Songklanakarin J. Sci. Technol.
44 (5), 1201–1208, Sep. – Oct. 2022

http://www.sjst.psu.ac.th

*Original Article*

# The SARS-CoV-2 infections in Thailand: Analysis of spike mutations complemented by protein structure insights

Sirawit Ittisoponpisan[1*], Shalip Yahangkiakan[1], Michael J E Sternberg[2], and Alessia David[2]

[1] *Center for Genomics and Bioinformatics Research, Division of Biological Science, Faculty of Science, Prince of Songkla University, Hat Yai, Songkhla, 90110 Thailand*

[2] *Structural Bioinformatics Group, Centre for Integrative Systems Biology and Bioinformatics, Department of Life Sciences, Sir Ernst Chain Building, Imperial College London, South Kensington, London, England*

Received: 10 July 2022; Revised: 20 September 2022; Accepted: 24 September 2022

## Abstract

Thailand was the first country outside China to officially report COVID-19 cases. With a large number of SARS-CoV-2 sequences collected from patients, the effects of many genetic variations, especially those unique to Thai strains, are yet to be elucidated. In this study, we analyzed 439,197 sequences of the SARS-CoV-2 spike protein collected from NCBI and GISAID databases. 595 sequences were from Thailand and contained 52 amino acid mutations, of which 6 had not been observed outside Thailand (p.T51N, p.P57T, p.I68R, p.S205T, p.K278T, p.G832C). These mutations were not predicted to be of concern. We demonstrate that p.D614G became the prevalent strain during the second outbreak, and the most common spike mutations detected in Thailand (p.A829T, p.S459F and p.S939F) do not appear to cause any major structural change to the spike trimer or the spike-ACE2 interaction. Among the spike mutations identified in Thailand was p.N501T. This mutation was not predicted to increase SARS-CoV-2 binding, in contrast to the spike mutation of interest p.N501Y. In conclusion, Thailand-specific mutations are unlikely to increase the fitness of SARS-CoV-2. The insights obtained from this study could aid in prioritizing SARS-CoV-2 variants and in strain surveillance.

**Keywords**: SARS-CoV-2, mutation analysis, Angiotensin-converting enzyme 2, spike protein, Thailand

## 1. Introduction

The local transmissions of COVID-19 in Thailand developed into the first wave in March 2020, resulting in a declaration of a state of emergency, followed by a lockdown nationwide (WHO Thailand, 2020a). As part of the new regulations aiming to contain the outbreak, many international flights were banned, and a 14-day state quarantine measure was enforced (Department of Disease Control, 2020a), rendering Thailand into virtual isolation. The lockdown was extensively lifted in July 2020 as the infection tally was flattened to about 3,000 cases (Department of Disease Control, 2020b; WHO Thailand, 2020b).

*Corresponding author
Email address: sirawit.i@psu.ac.th

Despite the government's strict travel regulations in an effort to control cross-border transmission, there had been reports of local transmission cases due to illegal border crossing and not complying with the 14 day-quarantine measure, importing new strains from abroad (Department of Disease Control, 2020d). In December 2020, the second wave of infections occurred among workers in a seafood market in Samut Sakhon province and later quickly spread to the surrounding provinces (WHO Thailand, 2020c).

The spike protein is one of the most widely studied proteins of SARS-CoV-2 as it interacts with the human ACE2 (hACE2), thus facilitating viral cell entry (Lan *et al.*, 2020; Shang, Wan, *et al.*, 2020). Analysis of the 3D structure of the spike protein, alone and bound to hACE2, can help researchers understand SARS-CoV-2 evolution and transmission and could guide therapeutic research (Waman *et al.*, 2021).

The rapid growth in COVID-19 cases worldwide and the global sequencing efforts resulted in a large number of viral sequences being collected from patients. Many of these have been made freely available through two major databases: NCBI and GISAID (Bogner, Capua, Cox, & Lipman, 2006). At the time of this study, over 600,000 SARS-CoV-2 sequences were available for analysis.

To date, studies on the effect of SARS-CoV-2 mutations identified in Thailand are limited (Buathong *et al.*, 2021; Joonlasak *et al.*, 2021; Okada *et al.*, 2020; Puenpa *et al.*, 2020). Therefore, in this study, we aim to provide structural insights into SARS-CoV-2 spike mutations identified in Thailand to obtain insights into SARS-CoV-2 Thai outbreaks and to identify mutations of interest which may require additional investigations.

## 2. Materials and Methods

### 2.1 SARS-CoV2 spike protein sequences

The spike protein sequences were retrieved on 7 February 2021 from two databases: NCBI (https://www.ncbi.nlm.nih.gov/sars-cov-2/) and GISAID (https://www.gisaid.org/) (Bogner *et al.*, 2006). All sequences that were 1,273 amino acids long, which is the same length as the original and canonical Wuhan reference sequence (YP_009724390.1), were selected from these databases. As a result, 53,056 spike protein sequences from NCBI (including the reference YP_009724390.1) and 386,176 sequences from GISAID were retrieved. The sequences were then further screened by comparing them with the Wuhan reference sequence (YP_009724390.1). Any sequence that resulted in less than 98% identity or more than three consecutive amino acid substitutions was removed from the dataset. This was done to eliminate the possibility of the sequence containing insertion/deletion mutations. By applying these criteria, six sequences from NCBI and 29 sequences from GISAID were removed. The final dataset consisted of 53,050 sequences from NCBI and 386,147 from GISAID (439,197 total).

### 2.2 Phylogenetic analysis

The phylogenetic tree was generated using the Nextstrain web server (Hadfield *et al.*, 2018). The tree was created based on the whole genome alignment of SARS-CoV-2 strains available on the GISAID database and was customised to cover only strains collected in Thailand from 20 December 2019 to 7 February 2021 ("Genomic Epidemiology of Novel Coronavirus - Thailand-Focused Subsampling," 2021). As Nextrain only displays complete strains on a fixed timeline, some mutations could not be mapped directly onto this tree due to incomplete strain information (i.e., the strain might be partially sequenced, or the date of collection was unclear). Therefore, the locations of some of these mutations were estimated based on their clade information and co-existing mutations on the strain.

### 2.3 Structural analysis

The following 3D coordinate files were extracted from the ProteinDataBank and used for mutations analysis: 6XR8 (resolution: 2.90 Å, released: 2020-07-22) for the complete structure of the SARS-CoV-2 spike protein complex (at inactive conformation) and 6M0J (resolution: 2.45 Å, released: 2020-03-18) for the spike-hACE2 complex. The accessible surface area (ASA) calculation was performed using DSSP (Kabsch and Sander, 1983). For each amino acid, the RSA was calculated using the formula RSA = ASA/maxASA. The maxASA for each amino acid is defined according to Rost and Sander (1994). A residue was regarded as "surface residue" when its relative solvent accessibility (RSA) was ≥ 9%, otherwise "buried". A residue was defined as an "interface" if it was within 4 Angstrom distance from any other residues of a different chain (Bosshard, Marti, and Jelesarov, 2004; Jeffrey, 1997). In this study, interface residue calculation was performed on the trimer spike protein complex (PDB: 6XR8) and the spike-hACE2 complex (PDB: 6M0J).

Mutations occurring in the SARS-CoV-2 spike were identified by comparing 439,196 sequences in our dataset against the Wuhan SARS-CoV-2 spike reference sequence (YP_009724390). The *in silico* mutagenesis was performed using a modified version of the Missense 3D algorithm (Ittisoponpisan *et al.*, 2019) to account for amino acid substitutions at interface residues.

## 3. Results

### 3.1 Phylogenetic analysis of spike mutations found in Thailand

A total of 439,197 SARS-CoV-2 spike sequences were collected and analysed. 595 sequences were from Thailand and contained 52 spike mutations. Forty-one of these mutations could be mapped onto a phylogenetic tree (Figure 1). Two major clusters of strains from the first and second SARS-CoV-2 Thai outbreaks were identified. The first outbreak contained mutations from seven clades of SARS-CoV-2 strains. Although the mutation of interest p.D614G was already detected in many strains since March 2020, it became the prevalent strain in the second outbreak (clade GH, Figure 1).

Six SARS-CoV-2 spike mutations — p.T51N, p.P57T, p.I68R, p.S205T, p.K278T, and p.G832C — were found only in some Thai strains and not identified in genetic sequences from other countries. Notably, three mutations (p.I68R, p.S205T, and p.G832C) were found during the first outbreak, and one (p.K278T) was found during the second outbreak. The other two Thailand-specific mutations (p.T51N and p.P57T), which co-localized in one genomic sequence, could not be mapped onto the phylogenetic tree (ID: 708814, clade: GH, collected on 9 October 2020). This strain had a total of 11 spike mutations: eight of them — p.V47A, p.T51N, p.Q52K, p.F55L, p.P57T, p.V70L, p.S71P, and p.A688P — were not shared by any other Thai strains. When the whole genome of this record was used in a BLAST search against SARS-CoV-2 records on the GISAID web server, similar strains were found mainly in samples from Saudi Arabia (data not shown). This suggests that this strain was not a result of local transmission but may have been introduced into Thailand from overseas and later mutated to contain the two unique spike mutations p.T51N and p.P57T.

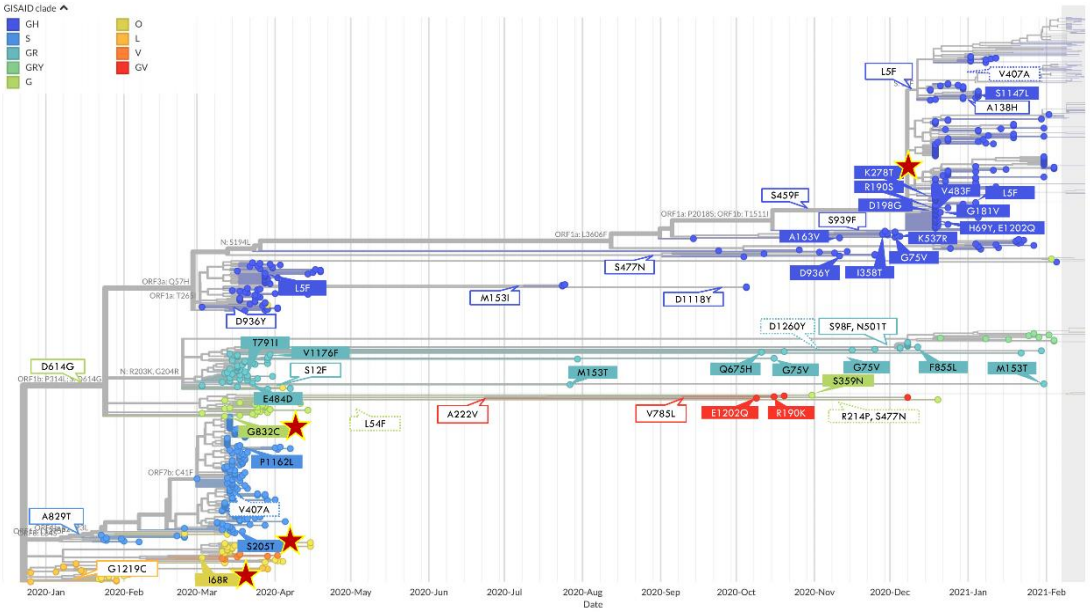


Figure 1. Phylogenetic analysis of SARS-CoV-2 strains collected in Thailand from December 2019 to February 2021. In the phylogenetic tree, each node represents a strain. White boxes with colour borders represent branches where spike mutations evolved and are shared among descendants. Colour boxes indicate newly evolved mutations found on each particular leaf node. Mutations that cannot be mapped directly onto this tree due to incomplete strain information are estimated and represented as white boxes with dotted borders. Mutations that were unique to Thai strains are indicated by the red asterisk (★).

## 3.2 Structural analysis of spike mutations found in Thailand

We first analysed the distribution of all mutations identified in the spike protein (Figure 2). Mutations, both in Thailand and worldwide, were distributed ubiquitously on the spike amino acid sequence, with the exception of 51 residues which did not contain any mutations. These residues are part of the receptor-binding (RBD) and Corona S2 superfamily (S2) domains.

Predictions of spike mutations detected in Thailand are shown in Table 1. No structural changes to the spike trimer and the spike-hACE2 interaction were identified in the six Thailand-specific mutations (Figure 3) with the exception of p.S205T, which was predicted to reduce the structural stability of the spike protein due to the drastic reduction of a surface cavity. Hence, it is likely that this mutation is deleterious to the virus. Recently, p.I68R was discovered in a mutagenesis study of SARS-CoV-2 strains in mice and was reported to be associated with the virus's ability to escape neutralising antibodies (Peter *et al.*, 2021). However, whether this spike mutation enhances the viral escape mechanism in humans is yet to be elucidated. Despite no structural damage found from p.K278T, the extra h-bond formed between the main chain of Thr278 and the side chain of Thr286 was due to the side chain repacking of Thr286 rather than the mutation from Lys278 to Thr278. Nevertheless, all these Thailand-specific mutations were rare (found in < 1% of the samples collected) and could imply that they randomly occurred without any selective pressure and are unlikely to enhance the virus's transmissibility.

We next examined additional spike mutations identified in Thailand. p.D614G was identified in 48.7% of all sequences analysed. This mutation was predicted to cause the loss of inter-chains H-bonds between Asp614 and Lys835 and Lys854 in the spike trimer, which could affect the spike trimer packing. One study using cryo-electron microscopy confirms that p.D614G disrupts an interprotomer contact, thus facilitating the spike "open conformation" (Yurkovetskiy *et al.*, 2020). This mutation is estimated to increase the transmissibility of SARS-CoV-2 by 20% (Volz *et al.*, 2021), thus explaining why it had become the dominant strain by February 2021 (second Thai outbreak), similarly to what was observed in other Southeast Asian countries (Mat Yassim *et al.*, 2021; Nyunt *et al.*, 2021).

The other most common spike mutations identified in Thailand were p.A829T (39.2% of all sequences), p.S459F (20.2%), p.L5F (4.7%), p.S939F (2.0%) and p.N501T (1.5%). p.A829T was the prevalent SARS-CoV-2 spike mutation during the first Thai outbreak, but it was no longer detected in the second outbreak. No clear structural damage was identified by our analysis. Mutation p.S459F was predicted to disrupt an H-bond formed within surface residues (an H-bond linking Arg457 and Ser459 was disrupted when the mutation was simulated on PDB 6XR8, while disruption of an H-bond formed between two consecutive residues Ser459 and Asn460 was detected when tested on PDB 4M0J), which is likely to be compensated by the interaction with nearby water molecules and therefore not structurally destabilising. Mutation p.L5F was a newly emerging mutation in the second Thai outbreak. It is located near the N-terminal of the spike protein. Unfortunately, this residue was not covered by the available 3D coordinates and no structural analysis could be performed. Mutation p.S939F was not predicted to affect the structure of the spike trimer or its interaction with ACE2. Accordingly, a previous study suggested that this mutation is not likely to

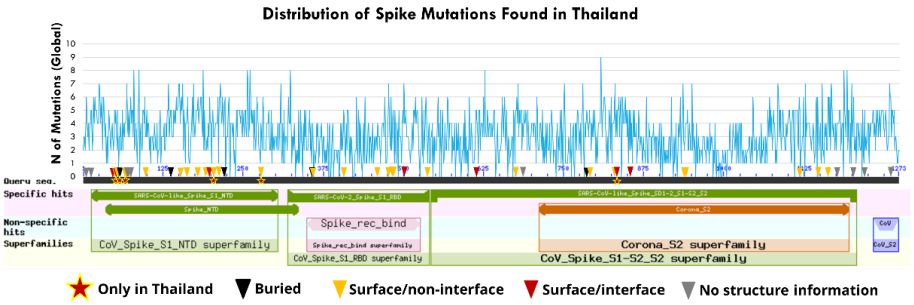


Figure 2. Distribution of mutations on the spike protein. Residues where Thai mutations were found are indicated by arrowheads. Mutations specific to Thailand are indicated by ★. Colour boxes depict spike domains.

Table 1. Analysis of spike mutations found in Thailand. The mutations that were found exclusively in Thailand are underlined and indicated by *. Deleterious predictions are shown in bold.

| Mutation | % | GISAID clade [1] | Location on spike | Structural consequences (Missense3D) | Note |
|---|---|---|---|---|---|
| L5F | 4.71 | GH | - | N/A (residue not contained in PDB) | |
| S12F | 0.50 | GR | - | N/A (residue not contained in PDB) | |
| V47A | 0.17 | GH | interface | No structural damage detected | 2 |
| T51N* | 0.17 | GH | surface | H-bond formed: Asn51 - His49 (no damage alert) | 2 |
| Q52K | 0.17 | GH | interface | No structural damage detected | 2 |
| L54F | 0.34 | G | surface | No structural damage detected | 3 |
| F55L | 0.17 | GH | buried | No structural damage detected | 2 |
| P57T* | 0.17 | GH | buried | H-bond formed: Thr57 - Gln271 (no damage alert) | 2 |
| I68R* | 0.34 | O | surface | No structural damage detected | |
| H69Y | 0.17 | GH | surface | No structural damage detected | |
| V70L | 0.17 | GH | - | N/A (residue not contained in PDB) | 2 |
| S71P | 0.17 | GH | - | N/A (residue not contained in PDB) | 2 |
| G75V | 0.17 | GR | - | N/A (residue not contained in PDB) | |
| S98F | 1.51 | GR | surface | H-bond disrupted: Ser98 - Lys97 (no damage alert) | |
| D138H | 0.34 | GH | buried | Buried salt bridge disrupted: Asp138 - Arg21 / buried to exposed switch | |
| M153I | 0.50 | GH | surface | No structural damage detected | |
| M153T | 0.34 | GR | surface | No structural damage detected | |
| A163V | 0.17 | GH | surface | No structural damage detected | |
| G181V | 0.17 | GH | surface | Disallowed phi psi angles / cavity contraction / H bond formed: Ser98 - Val181 | |
| R190S | 0.17 | GH | surface | Cavity contraction / H-bond disrupted: Arg190 - Lys206 | |
| R190K | 0.17 | GV | surface | Cavity contraction / H-bond disrupted: Arg190 - Lys206 | |
| D198G | 0.17 | GH | interface | No structural damage detected | |
| S205T* | 0.17 | S | surface | H-bond formed: Thr205 - Glu191 / Cavity contraction | |
| R214P | 0.17 | G | surface | Disallowed phi psi angles | 3 |
| A222V | 0.67 | GV | buried | No structural damage detected | |
| K278T* | 0.17 | GH | surface | H-bond formed: Thr278 - Thr286 (no damage alert) | |
| I358T | 0.17 | GH | buried | No structural damage detected | |
| S359N | 0.17 | G | surface | H-bond formed: Asn360 - Asn359 (no damage alert) | |
| V407A | 0.50 | S/GH/GR | surface | 6XR8 - cavity expansion<br>4M0J - no structural damage detected | 4 |
| S459F | 20.17 | GH | surface | 6XR8 - H-bond disrupted: Arg457 - Ser459 (no damage)<br>4M0J - H-bond disrupted: Ser459 - Asn460 (no damage) | |
| S477N | 1.18 | GH | surface | No structural damage detected (both 6XR8 and 4M0J) | |
| V483F | 0.17 | GH | surface | No structural damage detected (both 6XR8 and 4M0J) | |
| E484D | 0.17 | GR | surface | No structural damage detected (both 6XR8 and 4M0J) | |
| N501T | 1.51 | GR | interface (ACE) | 6XR8 and 4M0J - H-bond disrupted: Asn501- Gly496 and Asn501 - Tyr505, H-bond formed: Thr501- Gln498 (no damage alert)<br>4M0J only - H-bond formed: Thr501 - Gly502 (no damage alert) | |
| K537R | 0.17 | GH | surface | H-bond formed: Arg537 - Asn536 (no damage alert) | |
| D614G | 48.74 | G/GH/GR/ GRY/GV | interface | H-bond disrupted: Asp614 (chain A) - Lys835 (chain B) and Asp614 (chain A) - Lys854 (chain B) / buried charge replaced | |
| Q675H | 0.17 | GR | surface | No structural damage detected | |

Table 1. Continued.

| Mutation | % | GISAID clade [1] | Location on spike | Structural consequences (Missense3D) | Note |
|---|---|---|---|---|---|
| A688P | 0.17 | GH | - | N/A (residue not contained in PDB) | 2 |
| V785L | 0.50 | GV | buried | No structural damage detected | |
| T791I | 0.17 | GR | surface | Cavity contraction / H-bond disrupted: Thr791 - Pro792 | |
| A829T | 39.16 | S | surface | H-bond formed: Thr829 - Asp830 (no damage alert) | |
| G832C* | 0.34 | G | interface | No structural damage detected | |
| F855L | 0.17 | GR | interface | No structural damage detected | |
| D936Y | 0.50 | GH | surface | No structural damage detected | |
| S939F | 2.02 | GH | surface | No structural damage detected | |
| D1118Y | 0.34 | GH | surface | Cavity contraction | |
| S1147L | 0.17 | GH | surface | H-bond disrupted: Ser1147 - Pro1143 (no damage alert) | |
| P1162L | 0.17 | S | surface | No structural damage detected | |
| V1176F | 0.17 | GR | - | N/A (residue not contained in PDB) | |
| E1202Q | 0.34 | GV/GH | - | N/A (residue not contained in PDB) | |
| G1219C | 0.50 | L | - | N/A (residue not contained in PDB) | |
| D1260Y | 0.17 | GR | - | N/A (residue not contained in PDB) | 5 |

Note: 1) The column shows on the GISAID clade of strains in which the majority of variants can be found. Some variants are not directly mappable onto the phylogenetic tree due to 2) variant from GISAID ID: 708814, 3) incomplete genome information, 4) one record from GISAID ID 708814, one record had incomplete genome, one record had incomplete collection date, and 5) incomplete date of collection.

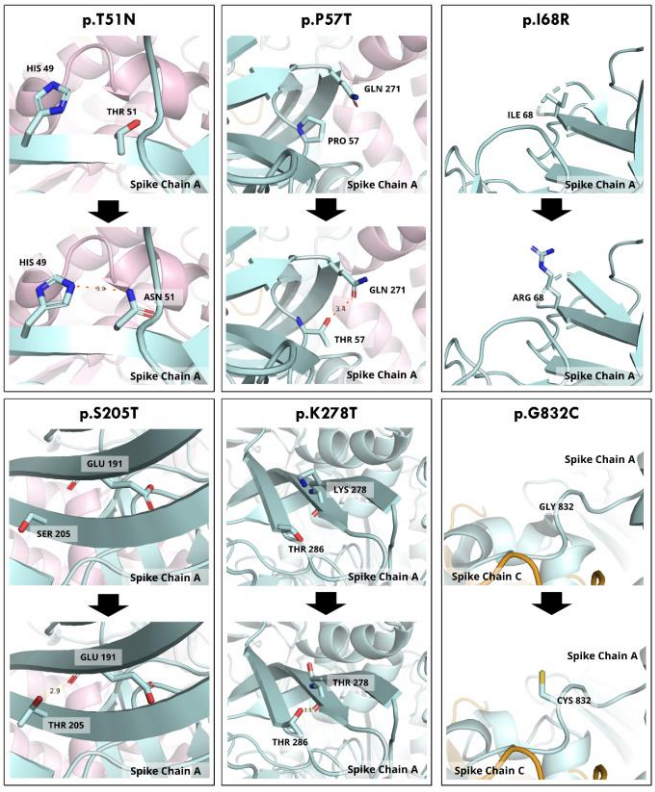


Figure 3. Structural analysis of Thailand-specific spike mutations. In each panel, the wild type structure is presented on top and the predicted mutant on the bottom. The PDB structure used in this analysis was 6XR8. Chains A, B, and C are represented in pale blue, light pink, and yellow, respectively. All spike mutations were simulated on Chain A. H-bonds are shown as orange dashed lines.

enhance SARS-CoV-2 infectivity (Li *et al.*, 2020). Mutation p. N501T was the only amino acid substitution occurring at the interface between the spike and the hACE2. Unlike p.N501Y, which has been proposed to enhance spike-hACE2 affinity (Laffeber, De Koning, Kanaar, and Lebbink, 2021; Shahhosseini, Babuadze, Wong, and Kobinger, 2021), our analysis showed that p.N501T did not result in major structural changes.

## 4. Discussion

At the time of this study, limited information was available on spike mutations identified in Thailand. Our phylogenetic analysis shows how the spike mutation of interest p.D614G, already present in March 2020, became the dominant strain during the second Thai outbreak in late December 2020 - February 2021. This mutation was detected in the majority of the strains collected from the seafood market (Department of Disease Control, 2020c, 2020e) and has been shown to have a higher transmissibility rate compared to the original Wuhan strain (Korber *et al.*, 2020; Volz *et al.*, 2021). The screening of mutations present in Thailand also identified mutation p.N501T. Aspargine 501 is a critical residue for spike-hACE2, and its substitution for tyrosine is a mutation of interest that had been widely studied. However, our analysis did not suggest that the asparagine-to-threonine substitution may cause any major structural change. Our results are supported by the in vitro analysis performed by Shang, Ye, *et al.* (2020), which did not show an increased spike-hACE2 binding affinity.

Interestingly, many highly conserved residues are clustered in the receptor-binding domain, which binds to the human ACE2 and triggers host cell invasion (Lan *et al.*, 2020; Shang, Wan, *et al.*, 2020; Shang, Ye, *et al.*, 2020), and the Coronavirus S2 domain, which mediates viral cell membrane fusion (Huang, Yang, Xu, Xu, & Liu, 2020). The lack of mutations in these regions may suggest the biological importance of those particular residues on protein folding and/or interaction. The conserved residues should be further investigated for their potential to serve as biomarkers (Zhang, & Guo, 2020) or drug targets against SARS-CoV-2 (Huang *et al.*, 2020).

One limitation of this study is that the Thai sequences available on NCBI and GISAID were from two sources: the state quarantines and domestic hospitals/research institutions. The strains collected in the state quarantine zones are likely imported into Thailand and less likely to cause local transmission. Hence, the number of mutations found in Thailand in this study could be an overestimation of the mutations that actually caused local transmissions. Nevertheless, it is always important to keep track of possible new strains as the mutation rate is very high in single-stranded RNA viruses compared to DNA viruses (Sanjuán, Nebot, Chirico, Mansky, and Belshaw, 2010). The high mutation rate poses a great challenge for developing vaccines (David A Steinhauer, Holland, 1987), highlighting why incorporating conserved residue information into structural analyses could be essential for discovering other alternative measures for COVID-19 diagnosis, treatment, and prevention.

Another limitation is that at the time of this study only the high-quality 3D coordinates of the spike trimer and the spike-hACE2 interaction were available. Recently, the spike protein was reported to bind to other human proteins, such as CD209 (Amraie *et al.*, 2020), HAVCR1 (Kane, 2020), and NRP1 (Cantuti-Castelvetri *et al.*, 2020), and to other mammalian proteins (Lam *et al.*, 2020). Therefore, in this study, the number of interface residues could have been underestimated, and it is possible that, in the future, additional mutations will be classified according to their effects on additional virus-host interactions. It is well established that interface residues are hot spots for disease-associated amino acid mutations (David, Razali, Wass, & Sternberg, 2012; David & Sternberg, 2015; Jubb *et al.*, 2017). The p.N501Y is an example that shows how viral spike mutations occurring at protein interface could increase viral transmissibility, and highlights that any SARS-CoV-2 mutation occurring at the spike trimer interface or spike-hACE2 interface should be carefully analysed.

## 5. Conclusions

In this study, we explored the structural effects of SARS-CoV-2 spike mutations identified in Thailand, highlighting how the second Thai outbreak was likely caused by the spike mutation of interest p.D614G. Additionally, we highlighted highly conserved residues on the spike protein, which could have implications for the development of biomarkers or drug targets. The insights obtained from this study could aid in prioritizing SARS-CoV-2 variants and help molecular biologists and virologists working on strain surveillance.

## Acknowledgements

This research was funded in whole, or in part, by the Wellcome Trust [Grant numbers 104955/Z/14/Z and 218242/Z/19/Z]. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. This Funder had no role in the conceptualization, design, data collection, analysis, decision to publish, or preparation of the manuscript.

## References

Amraie, R., Napoleon, M. A., Yin, W., Berrigan, J., Suder, E., Zhao, G., . . . Rahimi, N. (2020, June 23). CD209L/L-SIGN and CD209/DC-SIGN act as receptors for SARS-CoV-2 and are differentially expressed in lung and kidney epithelial and endothelial cells. *BioRxiv*. bioRxiv. doi:10.1101/2020.06.22.165803

Bogner, P., Capua, I., Cox, N. J., & Lipman, D. J. (2006, August 31). A global initiative on sharing avian flu data. *Nature*, *442*, 981. doi:10.1038/442981a

Bosshard, H. R., Marti, D. N., & Jelesarov, I. (2004). Protein stabilization by salt bridges: Concepts, experimental approaches and clarification of some mis understandings. *Journal of Molecular Recognition*, *17*(1), 1–16. doi:10.1002/jmr.657

Buathong, R., Chaifoo, W., Iamsirithawon, S., Wachara pluesadee, S., Joyjinda, Y., Rodpan, A., . . . Hemachudha, T. (2021). Multiple clades of SARS-CoV-2 were introduced to Thailand during the first quarter of 2020. *Microbiology and Immunology*, 1348-0421.12883. doi:10.1111/1348-0421.12883

Cantuti-Castelvetri, L., Ojha, R., Pedro, L. D., Djannatian, M., Franz, J., Kuivanen, S., . . . Simons, M. (2020). Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. *Science*, *370*(6518), 856–860. doi:10.1126/science.abd2985

David, A., Razali, R., Wass, M. N., & Sternberg, M. J. E. (2012). Protein-protein interaction sites are hot spots for disease-associated nonsynonymous SNPs. *Human Mutation*, *33*(2), 359–363. doi:10.1002/humu.21656

David A Steinhauer, Holland, J. J. (1987). Rapid Evolution of RNA Viruses. *Annual Review of Microbiology*, *41*(1), 409–431. doi:10.1146/annurev.mi.41.100187.002205

David, A., & Sternberg, M. J. E. (2015). The contribution of missense mutations in core and rim residues of protein-protein interfaces to human disease. *Journal of Molecular Biology*, *427*(17), 2886–2898. doi:10.1016/j.jmb.2015.07.004

Department of Disease Control. (2020a). Measures to control COVID-19 Version 5 updated on 2 April 2020. Retrieved from https://ddc.moph.go.th/viralpneumonia/eng/file/main/Measures_for_traverlers_from_aboard_v5.pdf

Department of Disease Control. (2020b). Special announcement of COVID-19 On 1 July 2020. Retrieved from https://ddc.moph.go.th/viralpneumonia/eng/file/news/news_no157_010763.pdf

Department of Disease Control. (2020c). Thailand situation update on 19 December 2020. Retrieved from https://ddc.moph.go.th/viralpneumonia/eng/file/situation/situation-no349-191263.pdf

Department of Disease Control. (2020d). Thailand situation update on 22 November 2020. Retrieved from https://ddc.moph.go.th/viralpneumonia/eng/file/situation/situation-no322-221163.pdf

Department of Disease Control. (2020e). Thailand situation update on 24 December 2020. Retrieved from https://ddc.moph.go.th/viralpneumonia/eng/file/situation/situation-no350-241263.pdf

Genomic epidemiology of novel coronavirus - Thailand-focused subsampling. (2021). Retrieved from https://nextstrain.org/community/faik/coni/Thailand?branchLabel=aa&c=GISAID_clade&d=tree&dmax=2021-02-07&fbclid=IwAR19RF7fisEmajtTO34mlTB7pDPzpvC-adGmXyLYMkdMikA_yDTVFLvcG5A

Hadfield, J., Megill, C., Bell, S. M., Huddleston, J., Potter, B., Callender, C., . . . Neher, R. A. (2018). NextStrain: Real-time tracking of pathogen evolution. *Bioinformatics*, *34*(23), 4121–4123. doi:10.1093/bioinformatics/bty407

Huang, Y., Yang, C., Xu, X. feng, Xu, W., & Liu, S. wen. (2020, September 1). Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. *Acta Pharmacologica Sinica*, *41*, 1141–1149. doi:10.1038/s41401-020-0485-4

Ittisoponpisan, S., Islam, S. A., Khanna, T., Alhuzimi, E., David, A., & Sternberg, M. J. E. (2019). Can predicted protein 3D structures provide reliable insights into whether missense variants are disease associated? *Journal of Molecular Biology*, *431*(11), 2197–2212. doi:10.1016/j.jmb.2019.04.009

Jeffrey, G. A. (1997). *An introduction to hydrogen bonding*. Oxford , England: Oxford University Press.

Joonlasak, K., Batty, E. M., Kochakarn, T., Panthan, B., Kümpornsin, K., Jiaranai, P., . . . Chantratita, W. (2021). Genomic surveillance of SARS-CoV-2 in Thailand reveals mixed imported populations, a local lineage expansion and a virus with truncated ORF7a. *Virus Research*, *292*, 198233. doi:10.1016/j.virusres.2020.198233

Jubb, H. C., Pandurangan, A. P., Turner, M. A., Ochoa-Montaño, B., Blundell, T. L., & Ascher, D. B. (2017, September). Mutations at protein-protein interfaces: Small changes over big surfaces have large impacts on human health. *Progress in Biophysics and Molecular Biology*, *128*, 3–13. doi:10.1016/j.pbiomolbio.2016.10.002

Kabsch, W., & Sander, C. (1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, *22*(12). doi:10.1002/bip.360221211

Kane, L. (2020). Review 1: "KIM-1/TIM-1 is a receptor for SARS-CoV-2 in lung and kidney" License: Creative Commons Attribution 4.0 International License (CC-BY 4.0). *Rapid Reviews COVID-19*. doi:10.1162/2e3983f5.919b71a6

Korber, B., Fischer, W. M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., . . . Montefiori, D. C. (2020). Tracking changes in SARS-CoV-2 spike: Evidence that D614G increases infectivity of the COVID-19 virus. *Cell*, *182*(4). doi:10.1016/j.cell.2020.06.043

Laffeber, C., De Koning, K., Kanaar, R., & Lebbink, J. H. (2021). Experimental evidence for enhanced receptor binding by rapidly spreading SARS-CoV-2 variants. *BioRxiv*, 2021.02.22.432357. doi:10.1101/2021.02.22.432357

Lam, S. D., Bordin, N., Waman, V. P., Scholes, H. M., Ashford, P., Sen, N., . . . Orengo, C. A. (2020). SARS-CoV-2 spike protein predicted to form complexes with host receptor protein orthologues from a broad range of mammals. *Scientific Reports*, *10*(1), 16471. doi:10.1038/s41598-020-71936-5

Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., . . . Wang, X. (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*, *581*(7807), 215–220. doi:10.1038/s41586-020-2180-5

Li, Q., Wu, J., Nie, J., Li, X., Huang, W., Correspondence, Y. W., . . . Wang, Y. (2020). *The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity*. doi:10.1016/j.cell.2020.07.012

Mat Yassim, A. S., Asras, M. F. F., Gazali, A. M., Marcial-Coba, M. S., Zainulabid, U. A., & Ahmad, H. F. (2021). COVID-19 outbreak in Malaysia: Decoding D614G mutation of SARS-CoV-2 virus isolated from an asymptomatic case in Pahang. *Materials Today: Proceedings*. doi:10.1016/j.matpr.2021.02.387

Nyunt, M. H., Soe, H. O., Aye, K. T., Aung, W. W., Kyaw, Y. Y., Kyaw, A. K., . . . Htun, Z. T. (2021). Surge of severe acute respiratory syndrome coronavirus 2 infections linked to single introduction of a virus strain in Myanmar, 2020. *Scientific Reports*, *11*(1), 1–6. doi:10.1038/s41598-021-89361-7

Okada, P., Buathong, R., Phuygun, S., Thanadachakul, T., Parnmen, S., Wongboot, W., . . . Maurer-Stroh, S. (2020, February 27). Early transmission patterns of coronavirus disease 2019 (COVID-19) in travellers from Wuhan to Thailand, January 2020. *Euro surveillance*, Vol. 25. European Centre for Disease Prevention and Control (ECDC). doi:10.2807/1560-7917.ES.2020.25.8.2000097

Peter, A. S., Roth, E., Schulz, S. R., Fraedrich, K., Steinmetz, T., Damm, D., . . . Jäck, H.-M. (2021). A pair of non-competing neutralizing human monoclonal antibodies protecting from disease in a SARS-CoV-2 infection model. *BioRxiv*, 2021.04.16.440101. doi:10.1101/2021.04.16.440101

Puenpa, J., Suwannakarn, K., Chansaenroj, J., Nilyanimit, P., Yorsaeng, R., Auphimai, C., . . . Poovorawan, Y. (2020). Molecular epidemiology of the first wave of severe acute respiratory syndrome coronavirus 2 infection in Thailand in 2020. *Scientific Reports*, *10*(1). doi:10.1038/s41598-020-73554-7

Rost, B., & Sander, C. (1994). Conservation and prediction of solvent accesibility in protein families. *Proteins: Structure, Function and Genetics*, *20*(3), 216–226.

Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M., & Belshaw, R. (2010). Viral Mutation Rates. *Journal of Virology*, *84*(19), 9733–9748. doi:10.1128/jvi.00694-10

Shahhosseini, N., Babuadze, G. (Giorgi), Wong, G., & Kobinger, G. P. (2021). Mutation Signatures and In Silico Docking of Novel SARS-CoV-2 Variants of Concern. *Microorganisms*, *9*(5), 926. doi:10.3390/microorganisms9050926

Shang, J., Wan, Y., Luo, C., Ye, G., Geng, Q., Auerbach, A., & Li, F. (2020). Cell entry mechanisms of SARS-CoV-2. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(21), 11727–11734. doi:10.1073/pnas.2003138117

Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., . . . Li, F. (2020). Structural basis of receptor recognition by SARS-CoV-2. *Nature*, *581*(7807), 221–224. doi:10.1038/s41586-020-2179-y

Volz, E., Hill, V., McCrone, J. T., Price, A., Jorgensen, D., O'Toole, Á., . . . Pybus, O. G. (2021). Evaluating the Effects of SARS-CoV-2 Spike Mutation D614G on Transmissibility and Pathogenicity. *Cell*, *184*(1). doi:10.1016/j.cell.2020.11.020

Waman, V. P., Sen, N., Varadi, M., Daina, A., Wodak, S. J., Zoete, V., . . . Orengo, C. (2021). The impact of structural bioinformatics tools and resources on SARS-CoV-2 research and therapeutic strategies. *Briefings in Bioinformatics*, *22*(2), 742–768. doi:10.1093/bib/bbaa362

WHO Thailand. (2020a). WHO Thailand situation report - 41. Retrieved from https://www.who.int/docs/default-source/searo/thailand/2020-04-3-tha-sitrep-41-covid19-final.pdf?sfvrsn=9e14aebc_0.pdf

WHO Thailand. (2020b). WHO Thailand situation report - 95. Retrieved from https://www.who.int/docs/default-source/searo/thailand/2020-07-06-tha-sitrep-95-covid19.pdf?sfvrsn=272b28fd_2

WHO Thailand. (2020c). WHO Thailand situation report - 102. Retrieved from https://cdn.who.int/media/docs/default-source/searo/thailand/2020_12_21_tha-sitrep-102-covid19_r02.pdf?sfvrsn=7e62038c_7&Status=Master

Yurkovetskiy, L., Wang, X., Pascal, K. E., Tomkins-Tinch, C., Nyalile, T. P., Wang, Y., . . . Luban, J. (2020). Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. *Cell*, *183*(3). doi:10.1016/j.cell.2020.09.032

Zhang, L., & Guo, H. (2020). Biomarkers of COVID-19 and technologies to combat SARS-CoV-2. *Advances in Biomarker Sciences and Technology*, *2*, 1–23. doi:10.1016/j.abst.2020.08.001