HOW CONCERNING IS A SARS-COV-2 VARIANT OF CONCERN? COMPUTATIONAL PREDICTIONS AND THE VARIANTS LABELING SYSTEM.

Dana Ashoor [1], Maryam Marzouq [1], Khaled Trabelsi [1], Sadok Chlif [2], Nasser Abotalib [1], Noureddine Ben Khalaf [1], Ahmed R. Ramadan [1], and M-Dahmani Fathallah [1, *]

[1] Department of Life Sciences, Health Biotechnology Program - King Fahad Chair for Health Biotechnology, College of Graduate Studies, Arabian Gulf University, PO Box 26671. Manama, Kingdom of Bahrain

[2] Department of Family and Community Medicine, College of Medicine and Medical Sciences, Arabian Gulf University, PO Box 26671, Manama, Kingdom of Bahrain.

*** Correspondence:**
M-Dahmani Fathallah
d.fathallah@agu.edu.bh
Tel.: Tel: +973 17 239602
Fax: +973 17 239484

ABSTRACT:

We herein report a study to evaluate the use of computational prediction of SARS-CoV-2 genetic variations in improving the current variants labeling system. First, we reviewed the basis of the system developed by the World Health Organization (WHO) for the labeling of SARS-CoV-2 genetic variants and the adaptations made to it by the United States Center of Diseases Control (CDC). We observed that the labeling system is based upon the virus' major attributes. However, we found that the labeling criteria of the SARS-CoV-2 variants derived from these attributes are not accurately defined and are used differently by the two health management agencies. Consequently, discrepancies exist between the labels given by WHO and CDC to same variants. Our observations suggest that giving the VOC label to a new variant is premature and might not be appropriate. Therefore, we carried out a comparative computational study to predict the effects of the mutations on the virus structure and functions of five VOCs. By linking these data to the criteria used by WHO and the CDC for variant labeling, we ascertained that comparative computational predictions of the impact of genetic variations are a better ground for rapid and more accurate labelling of SARS-CoV-2 variants. We propose to label all emergent variants VUM or VBM and to carry out computational predictive studies and thorough variants comparison, upon which more appropriate and informative labels can be attributed. Furthermore, harmonization of the variants labeling system would be globally beneficial to communicate about and fight COVID19 pandemic.

Keywords: SARS-CoV-2, variants of concern, Variant of Interest, Variant under Monitoring, computational prediction**.**

INTRODUCTION:

The new coronavirus SARS CoV-2 that emerged late 2019 is still causing a pandemic of Severe Acute Respiratory Syndrome or COVID19. The pandemic affected over 370 million people and claimed more than 5.5 million [1], thus posing a difficult challenge to the scientific and health care communities throughout the world [2-4]. Indeed, SARS CoV-2 is a positive RNA virus that is constantly evolving through the accumulation of various type of mutations [5-8]. Even though the majority of these mutations do not affect the virus infectious properties and have no real impact on the progress of the pandemic [9], some may enhanced specific viral attributes that give the virus a selective advantage [10]. Any new variant endowed with selective advantage(s) would favor the virus persistence and nurture the pandemic. Therefore, watching and predicting how the pandemic evolves and communicate it to the public is of paramount importance. The global surveillance of the pandemic is based on multidisciplinary approaches including epidemiological, genetic, structural, and clinical data [11-14]. This involves the use of a set of relevant criteria to categorize the variants. Toward this-end, all international and national health and sanitary authorities have set various strategies to control the evolution of the SARS-CoV-2 pandemic. In May 31, 2021, the World Health Organization (WHO) announced a labeling system to categorize the variants into different levels of priority to better organize the global monitoring and research, and ultimately organize the "infodemic" and communicate more effectively with the public about the adequate response to the emergence of new variants of SARS-CoV-2 (https://www.who.int/). WHO has first developed a system to facilitate naming SARS-CoV-2 variants in addition to the existing nomenclature systems for naming and tracking SARS-CoV-2 genetic lineages established by GISAID (https://www.gisaid.org/), Nextstrain (https://nextstrain.org/) and Pango (https://cov-lineages.org/).These nomenclatures are mostly used by the scientific research community. For practical reasons (particularly to ease the communication), the WHO has settled to name the SARS-CoV-2 emerging variants using the Greek letter alphabet sequence (α, β, γ, δ.…). Since the SARS-CoV-2 virus is showing high genetic variability [15-17], WHO has established a labeling system for the variants into variant of concern (VOC), variants of interest (VOI) and variant under monitoring (VUM). This labeling system is based upon definitions related to variant phenotypic attributes such as transmissibility, disease presentation, effect on current diagnostic tests and response to available vaccines. This system was set to prompt and harmonize the actions needed to control the spread of a given variant. While the variants that emerged sequentially were named α, β, γ and δ, the last emerging one was named *Omicron* which does not comply with the sequence of the Greek alphabetic order that call for naming this variant "*Nu*". WHO deemed that "Nu" is prone to confusion with "New". *Xi* was also not used because it is a common surname in Asia. Therefore, this variant was named "*Omicron*". Of interest is that the American Center for disease control has also adopted this system of variant classification and labelling but added more labels and labeling criteria and a different labels' change policy (https://www.cdc.gov/). Indeed, while keeping the labels, variant of concern (VOC), variant of interest (VOI) and variant under monitoring (VUM) (calling the later "VBM" for Variant Being Monitored). In addition, the CDC uses an extra label that is the "Variant Of High Consequences" (VOHC). As a result, some of the currently known variants are given different labels by each agency. Indeed, according to the CDC

there is no SARS-CoV-2 variant labeled VOI as of December 1, 2021. Furthermore, variants α, β and γ that are currently labelled VOC by WHO, have been deescalated to the VBM label by the CDC as of September 2021 (https://www.cdc.gov/). These discrepancies reflect different views on the labeling of SARS-CoV-2 variants and consequently the use of the labels to set public health actions. On another hand, in both systems, variants labels can change with more data accumulating for a particular variant.

In this work, we have undertaken an evaluation of the system developed by WHO and adapted by the CDC to label the SARS-CoV-2 variants. We carried out a review of the classifications criteria and analyzed how WHO and the CDC use these criteria to label the SARS-CoV-2 variants. Then we carried out an exhaustive comparative computational study of the S protein mutations that characterize the five SARS-CoV-2 VOCs. We concluded that computational predictions provide a good ground of evidences for a rapid and more accurate labeling system.

MATERIALS & METHODS

Data mining and information sources: We retrieved the genetic, epidemiological and clinical data on the variants available as of December 15, 2021 from primary and secondary sources including the GISAID (https://www.gisaid.org/) and the Variants (http://variants.org) data banks. We collected the information on variants naming and labelling from the following sources: WHO: https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/. CDC: https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html#anchor_1632154493691. and PANGO Lineages.

Variant sequences retrieval and solvent exposure analysis: SARS CoV-2 spike protein extracellular domain amino acid sequence was obtained from the National Center for Biological Information (NCBI) protein ID: YP_009724390.1. Variants-specific mutations were introduced to the collected sequence based on the list published on (https://covariants.org/). The sequence corresponding to the different variants (*Alpha, Beta, Gamma, Delta* and *Omicron*) were analyzed for solvent exposure and possible epitope residues using Sequential B-Cell Epitope Predictor server (BepiPred-2.0 server: http://www.cbs.dtu.dk/services/BepiPred/cite.p). BepiPred-2.0 is based on a random forest algorithm trained on epitopes annotated from antibody-antigen protein structures [18].

SARS CoV-2 spike protein furin cut site (FCS) loop modeling: For the loop modeling, Phyre2 web server [19] was used to generate *Alpha, Beta, Gamma, Delta* and *Omicron* variants 3D models of the extracellular spike monomers, the results were saved and visualized on PyMOL [20].

Model quality assessment: As a quality assessment for the generated model, crystalized model of the spike protein (PDB ID 6VXX 2.80A° version 2.4) was downloaded from the RCSB database (https://www.rcsb.org). The structure was cleaned of water and heteroatoms, the complex was split and a PBD file for a monomer chain was created (chain B) and saved using PyMOl software. This monomer chain B was used as reference model for the sequence generated models. To define the common contact map between the crystal structures and the generated models, CMView [21] with the following parameters (contact type: Ca; Distance cut-off: 8.0; Needleman-Wunsch alignment),

was used. Different contact maps were established between the crystal structure and the models and the common contact percentage was calculated. Higher common contacts indicate more structural similarity and hence the models are suitable for further analysis. In addition, Tm align (https://zhanggroup.org/TM-align/) [22] was used to calculate TM-score value for each model. The superimposition Root Mean Square Deviation (RMSD) was calculated using PyMOL. Low RMSD and TM-scoring between 0.5 and 1.0 indicate that the two compared structures (the crystal and the model) has about the same fold.

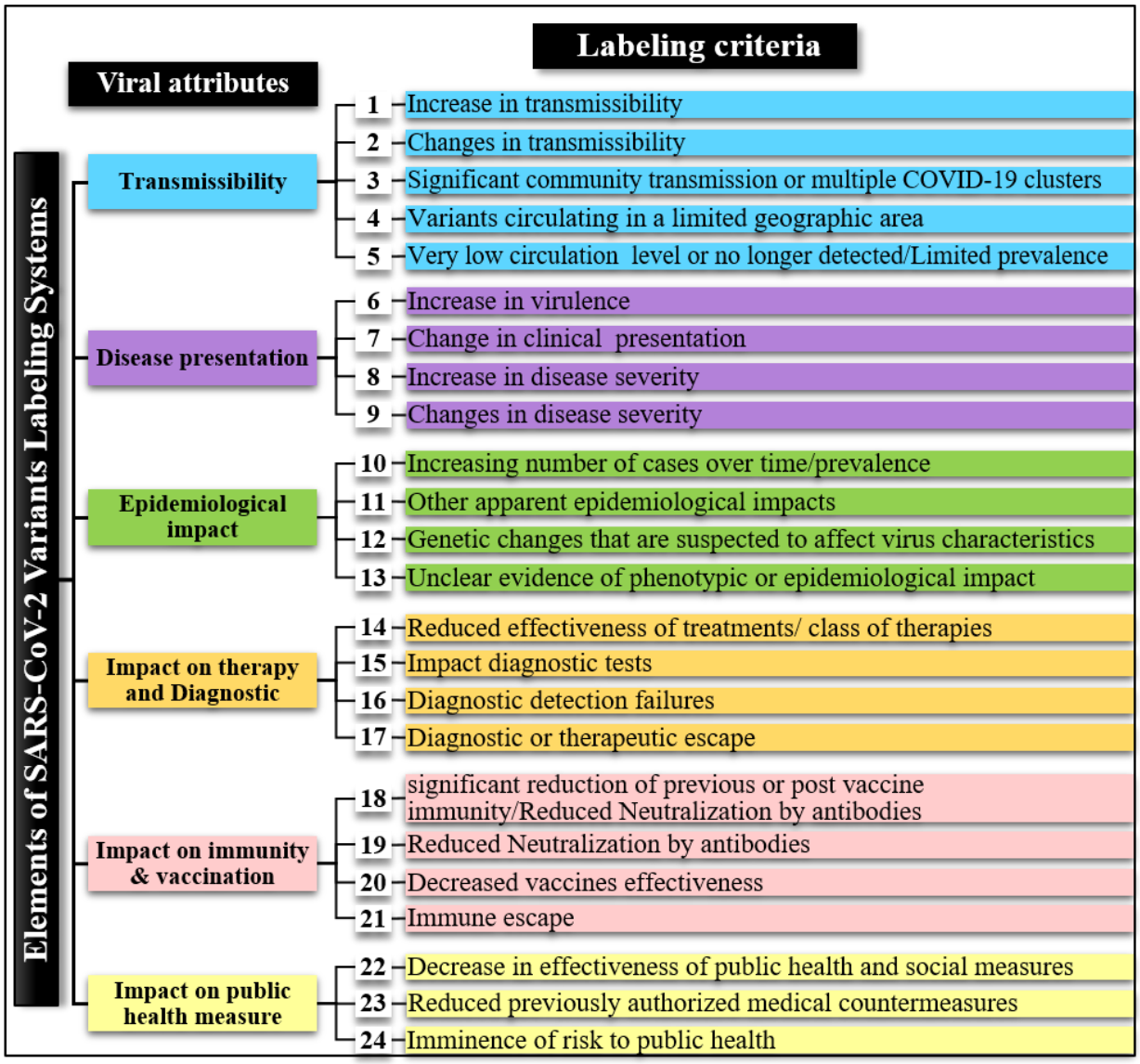
Mutational analysis: The effect of mutation on the interaction with ACE2 receptor: To analyze the effect of different RBD mutations on different SARS CoV-2 variants interaction with ACE2 receptor, PDB ID 6LZG structure was used as a model. The effect of single and accumulated mutations were evaluated by calculating changes in binding affinity ($\Delta\Delta G$) upon single or multiple mutations using MutaBind2 server (https://lilab.jysw.suda.edu.cn/research/mutabind2/) [23]. In addition, the server provide a structural model that was used to analyze polar interactions by LigPlot+ software [24].

RESULTS:

1. Study of SARS-CoV-2 variants labeling system

Expert groups at the world Health organization and the American Center of Disease Control have developed similar labeling systems to classify the new emergent SARS-CoV-2 variants. We retrieved a total of 24 different criteria from the working definitions elaborated by WHO and the CDC to give a particular label to a new variant. These criteria are derived from six viral attributes (Table 1). To each attribute, correspond a set of criteria that are formulated differently by each agency [see Table 1] who considers different criteria in their working definition of each label. Figure 1 shows that the two agencies use common (VOC, VOI) and different (VUM, VBM and VOHC) labels. Each agency uses generally different combination of criteria to give a variant a specific label except in four instances shown in Figure 1 where they use overlapping criteria for the same label (VOC and VOI). Consequently, there are discrepancies in the labels of currently active variants. Indeed, while WHO is currently labeling variants α, β ,γ, δ and *Omicron* as VOC, for the CDC, only variant *Omicron* is labeled VOC. Indeed, the CDC declassified variants α, β ,γ, and δ to VBM. In addition, there are no variants currently labelled VOI by the CDC. Figure 1 illustrates the differential use by WHO and the CDC of the viral attributes-derived criteria to label the SARS-CoV-2 genetic variants.

Table 1: SARS-CoV-2 variants attributes and derived criteria used by the World Health organization (WHO) and the American Center of Disease Control (CDC) for the labeling of emergent variants.



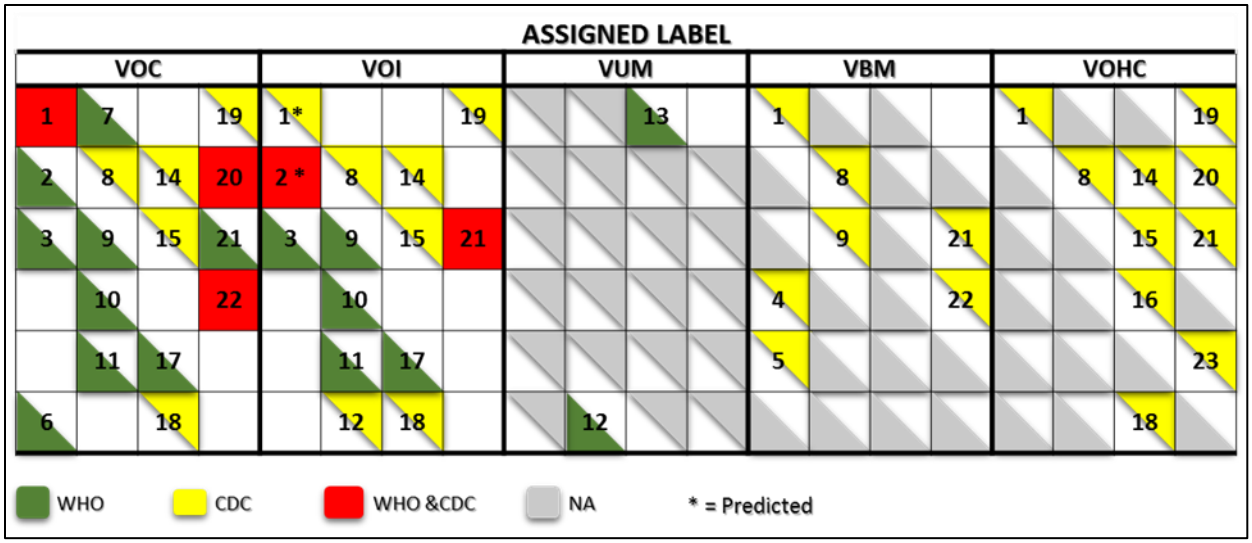


Figure 1: Matrix showing the commonalities and discrepancies of the criteria used by WHO and the CDC to label SARS CoV-2 variants. The colors green and yellow indicate the criteria used by WHO and the CDC respectively. The red indicates the criteria used by both agencies for a same label. The numbers correspond to the labeling criteria displayed in Table 1.

2. Attributes of the SARS-CoV-2 major Variants Of Concern

2.1. Epidemiological impact and Transmissibility:

While epidemiological data are available for variants α, β, γ and δ, data on the *Omicron* variant are limited and incomplete. This is mainly because not enough time has elapsed since the emergence of this variant to allow enough data accumulation and meaningful analysis. Table 2 summarizes the initial pieces of data available as of December 25, 2021 on the *Omicron* variant. For this variant, the early estimation of transmissibility increase is in the range of 3 to six. However, the previous data observed with the transmissibility of variant delta that became dominant worldwide and the currently observed high spreading /incidence of the *Omicron* variant makes the prediction of its transmissibility a 100-fold higher than the Delta variant plausible [25].

Table 2: Global Spreading of SARS-CoV-2 *Omicron* variant (B.1.1.529)

Date	Transmissibility	Incidence	Country	Reference
November 9 - 14, Detection of (B.1.1.529) variant	-	70	First 3 Botswana Hong Kong South Africa	Rao et al. 2021 https://doi.org/10.47488/dhrp.v1iS5.35 submission of the genomic sequence to GitHub
November 26 Classification as a Variant of Concern	High to very high		8	https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern
November 27	Prediction: 100-fold higher than the Delta variant	113	8	Rao et al. 2021 https://doi.org/10.47488/dhrp.v1iS5.35 https://bnonews.com/index.php/2021/11/omicron-tracker/
December 2	-	390	31	https://www.ecdc.europa.eu/en/news-events/epidemiological-update-omicron-variant-concern-voc-data-2-december-2021
December3	Three times higher Then other variant	486	38	https://www.nature.com/articles/d41586-021-03614-z https://www.ecdc.europa.eu/en/news-events/epidemiological-update-omicron-data-3-december-2021
December 4	-	689	42	https://www.gisaid.org/hcov19-variants/
December 7	three to six times higher than Delta	959	42	https://www.nature.com/articles/d41586-021-03614-z https://www.gisaid.org/hcov19-variants/
December 21	Six times	17,514	78	https://www.gisaid.org/hcov19-variants/
December 22	**Six times**	**20,322**	**79***	**https://www.gisaid.org/hcov19-variants/**

* Currently the omicron has gained global presence.

2.2 Disease presentation and Impact on therapy and Diagnostic

COVID-19 has three clinical presentation forms mild, moderate and severe. WHO and the CDC use four labeling criteria related to clinical presentation and four others pertaining to the impact a given variant on therapy and diagnostic. The two agencies use different formulation for all of these criteria (see Table 1) and use different combination of these criteria to label SARS-CoV-2 variants (see Figure1). There is no explicit mention to the disease forms in the variant labeling criteria related to disease presentation in both WHO and CDC variant-labeling usage.

2.3. Impact on immunity, vaccination and public health measures

WHO and the CDC use also different formulation for the criteria related to these attributes and use them differently to attribute of a specific label (Table1, Figue1). For instance, to WHO immune escape is a criterion used to label a variant VOI and VOC. However, the CDC uses it for labeling a variant VOI, VBM and VOHC. *Omicron* has been shown to have extensive but incomplete escape of Pfizer BNT162b2 vaccine [26] thus fulfilling the criterion of decreased vaccine effectiveness. For the criterion related to the evaluation of an imminent risk to public health a variant can pose, only the CDC uses it for the VOHC label. Table 3 show how the five

major SARS-CoV-2 variants fulfill the criteria used by WHO and the CDC for variants labeling regardless of the actual labels given by each agency.

Table 3: Actual application of the criteria formulated by WHO and the CDC to the five major SARS-CoV-2 variants.

Labeling Criteria

Variants	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
α																								
β																								
γ																								
δ																								
ο							#							#				#			#		#	

 # based on prediction

3. Genetic variations data:

Figure 2 shows a comparative mapping of the mutation profile of the *Omicron* variant with those of the *Alpha*, *Beta*, *Gamma* and *Delta* variants. Most of the mutations affect the SARS-CoV-2 S protein. *Omicron* display 63 different mutations as compared to the Wuhan Strain. Thirty-six mutations occurred in the S spike protein and fifteen are clustered in the RBD region. The representation of the mutations in the S protein 3D models of SARS-CoV-2 five VOCs shows that most of the mutations map to the solvent exposed regions (Figure 3 and Supplementary Table 1).

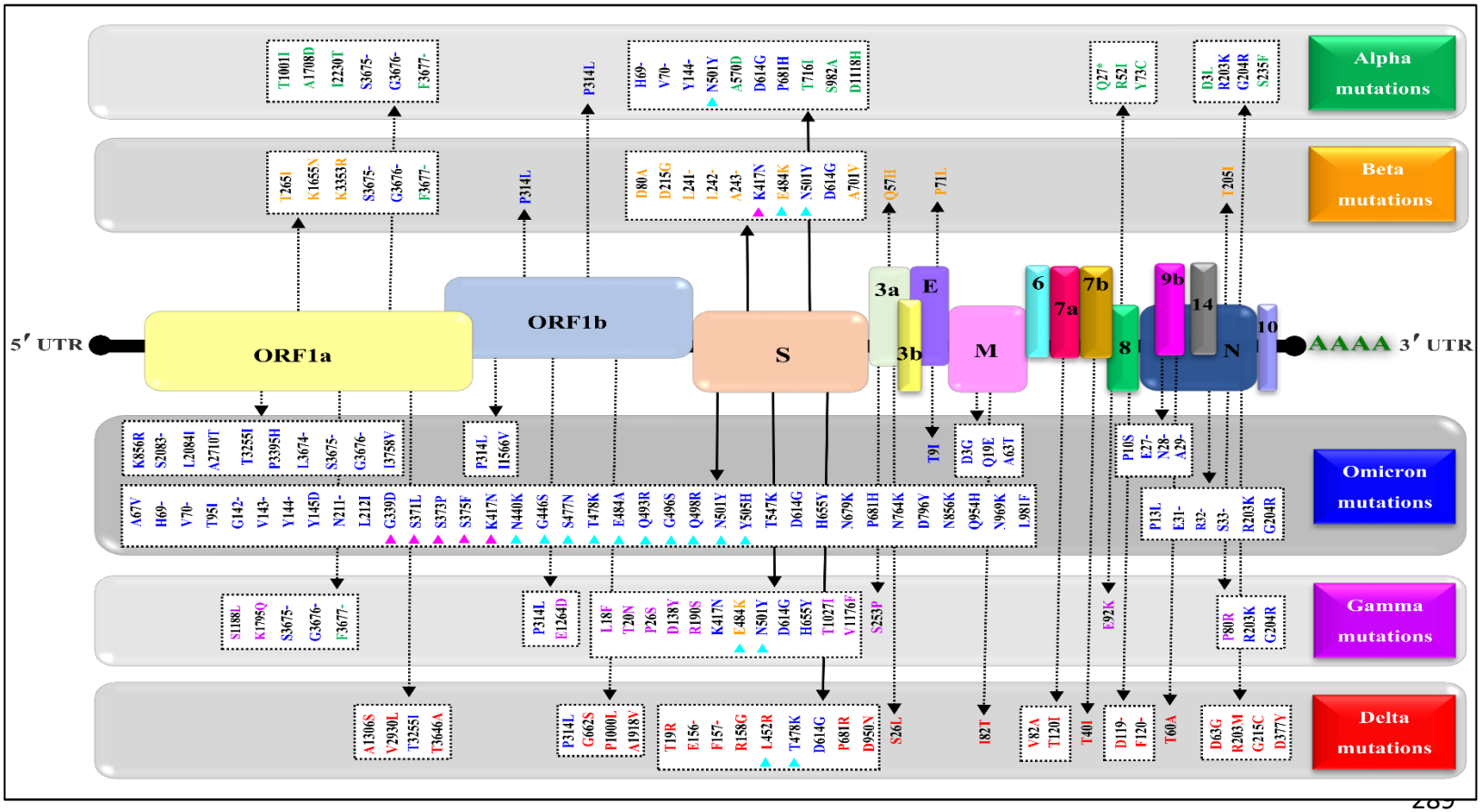


Figure 2. Mapping of the nonsynonymous mutations' characteristics of the SARS-CoV-2 five variants. A typical genomic organization of SARS-CoV-2 contains the following: 5' end UTR; Open reading frames: ORF 1a and ORF 1b; the structural genes coding for the Spike (S) protein, the Envelope (E), the Membrane (M), and the Nucleocapsid. The accessory genes such as (3a, 3b, 6, 7a, 7b, 8, 10 &14) are distributed among the structural genes. The 3' end UTR and follows the poly (A) tail. The green, yellow, blue purple and red show respectively the synonymous mutations characteristic of variant the Alpha (α) variant (B.1.1.7), Beta (β) variant (B.1.351), Omicron (o) variant (B.1.1.529), Gamma (γ) Variant (P.1) and Delta (δ) variant (B.1.617.2). The (-) represents the deletion, (*) represents the stop codon, magenta triangles indicate variations in the receptor-binding domain (RBD) and cyan triangles indicate variations in the receptor-binding motif (RBM).

2. Effect of Omicron S protein mutations on the immunogenicity:

Non-synonymous mutations of the different SARS-CoV-2 variants caused changes on the epitope probability and antibody exposure hence immunogenicity. Most of the epitope changes are noticeable in the S1 domain of the spike protein. Figure 4 shows the percentage of the exposed epitope on the different variants. Detailed probability and exposure states of each residue as predicted by BepiPred are listed in supplementary Table 1.

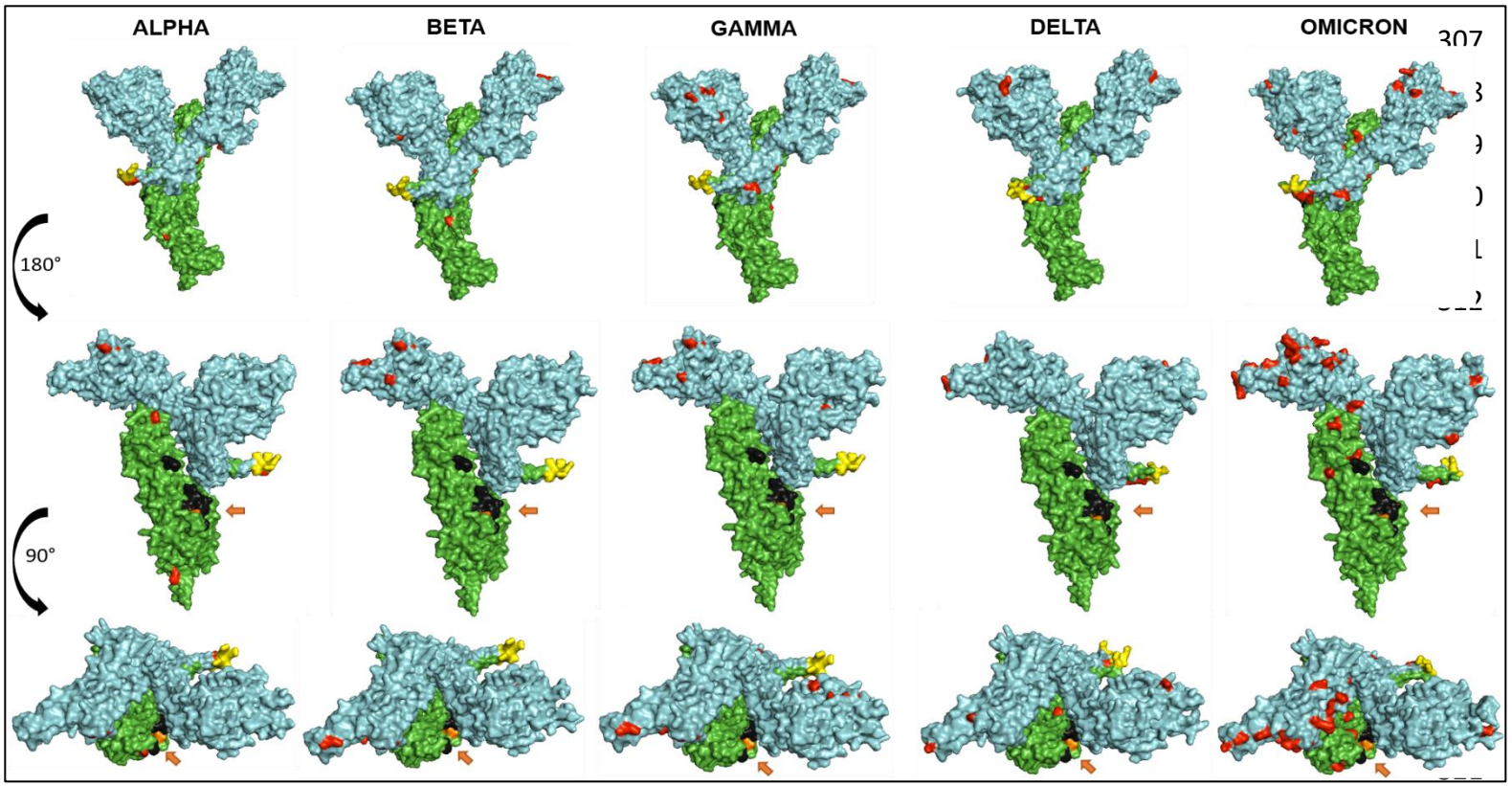


Figure 3. Representation of the surface models of SARS CoV-2 variant S spike protein (monomer). Side views (upper and middle rows), Top view (third row). Colors code is as follow: Mutations (Red), S1 subunit (Cyan), S2 subunit (Green), Furin Cleavage Site (Yellow), Fusion peptides FP1 and FP2 (Black), and the arrows show the TMPRSS2 cleavage site (Orange).

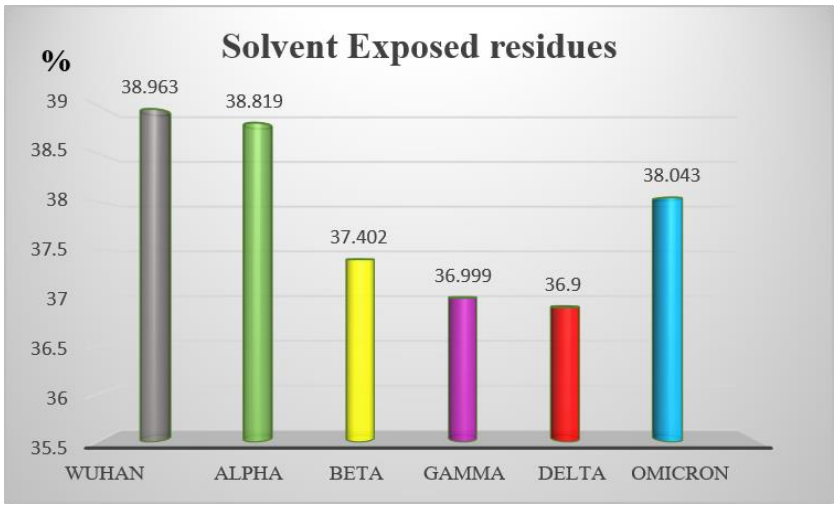


Figure 4. The percentage of solvent exposed residues in the S protein of the five SARS-CoV-2 VOCs.

3. Effect of Omicron S protein mutations on the 3D structure of the molecule: Loop modeling and quality assessment:

Computational methods allow the building of accurate protein models of the SARS-CoV-2 S protein based on data input and the alignment with experimentally solved multiple templates molecules. We used the Phyre2 web server to generate 3D models for the extracellular domain of the variants α, β ,γ ,δ and O monomers spike protein [Fig 5]. The structures were generated with 100% confidence and 84% coverage for *Alpha, Beta and Omicron*, and 83% coverage for *Gamma* and *Delta* by the single highest scoring template. The quality assessment of these models was obtained by Superimposition with the crystalized structure of the SARS-CoV-2 spike glycoprotein (closed state) 6VXX Chain B, showed an RMSD of 1.086 for *Alpha*, 1.085 for *Beta*, 1.079 for *Gamma*, 1.075 for *Delta* and 1.075 for *Omicron* (Fig. 5). The Tm-scores were all less than 1 and both RMSD and TM-scoring values are acceptable and indicates high similarity between the crystalized model and the generated one showing the same folds. Inanition, CMView common contact map gave a high similarity score ranging between 80% - 81.5% keeping in mind that the crystalized structure is missing few residues (gaps) including the furin cleavage loop. What is important here is that these generated models include the furin cleavage site that is lacked in all the crystalized models deposited on the protein database (PDB). With highly reliable Tm score and RMSD values, these computationally generated models are suitable to be used for further analysis especially for evaluating the effect of mutation on and around the furin cleavage site.

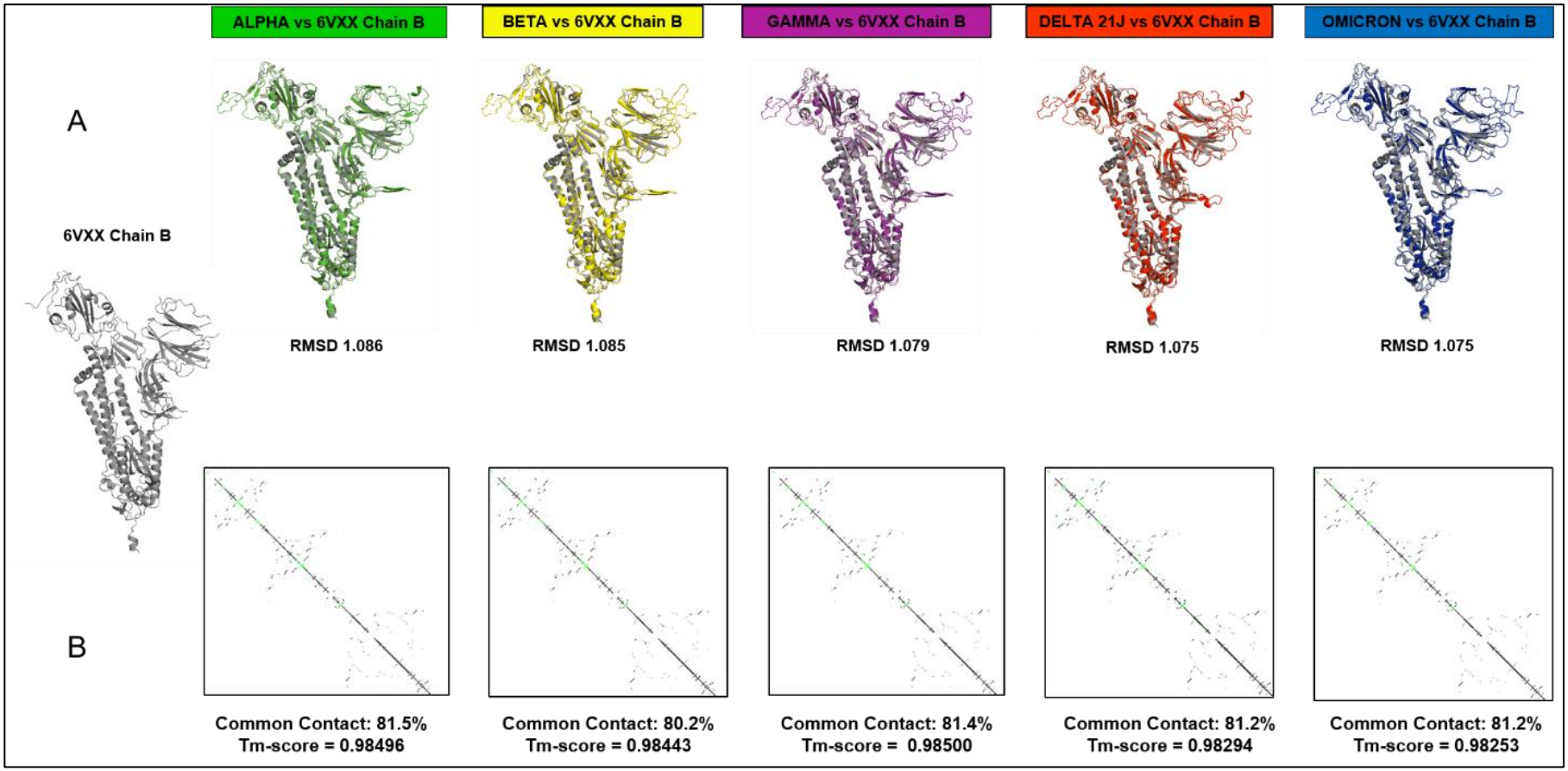


Figure 5. This figure represents the 3D models generated for the α, β ,γ ,δ and O variants, their structural analysis and the quality assessment by superimposition with the chain B of the crystalized model PBD 6VXX (A) Superimposition RMSD values (B) Calculated common contact percentages and Tm-scores.

4. The effect of contact residues' mutations on the SARS CoV-2 S protein/ACE2 complex thermodynamic stability:

Sequence analysis showed that *omicron* variant has the most mutated RBD with 15 different mutations out of which, nine mutations were in the contact residues with ACE2 receptor. The *Alpha* variant showed only one mutation in the contact residues (N501Y), *Beta* and *Gamma* showed three mutations all are of the contact residues while *Delta* showed two mutations none of which are in the contact residue with ACE2. Analyzing theses mutations as single mutation showed different effect on the complex of SARS-CoV2 RBD with ACE2 receptor. Where some mutations showed to be deleterious as single mutation, others have stabilizing effect. However, the combination of several mutations in the contact residue shows different effect; in the case of the *omicron* variant, the combination of 9 mutations in the contact residues showed to be not deleterious even though they slightly destabilize the complex with ACE2 with an accumulative $\Delta\Delta G$= 0.18. *Beta* and *Gamma* variant with three mutations in the contact residue showed an accumulative $\Delta\Delta G$ of 1.79 and 1.07 respectively, which highly destabilizes the complex and decrease the binding affinity. Delta has only two mutations that occurs out of the contact residues with a stabilizing effect on the complex and hence increasing affinity to ACE2 and an accumulative $\Delta\Delta G$= -0.33. The single contact residue mutation of the *Alpha* variant (N501Y) showed destabilizing effect that we reported previously [9] (Table 5).

Table 5. Comparison of the mutations profile of five SARS-Cov-2 VOCs S protein in the contact residues with the ACE2 receptor. Positive and negative signs correspond respectively to destabilizing (decreases in binding affinity) and stabilizing mutations (increases in binding affinity).

6LZG Contact residues	Alpha B.1.1.7	Beta B.1.351	Gamma P.1	Delta B.1.617.2	Omicron BA.1	Single mutations' ΔΔG	Effect of single mutation
K417		K417N			K417N	0.75	Destabilizing
			K417T			0.18	Destabilizing
G446					G446S	-0.18	Stabilizing
G447							
Y449							
-				L452R		-0.57	Stabilizing
Y453							
L455							
F456							
Y473							
A475							
G476							
S477					S477N	-0.15	Stabilizing
-				T478K		0.19	Destabilizing
E484		E484K	E484K		E484A	0.17	Destabilizing
G485							
F486							
N487							
Y489							
F490							
L492							
Q493					Q493R	0.25	Destabilizing
Y495							
G496					G496S	1.2	Destabilizing
F497							
Q498					Q498R	2.34	Destabilizing
P499							
T500							
N501	N501Y	N501Y	N501Y		N501Y	1.53	Destabilizing
G502							
V503							
Y505					Y505H	1.33	Destabilizing
Q506							
Accumulative ΔΔG	1.53	1.79	1.07	-0.33	0.18		
Effect of accumulative mutation	Destabilizing	Destabilizing	Destabilizing	Stabilizing	Destabilizing		

Table 6. Comparison of contact map percentage (Red) and TM-align (Blue) scoring between the different SARS CoV-2 VOCs. The arrowheads denote the variant used as a reference when calculating the data.

TM-align →

Contact map % ↑

	α	β	γ	δ	ο
α		0.99321	0.99811	0.99374	0.99244
β	96.1		0.99361	0.98905	0.99023
γ	98.4	96.9		0.99547	0.99007
δ	95.8	94.8	97.1		0.98892
ο	97.0	95.5	96.8	95.4	

5. Effect of mutations in the S protein of SARS-CoV-2 major variants on the polar interactions with the ACE2 receptor:

Using the crystalized structure PDB ID 6LZG that represent the interaction between SARS-CoV-2 receptor binding domain with ACE2 receptor as a reference, we generated models for the five variants on the LigPlot+ software and analyzed the polar interaction patterns. The interacting residues and type of interactions are listed in table 7. All the variants have one or more mutations on the contact residues except for the *Delta* variant. At a first glance to Table 7 and by comparison, one can spot the interaction pattern similarity between *Alpha* and *Delta* variants despite the fact that the *Alpha* variant showed only one mutation in the contact residue (N501Y) and *Delta* has two mutations out of the contact residues (L452R and T478K). Both *Alpha* and *Delta* variants were reported to be highly transmissible with *Delta* being 60% higher [27]. Delta shows that eight out of 9 main polar interactions are present with one missing polar interaction (Spike/ACE2: Asn487/Tyr83) and the addition of new two polar interactions (Gln493/Glu35 and Tyr505/Glu37) and one salt bridge (Glu484/Gln24). Similarly, *Alpha* is missing the same polar interactions but shows the same new salt bridge Glu484/Lys31 and no additional new polar interactions. The addition of new polar interactions and the Glu484/Lys31 salt bridge in the Delta variant could be the cause of the stabilizing effect on the complex with ACE2, which can explain the high transmissibility.

In addition to Glu484, other mutations on RBD binding hot spots residues, have been also linked to antibody binding and neutralization including mutations on Lys417, Gly446, Phe456, Asn501, Gly477 and Asp614 [28-31]. Interestingly, *omicron* variant compiles all these mutations (Table 5) indicating a potential higher degree of immune escape. Additionally, *omicron* has a consecutive mutation on residue Gln493, Gly496 and Gln498 that dramatically affect the interaction with ACE2. These residues are structurally in the receptor binding ridge and Gln493 is known to form with Leu455 the two receptor binding motif (RBM) stabilizing hot spots and to be a target for some therapeutic antibodies [32, 33]. This interaction disrupted profile may sum up into less stable complex with ACE2 ($\Delta\Delta G$= 0.18) (Table 7) and loss of antibody neutralization. Furthermore, the comparison of the interactions' differences between SARS-CoV-2 variants showed some interactions that are conserved in all the variants suggesting their importance in the complex stability.

Table 7. Contact residues' interaction pattern of the S protein/ ACE2 complex of different SARS CoV-2 variants. Polar interactions (P), hydrophobic interactions (H), Missing interactions are highlighted in red, new interactions highlighted in yellow, salt bridges are highlighted in green.

Spike RBD residue (PDB 6LZG-version 2.4)	Interaction residue on ACE2 isoform 1	SARS CoV-2 variants				
		Alpha	Beta	Gamma	Delta	Omicron
Lys417	1X Asp30(P) 1X Asp30(H)	1X Asp30(P) 2X Asp30(H)	Missing Missing	Missing Missing	1X Asp30(P) 2X Asp30(H)	Missing Missing
Tyr449	1X Gln42(P) 1X Asp38(P) 3X Asp38(H)	Missing 1X Asp38(P) 3X Asp38(H)	Missing 1X Asp38(P) 3X Asp38(H)	Missing 1X Asp38(P) 3X Asp38(H)	Missing 1X Asp38(P) 3X Asp38(H)	Missing 1X Asp38(P) 3X Asp38(H)
Tyr453	1X His34(H)	1X His34(H)	1X His34(P) 2X His34(H)	1X His34(P) 2X His34(H)	1X His34(H)	1X His34(P) 2X His34(H)
Leu455	1X His34(H)	1X His34(H) 1XAsp30(H)	3X His34(H)	3X His34(H)	1X His34(H) 1XAsp30(H)	3X His34(H)
Phe456	1X Thr27(H) 1X Asp30(H)	1X Thr27(H) 1X Asp30(H)	1X Thr27(H) Missing 1X Lys31(H)	1X Thr27(H) Missing 1X Lys31(H)	1X Thr27(H) 1X Asp30(H)	1X Thr27(H) missing 1X Lys31(H)
Ala475	1X Ser19(P) 2X Ser19(H) 1X Gln24(H)	1X Ser19(P) 2X Ser19(H) 1X Gln24(H) 1X Thr27(H)	1X Ser19(P) 2X Ser19(H) 1X Gln24(H) 1X Thr27(H)	1X Ser19(P) 2X Ser19(H) 1X Gln24(H)	1X Ser19(P) 2X Ser19(H) 1X Gln24(H) 1X Thr27(H)	1X Ser19(P) 2X Ser19(H) 1X Gln24(H) 1X Thr27(H)
Gly476	1X Ser19(H)	1X Ser19(H)	1X Ser19(H)	1X Ser19(H)	1X Ser19(H)	1X Ser19(H)
Glu484		Salt Lys31			Salt Lys31	
Phe486	1X Met82(H) 4X Tyr83(H)	1X Met82(H) 1X Tyr83(H)	1X Met82(H) 1X Tyr83(H)	1X Met82(H) 1X Tyr83(H)	1X Met82(H) 1X Tyr83(H)	1X Met82(H) 1X Tyr83(H)
Asn487	1X Gln24(P) 1X Tyr83(P) 6X Gln24(H) 3X Tyr83(H)	1X Gln24(P) Missing 4X Gln24(H) 2X Tyr83(H)	1X Gln24(P) Missing 4X Gln24(H) 2X Tyr83(H)	1X Gln24(P) Missing 4X Gln24(H) 2X Tyr83(H)	1X Gln24(P) Missing 4X Gln24(H) 2X Tyr83(H)	1X Gln24(P) Missing 4X Gln24(H) 2X Tyr83(H)
Tyr489	1X Thr27(H) 1X Phe28(H)	1X Thr27(H) 1X Phe28(H)	1X Thr27(H) 1X Phe28(H) 1X Lys31(H)	1X Thr27(H) 1X Phe28(H)	1X Thr27(H) 1X Phe28(H)	1X Thr27(H) 1X Phe28(H)
Gln493	1X His34(H) 1X Glu35(H)	1X Glu35(P) 1X His34(H) 2X Glu35(H)	1X Glu35(P) Missing 3X Glu35(H)	1X Glu35(P) Missing 2X Glu35(H)	1X Glu35(P) 1X His34(H) 2X Glu35(H)	1X His34(P) Salt Glu35 4X His34(H) 7X Glu35(H) 1X Asp38(H)
Gly496	1X Lys353(P)	Missing 1X Asp38(H)	Missing 1X Asp38(H)	Missing 1X Asp38(H)	1X Lys353(P) 1X Asp38(H) 3X Lys353(H)	1X Lys353(P) 1X Asp38(P) 5X Asp38(H)
Gln498	1X Gln42(P) 2X Tyr41(H) 2X Gln42(H) 1X Leu45(H)	1X Gln42(P) 6X Tyr41(H) 2X Gln42(H) 1X Leu45(H)	1X Gln42(P) 2X Tyr41(H) 2X Gln42(H) 3X Leu45(H)	1X Gln42(P) 2X Tyr41(H) 2X Gln42(H) 3X Leu45(H)	1X Gln42(P) 4X Tyr41(H) 2X Gln42(H) 1X Leu45(H)	Missing 1X Tyr41(H) 1X Gln42(H) 1X Leu45(H)
Thr500	1X Tyr41(P) 3X Tyr41(H) 1X Asn330(H) 2X Asp355(H) 2X Arg357(H)	1X Tyr41(P) 2X Tyr41(H) 1X Asn330(H) 2X Asp355(H) 2X Arg357(H)	1X Tyr41(P) 2X Tyr41(H) 1X Asn330(H) 1X Asp355(H) 2X Arg357(H)	1X Tyr41(P) 2X Tyr41(H) 1X Asn330(H) 1X Asp355(H) 2X Arg357(H)	1X Tyr41(P) 3X Tyr41(H) 1X Asn330(H) 2X Asp355(H) 2X Arg357(H)	1X Tyr41(P) 2X Tyr41(H) 1X Asn330(H) 2X Asp355(H) 2X Arg357(H)
Asn501	3X Tyr41(H) 1X Lys353(H)	5X Tyr41(H) 5X Lys353(H)	1X Asp38(P) 8X Tyr41(H) 12X Lys353(H)	1X Asp38(P) 9X Tyr41(H) 12X Lys353(H)	4X Tyr41(H) 3X Lys353(H)	1X Lys353(P) 9X Tyr41(H) 16X Lys353(H)
Gly502	1X Lys353(P) 1X Lys353(H) 2X Gly354(H)	1X Lys353(P) 1X Lys353(H) 2X Gly354(H)	1X Lys353(P) 1X Lys353(H) 2X Gly354(H)	1X Lys353(P) 1X Lys353(H) 2X Gly354(H)	1X Lys353(P) 1X Lys353(H) 2X Gly354(H)	1X Lys353(P) 1X Lys353(H) 2X Gly354(H)
Tyr505	5X Lys353(H) 1X Gly354(H)	4X Lys353(H) 1X Gly354(H)	1X Glu37(P) 12X Lys353(H) 1X Gly354(H) 4X Glu37(H)	1X Glu37(P) 11X Lys353(H) 1X Gly354(H) 4X Glu37(H)	1X Glu37(P) 12X Lys353(H) missing 3X Glu37(H)	4X Lys353(H) 1X Gly354(H)
Total interactions	10P/ 56H	8P/ 1 salt / 63H	10P/ 83H	10P/80H	10P/ 1 salt/ 72H	10P/ 1 salt/ 85H
ΔΔG		**1.53**	**1.79**	**1.07**	**-0.33**	**0.18**
Effect		**Destabilizing**	**Destabilizing**	**Destabilizing**	**Stabilizing**	**Destabilizing**

DISCUSSION:

Despite significant global health care measures and social mitigation efforts along with the availability of a number of vaccines, the SARS-CoV-2, COVID19 pandemic entered its 3rd year and the numbers of cases are soaring worldwide (https://www.gisaid.org/, https://coronavirus.jhu.edu/map.html). This is mainly due to the emergence of new and more fitted viral variants that keep fueling this pandemic. To create awareness about any additional health issue a new variant may cause, WHO the world most influential health agency, has developed a labeling system to classify the new variants, in addition to the conventional nomenclature. The CDC has adopted this labeling system but introduced major changes. Among the labels used by WHO and the CDC is the label "Variant of Concern or VOC". The word "concern" is synonymous of anguish, anxiety and apprehension. It naturally enjoins fear and disturbance, which predispose to take immediate actions and make odd changes. Indeed, when a new variant is labeled VOC, countries are inclined to close their borders and take drastic counter measures. This was particularly striking when a number of countries banned travels from South Africa that was the first country to describe and report the genomic sequence of the *omicron* variant. While preventive and cautionary actions are mandatory to control pandemics, these should be based on solid scientific evidences. However, it takes time and coordinated efforts for the scientific community to generate the data needed to label accurately new variants according to criteria such as transmissibility, disease severity, and change in the epidemiological pattern, immune escape, and resistance to previously neutralizing antibodies, efficacy of existing therapies or vaccine efficiency. Meanwhile, the rapid pandemic progress requires timely response including a good communication system. In this study, we highlight some flaws in the labeling system developed by WHO mainly to ease communication with the national health authorities and the public. Indeed, in addition to the non-appropriate wording for labeling emergent variant, the combination of criteria used to define a variant label as shown in WHO and the CDC web sites, is not accurate. Indeed, in their definitions of the different labels, both agencies build a combination of criteria using the prefix "OR" but not "AND". This introduces a confusion that is amplified by the CDC use of more labels, and different formulation and different combinations of the criteria used to attribute a label to a new variant. According to WHO a label is supposedly assigned to a variant through a comparative assessment with the previous ones. Even though SARS-CoV-2 variants are primarily detected upon the virus genomic sequence changes and particularly mutations in important functional regions, the genetic variations do not clearly appear among the viral attributes WHO and the CDC use to formulate their definition of the different variants. The labeling system developed by WHO and the definitions of the criteria for each label used by both WHO and the CDC do not mention a comparative assessment of the genetic variations but rather retain clinical, epidemiological, transmissibility and other non-genetic viral attributes. Such parameters need a substantial time to be accurately determined. It seems that the current labeling system relies more on fear from emergent genetic variations "the super killer virus" than evidences about the impacts of mutations [34]. Thus, it seems premature and not appropriate to give a new variant the VOC label. Even when a greater transmissibility can be established, this does not necessarily mean greater severity. For example, the very transmissible H1N1 influenza virus variant was not as severe as many other influenza viruses and the epidemic has naturally faded away [35]. With reference to WHO variant labeling system H1N1 could have been labeled VOC while it has never been one. Therefore, to solve the dilemma between the necessity of rapidly labeling new variant and the lack of scientific evidences, a good way is to analyze thoroughly the genetic variations

observed to generate rapidly the best data on the potential attributes of the new viral variant. To achieve this analysis, the computational prediction of variant effect on protein stability, function and interaction is a very useful way to determine the variant "importance". Several methods can be used based on the availability or not of 3D structure of the protein. In this study, we show that computational prediction and *in silico* comparative studies between new variants and older more characterized ones, can provide good insights into the impact of the mutations and thus the potential behavior of the emergent variants. Indeed, the observation that most of the mutations in SARS-CoV-2 VOCs map to the solvent exposed regions along with the comparison of antigenicity predictions, represent a good basis to make a fair assumption on a potential immune escape and/or a likely reduction of antibodies neutralization and to anticipate an evaluation of vaccine effectiveness. In addition, the comparative analysis of the S protein 3D structure between the five SARS-CoV-2 VOCs showed an identical overall folding as demonstrated by the RMSD values and Tm-scores of each model. Meanwhile the comparison of contact residues maps gave a fine tuning of the structural divergence which suggest potential functional differences that need to be further analyzed.

Furthermore, computational studies of the mutations effect on the thermodynamic stability of the S protein /ACE2 complex and the comparative analysis of the pattern of polar interactions of the different variant with the ACE2 receptor, give good indications to predict transmissibility and virulence and draw some plausible epidemiological scenario. Indeed, the effect of mutations can be computed as single or combined mutations. The data we obtained with the combined mutations of the *omicron* variant that has a much higher number of mutations than the other variants, shows that this variant engages into a more stable interaction with ACE2 than the β and γ variants do and has less interactions than the delta variant.

For the other variant labeling criteria, several studies discussed the importance of the Glu484Lys mutation in the interaction with ACE2 and immune escape [32, 36]. This was also observed with the *Beta* and *Gamma* variant that carry Glu484Lys mutation. It was reported that the *Beta* variant have reduced antibody neutralization compared to the *Delta* variant. Besides *Beta*'s resistance to neutralizing antibodies increased by 9.4 fold to convalescent plasma and 10.3 to 12.4 fold for sera from individuals who have been vaccinated [37]. In addition, it was suggested that new variants with the same mutation might bear new challenges for current vaccines or monoclonal antibody therapies [38, 39].

Meanwhile our analysis of the pattern of polar interactions between the different variants shows that *omicron* has an extra 13 hydrogen bonds as compared to variant *delta*. In addition, we noticed the presence in the more transmissible *alpha* and *delta* variants of an extra salt bridge between the S protein Glu 484 and ACE2 lysine 31 and the presence of an extra salt bridge between the S protein Gln 493 and ACE2 Glu 35. This observation combined to the observed high number of H bonds and the thermodynamic stability data, predicts a potential more efficient entry into host cells and enhanced transmissibility of the *omicron* variant, which has been confirmed since the description of *omicron* mutations profile. Nevertheless, relying on the study of a single mutation or the computing of one biophysical feature of a variant structure to predict how a viral attribute would evolve is not sufficient. Furthermore, multiple genes can control epidemiologically relevant viral attributes such as the mode of transmission and virulence. Therefore, it is recommended to integrate the data on multiple mutations with computing of various viral structural features to make the best predictions and attribute the right label to a given variant.

In conclusion, the system of SARS-CoV-2 labeling developed by WHO and amended by the CDC has some major flaws. Relying on the integrated biophysical and structural data generated from computational comparative predictions of the likely behavior of a new variant would help in the rapid and accurate labeling of emergent variants. Meanwhile, given our provisional and incomplete knowledge and the uncertain nature of the COVID19 pandemic, it would be wise to operate in epistemic self-abnegation, use the best tools and knowledge we have at hand and introduce revisions whenever new evidences become available.

REFERENCES:

1. Johns-Hopkins-University. *The Johns Hopkins University COVID-19 Dashboard*. 2022; Available from: https://coronavirus.jhu.edu/map.html.
2. Tutelyan, V.A., et al., *[COVID-19: new challenges for medical science and practical health].* Vopr Pitan, 2020. **89**(3): p. 6-13.
3. Malik, Y.S., et al., *Coronavirus disease pandemic (COVID-19): challenges and a global perspective.* 2020. **9**(7): p. 519.
4. World-Economic-Forum, *Challenges and Opportunities in the Post-COVID-19 World* 2020.
5. Zhao, J., et al., *Snapshot of the evolution and mutation patterns of SARS-CoV-2.* 2020.
6. Banoun, H., *Evolution of SARS-CoV-2: Review of Mutations, Role of the Host Immune System.* Nephron, 2021. **145**(4): p. 392-403.
7. Majumdar, P. and S. Niyogi, *SARS-CoV-2 mutations: the biological trackway towards viral fitness.* Epidemiol Infect, 2021. **149**: p. e110.
8. Singh, J., et al., *Evolutionary trajectory of SARS-CoV-2 and emerging variants.* Virol J, 2021. **18**(1): p. 166.
9. Ashoor, D., et al., *A Computational Approach to Evaluate the Combined Effect of SARS-CoV-2 RBD Mutations and ACE2 Receptor Genetic Variants on Infectivity: The COVID-19 Host-Pathogen Nexus.* 2021. **11**(722).
10. Gobeil, S., et al., *Effect of natural mutations of SARS-CoV-2 on spike structure, conformation and antigenicity.* 2021.
11. Agency, U.H.S. *SARS-CoV-2 variants of concern and variants under investigation in England (Technical briefing 27)*. 2021; Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1029715/technical-briefing-27.pdf.
12. England, P.H. *SARS-CoV-2 variants of concern and variants under investigation in England Technical briefing 12* 2021; Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/988619/Variants_of_Concern_VOC_Technical_Briefing_12_England.pdf.
13. France, S.P. *Nouveau variant détecté et sous surveillance en Bretagne 2021 [updated 2021-03-16]*. 2021; Available from: https://www.santepubliquefrance.fr/presse/2021/nouveau-variant-detecte-et-sous-surveillance-en-bretagne.
14. Campbell, F., et al., *Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021.* 2021. **26**(24): p. 2100509.
15. Dubey, A., et al., *Emerging SARS-CoV-2 Variants: Genetic Variability and Clinical Implications.* 2022. **79**(1): p. 1-18.
16. Yazdani, S., et al., *Genetic Variability of the SARS-CoV-2 Pocketome.* 2021. **20**(8): p. 4212-4215.
17. Toyoshima, Y., et al., *SARS-CoV-2 genomic variations associated with mortality rate of COVID-19.* 2020. **65**(12): p. 1075-1082.
18. Jespersen, M.C., et al., *BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes.* 2017. **45**(W1): p. W24-W29.

633 19. Kelley, L.A., et al., *The Phyre2 web portal for protein modeling, prediction and analysis.*
634 2015. **10**(6): p. 845-858.
635 20. DeLano, W., *The PyMOL Molecular Graphics System, Version 2.4. 0 Schrödinger, LLC.*
636 2002.
637 21. Vehlow, C., et al., *CMView: interactive contact map visualization and analysis.* 2011.
638 **27**(11): p. 1573-1574.
639 22. Zhang, Y. and J. Skolnick, *TM-align: a protein structure alignment algorithm based on the*
640 *TM-score.* Nucleic Acids Res, 2005. **33**(7): p. 2302-9.
641 23. Zhang, N., et al., *MutaBind2: predicting the impacts of single and multiple mutations on*
642 *protein-protein interactions.* 2020. **23**(3): p. 100939.
643 24. Laskowski, R.A. and M.B. Swindells, *LigPlot+: multiple ligand-protein interaction*
644 *diagrams for drug discovery.* J Chem Inf Model, 2011. **51**(10): p. 2778-86.
645 25. Rao, S. and M. Singh, *The Newly Detected B. 1.1. 529 (Omicron) Variant of SARS-CoV-2*
646 *With Multiple Mutations: Implications for transmission, diagnostics, therapeutics, and*
647 *immune evasion.* 2021. **1**(S5): p. 7-10.
648 26. Cele, S., et al., *SARS-CoV-2 Omicron has extensive but incomplete escape of Pfizer*
649 *BNT162b2 elicited neutralization and requires ACE2 for infection.* 2021.
650 27. Duong, D., *Alpha, Beta, Delta, Gamma: What's important to know about SARS-CoV-2*
651 *variants of concern?* 2021, Can Med Assoc.
652 28. Greaney, A.J., et al., *Comprehensive mapping of mutations in the SARS-CoV-2 receptor-*
653 *binding domain that affect recognition by polyclonal human plasma antibodies.* 2021.
654 **29**(3): p. 463-476. e6.
655 29. Liu, Z., et al., *Identification of SARS-CoV-2 spike mutations that attenuate monoclonal*
656 *and serum antibody neutralization.* 2021. **29**(3): p. 477-488. e4.
657 30. Montefiori, D.C., *SnapShot: SARS-CoV-2 antibodies.* 2021. **29**(7): p. 1162-1162. e1.
658 31. Supasa, P., et al., *Reduced neutralization of SARS-CoV-2 B. 1.1. 7 variant by convalescent*
659 *and vaccine sera.* 2021. **184**(8): p. 2201-2211. e7.
660 32. Greaney, A.J., et al., *Complete mapping of mutations to the SARS-CoV-2 spike receptor-*
661 *binding domain that escape antibody recognition.* Cell host & microbe, 2021. **29**(1): p.
662 44-57. e9.
663 33. Shang, J., et al., *Structural basis of receptor recognition by SARS-CoV-2.* 2020. **581**(7807):
664 p. 221-224.
665 34. Caini, S., et al., *Clinical characteristics and severity of influenza infections by virus type,*
666 *subtype, and lineage: A systematic literature review.* Influenza Other Respir Viruses,
667 2018. **12**(6): p. 780-792.
668 35. Grubaugh, N., M. Petrone, and E. Holmes, *We shouldn't worry when a virus mutates*
669 *during disease outbreaks.* Nat Microbiol, 2020. **5**(4): p. 529-530.
670 36. Makdasi, E., et al., *The neutralization potency of anti-SARS-CoV-2 therapeutic human*
671 *monoclonal antibodies is retained against viral variants.* 2021. **36**(10): p. 109679.
672 37. Liu, C., et al., *Reduced neutralization of SARS-CoV-2 B. 1.617 by vaccine and convalescent*
673 *serum.* 2021. **184**(16): p. 4220-4236. e13.
674 38. Zhou, W. and W. Wang, *Fast-spreading SARS-CoV-2 variants: challenges to and new*
675 *design strategies of COVID-19 vaccines.* Signal Transduction and Targeted Therapy,
676 2021. **6**(1): p. 226.

677 39. Krause, P.R., et al., *SARS-CoV-2 Variants and Vaccines.* N Engl J Med, 2021. **385**(2): p.
678 179-186.
679