

**ORIGINAL ARTICLE****Functional and Structural Characterization of SARS-Cov-2 Spike Protein: An *In Silico* Study****Hadi Sedigh Ebrahim-Saraie<sup>1</sup>, Behzad Dehghani<sup>2</sup>, Ali Mojtabehi<sup>3</sup>, Mohammad Shenagari<sup>3</sup>, Meysam Hasannejad-Bibalan<sup>3\*</sup>****OPEN ACCESS**

**Citation:** Hadi Sedigh Ebrahim-Saraie, Behzad Dehghani, Ali Mojtabehi, Mohammad Shenagari, Meysam Hasannejad-Bibalan. Functional and Structural Characterization of SARS-Cov-2 Spike Protein: An *In Silico* Study. *Ethiop J Health Sci.* 2021;31 (2):213. doi:<http://dx.doi.org/10.4314/ejhs.v31i2.2>

**Received:** October 28, 2020

**Accepted:** December 06, 2020

**Published:** March 1, 2021

**Copyright:** © 2021 Hadi S.E, et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This study was supported by Guilan University of Medical Sciences, Grant No. 2013.

**Competing Interests:** The authors declare that this manuscript was approved by all authors in its form and that no competing interest exists.

**Affiliation and Correspondence:**

<sup>1</sup>Razi Clinical Research Development Unit, Razi Hospital, Guilan University of Medical Sciences, Rasht, Iran

<sup>2</sup>Shiraz HIV/AIDS Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

<sup>3</sup>Department of Microbiology, School of Medicine, Guilan University of Medical Sciences, Rasht, Iran

\*Email: [meysam.hasannejad@gmail.com](mailto:meysam.hasannejad@gmail.com), [hasannejad@gums.ac.ir](mailto:hasannejad@gums.ac.ir)

**ABSTRACT**

**BACKGROUND:** Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the cause of the global outbreak of coronavirus disease 2019 (Covid-19), which has been considered as a pandemic by WHO. SARS-CoV-2 encodes four major structural proteins, among which spike protein has always been a main target for new vaccine studies. This *in silico* study aimed to investigate some physicochemical, functional, immunological, and structural features of spike protein using several bioinformatics tools.

**METHOD:** We retrieved all SARS-CoV-2 spike protein sequences from different countries registered in NCBI GenBank. CLC Sequence Viewer was employed to translate and align the sequences, and several programs were utilized to predict B-cell epitopes. Modification sites such as phosphorylation, glycosylation, and disulfide bonds were defined. Secondary and tertiary structures of all sequences were further computed.

**RESULTS:** Some mutations were determined, where only one (D614G) had a high prevalence. The mutations did not impact the B-cell and physicochemical properties of the spike protein. Seven disulfide bonds were specified and also predicted in several N-link glycosylation and phosphorylation sites. The results also indicated that spike protein is a non-allergen.

**CONCLUSION:** In summary, our findings provided a deep understanding of spike protein, which can be valuable for future studies on SARS-CoV-2 infections and design of new vaccines.

**KEYWORDS:** SARS-CoV-2, Spike, Postmodification, Mutation, Bioinformatics

**INTRODUCTION**

Coronaviridae is a family of enveloped, positive-sense single-stranded RNA viruses (ssRNA+) comprising coronaviruses for birds, bafiniviruses for fishes, and corona- and toroviruses for mammals (1). At the end of 2019, a series of pneumonia cases were reported from the Hubei Province of China with clinical presentations significantly resembling viral pneumonia (2). The etiology of infections was identified and confirmed as a novel coronavirus (2019-nCoV) belonging to  $\beta$ -coronavirus genera (2).

The resulting virus and disease are currently called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease 2019 (COVID-19), respectively (3).

There is limited information concerning the pathogenesis of COVID-19, and evidence has shown that the main mechanism is similar to SARS-CoV and MERS-CoV (4). The spike (S) protein of coronaviruses mediates viral entry into target cells. This entry is due to the binding of the surface unit (S1) of the S protein to a cellular receptor, known as angiotensin-converting enzyme 2 (ACE2). SARS-S and SARS-2-S share a high amino acid homology (>70%) (5). The interaction between viral proteins and cell membrane receptors is a critical step in the virus pathogenesis (6). The virus probably pass through major passages of the upper respiratory tract, especially nasal and larynx mucosa (7). The main target of virus entrance is lungs through the respiratory tract, but virus would also attack and enters other organs that express the type 2 transmembrane serine protease (TMPRSS2) and ACE2 receptor protein. The consequential of infection in host cells causes an excess release of pro-inflammatory cytokines that causes a cytokine storm (8).

COVID-19 patients exhibit various symptoms that cannot be easily distinguished from other respiratory diseases. Based on the severity of symptoms, this disease is classified into mild, moderate, severe, and critical (9). These symptoms, which may appear within a week after exposure to the virus, mainly include fever, cough, shortness of breath, chills, headache, muscle pain, and loss of taste or smell (10). The main reported complications associated with COVID-19 were pneumonia, heart injury, liver and kidney failure, and superinfections (11). Recent estimates showed that approximately half of died people with COVID-19 had a underlying diseases, where hypertension (46%) had the highest occurrence followed by diabetes (26%), cardiovascular

disease (21%), malignancy (11%), chronic obstructive pulmonary disease (COPD) (8%), kidney disease (7%), and liver disease (3%)(12).

To date, there exists no specific antiviral treatment recommended for COVID-19, and no vaccine is currently available (13). The current appropriate treatments include oxygen therapy (which is the major intervention), administration of antibiotics to prevent bacterial co-infections, fluid management, and supportive use of traditional medicine (14,15). Other carried out strategies were using antivirals (Lopinavir, Ritonavir, Ribavirin, Favipiravir (T-705), Remdesivir, Oseltamivir, Chloroquine, and Interferon), and convalescent plasma (16). However, still the treatment effectiveness is greatly varied, so future studies on SARS-CoV-2 genome organization can help design and develop effective antiviral drugs or inhibition approaches.

Over the past decades, bioinformatics has emerged as a powerful tool for analyzing bacterial and viral genomes, predicting the structure and function of proteins, and designing new vaccines (17,18). Due to the global health emergency declared for COVID-19 and the importance of any effort to control the outbreak, the present *in silico* study aimed to investigate some physicochemical, functional, immunological, and structural features of spike protein using several bioinformatics tools.

## MATERIALS AND METHODS

### Sequence alignment and phylogenetic tree:

All 52 SARS-CoV-2 spike protein sequences from different countries registered in NCBI GenBank (<http://www.ncbi.nlm.nih.gov/>) were retrieved from March to June 2020. The CLC Sequence Viewer Version Beta (Qiagen) was employed to analyze and detect the mutations in sequences. Phylogenetic tree was analyzed by UPGMA method (Bootstrap: 1000). The accession numbers of all sequences are displayed in Table 1.

Table 1: The accession numbers of all 52 sequences that were used in this study

Reference	NC_045512
China	MT281577, MT291829, MT291828, MT291826, MT291827, MT291830, MT291833, MT291832, MT291831, MT039874
USA	MT350236, MT350247, MT350269, MT350244, MT350237, MT350238, MT350252, MT350253, MT350254, MT350255
Iran	MT320891, MT281530
Japan	LC534418, LC534419, LC529905, LC528232, LC528233
India	MT050493, MT012098
Brazil	MT350282, MT126808
Nepal	MT072688
Italy	MT077125, MT066156
South Korea	MT304474, MT304475, MT304476, MT039890
Spain	MT292569, MT292570, MT292571, MT292572, MT292573, MT292574, MT292575, MT292576, MT292577, MT233523
Turkey	MT327745
South Africa	MT324062
Australia	MT007544

**Physicochemical properties:** ExPASy's ProtParam (<http://expasy.org/tools/protparam.html>) was used to predict all the properties of spike protein, including theoretical isoelectric point (pI), extinction coefficient, instability index, molecular weight, aliphatic index, and grand average hydropathy (GRAVY)(19).

**Post-modification changes:** Serine, threonine, and tyrosine phosphorylation sites were predicted by DISPHOS (<http://www.dabi.temple.edu/disphos/>) and NetPhos (<http://www.cbs.dtu.dk/services/NetPhos/>), glycosylation sites were predicted by NetNGlyc ([www.cbs.dtu.dk/services/NetNGlyc/](http://www.cbs.dtu.dk/services/NetNGlyc/)) and Nglyde (<http://bioapp.iis.sinica.edu.tw/Nglyde/run.php>), and disulfide bonds were predicted using Dianna (<http://clavius.bc.edu/~clotelab/DiANNA/>) and Scratch (<http://scratch.proteomics.ics.uci.edu/>) for spike protein.

**Immunoinformatics:** Chou and Fasman, Karplus and Schulz, Kolaskar and Tongaonkar,

Emini, Parker, and BepiPred (<http://www.iedb.org/>) methods were applied to predict the position of B-cell epitopes. Hydrophilicity, flexibility/mobility, accessibility, polarity, exposed surfaces, and turns features were determined by BcePred ([crdd.osdd.net/raghava/bcepred/](http://crdd.osdd.net/raghava/bcepred/)). ABCpred (<http://crdd.osdd.net/raghava/abcpred/>) software specified 16 meric B-cell epitope. Allergenic properties were estimated using AlgPred (<http://crdd.osdd.net/raghava/algpred/>) and VaxiJen (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen>) software which computed protective antigens and predicted subunit vaccines.

**Secondary and tertiary structures:** SOPMA software ([https://npsa-prabi.ibcp.fr/NPSA/npsa\\_sopma.html](https://npsa-prabi.ibcp.fr/NPSA/npsa_sopma.html)) and Phyre (<http://www.sbg.bio.ic.ac.uk/~phyre2/>) server were applied to calculate and confirm the secondary structure, respectively. I-TASSER (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>) was utilized to predict the tertiary structure, and the suggested models were then

refined by 3Drefine (<http://sysbio.mnet.missouri.edu/3Drefine/>). Finally, the refined models were assessed for stereochemistry, reliability, and quality by “Qmean” (<https://swissmodel.expasy.org/qmean/>), “ProSA-web” (<https://prosa.services.came.sbg.ac.at/prosa.ph>), “ERRAT” (<https://servicesn.mbi.ucla.edu/ERRAT/>), and “Rampage” (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>). The study design was approved by regional Ethics Committee of Guilan University of Medical Sciences (IR.GUMS.REC.1399.001)

## RESULTS

**Amino acid changes:** Analysis showed that spike protein was a highly conserved protein, and only one high frequent mutation (D614G) was detected in comparison with the reference sequence. Table 2 summarizes all mutations established in spike protein. In addition, the phylogenetic tree results are illustrated in Figure 1. The phylogenetic analysis showed two main clusters, the upper one containing eight sequences from Spain, USA, and South Africa, and the second one including other sequences and reference sequences. Two sequences from Iran and a sequence from USA were very close to the reference sequence. Interestingly, almost all sequences from South East Asia (China, Japan, South Korea) were close to each other, and the majority of USA sequences were almost located in upper cluster.

Table 2: The list of substitutions were found in the present study and the high prevalent mutation is bolded

Mutations	Frequency No. (%)
S50L	1(1.9%)
N74K	1(1.9%)
S221w	1(1.9%)
S247R	1(1.9%)
R408I	1(1.9%)
<b>D614G</b>	<b>13(25%)</b>
V772I	1(1.9%)
A930V	1(1.9%)
V1065L	2(3.8%)

**ProtParam analysis:** ProtParam analysis indicated that spike protein is an acidic peptide due to the high percentage of its acidic amino acids (Theoretical pI: 6.2). The instability index, an estimate of the stability of a protein in a test tube, was 33.01 and showed that the spike was a stable peptide. Aliphatic index, a positive factor for the increased thermostability of proteins, this factor was 84.67 which revealed that this peptide was a thermostable one. GRAVY is a hydrophathy index which augmented with the increase in the positive score. Thus, the peptide was also a hydrophilic one (-0.079).

**Postmodification and disulfide bond results:** Table 3 shows the postmodification and disulfide sites prediction; based on our results, the spike was highly phosphorylated, and four conserved positions were further suggested. Glycosylation prediction by two online software showed seven positions (61,74, 234, 282, 616, 709, and 1195); and results showed the prediction of possible disulfide bonds by Dianna and Scratch, which determined several cysteines.

Table 3: Postmodification and disulfide bond prediction results using several software

Software	NETPHOS	DISPHOS	Combined results
<b>Phosphorylation sites</b>	46, 50, 76, 95, 151, 170, 250, 255, 302, 313, 359, 376, 415, 459, 469, 477, 523, 555, 572, 632, 637, 659, 680, 686, 730, 735, 813, 816, 937, 939, 1003, 1105, 1147, 1155, 1196	250, 423, 572, 612, 680, 686, 811, 814, 817, 1048, 1262	250, 572, 680, 686
Software	NetNGlyc 1.0 Server	Nglyde	Combined results
<b>Glycosylation Sites</b>	61,74,234,282,616,709,717,1159,1195	61,74,122,149,234,282,343,603,616,657,709,1195	61,74,234,282,616,709,1195
Software	Dianna	Scratch	Combined results
<b>Disulfide Bonds</b>	15 – 1241, 131 – 391, 136 – 662, 166 – 1237, 291 – 671, 301 – 336, 361 – 488, 379 – 743, 432 – 1236, 480 – 1249, 525 – 1248, 538 – 1044, 590 – 617, 649 – 1242, 738 – 1244, 749 – 1127, 760 – 1255, 841 – 1033, 852 – 1251, 1083 - 1254	291-301, 1033-1044, 480-488, 379-391, 841-852, 1083-1127, 15-1255, 525-538, 336-361, 662-671, 136-166, 131-1254, 590-617, 738-749, 1236-1248	15, 131, 136, 166, 291, 301, 336, 361,379, 391, 480, 488, 525, 538,590, 617, 662, 671, 738, 749, 841, 852, 1033, 1044, 1083, 1127, 1254, 1255, 1236, 1248

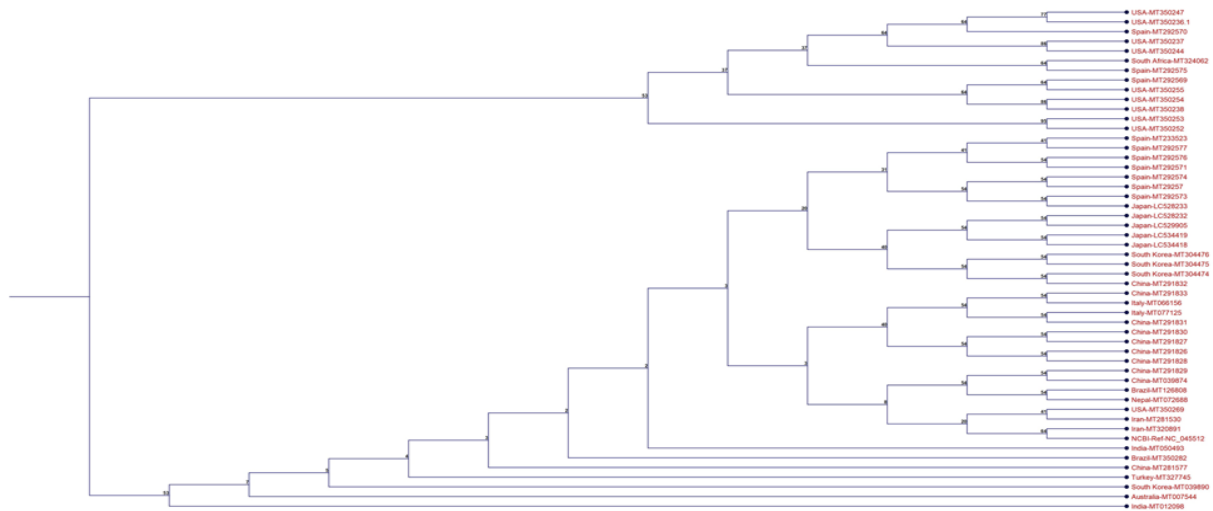
**Secondary and tertiary structure prediction:**

The secondary structure results using SOPMA showed that random coil was the major structure with 43.9% and after that Alpha helix, extended strand and Beta turn by 29.3%, 23.3% and 3.5%

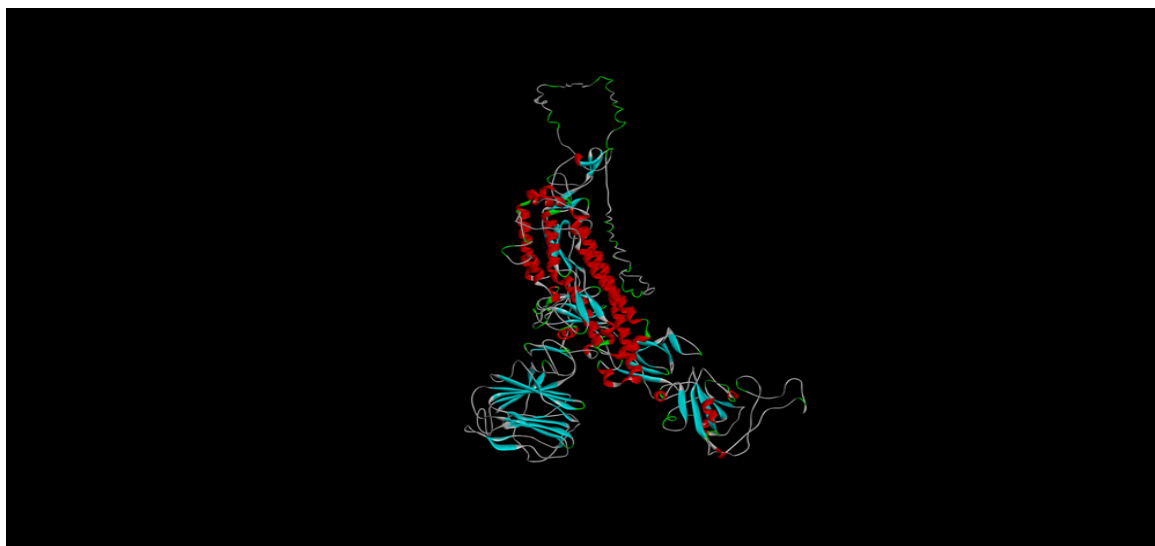
respectively. Table 4 presents the qualification results of the refined models suggested by 3D-refine. Figure 2 illustrates the tertiary structure of spike protein.

Table 4: The final results of 4 servers used to define the best tertiary structure for spike protein; selected structure in bold writing

Server	Models	Qmean	ERRAT	ProSA-web	Rampage (Ramachandran plot)
I-TASSER	without refinement	-10.01	71.371	not calculated	956 (75.2%)
3D-Refine	1	-6.61	75.4772	-10.8	1010 (79.5%)
	2	-6.03	74.7076	-11.07	1023 (80.5%)
	3	-5.75	63.1799	-11.15	1039 (81.7%)
	4	-5.75	60.6695	-11.17	1045 (82.2%)
	<b>5</b>	<b>-5.62</b>	<b>61.5063</b>	<b>-11.27</b>	<b>1047 (82.4%)</b>



**Figure 1:** Phylogenetic tree of Spike protein sequences using neighbor joining method. The phylogenetic tree was constructed by the UPGMA method. The numbers at the forks show the numbers of occurrences of the repetitive groups to the right out of 1000 bootstrap samples.



**Figure 2:** The final 3D structure of spike protein constructed by I-TASSER and refined by 4 online tools



**Immunoinformatics results:** The combination of all predictions provided by four software specified four B-cell epitopes (249-259, 674-687, 807-816, and 1254-1265). The VaxiJen score was 0.4646, indicating that this protein is a probable antigen. AlgPred results suggested that spike protein is a non-allergen protein.

## DISCUSSION

The results of the present study showed that the spike protein was highly conserved, and high prevalence mutation was detected only in one site (D614G). The mentioned mutation was observed amongst 90% of the sequences from USA, which was the highest rate among all countries. Analysis of other features of spike protein revealed that this mutation did not have any effect on post-modification sites, B-cell epitopes, and psychochemical properties. Banerjee et al. defined several mutations in the spike protein sequences from the USA, South America, China, and European countries (Banerjee et al. 2020) (20). Similar to our findings, substitution in amino acid 614 (D614G) was the most prevalent mutation (25%).

Previous studies suggested a region, KRSEIEDLLFNKV, as a potential Achilles' heel for controlling the life cycle of SARS-CoV-2. This site is exposed and this region is required for proteolytic activation cleavage (21,22). In addition, it is a well-conserved region located on the surface of the virus. Similar to previous investigations, our findings showed that KRSEIEDLLFNKV was completely conserved among all selected sequences from all regions. Interestingly, prediction of post-modification sites revealed that this region was phosphorylated. It was further predicted as a B-cell epitope, confirming its importance as a possible candidate for designing new vaccines.

Spike proteins contain a receptor binding domain (RBD) positioned between amino acids 331 and 524. Mutations in this region may critically impact virus entry and attachment to ACE2 receptor (23). In one sequence, we detected a substitution in this region, indicating that this domain is highly conserved and could be a new target for inhibiting virus attachment.

Contrary to our predictions, Banerjee et al. specified four mutations (348, 476, 483, and 520) with very low prevalence (20). The difference between the two studies regarding the number of the mutations might be ascribed to the different sets of sequences and study methods.

Korber et al. focused on D614G substitution as an urgent concern, proposing that this mutation began spreading in Europe in early February 2020 (24). Although they were not able to define the origin of this mutation, there existed certain hypotheses as to its Chinese or European origin. The potential impacts suggested for this mutation are increased viral transmission, infected spike, enhanced receptor binding, and ADE (antibody-dependent enhancement) antibody elicitation (24). In agreement with Korber's study, our results indicated the spread of D614G substitution. Moreover, almost all sequences from North America (USA) and three sequences from Europe (Spain) harbored this mutation. Interestingly, this mutation was not detected in the sequences from China and South East Asia (Japan and South Korea).

Our analysis described spike protein as acidic, thermostable, and hydrophilic. However, because spike requires some post-modification processes, it seems yeast, and mammalian cells can better express this protein. Similar to our ProtParam prediction, Walls et al and Ou et al used different cell lines to express spike protein, which showed its stability in mammalian cells (25,26). Likewise, Zhang et al. expressed spike protein in *Escherichia coli*; they confirmed that *E. coli* was an appropriate host for the expression of spike (27).

Phosphorylation prediction showed four completely conserved sites among the selected sequences. Previous studies suggested some functions for protein phosphorylation in coronavirus. Petit et al proposed that phosphorylation is vital in the retention of spike protein at cell surfaces (28). Furthermore, Davidson et al stated that the phosphorylation sites on the spike glycoprotein might be necessary for assembling the trimer (29). Therefore, it can be concluded that blocking the

phosphorylation process could be an effective approach to disturb the spike protein function.

Fung et al. defined the vital role of glycosylation in antigenicity, fusogenic, and immunomodulatory activities of the spike protein (30). Glycosylation prediction by NetNGlyc and Nglyde determined seven positions. Of these, except in the position from a Brazil sequence (74), which showed a substitution, other sites were highly conserved and seemingly highly vital to spike protein function. Shajahan et al. and Watanabe et al., using the high-resolution mass spectrometry, revealed 22 glycosylation sites for spike protein (31,32). Seven positions mentioned in our findings were similar to the foregoing studies. Similar to this study, Kumar et al. used bioinformatics tools to compare the glycosylation features of 2019-nCoV and SARS-CoV (33). In spite of different sequences used in the present study and Kumar's (Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1 2019-nCoV-MN908947.3), the final predicted positions were completely similar. This shows that these positions remained unchanged during the spread of the novel coronavirus. By considering the critical role of glycosylation in the pathogenesis of SARS-CoV-2, these positions might be considered as new targets for novel inhibitors.

It has been proposed that disulfide bonds are required for a proper folding and trimerization of coronavirus spike protein (30). Dianna and Scratch results showed numerous positions for disulfide bonds that were completely conserved in all analyzed sequences. Dianna uses a support vector machine (SVM) with degree 2 polynomial kernel for the spectrum representation, and Scratch works based on 2D recurrent neural network, support vector machine, graph matching, and regression algorithms. Both online software are well-known and were previously employed in numerous studies to define disulfide bonds. Ibrahim et al. made use of a combined molecular docking and structural bioinformatics; they detected 13 disulfide bonds in four distinct regions and suggested that these regions were involved in cell attachment (34). Despite the different sequences and methods used in

Ibrahim's study and the present one, all predicted bonds were similar.

Previous studies confirmed the humoral immune response against spike protein in infected patients (Tay et al. 2020; Temperton et al. 2005; Yuchun et al. 2004) (35-37). It was shown that neutralizing antibody responses to the spike protein began by week two and in most patients developed by week three. Immunoinformatics analysis of spike protein by certain online databases suggested four regions that confirmed the possible potential of this protein for inducing humeral immune system. Interestingly, no mutation was detected in these regions; hence, they could be proper regions for the production of new vaccines. In addition, estimating allergenic characteristics showed that spike protein could not provoke allergenic reactions in humans.

Ahmed et al. used bioinformatics approaches to define B-cell epitopes in different proteins of SARS-CoV-2 (38). They were able to define 23 B-cell epitopes for spike protein; our prediction, on the other hand, showed new regions, a difference possibly attributable to the different sequences used in both studies.

Moreover, through bioinformatics analysis and machine learning, Grifoni et al. and Fast et al. analyzed spike protein to define immunological properties. Compared with our prediction (39,40), the two sites (674-687 and 807-816) were similar.

As a major limitation of this study, information on the COVID-19 crisis is constantly changing, and day-by-day number of new sequences in online databases are updated, therefore our study may not present a comprehensive view of spike protein. However, as a preliminary study, our results provide an insight for further works.

In summary, the results of the present study provided a comprehensive understanding of spike protein which can be used for further studies. This protein is a highly capable epitope on the SARS-CoV-2 surface which included several features appropriate in a vaccine construct. Other features of spike protein could be employed to express this protein, and post-modification sites could be utilized as new targets for SARS-CoV-2 inhibitors. Meanwhile,



it is not easy to forecast any realistic scenario, but mutations in spike protein suggest potential impacts on the pathogenesis of the virus in near future.

## REFERENCES

- Payne S. Family Coronaviridae. *Viruses*. 2017;149.
- Wang L, Wang Y, Ye D, Liu Q. Review of the 2019 novel coronavirus (SARS-CoV-2) based on current evidence. *International Journal of Antimicrobial Agents*. 2020; 55(6):105948.
- Lai CC, Shih TP, Ko WC, Tang HJ, Hsueh PR. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *International Journal of Antimicrobial Agents*. 2020;55(3):105924.
- Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *The New England Journal of Medicine*. 2020;382:1708.
- Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*. 2020;181(2):271-80.e8.
- Maginnis MS. Virus-Receptor Interactions: The Key to Cellular Invasion. *Journal of Molecular Biology*. 2018;430(17):2590-611.
- Di Gennaro F, Pizzol D, Marotta C, Antunes M, Racalbuto V, Veronese N, et al. Coronavirus Diseases (COVID-19) Current Status and Future Perspectives: A Narrative Review. *International Journal of Environmental Research and Public Health*. 2020;17(8): 2690.
- Keni R, Alexander A, Nayak PG, Mudgal J, Nandakumar K. COVID-19: Emergence, Spread, Possible Treatments, and Global Burden. *Frontiers in Public Health*. 2020;8:216.
- Abbasi-Oshaghi E, Mirzaei F, Farahani F, Khodadadi I, Tayebinia H. Diagnosis and treatment of coronavirus disease 2019 (COVID-19): Laboratory, PCR, and chest CT imaging findings. *International Journal Of Surgery (London, England)*. 2020;79:143-53.
- Lee Y, Min P, Lee S, Kim SW. Prevalence and Duration of Acute Loss of Smell or Taste in COVID-19 Patients. *Journal of Korean Medical Science*. 2020;35(18):e174.
- Renu K, Prasanna PL, Valsala Gopalakrishnan A. Coronaviruses pathogenesis, comorbidities and multi-organ damage - A review. *Life Sciences*. 2020;255:117839.
- Javanmardi F, Keshavarzi A, Akbari A, Emami A, Pirbonyeh N. Prevalence of underlying diseases in died cases of COVID-19: A systematic review and meta-analysis. *PloS One*. 2020;15(10):e0241265.
- Zhang J, Xie B, Hashimoto K. Current status of potential therapeutic candidates for the COVID-19 crisis. *Brain, Behavior, and Immunity*. 2020;87:59-73.
- Nicola M, O'Neill N, Sohrabi C, Khan M, Agha M, Agha R. Evidence based management guideline for the COVID-19 pandemic - Review article. *International Journal Of Surgery (London, England)*. 2020;77(206-16).
- Mirzaie A, Halaji M, Dehkordi FS, Ranjbar R, Noorbazargan H. A narrative literature review on traditional medicine options for treatment of corona virus disease 2019 (COVID-19). *Complementary Therapies in Clinical Practice*. 2020;40(101214).
- Wu R, Wang L, Kuo HD, Shannar A, Peter R, Chou PJ, et al. An Update on Current Therapeutic Drugs Treating COVID-19. *Current Pharmacology Reports*. 2020;1-15.
- Tayebbeh H, Behzad D, Zahra M, Tahereh A, Zahra H, Javad M, et al. Association of Mutations in the NS5A-PKRBD Region and IFNL4 Genotypes with Hepatitis C Interferon Responsiveness and its Functional and Structural Analysis. *Current Proteomics*. 2020;17:1-12.
- Dehghani B, Hasanshahi Z, Hashempour T, Motamedifar M. The possible regions to design Human Papilloma Viruses vaccine in Iranian L1 protein. *Biologia*. 2020;75(5):749-59.
- Moattari A, Dehghani B, Khodadad N, Tavakoli F. In Silico Functional and Structural Characterization of H1N1 Influenza A Viruses Hemagglutinin, 2010-2013, Shiraz, Iran. *Acta Biotheoretica*. 2015;63(2):183-202.
- Banerjee AK, Begum F, Ray U. Mutation Hot Spots in Spike Protein of COVID-19. *Preprints*. 2020.doi: 10.20944/preprints202004.0281.v1
- Robson B. Computers and viral diseases. Preliminary bioinformatics studies on the design of a synthetic vaccine and a preventative peptidomimetic antagonist against the SARS-CoV-2 (2019-nCoV, COVID-19) coronavirus. *Computers in Biology and Medicine*. 2020;119(103670).
- Robson B. COVID-19 Coronavirus spike protein analysis for synthetic vaccines, a peptidomimetic antagonist, and therapeutic drugs, and analysis of

- a proposed achilles' heel conserved region to minimize probability of escape mutations and drug resistance. *Computers in Biology and Medicine*. 2020;103749.
23. He Y, Li J, Jiang S. A single amino acid substitution (R441A) in the receptor-binding domain of SARS coronavirus spike protein disrupts the antigenic structure and binding activity. *Biochemical and Biophysical Research Communications*. 2006;344(1):106.
  24. Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al. Spike mutation pipeline reveals the emergence of a more transmissible form of SARS-CoV-2. *BioRxiv*. 2020;2020.04.29.069054.
  25. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell*. 2020;181(2):281-92.e6.
  26. Dehghani B, Hashempour T, Hasanshahi Z. Interaction of human herpesvirus 8 viral interleukin-6 with human interleukin-6 receptor using in silico approach: the potential role in HHV-8 pathogenesis. *Current Proteomics*. 2020;17(2):107-16.
  27. Zhang P, Gao Q, Wang T, Ke Y, Mo F, Jia R, et al. Evaluation of recombinant nucleocapsid and spike proteins for serological diagnosis of novel coronavirus disease 2019 (COVID-19). *MedRxiv*. 2020;2020.03.17.20036954.
  28. Petit CM, Melancon JM, Chouljenko VN, Colgrove R, Farzan M, Knipe DM, et al. Genetic analysis of the SARS-coronavirus spike glycoprotein functional domains involved in cell-surface expression and cell-to-cell fusion. *Virology*. 2005;341(2):215-30.
  29. Davidson AD, Williamson MK, Lewis S, Shoemark D, Carroll MW, Heesom K, et al. Characterisation of the transcriptome and proteome of SARS-CoV-2 using direct RNA sequencing and tandem mass spectrometry reveals evidence for a cell passage induced in-frame deletion in the spike glycoprotein that removes the furin-like cleavage site. *BioRxiv*. 2020;2020.03.22.002204.
  30. Fung TS, Liu DX. Post-translational modifications of coronavirus proteins: roles and function. *Future Virology*. 2018;13(6):405-30.
  31. Shajahan A, Supekar NT, Gleinich AS, Azadi P. Deducing the N- and O- glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. *Glycobiology*. 2020;30(12):981-8
  32. Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M. Site-specific analysis of the SARS-CoV-2 glycan shield. *BioRxiv*. 2020;2020.03.26.010322.
  33. Kumar S, Maurya VK, Prasad AK, Bhatt MLB, Saxena SK. Structural, glycosylation and antigenic variation between 2019 novel coronavirus (2019-nCoV) and SARS coronavirus (SARS-CoV). *VirusDisease*. 2020;31(1):13-21.
  34. Ibrahim IM, Abdelmalek DH, Elshahat ME, Elfiky AA. COVID-19 spike-host cell receptor GRP78 binding site prediction. *The Journal of Infection*. 2020;80(5):554-62.
  35. Tay MZ, Poh CM, Rénia L, MacAry PA, Ng LFP. The trinity of COVID-19: immunity, inflammation and intervention. *Nature Reviews Immunology*. 2020; 20(6):363-74.
  36. Nie Y, Wang G, Shi X, Zhang H, Qiu Y, He Z, et al. Neutralizing antibodies in patients with severe acute respiratory syndrome-associated coronavirus infection. *The Journal of Infectious Diseases*. 2004;190(6):1119-26.
  37. Temperton NJ, Chan PK, Simmons G, Zambon MC, Tedder RS, Takeuchi Y, et al. Longitudinally profiling neutralizing antibody response to SARS coronavirus with pseudotypes. *Emerging Infectious Diseases*. 2005;11(3):411-6.
  38. Ahmed SF, Quadeer AA, McKay M. Preliminary Identification of Potential Vaccine Targets for the COVID-19 Coronavirus (SARS-CoV-2) Based on SARS-CoV Immunological Studies. *Viruses*. 2020;12(3):254.
  39. Grifoni A, Sidney J, Zhang Y, Scheuermann RH, Peters B, Sette A. A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. *Cell Host Microbe*. 2020;27(4):671-80.e2.
  40. Fast E, Chen B. Potential T-cell and B-cell Epitopes of 2019-nCoV. *BioRxiv*. 2020;2020.02.19.955484.