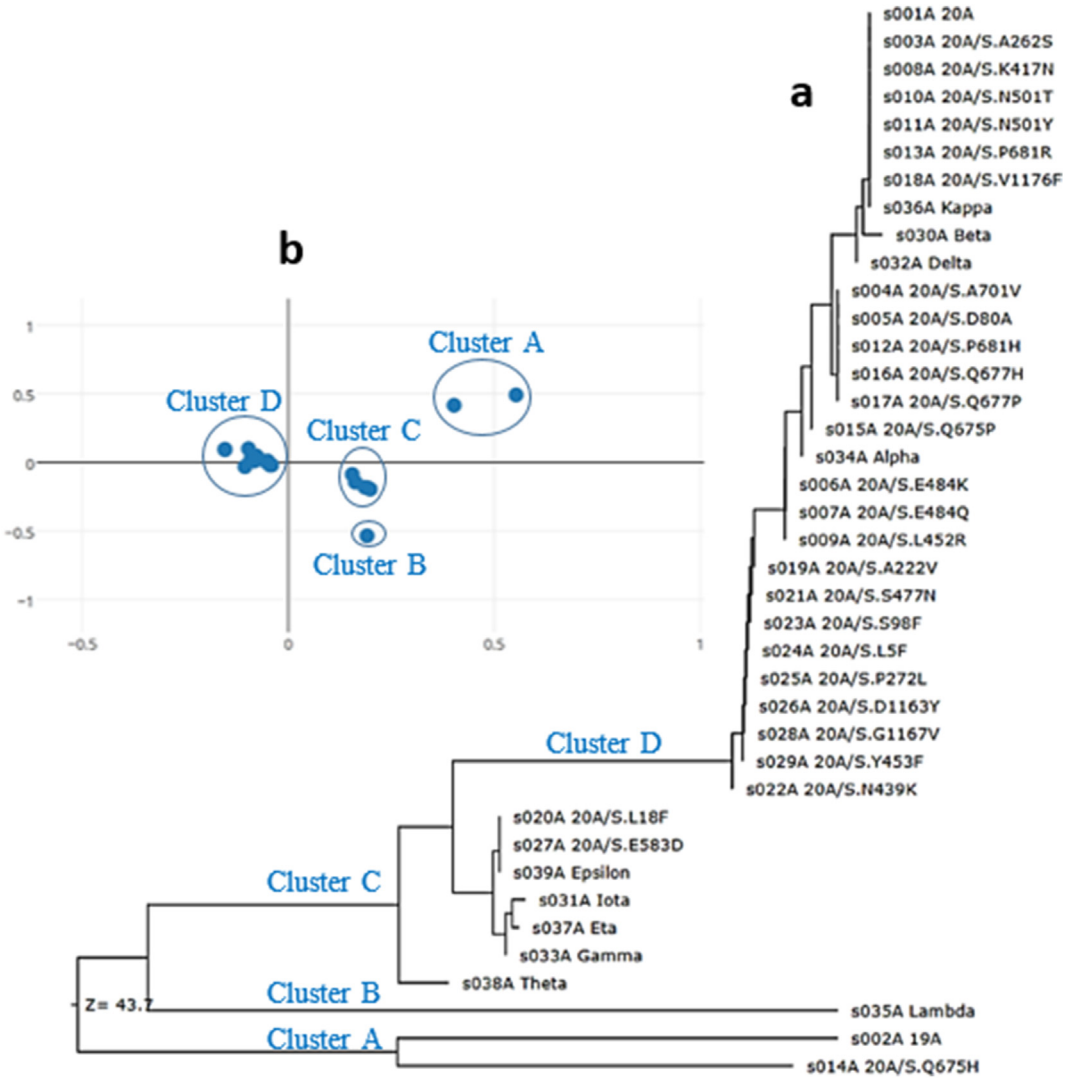


Fig. 3. a: Structural Similarity Dendrogram of 39 studied structures. The dendrogram is derived by average linkage clustering of the structural similarity matrix (Dali Z-score = 43.7).
b: Correspondence Analysis is a multidimensional scaling method. It positions data points with the most similar structural neighbourhoods as clusters.

4. Discussion

A new coronavirus, SARS-CoV-2, associated with human respiratory disease, COVID-19, was first reported in Wuhan, China [15,21]. It is well known that the spike protein of SARS-CoV-2 plays a crucial role in its binding ability with human Angiotensin Converting Enzyme-2 (ACE2) receptor on the cell surface, which ultimately cause its pathogenicity when to get inside the human cells. It has been reported that many mutations played a critical role in changing the virus molecular stability and flexibility and hence increasing or decreasing its transmissibility. For example, one of the most well-studied and dominant emerging mutations in the spike protein of the SARS-CoV-2 virus is D614G. Previous reports hypothesized that the mutation D614G impacts viral fitness and infectivity [8–10]. This study aimed to predict the effect of molecular dynamics changes of notable variants on the wild type spike protein. The results from DynaMut thermodynamics prediction indicated that 18 variants mentioned in the results section showed stabilizing tendency compared with the wild type spike protein. The majority of these variants have been found in the variants listed in Table 1. For example, South African; B.1.351 (20H/501Y.V2) or Beta, the UK; B.1.1.7 (20I/501Y.V1) or Alpha, multiple countries, e.g. the UK \ Nigeria; B.1.525 21)D/484K.V3) or Eta, India; B.1.617.2 (21A/478K.V1) or Delta and the US-Californian; B.1.429 (21C/452R) or Epsilon, agreed with previous reports in terms of increased transmission and infectivity, and reduced antibody neutralization [22–26]. The variant N501Y showed the highest frequency in the RBD with around 711234 counts as displayed on the SARS-CoV-2 mutation tracker of genome-wide analysis. The variant Q677H found among Eta variants has the highest stabilizing ability with decreased flexibility which might play a crucial biological function as it is located in the Pre-Furin cleavage site.

Interestingly, the NMA analysis results show that some variants are stabilizing but with an increased flexibility effect, for example, D080A, A701V, and D614G that might have an impact on improving the viral fitness and transmissibility by increasing its ability to bind with its receptor. Moreover, the results from DynaMut thermodynamic predictions indicated 10 variants had destabilized the spike protein (e.g., E484K, K417 N, V1176F, E484Q, Q675P, and Q675H), two of them, the E484K and the V1176F, have been found among Gamma and Theta variants. As these mutations are expected to destabilize the spike protein, they are expected to reduce antibody neutralization, which agrees with previous reports [24,27]. The mutation Q675H located in the Pre-Furin cleavage site in the spike

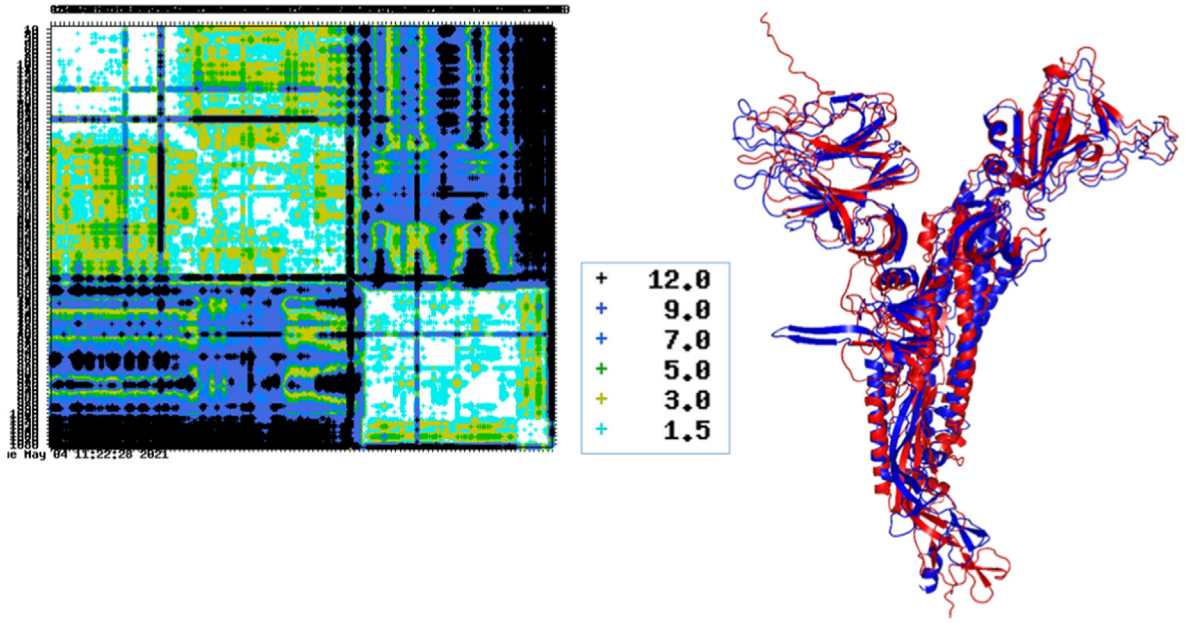


Fig. 4. 2D-difference distance matrix by SuperPose v.1.0 (on the left hand side) with the 3D visualization by the TM-align, v.20190822,for the spike protein with variant D614G (clade 20A; in blue) versus the wild type (clade 19A; in red). The small box represents the superimposition marker in the angstrom unit (Å).

Table 3
Disulfide score (Sss) prediction by MAESTRO.

Disulfide bond interaction by MAESTRO	S_{ss}	Critical Amino Acid Sequence	High mutation by Phyre2
C743 with C749	−4.610	CGDSTEC	**C743, C749**
C738 with C760	−4.536	CTMYICGDSTECSNLLLQYGSFC	**C738, C743, C749, C760**
C538.with C590	−4.528	CVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPC	**C538, G550, C590**
C480 with C488	−4.352	CNGVEGFNC	**C480, C488**
C1082 with C1126	−4.320	CHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNC	**C1082, P1090, C1126**
C291.with C301	−3.996	CALDPLSETKC	**C291, C301**
C617.with C649	−3.934	CTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGC	**C617, C649**
C662.with C671	−3.252	CDIPIGAGIC	**C662, C671**
C131 with C166	−3.193	CEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNC	**C131, C136, P139, C166**
C391with C525	−2.653	CFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGC …	**C391, F400, C432, C525**

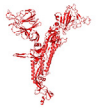
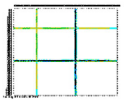
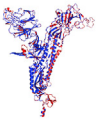
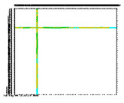
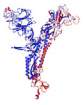
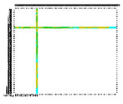
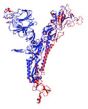
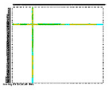
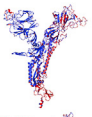
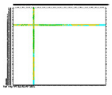
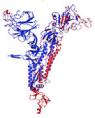
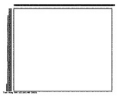
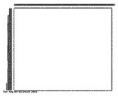
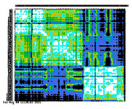
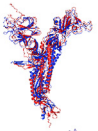
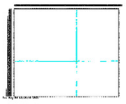
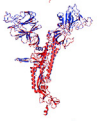
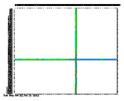
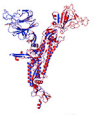
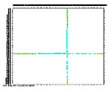
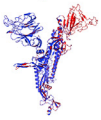
protein, with the highest destabilizing ability, has appeared in Bangladesh in late 2020 [28]. However, some variants that have destabilized the protein with increased flexibility (e.g., E484K, K417 N, E484Q, Q675P, and Q675H) might impact the virus's ability to escape the antibodies neutralization by changing the antigenicity drift of the protein 3D structure; this is in agreement with a previous study [23–28].

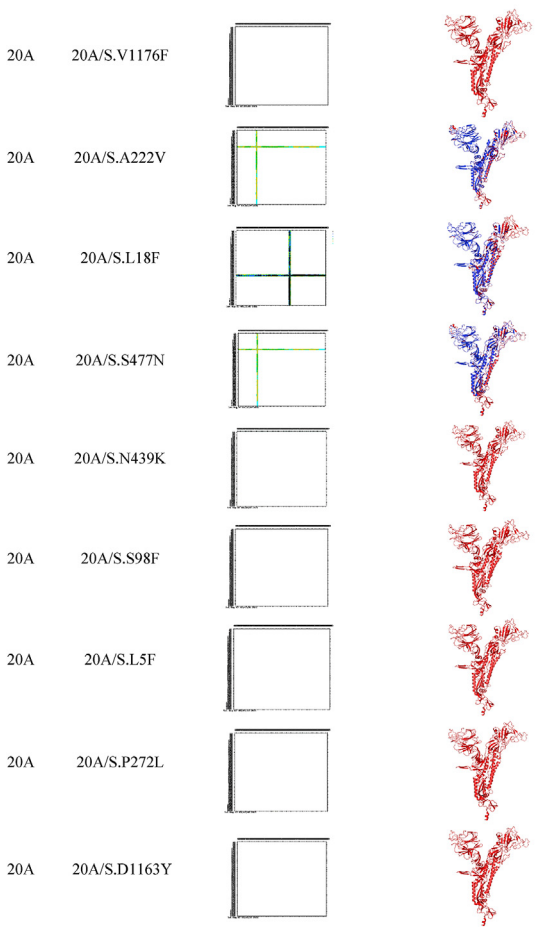
The results of MAESTRO in this study revealed the cysteine disulphide partners in the receptor binding domain (RBD); C391 with C525 where the mutations K417 N, N439K, L452R, Y453F, S477F, N501Y, and N501T among the top ten global frequency ranking are located, all collectively are expected to decrease the flexibility of the 3D structure of the RBD of the spike protein due to a notable change in the interatomic interaction. On the other hand, the partners C480 with C488, where the mutations E484K and E484Q are located, are both collectively expected to increase the flexibility of the 3D structure of the RBD.

It is well-noticed that the mutation D614G is found in all clades reported in the Nextstrain database as mentioned in the introduction; this has led us to question if the current emergency use authorization (EUA)-vaccines have taken into consideration the vaccine design to be effective against the D614G as a basal and dominant mutation or even other mutations of interest. Even after mass vaccination, some countries have shown a significant resurgence in Manaus, Brazil [29]. The results of our analysis on DALI, SuperPose v 1.0 and TM-score servers have revealed that the single dominant mutation the D614G has changed the whole structural similarity matrix (Dali Z-scores) (Figs. 2 and 3a&b) of the spike protein when compared with the wild type (clade 19A) giving a change of more than 20 Å on the SuperPose v1.0 as shown in Fig. 4. The DALI, SuperPose and TM-scores showed that the subclades with D614G have the same structural similarity compared with the single mutation of D614G (clade 20A). From all of these results the main superimposition change that happened at the overall 3D structure of the spike protein was due to the presence and effect of D614G only. The DynaMut analysis of single mutations against the wild type shown that the increasing flexibility mutations have structural superimposition change up to 12 Å when compared with the reference sequence of clade 20A; meanwhile, the decreasing flexibility of single mutations have shown trivial to no structural superimposition change when compared with the reference sequence of clade 20A.

Interestingly, the mutation 20A/S.Q675H showed a profound superimposition change with up to 20 Å even when compared with a reference sequence of clade 20A, as shown in Table 4. At last, the subclades Alpha, Beta, Gamma, Delta, Epsilon, Iota, Lambda, Kappa, Eta, and Theta showed profound superimposition change at 20 Å when compared with clade 19A. These subclades are derived from the clade 20A; this has been confirmed by the heat map or structural similarity matrix by DALI Z-score, as shown in Fig. 2 and

Table 4
2D-difference distance matrix by SuperPose v.1.0 with 3D visualization by the TM-align for the spike protein variants.

Protein-1	Protein-2	Outcome of SuperPose v.1.0	Visualization (Protein-1 in blue and Protein-2 in red)
20A	20A/S.A262S		
20A	20A/S.A701V		
20A	20A/S.D080A		
20A	20A/S.E484K		
20A	20A/S.E484Q		
20A	20A/S.K417N		
20A	20A/S.L452R		
20A	20A/S.N501T		
20A	20A/S.N501Y		
20A	20A/S.P681H		
20A	20A/S.P681R		
20A	20A/S.Q675H		
20A	20A/S.Q675P		
20A	20A/S.Q677H		
20A	20A/S.Q677P		



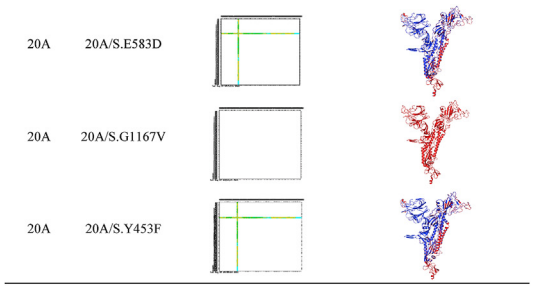


Table 5
2D-difference distance matrix by SuperPose v.1.0 with 3D visualization by the TM-align for the spike protein variants, including the WHO-named variants.

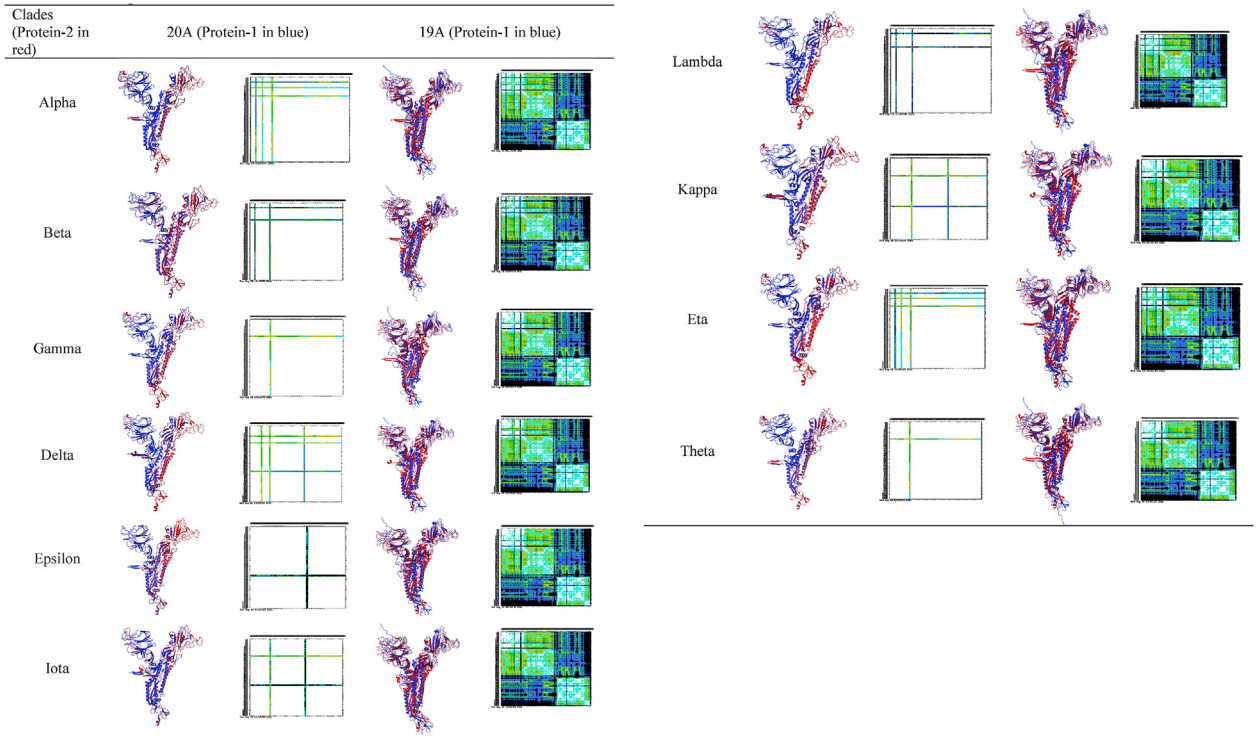


Table 5. It is noteworthy that the clade Lambda has a novel deletion, S.Δ247-253, located at the N-terminal domain in addition to L452Q and F490S mutations in the RBD of the spike protein; this might explain why the conformational change in Lambda has included a formation of new a Beta sheet at the RBD. Hence a role in affinity of the spike protein to ACE2 or the susceptibility to antibodies neutralization is expected (Cluster B in Fig. 3a). In addition to Cluster B, Cluster D (20A) included Alpha, Beta, Delta, Kappa; the Cluster C 20A/S.L18F or 20A/S.D583E included Gamma, Epsilon, Iota, Eta; and Cluster A included only the 20A/S.Q675H as shown in Figs. 2 and 3a&b. All these set of evidences imply the crucial role of the mutational conformational change in the Furin cleavage site, the RBD, as well as the overall 3D structure of the spike protein of SARS-CoV-2 in any future vaccine updates.

5. Conclusion

Previous studies have reported that D614G impacts the SARS-CoV-2 viral fitness and transmissibility by increasing the 'up' conformation in the RBD region of the spike protein. The thermodynamics prediction results in this study have exhibited the stabilizing role of D614G and other variants and their role in increasing the spike protein. The stabilizing and increasing flexibility phenomena have triggered us to link them with the current vaccine efficacy and antibody neutralization, as the current vaccines are based on the wild type D614, clade 19A. The results of this study have also revealed the impact of another critical associated mutation with D614G, which is Q675H; that has a considerable superimposition change at the 3D structure before the Furin cleavage site of the spike protein of SARS-CoV-2, where it has a thermodynamic effect in destabilizing the spike protein with increasing its flexibility. Overall, and based on the findings of this study, two aspects should attract the attention of vaccine developers all over the world, which we here call the components of Haddad-Zyoud new hypothesis of COVID-19 vaccine update. First, the role of the emerging variants, especially the D614G, on the overall 3D structure of the spike protein should lead to change the way how vaccine developers evaluate the neutralization activities of the antibodies that have been previously produced by the immune system of the vaccinated people by the current vaccine. Second, the role of the emerging variants on the 3D structure of the Furin cleavage site and the RBD of the spike protein should change the way how can vaccine developers understand the conformational change by SARS-CoV-2 to enhance its free binding energy with ACE2 receptor and hence its ability to take off the S1 domain to enter the human cells.

Funding

"This research received a seed grant # SAMS 01\2020 from the Scientific Research Council at the German Jordanian University."

Author contributions

Walid Al-Zyoud & Hazem Haddad have contributed equally to Conceptualization, Methodology, Software, Data curation, Writing-Original draft preparation. Visualization, Investigation: Supervision, Software, Validation. Writing Reviewing and Editing.

All authors have read and agreed to the published version of the manuscript.

Declaration of competing interest

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.

Appendix

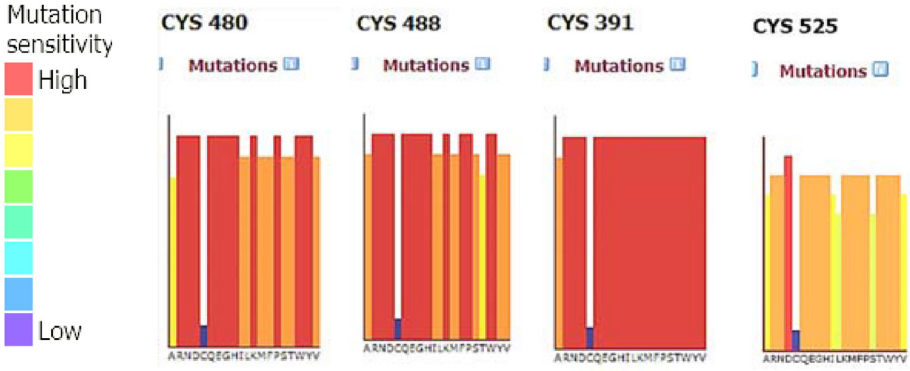


Fig. S. The mutational sensitivity by Phyre2 of selected Cysteines which are expected to have significant impact.

Table S
Variants standard naming with chosen Accession of GSAID.

Variants Standard Naming	GSAID Accession
B.1.1.7 (20I/501Y.V1) or Alpha	hCoV-19/Poland/PZH-GUM-1840/2021|EPI_ISL_2897131|2021-07-01
B.1.351 (20H/501Y.V2) or Beta	hCoV-19/South Africa/N00390/2020|EPI_ISL_712081|2020-10-08
P.1 (20J/501Y.V3) or Gamma	hCoV-19/Brazil/AM-FIOCRUZ-20842882CA/2020|EPI_ISL_833137|2020-12-04
B.1.617.2 (21A/478K.V1) or Delta	hCoV-19/Spain/CT-HUJ23-23027/2021| EPI_ISL_3189941
B.1.429 (21C/452R) or Epsilon	hCoV-19/Mexico/MIC-InDRE_FB16450_S3431/2021|EPI_ISL_2736825|2021-05-23
B.1.525 (21D/484K.V3) or Eta	hCoV-19/Canada/QC-L00366912001/2021|EPI_ISL_3459145|2021-07-10
P.3 (21E) or Theta	hCoV-19/Hong Kong/HKPU-11464/2021|EPI_ISL_2860802|2021-03-31
B.1.526 (21F/253G.V1) or Iota	hCoV-19/USA/NY-IVY-ZZX9KY4S/2021
B.1.617.1 (21B/452R.V3) or Kappa	hCoV-19/India/RJ-CCMB-CIA6287/2021|EPI_ISL_3453835|2021-03-27
C.37 (21G/452Q.V1) or Lambda	hCoV-19/USA/OH-GD-SID-21060608511/2021|EPI_ISL_3461191|2021-06-06

References

[1] L.M. Hartwell, H. L, L. Hood, M.L. Goldberg, A.E. Reynolds, Silver, Genetics: from Genes to Genomes, McGraw Hill, Boston, 2011. https://www.amazon.com/Genetics-Genomes-Leland-Hartwell-Dr/dp/1259700909.

[2] SS Abdool Karim, T. de Oliveira, New SARS-CoV-2 variants — clinical, public health, and vaccine implications, N. Engl. J. Med. (2021), https://doi.org/10.1056/NEJMc2100362. NEJMc2100362.

[3] SARS-CoV-2 Variants of Concern, CDC, 2021. March 29, https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/variant-surveillance/variant-info.html#Consequence.

[4] A. Rambaut, E.C. Holmes, Á. O'Toole, V. Hill, J.T. McCrone, C. Ruis, L. du Plessis, O.G. Pybus, A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology, Nat. Microbiol. 5 (2020) 1403–1407, https://doi.org/10.1038/s41564-020-0770-5.

[5] Tracking the International Spread of SARS-CoV-2 Lineages B.1.1.7 and B.1.351/501Y-V2 - SARS-CoV-2 Coronavirus/nCoV-2019 Genomic Epidemiology - Virological, 2021. March 29, 2021, https://virological.org/t/tracking-the-international-spread-of-sars-cov-2-lineages-b-1-1-7-and-b-1-351-501y-v2/592.

[6] H. Tegally, E. Wilkinson, M. Giovanetti, A. Iranzadeh, V. Fonseca, J. Giandhari, D. Doolabh, S. Pillay, E.J. San, N. Msomi, K. Mlisana, A. von Gottberg, S. Walaza, M. Allam, A. Ismail, T. Mohale, A.J. Glass, S. Engelbrecht, G. Van Zyl, W. Preiser, F. Petruccione, A. Sigal, D. Hardie, G. Marais, M. Hsiao, S. Korsman, M.-A. Davies, L. Tyers, I. Mudau, D. York, C. Maslo, D. Goedhals, S. Abrahams, O. Laguda-Akingba, A. Alisoltani-Dehkordi, A. Godzik, C.K. Wibmer, B.T. Sewell, J. Lourenço, L.C.J. Alcantara, S.L.K. Pond, S. Weaver, D. Martin, R.J. Lessells, J.N. Bhiman, C. Williamson, T. de Oliveira, Emergence and Rapid Spread of a New Severe Acute Respiratory Syndrome-Related Coronavirus 2 (SARS-CoV-2) Lineage with Multiple Spike Mutations in South Africa, MedRxiv, 2020, p. 2020, https://doi.org/10.1101/2020.12.21.20248640, 12.21.20248640.

[7] Pango lineages, Global Report Investigating Novel Coronavirus Haplotypes, 2021. March 30, 2021, https://cov-lineages.org/index.html.

[8] B. Korber, W.M. Fischer, S. Gnanakaran, H. Yoon, J. Theiler, W. Abfalterer, N. Hengartner, E.E. Giorgi, T. Bhattacharya, B. Foley, K.M. Hastie, M.D. Parker, D.G. Partridge, C.M. Evans, T.M. Freeman, T.I. de Silva, A. Angyal, R.L. Brown, L. Carrilero, L.R. Green, D.C. Groves, K.J. Johnson, A.J. Keeley, B.B. Lindsey, P.J. Parsons, M. Raza, S. Rowland-Jones, N. Smith, R.M. Tucker, D. Wang, M.D. Wyles, C. McDanal, L.G. Perez, H. Tang, A. Moon-Walker, S.P. Whelan, C.C. LaBranche, E.O. Saphire, D.C. Montefiori, Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus, Cell 182 (2020) 812–827, https://doi.org/10.1016/j.cell.2020.06.043, e19.

[9] M. Becerra-Flores, T. Cardozo, SARS-CoV-2 viral spike G614 mutation exhibits higher case fatality rate, Int. J. Clin. Pract. 74 (2020), https://doi.org/10.1111/ijcp.13525.

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

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

[12] J. Hadfield, C. Megill, S.M. Bell, J. Huddleston, B. Potter, C. Callender, P. Sagulenko, T. Bedford, R.A. Neher, Nextstrain: real-time tracking of pathogen evolution, Bioinformatics 34 (2018) 4121–4123, https://doi.org/10.1093/bioinformatics/bty407.

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

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

[15] F. Wu, S. Zhao, B. Yu, Y.M. Chen, W. Wang, Z.G. Song, Y. Hu, Z.W. Tao, J.H. Tian, Y.Y. Pei, M.L. Yuan, Y.L. Zhang, F.H. Dai, Y. Liu, Q.M. Wang, J.J. Zheng, L. Xu, E.C. Holmes, Y.Z. Zhang, A new coronavirus associated with human respiratory disease in China, Nature 579 (2020) 265–269, https://doi.org/10.1038/s41586-020-2008-3.

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

[17] J. Laimer, H. Hofer, M. Fritz, S. Wegenkittl, P. Lackner, Maestro - multi agent stability prediction upon point mutations, BMC Bioinf. 16 (2015), https://doi.org/10.1186/s12859-015-0548-6.

[18] L. Holm, DALI and the persistence of protein shape, Protein Sci. 29 (2020) 128–140, https://doi.org/10.1002/pro.3749.

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

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

[21] Y.C. Liu, R.L. Kuo, S.R. Shih, COVID-19: the first documented coronavirus pandemic in history, Biomed. J. (2020), https://doi.org/10.1016/j.bj.2020.04.007.

[22] Y. Liu, J. Liu, K.S. Plante, J.A. Plante, X. Xie, X. Zhang, Z. Ku, Z. An, D. Scharton, C. Schindewolf, V.D. Menachery, P.-Y. Shi, S.C. Weaver, The N501Y spike substitution enhances SARS-CoV-2 transmission, BioRxiv Prepr. Serv. Biol. (2021), https://doi.org/10.1101/2021.03.08.434499.

[23] X. Deng, M.A. Garcia-Knight, M.M. Khalid, V. Servellita, C. Wang, M. Kate Morris, A. Sotomayor-González, K.R. Reyes, A.S. Gliwa, N.P. Reddy, C. Sanchez, S. Martin, S. Federman, J. Cheng, J. Balcerek, J. Taylor, A. Streithorst, S. Miller, G. Renuka Kumar, B. Sreekumar, P.-Y. Chen, U. Schulze-Gahmen, T.Y. Taha, J. Hayashi, S. McMahon, P. V Lidsky, Y. Xiao, P. Hemarajata, N.M. Green, A. Espinosa, C. Kath, M. Haw, J. Bell, C. Hanson, D.A. Wadford, C. Anaya, D. Ferguson, P.A. Frankino, H. Shivram, S.K. Wyman, R. Andino, C.Y. Chiu, Transmission, Infectivity, and Antibody Neutralization of an Emerging SARS-CoV-1 2 Variant in California Carrying a L452R Spike Protein Mutation 2, MedRxiv, 2021, p. 2021, https://doi.org/10.1101/2021.03.07.21252647, 03.07.21252647.

[24] S. Jangra, C. Ye, R. Rathnasinghe, D. Stadlbauer, F. Krammer, V. Simon, L. Martinez-Sobrido, A. García-Sastre, M. Schotsaert, H. Alshammary, A.A. Amoako, M.H. Awawda, K.F. Beach, M.C. Bermúdez-González, R.L. Chernet, L.Q. Eaker, E.D. Ferreri, D.L. Floda, C.R. Gleason, G. Kleiner, D. Jurczyszak, J.C. Matthews, W.A. Mendez, L.C.F. Mulder, K.T. Russo, A.-B.T. Salimbangon, M. Saksena, A.S. Shin, L.A. Sominsky, K. Srivastava, SARS-CoV-2 spike E484K mutation reduces antibody neutralization, The Lancet Microbe (2021), https://doi.org/10.1016/s2666-5247(21)00068-9, 0.

[25] N. Davies, S. Abbott, R. Barnard, C. Jarvis, A. Kucharski, J. Munday, C. Pearson, T. Russell, D. Tully, A. Washburne, T. Wenseleers, A. Gimma, W. Waites, K. Wong, K. van Zandvoort, J. Silverman, K. Diaz-Ordaz, R. Keogh, R. Eggo, S. Funk, M. Jit, K. Atkins, W.J. Edmunds, Estimated Transmissibility and Impact of SARS-CoV-2 Lineage B.1.1.7 in England, Science, 2021, https://doi.org/10.1101/2020.12.24.20248822, 80-.eabg3055.

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

[27] Z. Jia, W. Gong, Will mutations in the spike protein of SARS-CoV-2 lead to the failure of COVID-19 vaccines? J. Kor. Med. Sci. 36 (2021) e124, https://doi.org/10.3346/jkms.2021.36.e124.

[28] S.Z. Afrin, S.K. Paul, J.A. Begum, S.A. Nasreen, S. Ahmed, F.U. Ahmad, M.A. Aziz, R. Parvin, M.S. Aung, N. Kobayashi, Extensive genetic diversity with novel mutations in spike glycoprotein of SARS-CoV-2, Bangladesh in late 2020, New Microbes New Infect 41 (2021) 100889, https://doi.org/10.1016/j.nmni.2021.100889.

[29] C. Aschwanden, Five reasons why COVID herd immunity is probably impossible, Nature 591 (2021) 520–522, https://doi.org/10.1038/d41586-021-00728-2.