



(RESEARCH ARTICLE)



Emerging XBB variants in Vietnam show high affinity for hACE2

Quan Ke Thai ^{1,*} and Phuoc Huynh ²

¹ Faculty of Natural science Education, Saigon University, 273 An Duong Vuong, Ward 3, District 5, Ho Chi Minh City, Vietnam, 700000.

² VNU HCMC University of Science, 227 Nguyen Van Cu, Ward 4 District 5, Ho Chi Minh city, Vietnam, 700000

GSC Biological and Pharmaceutical Sciences, 2023, 23(03), 231–236

Publication history: Received on 19 May 2023; revised on 27 June 2023; accepted on 29 June 2023

Article DOI: <https://doi.org/10.30574/gscbps.2023.23.3.0256>

Abstract

Since the beginning of 2022, the Omicron variant of the SARS-CoV-2 virus responsible for the COVID-19 pandemic has been the dominant variant. Numerous subvariants of Omicron have been identified, including BA.2.75, BA.4, and BA.5, which are classified as Variants of Concern (VOCs) due to their potential to cause severe illness. Additionally, an emerging variant called XBB has recently caused outbreaks and is believed to be highly lethal. To better understand the XBB variant, we compared the variation in the S protein to that of the other variants. Our initial findings indicate that the XBB variant has spread in Vietnam and carries unique nucleotide changes in the S gene. Specifically, we observed two mutations, G22317T (G252V) and C23123T (P521S), that have resulted in a new subvariant called XBB.1 on the phylogenetic tree. Our analysis also suggests that the XBB variant has a high affinity for hACE2, as indicated by an increase in the interface's number of residues and van der Waals energy. We found that XBB has conserved mutations in RBD that enhance its binding affinity for hACE2. In this report, we noted the mutation V83A, H146Q, and G252V in the NTD increased the binding free energy of the XBB spike protein in the complex with hACE2.

Keywords: SARS-CoV-2; Omicron; XBB variant; Spike protein; Spike gene; Vietnam.

1. Introduction

In late 2019, a previously unknown virus causing respiratory symptoms surfaced in Wuhan, Hubei Province, China. Molecular analysis identified it as a novel coronavirus, closely related to the one that caused the 2002-2003 SARS outbreak and was named SARS-CoV-2. Possessing distinct advantages over its predecessor, SARS-CoV-2 has rapidly spread worldwide and triggered the COVID-19 pandemic. A defining characteristic of SARS-CoV-2 is its high mutation rate and adaptability in exploiting host cells. Its single-stranded RNA genome encodes four structural genes: nucleocapsid protein (N), spike protein (S), envelope protein (E), and matrix protein (M) [1, 2]. Notably, the S gene, which codes for the spike protein, undergoes positive selection, leading to rapid evolution and diversification of the virus during transmission [3]. In contrast, the other structural genes exhibit lower mutation rates [4].

The rapid mutation rate of the S gene can be explained by its vital role in the life cycle of SARS-CoV-2. The S protein not only directly interacts with the hACE2 receptor on the human surface for viral membrane fusion but is also represented as an antigenic factor. Consequently, the S protein was a gold target for vaccine development, monoclonal antibodies, and therapeutic drug [5-7]. Studies on the impact of amino acid mutations on the protein's interaction efficiency with hACE2 and antibodies are increasingly famous [8]. The significance of the S gene and protein is understood, which has now classified new variants of SARS-CoV-2 based on the mutation information on the S protein. PANGO Lineages did the same by constructing a secondary database to evaluate the nucleotide changes of the S gene [9]. This shows a trend to assess changes in the S gene and S protein that are essential in preventing the COVID-19 pandemic. There are currently three main groups: VUM (Variants Under Monitoring), VOI (Variants of Interesting), and VOC (Variants of

* Corresponding author: Quan Ke Thai

Concern). In particular, the VOC variants are considered the most dangerous and successful variants of SARS-CoV-2 worldwide. Indeed, the first VOC variant, Alpha, which was more infectious than the wildtype (WT) [10], caused intense outbreaks in the UK and several other countries. The following VOC variant is Delta, with characteristic mutations on the S protein that has rapidly become the most dominant variant in the world and caused the deaths of millions of people worldwide, with the United States and India two countries being heavily affected. The success of Delta is attributed to the mutations in the RBD (Receptor Binding Domain) region K417N, L452R, and T478K acting as escape antibodies and increasing binding affinity [11-14], and in the FCS (Furin Cleavage Site) region, P681R increased the cleavage effectiveness [15, 16]. It is shown that mutations appearing on the S protein contribute directly to the infectious ability of SARS-CoV-2.

In late 2021, a new VOC variant called Omicron first appeared in South Africa, carrying over 30 amino acid mutations in the S protein [17]. Shortly after, many studies showed that Omicron was able to escape from mAbs and induce reinfection [18-20]. Studies have also demonstrated that Omicron has a better affinity for the hACE2 receptor than Delta to transmit more quickly. As a result, Omicron quickly replaced Delta to become the popular variant worldwide. With its rapid infectious rate and superior mutagenicity, Omicron rapidly proliferated many novel variants, of which BA.2, BA.2.75, BA.4, and BA.5 had prominent contagious phenotypes, which recently, are classified as a VOC variant by WHO and ECDC. Currently, an emerging subvariant of Omicron is the XBB variant causing new outbreaks worldwide. Many questions have been raised about XBB as well as transmitting infection compared to other subvariants of Omicron.

In our previous study, the XBB variant's appearance in Vietnam was not recorded until May 2022 [21]. And although it has documented changes in the S protein of variant BA.1, BA.2, and BA.3 with the hACE2 receptor in previous studies, there is currently no publication on XBB. Herein, to quickly evaluate the mutation in the S protein of XBB, which appeared in Vietnam, we conducted an *in-silico* approach to compare this variant to other variants classified as VOCs as BA.4 and BA.5, and VOI is BA.2.75.

2. Materials and methods

2.1. Data collection

The whole genomes of XBB variants were extracted from the GISAID database. The S sequences of Alpha and Delta were selected from our previous publication [21], and the first sequences of BA.2.75, BA.4, and BA.5 variants that appeared in Vietnam were taken from GISAID (accession number: EPI_ISL_16527063, EPI_ISL_13729175, EPI_ISL_15896945 respectively). The S gene was split by the MAFFT ver 2.0 [22] server based on the Wuhan-hu-1 sequence in GenBank (accession number: NC_045512.2). After that, the S gene sequences were performed multiple sequence alignment by the ClustalW program in MEGA11 software [23].

2.2. Phylodynamic analysis and haplotype network

Continuous, we analyze the phylodynamic of the S gene based on the Maximum-Likelihood (ML) phylogenetic tree in TreeTime (<https://treetime.biozentrum.unibas.ch/>) [24] with the GTR model and set up mutation rates at 0,0008/site/year (Base on Nextstrain validation). The cross-validation ML tree was built by IQtree ver 2.2.0 [25] with the GTR model and 10,000 bootstrap values. We used POPART [26] software to identify nucleotide mutation to construct the haplotype network of S gene sequences based on Median Joining Network methods.

2.3. Homology modelling and molecular docking

The S gene sequences were translated into protein sequences by MEGA11 software. Then, we used these protein sequences to build the homology three-dimensional (3D) S protein structure in the SWISS-MODELLING server [27] with the S protein (chain A) of crystal Spike – hACE2 complex in PDB (PDB ID: 7T9K) as the template structure. The quality of homology structures were estimated by the Structure Assessment tool, ProSA-web server [28] and ERRAT on the SAVES ver 6.0 package [29].

After building the homology structure, we used these protein models to dock with hACE2 receptor of template structure by HADDOCK (High Ambiguity Driven protein-protein DOCKing) server [30]. To advance the reliability of the docking process, we used Optimize run for bioinformatics and ambiguous restraints (AIRs) features of HADDOCK server. We also used the HDock server [31] with the hybrid algorithm docking as the cross-validation of the Spike-hACE2 complex.

2.4. Estimating binding affinity and binding free energy

The binding free energy (G_{bind}) and binding affinity ($\Delta G_{predicted}$) are the primary indexes to determine the effectiveness of protein-protein interactions. To calculate the G_{bind} of spike protein with hACE2 complex (Spike-hACE2 complex), we submitted the complex into the Hawkdock server for examination G_{bind} with Molecular Mechanics Generalized Born Surface Area (MM/GBSA) approach which is the high impact method for protein-protein interactions [32, 33]. Hawkdock server minimized the complex in ff02 force fields for 5000 steps, and the distance for van der Waals interactions is 12 Å. The total G_{bind} has indicated details in equation (1). Furthermore, to give more information about the complex, PRODIGY server was used to calculate binding affinity [34, 35].

$$G_{bind} = \Delta G_{Complex} - \Delta G_{Spike} - \Delta G_{hACE2} \quad (1)$$

2.5. Protein-protein interface network bonding

After molecular docking, these complex structures were examined by the PDBsum server to identify the network bonding of the interface. Then, to determine the positive effect of a mutation that appeared in S protein, we used equals that were calculated in equation (2) which was suggested in the Lei Xu *et al.* publication [36]. $\Delta\Delta E$ is below 0, meaning the mutation enhances the binding affinity of the spike with hACE, and the opposite $\Delta\Delta E$ is higher than 0.

$$\Delta\Delta E = \Delta E_{mutant} - \Delta E_{wildtype} \quad (2)$$

Finally, we used open-source Pymol software to visualize the complex and estimate the electrostatics of S protein based on the APBS plugin.

3. Results and discussion

3.1. Examination nucleotide changes of the S gene of XBB variants

The analysis of ten S gene sequences of the XBB variants reveals clustering of XBB and XBB.1 variants through the G22317T (G252V) and C23123T (P521S) mutations, with bootstrap values of 72 and 87, respectively (Figure 1A). We have also observed high substitution rates of XBB variants since their emergence in Vietnam in November 2022. The haplotype network (Figure 1B) demonstrates that the G22317T mutation in the S gene led to the cleavage of the XBB variant to XBB.1, followed by further mutations at C23123T, forming a new haplotype group in December 2022. These results indicate that the XBB and XBB.1 variants have the potential to infect the Vietnamese population. The high mutation rate recorded underscores the need to monitor the frequency of the XBB and XBB.1 variants in Vietnam.

To assess the impact of S gene mutations on protein sequences, we compared the mutations characteristic of the XBB variant with the variants identified by the WHO as Variants of Concern (VOCs) currently present in Vietnam, including BA.2.75, BA.4, and BA.5. We found that XBB has five specific mutations in the N-Terminal Domain (NTD) region, namely T21810C (V83A), C22000A (H146Q), C22109G (Q183E), G22317T (G252V) and T22200A (V213E). In the Receptor-Binding Domain (RBD) region, we also note three mutations G22599A (R346T), C22664A (L368I), and T23019C (F486S) (Figure 2), as the specific mutation which split the XBB from BA.2 variant [37]. The XBB variant exhibits more amino acid mutations, with 42 mutations. In the S1 subunit, the NTD region recorded 9 mutations and three deletion mutations, the RBD region had 23 mutations, and the fusion peptide (FP) and connecting peptide (CP) subunits contained 5 mutations, whereas the S2 subunit recorded 5 mutations. The positions of spike protein mutations are illustrated in Figure 2.

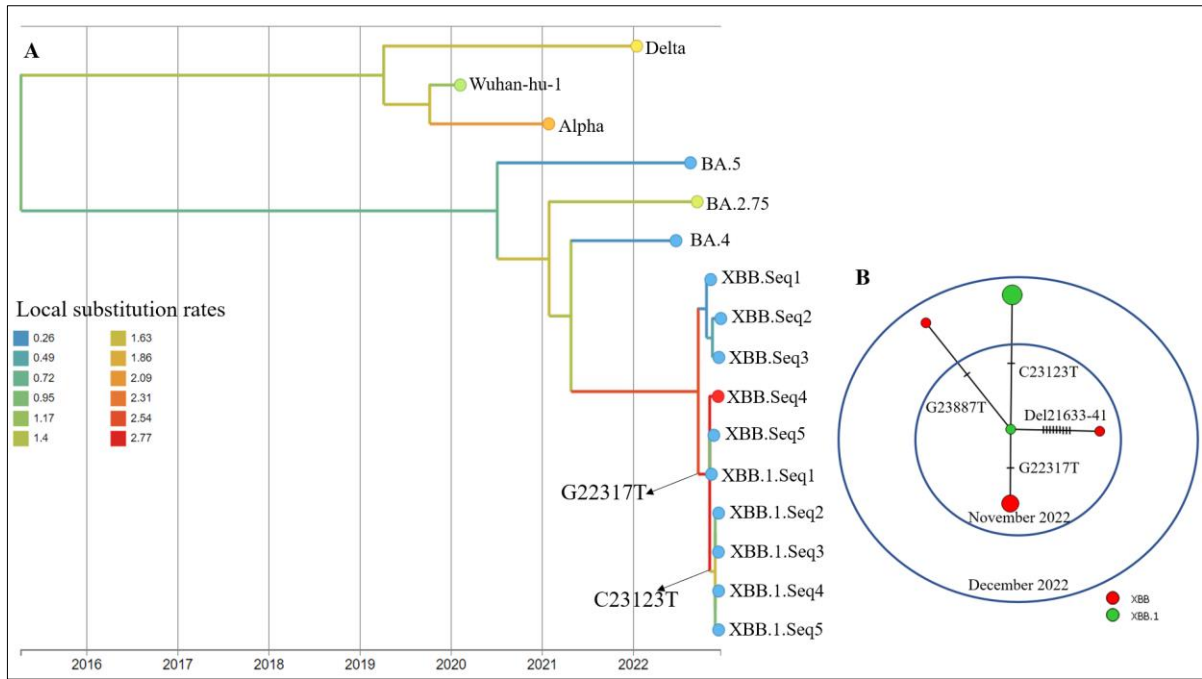


Figure 1 The Maximun-Likelihood phylogenetic tree and haplotype network of the S gene of XBB variants

(A. The branch color of ML phylogenetic tree displayed for the local substitution rates of S gene. B. The node color in the haplotype network is present for the XBB variant, the red for XBB and green for XBB.1.)

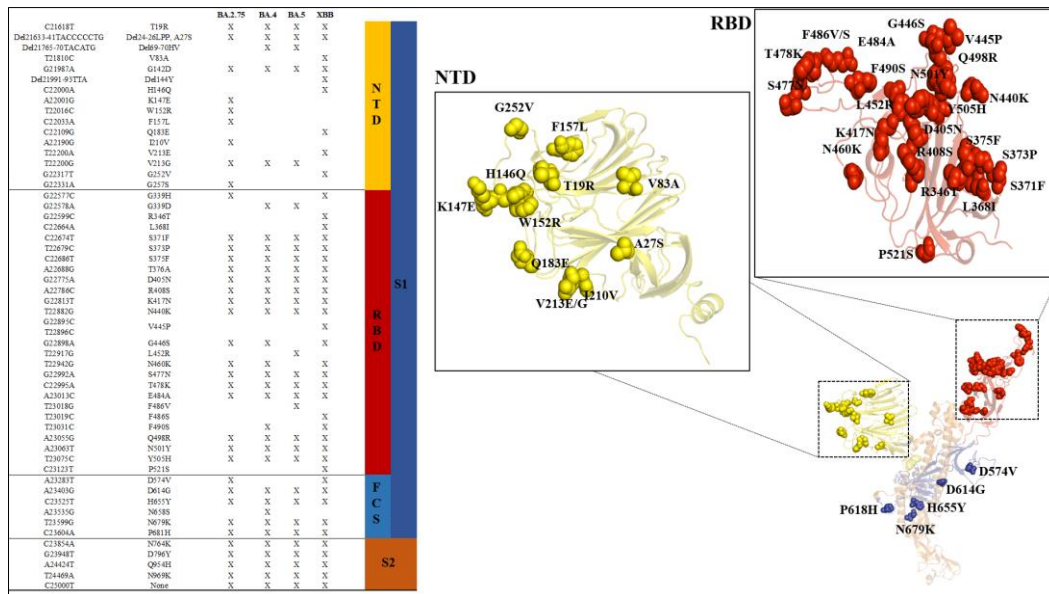


Figure 2 The mutation profile of BA.2.75, BA.4, BA.5 and XBB variants

(In this figure, the color of the spike protein domain in the table corresponding to the 3D protein models)

3.2. Homology spike protein models and protein-protein docking

To generate 3D protein structures, we used the SWISS-MODELLING server to build the homology S protein structures. These were subsequently evaluated for validation using the Structure Assessment tool, ProSA-web server, and ERRAT of the SAVES ver 6.0 package. Analysis of the homology models revealed that all models had more than 90% acceptable positions on the Ramachandran plot, with an overall quality factor of over 90% on ERRAT. The homology structures were then superimposed to check the angle of the C-alpha of the spike protein with the template structure. The RMSD (Root-mean-square deviation) of the homology structures was found to be 0.01 – 0.365 (Å), indicating successful amino

acid mutation of the spike protein. All homology models passed our evaluation criteria and are suitable for downstream analysis. The detailed index of homology models is displayed in Table 1.

Table 1 The quality of homology models were built by SWISS-MODELLING

Models	Ramachandran plot (%)	ProSA-web (Z-scores)	ERRAT (%)	RMSD (Å)
BA.2.75	95.62	-12.37	90.4714	0.099
BA.4	95.08	-12.26	91.4373	0.112
BA.5	95.12	-12.76	90.9091	0.102
XBB	91.32	-12.48	90	0.365

To investigate the interaction between S protein and hACE2, we performed molecular docking on the HADDOCK server, with the HDCK server used for cross-validation. We selected the best docking model generated by HADDOCK based on the HADDOCK score, while the HDCK Confidence was used to assess the spike protein's ability to bind to the hACE2 receptor. A HDCK Confidence value above 0.7 indicates direct interaction between the two proteins, while a value below 0.5 implies that the two proteins cannot bind. Our molecular docking results revealed that the complexes mimicking the binding of the spike protein to hACE2 exhibited HDCK Confidence values exceeding 0.9, which were even better than those of the WT variant. Moreover, the fluctuations of the simulation Spike-hACE2 complex were below 2.0 (Å). However, the RMSD coefficient established by HADDOCK indicated that the association simulations of Omicron subvariants showed higher variability than the WT variant, possibly due to the large number of protein mutations that affect the protein model's stability during simulation. Overall, our results suggest that the interaction simulations of Spike-hACE2 complex were successfully built. These interaction models will be utilized to evaluate the efficiency of protein-protein interaction. The HADDOCK scores, RMSD, and HDCK Confidence parameters are presented in Table 2.

Table 2 The HADDOCK scores, RMSD index, and HDCK Confidence of Spike-hACE2 complex

	WT	BA.2.75	BA.4	BA.5	XBB
HADDOCK score	-147.1 ± 1.2	-147.0 ± 0.6	-152.6 ± 1.3	-134.5 ± 1.6	-136.9 ± 4.3
RMSD	0.5 ± 0.3	0.8 ± 0.7	1.0 ± 0.6	1.7 ± 1.0	1.0 ± 0.7
HDCK Confidence	0.7359	0.9742	0.9761	0.9535	0.9474

3.3. Estimate binding free energy and binding affinity of Spike-hACE2 complex

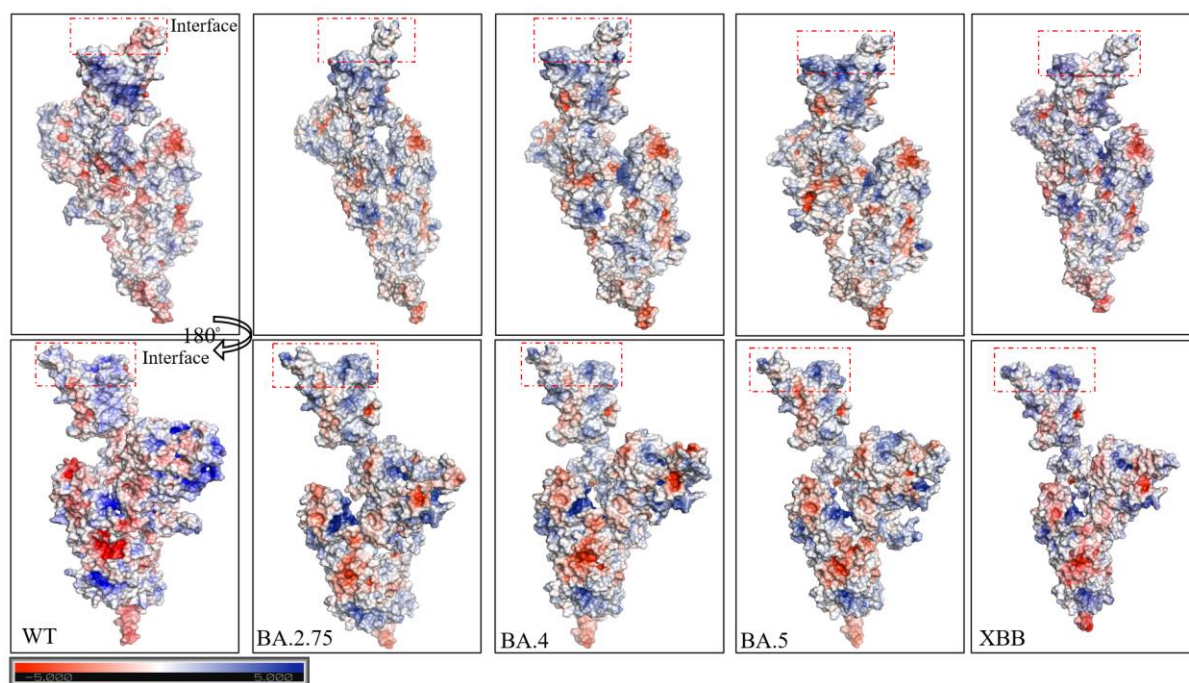
To assess the efficacy of the Spike-hACE2 complex binding, we utilized the binding free energy (G_{bind}) and binding affinity ($\Delta G_{predicted}$) parameters. We noted BA.2.75, BA.4, BA.5, and XBB enhanced the $\Delta E_{electrostatics}$ energy interactions. Analysis of G_{bind} using the Hawkdock server showed that variants BA.5 and XBB had high binding free energy, -104.31 and -103.13, respectively. In addition, the dissociation constant (K_d) of XBB was also highest when compared with BA.2.75, BA.4, and BA.5, significantly superior to the WT variant indicating that the XBB has binding most closely associated with hACE2. The highlight of this result is that we analyzed the XBB Spike-hACE2 complex, and it was shown that this variant had the highest $\Delta G_{predicted}$ at -16.0 and increase the $\Delta E_{van der Waals}$ energy interactions compared to BA.2.75, BA.4, and BA.5. Strong enhancement of $\Delta E_{van der Waals}$ has not been mentioned in previous studies on BA.1, BA.2, BA.3 [38] and others variants [39] [40] [41]. Our results show that the XBB variant appearing in Vietnam has solid infectious potential and a better affinity for the hACE2 receptor.

Table 3 The binding free energy (G_{bind}) and binding affinity ($\Delta G_{predicted}$) of Spike-hACE2 complex

	WT	BA.2.75	BA.4	BA.5	XBB
$\Delta E_{van\ der\ Waals}$	-119.43	-121.39	-115.92	-133.7	-139.92
$\Delta E_{electrostatics}$	-630.99	-1815.6	-1577.41	-1779	-1430.34
$\Delta G_{Polar\ Solvation}$	699.05	1858.7	1613.61	1825.28	1485.31
$\Delta G_{Nonpolar\ Solvation}$	-16.06	-15.21	-14.66	-16.89	-18.18
ΔG_{bind}	-67.44	-93.49	-89.4	-104.31	-103.13
$\Delta G_{predicted}$	-12.2	-13.5	-14.4	-15.1	-16.0
K_d	1.2E-09	1.3E-10	2.5E-11	8.1E-12	1.7E-12

3.4. Protein-protein network bonding

To evaluate the changes affecting the protein structure in more detail, we analyzed the S protein's interaction surface with hACE2 (interface). We analyzed the charge distribution on the S protein surface, showing that the variants enhanced the positive charge expression on the interface (Figure 2). This explains why the entire evaluated variant strongly enhanced $\Delta E_{electrostatics}$ of Spike-hACE2 complex. Here, in terms of the interface of the complex, we find that the surface of the interface has a slight change in favorable charge distribution when compared with BA.2.75, BA.5, and BA.5, which is similar to the decrease in the electrostatics energy of XBB (Table 3) when compared to the other variants.

**Figure 3** The electrostatic potential distribution of spike protein models

(Protein surface is coloured according to the electrostatic potential. Color scale ranges from $-5.000T/e$ (red) to $+5.000T/e$ (blue) as reported by the bar at the bottom of figure)

In addition, looking at the amino acids of S protein which are directly interacting with hACE2, we found that the XBB variant has the highest number of residues participating in the interface, 27 residues. In contrast, the variants BA.2.75 had 21 residues, BA.4 had 23 residues, BA.5 had 25 residues, and WT had 21 residues, respectively (Table 3). The enhancement of amino acids involved in the interaction directly affects the area of the interface. Therefore, the XBB variant has the largest interface area compared to the remaining variants in the study (Table 3). The enhanced expression of residues in the interface also increases the binding of the spike protein to hACE2. Accordingly, the XBB variant has the most significant number non-bonded contacts at 184 (Table 3). The increased in non-bonded contacts

explains why the XBB variant had the highest $\Delta E_{van\ der\ Waals}$ compared with BA.2.75, BA.4, and BA.5 (Table 2). The increase in non-bond contacts was contributed by D405N and K417N, which are not present in BA.2.75 and BA.4.

Table 3 Interface statistics of Spike-hACE2 complex

	Interface area (\AA^2)	Interface residues	Non-bonded contacts
WT	962	21	116
BA.2.75	1048	21	127
BA.4	1072	23	156
BA.5	1127	25	175
XBB	1226	27	184

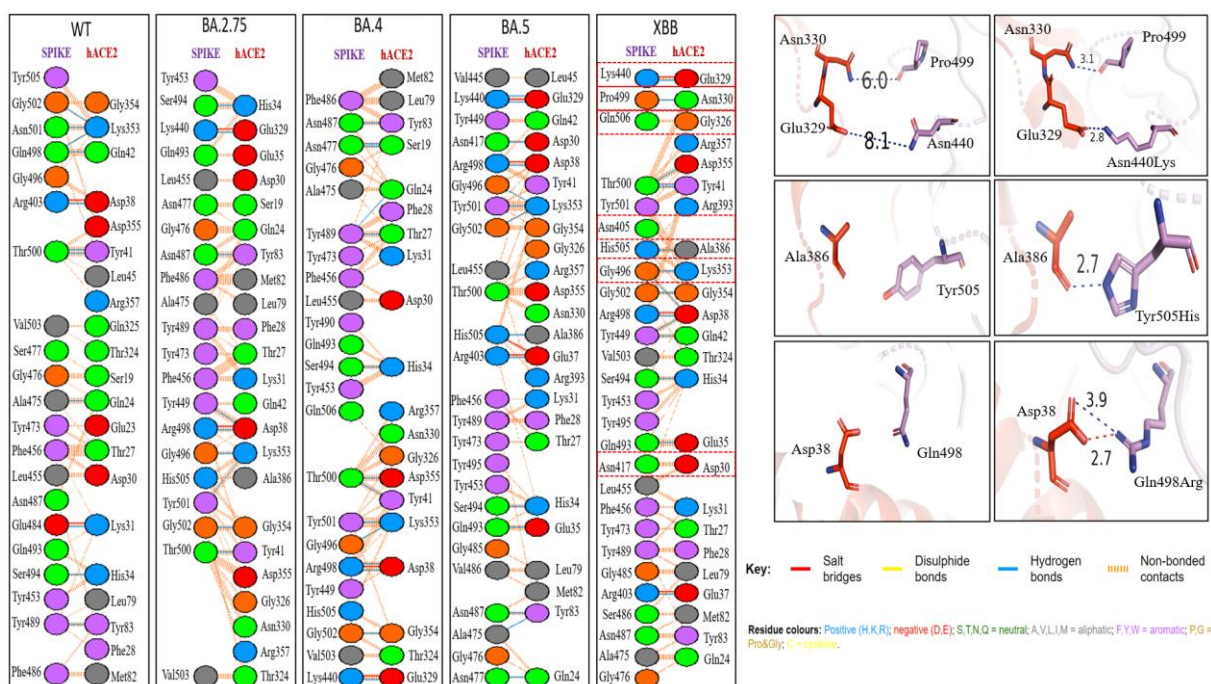


Figure 4 The network bonding interface of Spike-hACE2 complex

(The networks bonding of interface were determined by PDBsum server. The residues colours display for characteristics of residues and the bar colours present for the type of bonding according to figure. Amino acid interactions of spike and hACE2 visualized by Pymol software with hydrogen bonding shown in blue dashes and red salt bridge interactions with the respective distances shown above.)

The analyzed free energy changes per residue found that the XBB variant preserved most mutations with $\Delta\Delta E < 0$. In the NTD, there were three new mutations, V83A, H146Q and G252V, which enhance G_{bind} of the Spike-hACE2 complex (Figure 4), showing that these mutations affect spike adhesion to hACE2. Interestingly, the experimental study of Tamura *et al.* (2022) [42] found that the V83A mutation enhances the transmissibility and fusion capacity of XBB 3.3 times, which is consistent with our prediction based on the V83A mutation. We also checked another new mutation in NTD of XBB, G22317T (G252V), that increased G_{bind} of Spike-hACE2 complex, which was not determined in previous studies [37] [42] [43]. In the RBD region, the XBB variant in Vietnam recorded three new mutations, R346T, L368I, and P521S, in which L368I and P521S were recorded with $\Delta\Delta E = 0$. R346T and V445P has an $\Delta\Delta E > 0$, which implies a reduction in the binding free energy of the Spike-hACE2 complex. The reduced binding free energy of R346T and V445P may be due to the contribution of these two mutations to the ability to evade neutralizing antibodies [44] [45]. Other matching types in the F486V and F486S mutations suggested that BA.5 and XBB trade-off immune clearance by decreasing affinity for hACE2 [46] [42]. Likewise, the K417N mutation in variants BA.2.75 and BA.4 has $\Delta\Delta E > 0$, but the main effect of this mutation is to help the virus escape mAbs [47]. The decrease of K417N because of the substitution of Lys for Asn broke a salt bridge and a hydrogen bond [38]. But our results show that the K417N mutants of BA.5 and XBB

are enhanced compared with BA.2 and BA.2.75. This may be due to the enhancement of the interface surface (Table 3), enabling K417N to form non-bonded contacts with Asp30 of hACE2 (Figure 4). In particular, we further noted the appearance of a new hydrogen bond of XBB compared with BA.5 at the Pro499 position (Figure 3). Accordingly, the N440K mutation changed Asn440 residues with the uncharged side chain to Lys440 residues with a positive charge to form hydrogen bonds and salt bridge interactions with Glu329 of hACE2 (Figure 3). We also noted the Q498R and N501Y as mutations that positively affected the interaction efficiency of spike protein with hACE2 previously mentioned in other studies (Figure 3). The remaining XBB variants and BA.2.75, BA.4, and BA.5 have preserved favorable mutations in the FCS (D614G and N679K) and S2 subunit (N764K, D796Y, and N969K).

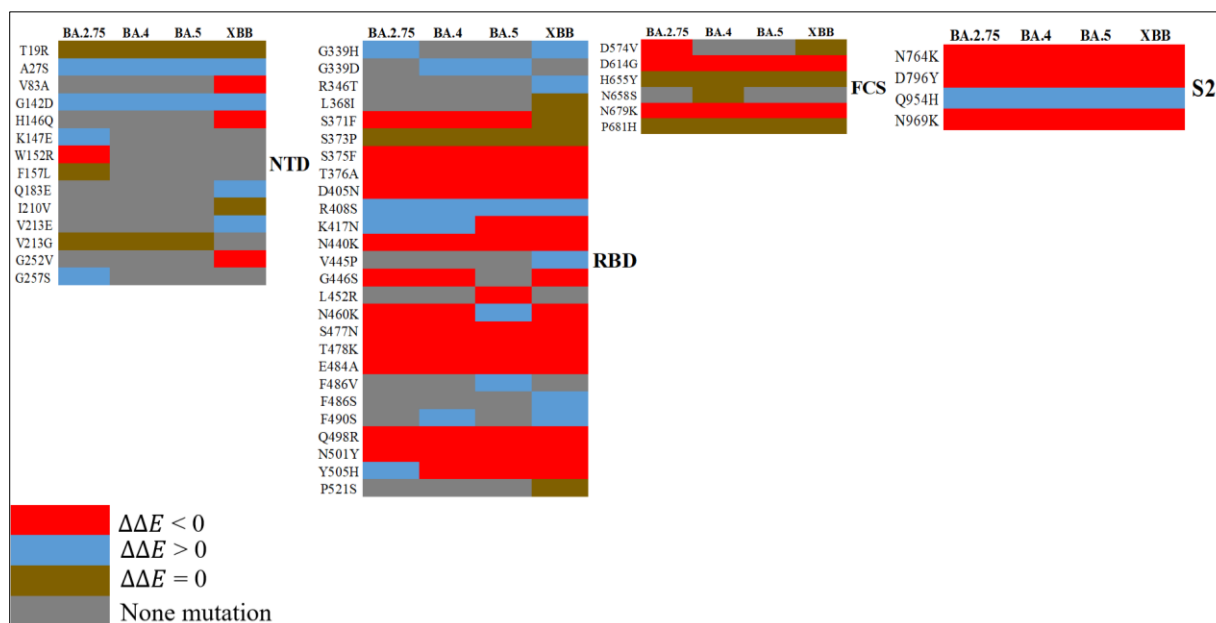


Figure 5 The $\Delta\Delta E$ per residues of spike protein

(This results present for the positive mutation for spike protein with the $\Delta\Delta E < 0$ in red colors. The $\Delta\Delta E$ means negative effect was colored in blue. The $\Delta\Delta E = 0$ shows that no changes in the binding free energy were displayed in brown color. And the mutations are not present in the variants were painted by grey color.)

4. Conclusion

Herein, we found that mutations on the S protein of XBB, BA.2.75, BA.4, and BA.5 variants preserve mutations that enhance affinity for hACE2 as S371F, S375F, T376A, D405N, N440K, G446S, N460K, S477N, T478K, E484A, Q498R, N501Y, Y505H and as well as mutations that positively affect Spike-hACE2 complex are D614G, N679K, N764K, D796Y. In this study, we highlight the mutation in NTD, V83A, H146Q, and G252V (new mutation) could increase the binding free energy for spike protein in complex with hACE2. Our results also showed that the appearance of new mutations in the S protein of the XBB variant enhanced the contact area of the protein with hACE. The result is an increase in adhesion efficiency through an increase in van der Waals interaction. Although our study has many limitations due to the approach only using various computational tools, these initial results will provide helpful information for further experimental investigations.

Compliance with ethical standards

Acknowledgments

We would like to thank Prof. Dr. Jürgen Pleiss for fruitful discussions about the results.

Disclosure of conflict of interest

The authors declare no conflict of interest.

References

- [1] Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020;579(7798):265-9.
- [2] Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;579(7798):270-3.
- [3] Berrío A, Gartner V, Wray GA. Positive selection within the genomes of SARS-CoV-2 and other Coronaviruses independent of impact on protein function. *PeerJ*. 2020;8:e10234.
- [4] Chaw SM, Tai JH, Chen SL, Hsieh CH, Chang SY, Yeh SH, et al. The origin and underlying driving forces of the SARS-CoV-2 outbreak. *J Biomed Sci*. 2020;27(1):73.
- [5] Martínez-Flores D, Zepeda-Cervantes J, Cruz-Reséndiz A, Aguirre-Sampieri S, Sampieri A, Vaca L. SARS-CoV-2 Vaccines Based on the Spike Glycoprotein and Implications of New Viral Variants. *Front Immunol*. 2021;12:701501.
- [6] Tabynov K, Turebekov N, Babayeva M, Fomin G, Yerubayev T, Yespolov T, et al. An adjuvanted subunit SARS-CoV-2 spike protein vaccine provides protection against Covid-19 infection and transmission. *npj Vaccines*. 2022;7(1):24.
- [7] Candido KL, Eich CR, de Farina LO, Kadowaki MK, da Conceicao Silva JL, Maller A, et al. Spike protein of SARS-CoV-2 variants: a brief review and practical implications. *Braz J Microbiol*. 2022.
- [8] Khan A, Zia T, Suleman M, Khan T, Ali SS, Abbasi AA, et al. Higher infectivity of the SARS-CoV-2 new variants is associated with K417N/T, E484K, and N501Y mutants: An insight from structural data. *J Cell Physiol*. 2021;236(10):7045-57.
- [9] O'Toole A, Pybus OG, Abram ME, Kelly EJ, Rambaut A. Pango lineage designation and assignment using SARS-CoV-2 spike gene nucleotide sequences. *BMC Genomics*. 2022;23(1):121.
- [10] Washington NL, Gangavarapu K, Zeller M, Bolze A, Cirulli ET, Schiabor Barrett KM, et al. Emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States. *Cell*. 2021;184(10):2587-94.e7.
- [11] Han P, Li L, Liu S, Wang Q, Zhang D, Xu Z, et al. Receptor binding and complex structures of human ACE2 to spike RBD from omicron and delta SARS-CoV-2. *Cell*. 2022;185(4):630-40.e10.
- [12] Di Giacomo S, Mercatelli D, Rakhimov A, Giorgi FM. Preliminary report on severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Spike mutation T478K. *Journal of medical virology*. 2021;93(9):5638-43.
- [13] Cherian S, Potdar V, Jadhav S, Yadav P. SARS-CoV-2 Spike Mutations, L452R, T478K, E484Q and P681R, in the Second Wave of COVID-19 in Maharashtra, India. 2021;9(7).
- [14] Liu Y, Liu J, Plante KS, Plante JA, Xie X, Zhang X, et al. The N501Y spike substitution enhances SARS-CoV-2 infection and transmission. *Nature*. 2022;602(7896):294-9.
- [15] Kumar V, Singh J, Hasnain SE, Sundar D. Possible Link between Higher Transmissibility of Alpha, Kappa and Delta Variants of SARS-CoV-2 and Increased Structural Stability of Its Spike Protein and hACE2 Affinity. *International journal of molecular sciences*. 2021;22(17):9131.
- [16] Fan L-Q, Hu X-Y, Chen Y-Y, Peng X-L, Fu Y-H, Zheng Y-P, et al. Biological Significance of the Genomic Variation and Structural Dynamics of SARS-CoV-2 B.1.617. *Frontiers in microbiology*. 2021;12:750725-.
- [17] Viana R, Moyo S, Amoako DG, Tegally H, Scheepers C, Althaus CL, et al. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature*. 2022;603(7902):679-86.
- [18] Tan ST, Kwan AT, Rodríguez-Barraquer I, Singer BJ, Park HJ, Lewnard JA, et al. Infectiousness of SARS-CoV-2 breakthrough infections and reinfections during the Omicron wave. *Nature Medicine*. 2023.
- [19] Cele S, Jackson L, Khoury DS, Khan K, Moyo-Gwete T, Tegally H, et al. Omicron extensively but incompletely escapes Pfizer BNT162b2 neutralization. *Nature*. 2022;602(7898):654-6.
- [20] Carreño JM, Alshammari H, Tcheou J, Singh G, Raskin AJ, Kawabata H, et al. Activity of convalescent and vaccine serum against SARS-CoV-2 Omicron. *Nature*. 2022;602(7898):682-8.
- [21] Quan Ke Thai, Phuoc Huynh, Yen Le Thi, Huyen Nguyen Thi Thuong. Genetic diversity of SARS-CoV-2 Omicron variants' spike gene in Vietnam. *International Journal of Biosciences*. 2022;21(3):166-76.

- [22] Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*. 2013;30(4):772-80.
- [23] Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular biology and evolution*. 2021;38(7):3022-7.
- [24] Sagulenko P, Puller V, Neher RA. TreeTime: Maximum-likelihood phylodynamic analysis. *Virus Evolution*. 2018;4(1):vex042.
- [25] Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*. 2020;37(5):1530-4.
- [26] Leigh JW, Bryant D. popart: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*. 2015;6(9):1110-6.
- [27] Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research*. 2018;46(W1):W296-W303.
- [28] Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*. 2007;35(suppl_2):W407-W10.
- [29] Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci*. 1993;2(9):1511-9.
- [30] Dominguez C, Boelens R, Bonvin AMJJ. HADDOCK: A Protein-Protein Docking Approach Based on Biochemical or Biophysical Information. *Journal of the American Chemical Society*. 2003;125(7):1731-7.
- [31] Yan Y, Tao H, He J, Huang S-Y. The HDock server for integrated protein-protein docking. *Nature Protocols*. 2020;15(5):1829-52.
- [32] Weng G, Wang E, Wang Z, Liu H, Zhu F, Li D, et al. HawkDock: a web server to predict and analyze the protein-protein complex based on computational docking and MM/GBSA. *Nucleic Acids Research*. 2019;47(W1):W322-W30.
- [33] Chen F, Liu H, Sun H, Pan P, Li Y, Li D, et al. Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein-protein binding free energies and re-rank binding poses generated by protein-protein docking. *Physical Chemistry Chemical Physics*. 2016;18(32):22129-39.
- [34] Vangone A, Bonvin AMJJ. Contacts-based prediction of binding affinity in protein-protein complexes. *eLife*. 2015;4:e07454.
- [35] Xue LC, Rodrigues JP, Kastiris PL, Bonvin AM, Vangone A. PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics*. 2016;32(23):3676-8.
- [36] Xu L, Li Y, Li D, Xu P, Tian S, Sun H, et al. Exploring the binding mechanisms of MIF to CXCR2 using theoretical approaches. *Physical Chemistry Chemical Physics*. 2015;17(5):3370-82.
- [37] Qu P, Faraone JN, Evans JP, Zheng YM, Carlin C, Anghelina M, et al. Extraordinary Evasion of Neutralizing Antibody Response by Omicron XBB.1.5, CH.1.1 and CA.3.1 Variants. *bioRxiv : the preprint server for biology*. 2023.
- [38] Kumar S, Karuppanan K, Subramaniam G. Omicron (BA.1) and sub-variants (BA.1.1, BA.2, and BA.3) of SARS-CoV-2 spike infectivity and pathogenicity: A comparative sequence and structural-based computational assessment. *Journal of medical virology*. 2022;94(10):4780-91.
- [39] Khan A, Zia T, Suleman M, Khan T, Ali SS, Abbasi AA, et al. Higher infectivity of the SARS-CoV-2 new variants is associated with K417N/T, E484K, and N501Y mutants: An insight from structural data. (1097-4652 (Electronic)).
- [40] Celik I, Khan A, Dwivany FM, Fatimawali, Wei D-Q, Tallei TE. Computational prediction of the effect of mutations in the receptor-binding domain on the interaction between SARS-CoV-2 and human ACE2. *Molecular Diversity*. 2022;26(6):3309-24.
- [41] Suleman M, Yousafi Q, Ali J, Ali SS, Hussain Z, Ali S, et al. Bioinformatics analysis of the differences in the binding profile of the wild-type and mutants of the SARS-CoV-2 spike protein variants with the ACE2 receptor. (1879-0534 (Electronic)).
- [42] Tamura T, Ito J, Uriu K, Zahradnik J, Kida I, Nasser H, et al. Virological characteristics of the SARS-CoV-2 XBB variant derived from recombination of two Omicron subvariants. *bioRxiv : the preprint server for biology*. 2022:2022.12.27.521986.

- [43] Lasrado N, Collier AY, Miller J, Hachmann NP, Liu J, Sciacca M, et al. Waning Immunity Against XBB.1.5 Following Bivalent mRNA Boosters. *bioRxiv : the preprint server for biology*. 2023.
- [44] Cao Y, Jian F, Wang J, Yu Y, Song W, Yisimayi A, et al. Imprinted SARS-CoV-2 humoral immunity induces convergent Omicron RBD evolution. *Nature*. 2022.
- [45] Wang Q, Iketani S, Li Z, Liu L, Guo Y, Huang Y, et al. Alarming antibody evasion properties of rising SARS-CoV-2 BQ and XBB subvariants. *Cell*. 2023;186(2):279-86 e8.
- [46] Tuekprakhon A, Nutalai R, Djokaite-Guraliuc A, Zhou D, Ginn HM, Selvaraj M, et al. Antibody escape of SARS-CoV-2 Omicron BA.4 and BA.5 from vaccine and BA.1 serum. *Cell*. 2022;185(14):2422-33 e13.
- [47] Planas D, Saunders N, Maes P, Guivel-Benhassine F, Planchais C, Buchrieser J, et al. Considerable escape of SARS-CoV-2 Omicron to antibody neutralization. *Nature*. 2022;602(7898):671-5.