http:// ijp.mums.ac.ir
Original Article (Pages: 16486-16497)

# Evaluation of Mutations in SARS-CoV-2 N and S Genes on the Proteins Stability, Immunogenicity, and Pathogenicity in Iranian Patients from Golestan Province

Bahman Aghcheli[1], Alireza Tahamtan[2], Hadi Razavi Nikoo[3], Zahra Bazi[4], Mohammad Reza Kalani[5], *Abdolvahab Moradi[6]

[1] Department of Microbiology, School of Medicine, Golestan University of Medical Sciences, Gorgan, Iran.
[2] Department of Microbiology, School of Medicine, Golestan University of Medical Sciences, Gorgan, Iran.
[3] Department of Microbiology, School of Medicine, Golestan University of Medical Sciences, Gorgan, Iran.
[4] Department of Biotechnology, Golestan University of Medical Sciences, Gorgan, Iran.
[5] Senior Scientist, Enterprise-TTM, University of Pittsburgh Medical Center, Pittsburgh, PA.
[6] Department of Microbiology, School of Medicine, Golestan University of Medical Sciences, Gorgan, Iran.

***Abstract***

***Background:*** Natural selection such as mutations is known as a constant process for viral fitness and selective adaptation. Understanding the effects of each mutation, especially on structural proteins in the viral life cycle, is important in tracking the viruses behavior. Here, we evaluated the effects of mutations in SARS-CoV-2 nucleoprotein (N) and spike (S) genes on the protein stability, immunogenicity, and pathogenicity in Iranian COVID-19 patients from Golestan province.

***Methods:*** In this study, 8 SARS-CoV-2 RNA samples were enrolled from referral hospitals in Golestan province. These samples were confirmed using a real-time RT-PCR assay targeting the SARS-CoV-2 nucleoprotein (N) and ORF1ab genes (Pishtazteb, Iran). Next-generation sequencing (NGS) was done on samples and subsequent sequences were retrieved from Global Initiative on Sharing All Influenza Data (GISAID) EpiCoV database. Structural analysis was performed between wild type (Wuhan accession number: NC_045512.2) and mutant N and S proteins to evaluate their stability, immunogenicity, and pathogenicity via bioinformatics servers such as Dynamut, Prodigy, IEDB, and software's (Mega XI and Pymol II.V.II visualizer).

***Results:*** Amino acid codon changes in N and S proteins show that mutations could alter the translation efficiency. Normal Mode Analysis (NMA) by Dynamut server shows that stability and flexibility are changed by the mutations of these proteins. Immunogenicity analysis indicates the potential effects of some mutations such as P681H, Q675R, L699I, and D3L on immune escape. Interaction complex binding energy and affinity are higher in the mutant type compared to the Wuhan wild type, indicating higher pathogenicity.

***Conclusion:*** The results indicate that there are some important mutations in N and S genes that affect the virus behavior in the infectivity. Regarding the sample size limitation and various mutations in SARS-CoV-2 variants, other studies using whole-genome sequencing with larger sample sizes will be required. Therefore, continuous monitoring of the SARS-CoV-2 genome seems important.

***Key Words:*** Bioinformatics Analysis, COVID-19, Mutations, SARS-CoV-2, Spike, Nucleoprotein.

* Please cite this article as: Aghcheli B, Tahamtan A, Razavi-Nikoo H, Bazi Z, Kalani MR, Moradi A. Evaluation of Mutations in SARS-CoV-2 N and S Genes on the Proteins Stability, Immunogenicity, and Pathogenicity in Iranian Patients from Golestan Province. Int J Pediatr 2022; 10 (8):16486-16497. DOI: **10.22038/ijp. 2022.64880.4906**

***Corresponding Author:***

Abdolvahab Moradi, Department of Microbiology, School of Medicine, Golestan University of Medical Sciences, Gorgan, Iran. Email: abmoradi@gmail.com

Received date: Apr.11,2022; Accepted date:May.13,2022

## 1- INTRODUCTION

Coronavirus disease 2019 (COVID-19), is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The SARS-CoV-2 genome is approximately 30 kb in size and encodes 12 structural proteins and 16 non-structural proteins (1).

Mutations in structural proteins (spike, nucleoprotein, envelope, and membrane) enable the adaptation of the virus to the host (2). Changes in viral protein structure play a key role in functions, virulence, infectivity, and transmissibility (3, 4).

Mutations in the spike gene can increase angiotensin converting enzyme 2 (ACE-2) receptor binding affinity, viral replication or infectivity, disease severity, higher transmissibility, resistance to neutralizing antibodies or immune escape, and reinfection risk (5-8), as well as mutations on the N-protein that make alter virulence and virus spread (9-11).

A study by Mohammad et al. showed that mutations lead to the evolution of the viral genome, allowing them to better adapt and survive in the human host for their active reproduction. Such mutations are achieved by modifications in the epitopes of viral genes, making them more infective; and help in escaping the immune responses of the host (12).

Mutation in structural proteins likely changes the phenotype of proteins, disrupting stability, structural folding, macromolecular binding, and ablation of posttranslational modification sites (13-16).

Another study by Al-Zyoud and Haddad showed that spike protein mutations could increase its stability and flexibility and unify the superimposition which might impact the viruses ability to escape the antibodies neutralization by changing the antigenicity drift of the protein three-dimensional (3D) structure from the wild type (17).

The receptor binding domain (RBD) undergoes hinge-like conformational movements that transiently hide or expose determinants that facilitate engagement of the host cell receptor (18). The conformational state of the spike protein was found more stable in up conformation and the down conformational state reduces the number of exposed (electrostatic) epitopes (19, 20).

Intrinsically disordered regions (IDRs) are abundant in the SARS-CoV-2 viral proteins, including the S1 subunit of spike and nucleoprotein (21). Mutations in the mentioned regions are critical for immune evasion and antibody escape, suggesting potential additional implications for vaccines and monoclonal therapeutic strategies (22). The IDRs play a special role in increasing binding affinity and enhancing the binding allosteric that enables the N protein to bind RNA with a high cooperativity (23).

A study by Azad GK Azad showed that a high rate of mutations are located in intrinsically disordered regions (IDRs) of N-protein and multiple mutations lead to considerable alterations in the function of this protein (24).

SARS-CoV-2 nucleoprotein can be directly activated by Smad3 and enhance TGF-β-mediated gene expression (25, 26). SARS-CoV-2 is capable of suppressing the type-I IFN innate immune pathway possibly due to the role of N-protein in signaling events (27, 28).

Studying the effect of mutations on structural proteins can help us to understand the viruses virulence, antibody escape, and infectivity. We studied the effects of each mutation in the N and S structural proteins on stability, immunogenicity, and pathogenicity. It can be a new perspective on virus behavior and

an effective step in the management and treatment of COVID-19 disease.

## 2- METHODS

### 2-1. Sample collection

In this study, 8 SARS-CoV-2 RNA samples, confirmed by real-time RT-PCR assay targeting the SARS-CoV-2 nucleoprotein (N) and ORF1ab genes (Pishtazteb, Iran), were enrolled from referral hospitals in Golestan province between March 2020 and September 2021. The range of CT (Cycle of Threshold) values in all COVID-9 patients was between 12-15.

### 2-2. Whole genome sequencing and genome assembly

The SARS-CoV-2 whole-genome sequencing was done through the clean plex® SARS-CoV-2 research and surveillance panel (Paragon genomics, Inc.). The samples were paired-end sequenced with Illumina MiSeq by 300-cycle MiSeq v2 reagent kits (Illumina, Inc.).

### 2-3. Multiple sequence alignment

The SARS-CoV-2 whole genome sequences from Gorgan recorded in the GISAID EpiCoV database were retrieved. These sequences aligned and compared with the first sequence, reported in Wuhan China (NC_045512.2) with Mega XI software and the Clustal W algorithm. Nucleoprotein and spike nucleotide sequences were translated into protein sequences.

### 2-4. N and S PDB protein construction (Homology modeling)

The Swiss-model server (https://swissmodel.expasy.org/interactive) was used for N and S PDB protein construction.

### 2-5. N and S PDB protein structure validation

SAVES server (https://saves.mbi.ucla.edu) a complete package including (ERRAT, VERIFY3D, PROVE, PROCHECK, and WHAT CHECK) was used to validate or evaluate the overall quality of the structure of the generated PDB protein.

### 2-6. N and S protein structure stability and flexibility

The impact of mutations on the molecular stability and flexibility of protein structures was estimated using the DynaMut web server (29). This server utilized other servers such as Site-Directed Mutator (SDM), mCSM, or DUET (combine SDM and mCSM), for stability prediction (30-33).

### 2-7. N and S proteins immunogenicity analysis

The immunogenicity of the mutations was compared with the native S and N proteins using the immune epitope database (IEDB).

This server is a comprehensive immunoinformatics database for examining T cells and B cells antigens. In addition, linear and discontinuous B cells epitopes were predicted by Antigen Sequence Properties and Disco Tope tools, respectively (34).

### 2-8. Molecular docking spike/ACE2 and nucleoprotein/Smad3

Servers such as LZer Dock, Hawk Dock, and Clus pro are used for binding prediction (35-37).

YASARA server for all input structures before docking is conducted to the relaxation energy minimization, deletion of water, adding the polar hydrogen atoms of protein and ligands, along with the repack process for backbone and side chains clashes fixation, respectively (38).

### 2-9. Binding affinity of spike/ACE2 and nucleoprotein/Smad3

A user-friendly online server such as Prodigy is used for the prediction of

binding affinity in protein-protein complexes (39).

## 3- RESULTS

### 3-1. N and S protein mutations

Nucleotide sequences are translated to an amino acid with MEGA software. N and S mutations are listed in **Table 1**. In a total of 8 samples, mutations are seen on 5 samples. Our findings also suggest that the translation efficiency of these proteins may be affected by altering the amino acid codon, for example, Q675R, D3L, and S194L.

**Table-1:** Effect of some mutations in spike and nucleoprotein on translation efficiency

| Sample ID | Amino acid change | Translation alteration (%) | Position |
|---|---|---|---|
| Spike | | | |
| Gorgan 80 | G142S | 10.8 : 11.9 | N-terminal domain |
| Gorgan 25 | V483A | 10.9: 18.6 | Receptor binding motif |
| Gorgan I15 | Q675R | 11.9: 34.6 | Near cleavage site |
| Gorgan 117 | L699I | 12.8 : 15.7 | Cleavage site-Fusion peptide |
| Gorgan E312 | D614N | 16.7: 22.3 | Outer receptor binding domain |
| Nucleoprotein | | | |
| Gorgan 80 | N192K | 19.5: 32.9 | Linker |
| Gorgan 25 | S194L | 7.2 : 14.7 | Linker |
| Gorgan I15 | M234I | 15.7 : 22.3 | Linker |
| Gorgan117 | D3L | 6.9: 22.3 | N-terminal domain |
| Gorgan E312 | R185C | 4.7: 9.9 | Linker |

### 3-2. N and S protein PDB validation

We selected the best structure of N and S PDB constructed by swiss model criteria such as Molprobity indexes, Ramachandran plot, QMEAN (Qualitative Model Energy Analysis) Z-score, GMQE, as well as QMEANDisCo Local model evaluation for wild and mutant spike proteins. ERRAT validation revealed the overall quality factor of these model proteins to be above 97. Ramachandran plot analysis for all structure S and N models by PROCHECK revealed that 92.3% of its residues were in the most favored regions, 7.4% were in additional allowed regions, 0.3% in generously allowed regions, and 0% in disallowed regions (**Fig 1**). A model would be considered as good quality and high-reliability if it has over 90% of its residues in the most favored regions. All of these validation scores suggested that this model was highly reliable to use for further analysis.

### 3-3. Molecular docking of spike/ACE2 and nucleoprotein/Smad3

To understand the strength of the binding, specificity (affinity), and the possible role of mutations in facilitating the binding and entry of the virus into the cell, molecular docking was performed with the servers mentioned in the method section. Gorgan 25, Gorgan117, Gorgan E312, Gorgan I15, and Gorgan 80 higher binding interaction with the receptor rather than Wuhan ref. (NC_045512.2) (**Table 2**).

### 3-4. Effect of N and S protein mutations on affinity by prodigy server

**RBD/ACE2 and N/Smad3 complex:** Gorgan 25, Gorgan117, Gorgan E312, Gorgan I15, and Gorgan 80 have a higher binding affinity, respectively according to the prodigy server (**Table 3**). The increase in binding affinities is correlated with interfacial interactions, such as hydrogen bonding, salt bridges, and hydrophobic contacts.

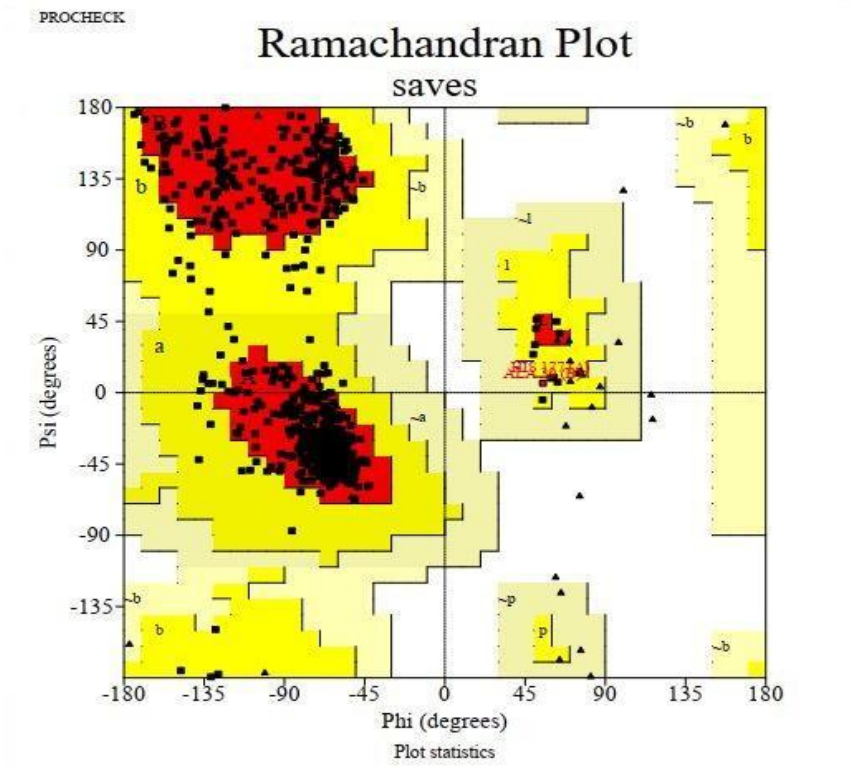


**Fig. 1:** An example of Ramachandran plot for N and S; proteinsRamachandran plot analysis revealed that 92.3% of its residues were in the most favored regions, 7.4% were in additional allowed regions, 0.3 % in generously allowed regions, and 0% in disallowed regions

**Table-2:** RBD-ACE2 and Nucleoprotein-Smad3 complex docking scores (Binding energy)

| Variable | Sample Servers | Wuhan | Gorgan 25 | Gorgan 117 | Gorgan I15 | Gorgan 75 | Gorgan 76 | Gorgan E312 | Gorgan 80 | Gorgan 83 |
|---|---|---|---|---|---|---|---|---|---|---|
| RBD-ACE2 complex | LZer Dock | 189 | 372 | 280 | 215 | 189 | 189 | 225 | 201 | 189 |
| | Clus pro Dock | -676.2 | -699.8 | -684.8 | -679.5 | -676.2 | -676.2 | -681.5 | -678.2 | -676.2 |
| | Hawk Dock | -6190.5 | -6589.2 | -6470.6 | -6245.4 | -6190.5 | -6190.5 | -6255.4 | -6210.5 | -6190.5 |
| N-Smad3 complex | LZer Dock | 233 | 433 | 375 | 360 | 233 | 233 | 366 | 341 | 233 |
| | Clus pro Dock | -872.8 | -919.6 | -914.9 | -905.3 | -872.8 | -872.8 | -908.5 | -901.8 | -872.8 |
| | Hawk Dock | -4850.2 | -5671.8 | -5400.7 | -5260.9 | -4850.2 | -4850.2 | -5292.4 | -4901.2 | -4850.2 |

### 3-5. N and S protein stability and flexibility

Mutations in N and S proteins had different behaviors on stability and flexibility. Four manners are conceivable: 1- decreased stability and flexibility (L699I and G204R) 2- increased stability and flexibility (S982A and M234I) 3- increased stability but decreased flexibility (T732I and N192K) 4- decreased stability but increased flexibility (I100 and D3L) (**Table 4**).

Stability and flexibility are affected by the number of interacting amino acids, number of interactions, and distance between atom residue interactions on the secondary structure. Prediction of interatomic interactions in the N and S protein is shown in **Fig. 2**.

**Table-3:** Prediction affinity of RBD-ACE2 and N-Smad3 complex by prodigy server (Binding affinity)

| Variable | Gorgan E312 | Gorgan 117 | Gorgan 25 | Gorgan I15 | Gorgan 75 | Gorgan 76 | Gorgan 80 | Gorgan 83 | Wuhan |
|---|---|---|---|---|---|---|---|---|---|
| RBD-ACE2 complex | | | | | | | | | |
| Binding energy (kcal/mol) | -14.1 | -14.3 | -14.8 | -13.9 | -13.4 | -13.4 | -13.6 | -13.4 | -13.4 |
| Kd (M) | 1.5E-10 | 1.4E-10 | 3.6E-11 | 2.3E-10 | 3.8E-10 | 3.8E-10 | 3.8E-10 | 3.8E-10 | 3.8E-10 |
| Charged-charged | 5 | 7 | 8 | 4 | 3 | 3 | 3 | 3 | 3 |
| Charged-polar | 5 | 5 | 5 | 5 | 6 | 6 | 6 | 6 | 6 |
| Charged-apolar | 20 | 20 | 20 | 20 | 18 | 18 | 21 | 18 | 18 |
| Polar-polar | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| Polar-apolar | 23 | 22 | 24 | 22 | 22 | 22 | 21 | 22 | 22 |
| Apolar-apolar | 11 | 12 | 12 | 11 | 10 | 10 | 10 | 10 | 10 |
| Total number of ICs | 70 | 72 | 75 | 68 | 65 | 65 | 67 | 65 | 65 |
| N-Smad3 complex | | | | | | | | | |
| Binding energy (kcal/mol) | -11.8 | -12.5 | -12.8 | -11.5 | -10.5 | -10.5 | -11 | -10.5 | -10.5 |
| Kd (M) | 2.5E-09 | 3.8-09 | 5.2E-09 | 1.5E-09 | 7.7E-08 | 7.7E-08 | 1.3E-09 | 7.7E-08 | 7.7E-08 |
| Charged-charged | 12 | 13 | 13 | 10 | 8 | 8 | 8 | 8 | 8 |
| Charged-polar | 25 | 28 | 30 | 23 | 18 | 18 | 21 | 18 | 18 |
| Charged-apolar | 32 | 19 | 22 | 38 | 22 | 22 | 41 | 22 | 22 |
| Polar-polar | 13 | 22 | 21 | 11 | 13 | 13 | 5 | 13 | 13 |
| Polar-apolar | 18 | 28 | 35 | 12 | 21 | 21 | 16 | 21 | 21 |
| Apolar-apolar | 10 | 9 | 13 | 12 | 17 | 17 | 9 | 17 | 17 |
| Total number of ICs | 110 | 119 | 134 | 106 | 97 | 97 | 100 | 97 | 97 |

ICs—number of interfacial contacts

**Table-4:** Prediction effect of some mutations in spike and nucleoprotein on stability and flexibility with Dynamut server

| Mutations | ΔΔG (kcal/mol) | ΔΔG ENCoM | ΔΔG mCSM | ΔΔG SDM | ΔΔG DUET | ΔΔE SVib ENCoM | stability | flexibility |
|---|---|---|---|---|---|---|---|---|
| Spike | | | | | | | | |
| D614N | -0.179 | -0.07 | -0.47 | -0.83 | -0.69 | 0.088 | D | I |
| L699I | -0.135 | 0.086 | -0.463 | -0.06 | -0.161 | -0.107 | D | D |
| V483A | -0.009 | -0.163 | -0.358 | -0.750 | -0.344 | -0.203 | D | D |
| G142S | -0/505 | -0/559 | -1/656 | -2/330 | -1/747 | 0/699 | D | I |
| Q675R | -0.321 | -0.158 | -0.112 | 0.562 | -0.248 | 0.123 | D | I |
| Nucleoprotein | | | | | | | | |
| S194L | -0.647 | -0.035 | -0.446 | -2.270 | -0.313 | 0.044 | D | I |
| N192K | 0.652 | 0.194 | 0.093 | -0.860 | 0.336 | -0.242 | I | D |
| M234I | 0.237 | -0.045 | 0.053 | 0.440 | 0.692 | 0.057 | I | I |
| D3L | -0.548 | -0.035 | 0.194 | -0.190 | -0.380 | 0.044 | D | I |
| R185C | 1.292 | 0.834 | 2.110 | 0.330 | 1.721 | 1.042 | I | I |

I: Increase, D: Decrease

ΔΔG: Negative sign related to destabilizing and reverse manner on the stability and flexibility, ΔΔE SVib ENCoM (Δ vibrational entropy energy), mCSM (Mutation cutoff scanning matrix), SDM (Site-directed mutator), and DUET (SDM + mCSM).

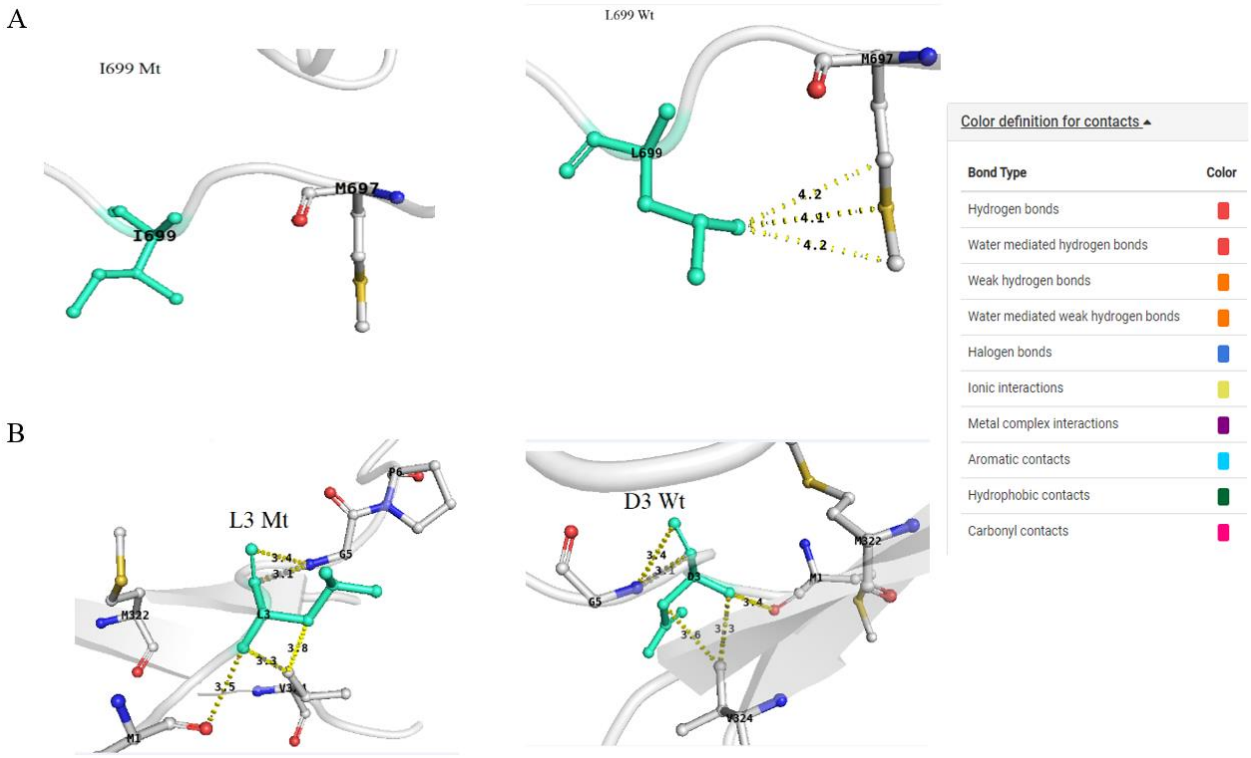


**Fig. 2:** DynaMut prediction of interatomic interactions for L699I spike (A) and D3L nucleoprotein (B) mutation; Wild-type and mutant residues are colored in light green and are also represented as sticks alongside the surrounding residues which are involved in any type of interaction. (W.t: wild type, M.t: Mutant).

### 3-6. N and S protein immunogenicity analysis

The effect of N and S protein mutations on acquired immunity and the possibility of the virus escaping antibody-mediated detection, as well as its ability to detect and deliver to T cells, especially helper and cytotoxic T cells, is shown in **Table 5**. Higher scores indicate greater immunogenicity (Q675R and D3L).

## 4- DISCUSSION

The formation of a stable binding interface between the receptor binding domain of SARS-CoV-2 spike protein and the ACE2 is essential for viral entry. The SARS-CoV-2 nucleoprotein is one hotspot for coding mutations, particularly at amino acid residues 203-205 within its serine-rich (SR) domain (40, 41).

SARS-COV-2 needs to hijack the translational machinery of the host to efficiently replicate and produce new virions. In this context, the codon usage of viral proteins should potentially resemble the host cell in order to adapt to the tRNA pools that drive an optimal translation (42). Our findings identified mutations in structural protein's tendency to increase translation efficiency, replication, and transcription. For example, translation efficiency rates for S194L, N192K, D3L, T362I, Q675R, and R203K mutations were doubled.

Point mutations affect the 3D structure, folding, stability of spike protein, and thermodynamic activity (43). Some mutations decrease or increase the binding affinity of spike protein with the receptor. For example, P681H, Q675R, V483A, R203K, and T362I mutations increased the affinity of RBD–ACE2 and N-Smad3 complexes. This finding is consistent with those of the other studies (44-46).

**Table-5:** Prediction effect of each spike and nucleoprotein mutation on immunogenicity

| Sample ID | Mutation | Disco tope | ASP | MHC-I | MHC-II |
|---|---|---|---|---|---|
| Spike | | | | | |
| Gorgan I15 | Q675R | W: -9.34 | W: 0.65 | W: 1.10 | W: 66 |
| | | M: -11.77 | M: 0.61 | M: 0.53 | M: 65 |
| Gorgan117 | L699I | W: -7.95 | W: 0.63 | W: 1.10 | W:49 |
| | | M:-9.19 | M:0.44 | M:0.51 | M:21 |
| Gorgan 25 | V483A | W: -5.16 | W: 0.52 | W: 1.10 | W:3.70 |
| | | M: -7.74 | M:0.34 | M:0.37 | M:3.10 |
| Gorgan E312 | D614N | W: -17.62 | W: 0.41 | W: 1.10 | W:85 |
| | | M: -19.57 | M: -0.47 | M:0.34 | M:70 |
| Gorgan 80 | G142S | W: -14.37 | W: 0.58 | W: 1.10 | W:70 |
| | | M: -17.81 | M:0.15 | M:0.18 | M:60 |
| Nucleoprotein | | | | | |
| Gorgan I15 | M234I | W: 3.703 | W: 0.77 | W: -6.05 | W:70 |
| | | M: -2.15 | M: 0.70 | M:-8.75 | M: 43 |
| Gorgan117 | D3L | W: -0.465 | W:1.84 | W:-6.05 | W:2.1 |
| | | M: -2.495 | M: 0.72 | M:-3.25 | M:1.7 |
| Gorgan 25 | S194L | W: 10.95 | W:1.61 | W:-6.05 | W:90 |
| | | M: 9.338 | M: 0.89 | M:-8.90 | M:88 |
| Gorgan E312 | R185C | W: 6.293 | W:1.40 | W:-6.05 | W:79 |
| | | M: 2.134 | M:1.04 | M:-7.06 | M:70 |
| Gorgan75 | N192K | W: 11.16 | W:1.36 | W:-6.05 | W:85 |
| | | M: 10.12 | M:1.22 | M:-6.18 | M:76 |

Disco tope: Predicting discontinuous epitopes from 3D structures of proteins; ASP (Antigen sequence properties): Predicting Linear B cell epitopes using antigen sequence properties; MHC-I (Major histocompatibility complex I): Presenting antigens for TCD8 cells; MHC-II (Major histocompatibility complex-I): Presenting antigens for TCD4 and TCD8 cells. Higher scores indicate greater immunogenicity.

Protein stability is a basic property that influences protein function, activity, and regulation. Free energy related to protein unfolding is a key index of protein stability. Therefore, by analyzing the effect of a mutation on free energy, it is possible to determine precisely its effect on the stability of the protein. Our findings show that the mutations have different behaviors in stability and flexibility. For instance, P681 increases stability and decreases the flexibility of the spike protein, but A570D mutation has a reverse manner, and also NP mutations have a different behavior, which was consistent with the findings of Socher et al. (47).

The presence of interfacial residue, intermolecular contacts such as hydrogen bonding, salt bridge, and non-hydrogen bonded interactions might be the reason for its higher binding affinity. The four mutations in NP that exhibited the highest values for ΔΔG and ΔΔSVibENCoM identified in our study are S194L, D3L, R203K, and R185C. C-terminal domain (CTD) mutation T362I is essential for RNA binding.

Mutations in the furin cleavage site (FCS) such as P681H mutation are important in the fusion. In this mutation, proline was replaced with histidine, and distance interaction between wild type and mutant differed. Additionally, this mutation could potentially confer the replication advantage through increased cleavage efficacy by furin and adaptation to resist the acquired immunity. Our findings are consistent with other studies (48, 49). The

IEDB server predicted T362I (Threonine replaced with Isoleucine) in N protein to be a probable virus evading from the immune system. These mutations in the C-terminal domain could probably affect its interaction with RNA, which might be translated into viral RNA packaging and stability.

## 5- CONCLUSION

Mutations in S and N genes revealed in this study can interfere with various aspects of protein functions, including their stability, immunogenicity, and pathogenicity. Our identified mutations suggest that the virus is attempting to optimize its replication inside the host cells by tuning both flexibility, binding, and also escaping from the immune system. In addition, our findings could be useful for future works on monoclonal antibodies, miRNA design, or vaccines targeting S and N protein of SARS-CoV-2, as well as the diagnostic tools. Due to the sample size limitation in this study and various mutations in different variants, other studies using whole-genome sequencing and larger sample sizes for tracking the virus's behavior are needed. We hope that, with more effort and cooperation, support from the scientific community, and information sharing, the overcoming of COVID-19 will come soon.

## 6- DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

## 7- ETHICAL CONSIDERATIONS

This study was approved by the Ethics Committee of the Golestan University of Medical Sciences (IR.GOUMS.REC.1398.386)

## 8- AUTHOR CONTRIBUTIONS

AM and BA conceptualized and designed the study, drafted the initial manuscript, and reviewed and revised the manuscript. AT designed the data collection instruments, coordinated and supervised data collection, and critically reviewed the manuscript. ZB and MK collected data and carried out the initial analyses. HRN reviewed and revised the manuscript. All authors approved the final manuscript as submitted and agreed to be accountable for all aspects of the work.

## 9- ACKNOWLEDGMENTS

The authors would like to thank Golestan University of Medical Sciences and the SARS-CoV-2 laboratory, and the nursing team that is fighting against the illness. This manuscript was retrieved from a part of the Ph.D. dissertation of the corresponding author.

## 10- CONFLICTS OF INTEREST

None

## 11- FUNDING

No specific funding was obtained for this work.

## 12- REFERENCES

1. Hardenbrook NJ, Zhang P. A structural view of the SARS-CoV-2 virus and its assembly. Current Opinion in Virology. 2022; 52:123-34.

2. Rochman ND, Wolf YI, Faure G, Mutz P, Zhang F, Koonin EV. Ongoing global and regional adaptive evolution of SARS-CoV-2. Proceedings of the National Academy of Sciences. 2021; 118(29):e2104241118.

3. Li Q, Wu J, Nie J, Zhang L, Hao H, Liu S, et al. The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. Cell. 2020; 182(5):1284-94.e9.

4. Nguyen TT, Pathirana PN, Nguyen T, Nguyen QVH, Bhatti A, Nguyen DC, et al.

Genomic mutations and changes in protein secondary structure and solvent accessibility of SARS-CoV-2 (COVID-19 virus). Scientific Reports. 2021; 11(1):3487.

5. Hirabara SM, Serdan TDA, Gorjao R, Masi LN, Pithon-Curi TC, Covas DT, et al. SARS-COV-2 Variants: Differences and Potential of Immune Evasion. Frontiers in cellular and infection microbiology. 2021; 11:781429.

6. Altmann DM, Boyton RJ, Beale R. Immunity to SARS-CoV-2 variants of concern. Science (New York, NY). 2021; 371(6534):1103-4.

7. Mistry P, Barmania F, Mellet J, Peta K, Strydom A, Viljoen IM, et al. SARS-CoV-2 Variants, Vaccines, and Host Immunity. Frontiers in immunology. 2021; 12:809244.

8. Focosi D, Maggi F. Neutralizing antibody escape of SARS-CoV-2 spike protein: Risk assessment for antibody-based Covid-19 therapeutics and vaccines. Reviews in medical virology. 2021; 31(6):e2231.

9. Tomaszewski T, DeVries RS, Dong M, Bhatia G, Norsworthy MD, Zheng X, et al. New Pathways of Mutational Change in SARS-CoV-2 Proteomes Involve Regions of Intrinsic Disorder Important for Virus Replication and Release. bioRxiv. 2020:2020.07.31.231472.

10. Woo J, Lee EY, Lee M, Kim T, Cho YE. An in vivo cell-based assay for investigating the specific interaction between the SARS-CoV N-protein and its viral RNA packaging sequence. Biochemical and biophysical research communications. 2019; 520(3):499-506.

11. Oliveira SC, de Magalhães MTQ, Homan EJ. Immunoinformatics Analysis of SARS-CoV-2 Nucleocapsid Protein and Identification of COVID-19 Vaccine Targets. Frontiers in immunology. 2020; 11:587615.

12. Mohammad T, Choudhury A, Habib I, Asrani P, Mathur Y, Umair M, et al. Genomic Variations in the Structural Proteins of SARS-CoV-2 and Their Deleterious Impact on Pathogenesis: A Comparative Genomics Approach. Front Cell Infect Microbiol. 2021; 11:765039.

13. Rodriguez-Rivas J, Croce G, Muscat M, Weight M. Epistatic models predict mutable sites in SARS-CoV-2 proteins and epitopes. Proceedings of the National Academy of Sciences. 2022; 119(4):e2113118119.

14. Das JK, Thakuri B, MohanKumar K, Roy S, Sljoka A, Sun G-Q, et al. Mutation-Induced Long-Range Allosteric Interactions in the Spike Protein Determine the Infectivity of SARS-CoV-2 Emerging Variants. ACS Omega. 2021; 6(46):31305-20.

15. Fung TS, Liu DX. Post-translational modifications of coronavirus proteins: roles and function. Future virology. 2018; 13(6):405-30.

16. Mohammad T, Choudhury A, Habib I, Asrani P, Mathur Y, Umair M, et al. Genomic Variations in the Structural Proteins of SARS-CoV-2 and Their Deleterious Impact on Pathogenesis: A Comparative Genomics Approach. Front Cell Infect Microbiol. 2021; 11:765039-.

17. Al-Zyoud W, Haddad H. Dynamic's prediction of emerging notable spike protein mutations in SARS-CoV-2 implies a need for updated vaccines. Biochimie. 2021; 191:91-103.

18. Chambers JP, Yu J, Valdes JJ, Arulanandam BP. SARS-CoV-2, Early Entry Events. J Pathog. 2020; 2020:9238696-.

19. Giron CC, Laaksonen A, Barroso da Silva FL. Up State of the SARS-COV-2 Spike Homotrimer Favors an Increased Virulence for New Variants. Front Med Technol. 2021; 3:694347-.

20. Mori T, Jung J, Kobayashi C, Dokainish HM, Re S, Sugita Y. Elucidation of interactions regulating conformational stability and dynamics of SARS-CoV-2 S-protein. Biophysical Journal. 2021; 120(6):1060-71.

21. Quaglia F, Salladini E, Carraro M, Minervini G, Tosatto SCE, Le Mercier P. SARS-CoV-2 variants preferentially emerge at intrinsically disordered protein sites helping immune evasion. The FEBS journal. 2022.

22. Tenchov R, Zhou QA. Intrinsically Disordered Proteins: Perspective on COVID-19 Infection and Drug Discovery. ACS Infectious Diseases. 2022; 8(3):422-32.

23. Bai Z, Cao Y, Liu W, Li J. The SARS-CoV-2 Nucleocapsid Protein and Its Role in Viral Structure, Biological Functions, and a Potential Target for Drug or Vaccine Mitigation. Viruses. 2021; 13(6):1115.

24. Azad GK. Identification and molecular characterization of mutations in nucleocapsid phosphoprotein of SARS-CoV-2. PeerJ. 2021; 9:e10666.

25. Zhao X, Nicholls JM, Chen YG. Severe acute respiratory syndrome-associated coronavirus nucleocapsid protein interacts with Smad3 and modulates transforming growth factor-beta signaling. The Journal of biological chemistry. 2008; 283(6):3272-80.

26. Ferreira-Gomes M, Kruglov A, Durek P, Heinrich F, Tizian C, Heinz GA, et al. SARS-CoV-2 in severe COVID-19 induces a TGF-β-dominated chronic immune response that does not target itself. Nature Communications. 2021; 12(1):1961.

27. Mu J, Fang Y, Yang Q, Shu T, Wang A, Huang M, et al. SARS-CoV-2 N protein antagonizes type I interferon signaling by suppressing phosphorylation and nuclear translocation of STAT1 and STAT2. Cell Discov. 2020; 6:65-.

28. Chen K, Xiao F, Hu D, Ge W, Tian M, Wang W, et al. SARS-CoV-2 Nucleocapsid Protein Interacts with RIG-I and Represses RIG-Mediated IFN-β Production. Viruses. 2020; 13(1):47.

29. Rodrigues CH, Pires DE, Ascher DB. DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. Nucleic acids research. 2018; 46(W1):W350-w5.

30. Worth CL, Preissner R, Blundell TL. SDM--a server for predicting effects of mutations on protein stability and malfunction. Nucleic acids research. 2011; 39(Web Server issue):W215-W22.

31. Chen C-W, Lin M-H, Liao C-C, Chang H-P, Chu Y-W. iStable 2.0: Predicting protein thermal stability changes by integrating various characteristic modules. Comput Struct Biotechnol J. 2020; 18:622-30.

32. Seifi M, Walter MA. Accurate prediction of functional, structural, and stability changes in PITX2 mutations using in silico bioinformatics algorithms. PLOS ONE. 2018; 13(4):e0195971.

33. Laimer J, Hofer H, Fritz M, Wegenkittl S, Lackner P. MAESTRO - multi agent stability prediction upon point mutations. BMC Bioinformatics. 2015; 16(1):116.

34. Kim Y, Ponomarenko J, Zhu Z, Tamang D, Wang P, Greenbaum J, et al. Immune epitope database analysis resource. Nucleic acids research. 2012; 40(Web Server issue):W525-30.

35. Christoffer C, Bharadwaj V, Luu R, Kihara D. LZerD Protein-Protein Docking Webserver Enhanced With de novo Structure Prediction. Frontiers in molecular biosciences. 2021; 8:724947.

36. 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.

37. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, et al. The ClusPro web server for protein-protein docking. Nature protocols. 2017; 12(2):255-78.

38. Land H, Humble MS. YASARA: A Tool to Obtain Structural Guidance in Biocatalytic Investigations. Methods in molecular biology (Clifton, NJ). 2018; 1685:43-67.

39. Xue LC, Rodrigues JP, Kastritis PL, Bonvin AM, Vangone A. PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. Bioinformatics (Oxford, England). 2016; 32(23):3676-8.

40. Ye Q, West AMV, Silletti S, Corbett KD. Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. Protein science: a publication of the Protein Society. 2020; 29(9):1890-901.

41. Johnson BA, Zhou Y, Lokugamage KG, Vu MN, Bopp N, Crocquet-Valdes PA, et al. Nucleocapsid mutations in SARS-CoV-2 augment replication and pathogenesis. bioRxiv. 2021:2021.10.14.464390.

42. Hernandez-Alias X, Schaefer MH, Serrano L. Translational adaptation of human viruses to the tissues they infect. bioRxiv. 2020:2020.04.06.027557.

43. Chen C, Boorla VS, Banerjee D, Chowdhury R, Cavener VS, Nissly RH, et al. Computational prediction of the effect of amino acid changes on the binding affinity between SARS-CoV-2 spike RBD and human ACE2. Proceedings of the National Academy of Sciences. 2021; 118(42):e2106480118.

44. Wu H, Xing N, Meng K, Fu B, Xue W, Dong P, et al. Nucleocapsid mutations R203K/G204R increase the infectivity, fitness, and virulence of SARS-CoV-2. Cell Host Microbe. 2021; 29(12):1788-801.e6.

45. Davies NG, Jarvis CI, van Zandvoort K, Clifford S, Sun FY, Funk S, et al. Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. Nature. 2021; 593(7858):270-4.

46. 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.

47. Socher E, Conrad M, Heger L, Paulsen F, Sticht H, Zunke F, et al. Mutations in the B.1.1.7 SARS-CoV-2 Spike Protein Reduce Receptor-Binding Affinity and Induce a Flexible Link to the Fusion Peptide. Biomedicines. 2021; 9(5).

48. Lubinski B, Fernandes MHV, Frazier L, Tang T, Daniel S, Diel DG, et al. Functional evaluation of the P681H mutation on the proteolytic activation of the SARS-CoV-2 variant B.1.1.7 (Alpha) spike. bioRxiv: the preprint server for biology. 2021:2021.04.06.438731.

49. Mohammad A, Abubaker J, Al-Mulla F. Structural modeling of SARS-CoV-2 alpha variant (B.1.1.7) suggests enhanced furin binding and infectivity. Virus Res. 2021; 303:198522-.