Original Article

The hidden facts behind the variation of the spike region in SARS-Cov2 and hypothetical binding receptors to ICAM-1

Paywast J. Jalal[1]*, Sazan Q. Maulud[2], Shad A. Mohammed[1]

[1]Biology Department, College of Science, University of Sulaimani, Sulaimani, Iraq. [2]Department of Biology, College of Education, Salahaddin University-Erbil, Erbil, Kurdistan Region, Iraq.

Correspondence: Paywast J. Jalal, Biology Department, College of Science, University of Sulaimani, Sulaimani, Iraq. paywast.jalal@univsul.edu.iq

ABSTRACT

Since the novel coronavirus appeared in January 2020, it has brought global attention to this new virus due to its high infectivity rate and increasing fatality rate to more than 7%. For this purpose, in this study, we aimed to address the evolution and the genetic variation of the spike region of the SARS-CoV2 extensively, besides addressing some hypothetical binding with two cell receptors (CCR3 and ICAM-1) using computational tools, as the spike glycoproteins mediate the viral entry. The current study performed successive computational work, starting with multiple sequence alignment and phylogenetic tree construction, followed by alignment prediction with some cell receptors. The alignment analysis based on the spike glycoprotein of SARS-CoV2 showed that this novel coronavirus is closer to the Bat and Pangolin coronavirus and quite different from other human coronaviruses. Interestingly, the sequence and structural alignments show that 23 amino acid (amino acid) residues are inserted in the S1 subunit of the five spike regions, most of which are in the receptor-binding domain (RBD). None of these proteins has been detected in any previously identified human coronaviruses, and some of them are not shown in the sequence of Bat and Pangolin coronaviruses. They could be considered unique for this novel coronavirus. Some of these proteins are predicted to bind to 1IAM in the ICAM-1 receptor with a higher degree of affinity than binding to CCR3. The protein docking heuristic algorithms analysis revealed that some of the inserted amino acids in the spike region have a high affinity for binding to the ICAM-1 receptor through hydrophilic bonds and Beta-sheet bonding. Despite the need for more analysis, this study revealed that the hypothetical binding between the novels inserted amino acid in the SARS-CoV2 Spike RBD might be the right candidate for developing antiviral treatment against COVID-19. Moreover, the genomic region encoding these inserts only in the novel coronavirus represents an ideal candidate for distinguishing SARS-CoV2 from other coronavirus family members.

Keywords: SARS-CoV2, CCR3, ICAM-1, RBD and COVID-19

Introduction

The outbreak of SARS-like pneumonia that was announced by the Chinese authority in late December 2019 and later denoted by the WHO (World Health Organization) as a COVID-19 (Coronavirus Infectious Disease) made an alert all over the world as the most transmissible and infectious disease (1, 2). The SARS-CoV2 (Severe Acute Respiratory Syndrome-Coronavirus-2) was considered a zoonotic disease which is thought to be transmitted from Bat (3-5) and Pangolin (intermediate host) (2). Subsequently, it was carried from the wet seafood market in the Wuhan province, China (6-8), and human-to-human transmission was confirmed in the early days of its spread (5, 9). The virus is highly transmissible from infected humans through aerosol or bodily contact and contaminated objects (9, 10). Severe infectivity and mortality are lower than other coronaviruses (11). However, it is divided into two categories: most patients were symptomless and others with mild or severe infection (12). The patients with mild disease had a high fever (above 38 ^{0}C), dry cough, malaise, fatigue, and difficulty breathing (13). However, in severe cases, the symptom was raised to an acute respiratory syndrome, and diarrhoea progressed to septic shock and coagulopathy (11). From its emergence until the writing of this article, the novel coronavirus has infected more than 3 million people in 210 countries, and from this number, a rate of 5% has been recorded in a critical situation and a death rate of 7% (14). These numbers have been continuously changing and increasing gradually daily. Today's lifestyle significantly impacts one's quality of life. Quality of life is influenced by elements such as proper eating, physical

exercise, access to healthcare, stress reduction, and poverty alleviation (15, 16).

SARS-CoV2 is an enveloped virus with a positive- single-stranded RNA genome belonging to a Betacoronavirus genera in the Coronaviridae family. It is the seventh member of this family that is known to infect humans, causing severe infection just like previously identified strains, for example, SARS (Severe Acute Respiratory Syndrome) and MERS (Middle East Respiratory Syndrome), which are both members of the same genera (17). In 2003 and 2012, a high fatality rate of 10% and 35% was recorded, respectively (18, 19). The whole genome of SARS-CoV2, approximately 30 Kb, consists of six major open reading frames (ORF) divided into structural and non-structural proteins (20). One of the most crucial proteins in the structural proteins is the spike protein, which consists of 1273 amino acids and is responsible for viral attachment and entry (4, 21). The coronavirus spike glycoprotein consists of S1 and S2 subunits, which bind to the receptor and facilitate fusion to the membrane, respectively (22, 23). This glycoprotein is a crucial part of tissue tropism and host range; therefore, it is considered a good target for developing a vaccine (23, 24). It has been reported that the efficient spreading of SARS-CoV2 is related to the receptor-binding domain in the spike region that binds to human ACE2 (25). Moreover, others suggested that some receptors also link to the Spike of SARS-CoV2, for example, cell-surface receptor GRP78 (Glucose Regulated Protein 78) (26).

The spike glycoprotein of SARS-CoV2 was modelled based on the protein data and compared with different coronaviruses in the Bat, Pangolin and SARS-CoV to detect the variation in this region and the underlying evolutionary changes. In the current project, molecular docking was performed using specific software to address the hypothetical binding with CCR3 (expressed on eosinophil, mast cells, and Th2 cells) and ICAM-1 [intercellular adhesion molecule-1] (expressed on endothelial cells and cells of the immune system) receptors. The sequence and structural comparisons predicted different regions binding in these receptors with a different ratio. The results suggest a promising recognition of the SARS-CoV2 spike glycoprotein with the cell receptors that could open a new antiviral development concept.

Materials and Methods

Multiple sequence alignment and phylogenetic analysis

The NCBI database (www.ncbi.nlm.nih.gov) was the source for retrieving the human and animal coronavirus sequence. The full-length sequences of the novel coronavirus (n=178) were retrieved from the 11th of January 2020 until the 4th of April 2020. Moreover, we retrieved the full-length of other coronaviruses from both humans and different animals (n=28) published in the NCBI gene bank. From the full-length sequences of the SARS-CoV2, we only picked those from each country with a high mortality rate, and we excluded the sequences that were 100% similarity detected among their sequences. Finally, the start codon of each sequence spike was obtained from NCBI graphics of the following selected sequence, Human SARS-CoV2/2020 [MN985325.1 (USA-WA), MN908947.3 (Wuhan-Hu), MN994467.1 (USA-CA), MT066156.1 (ITA-INMI1), MT126808.1 (BRA-SP02), MT093571.1 (SWE), MT233523.1 (ESP- Valencia), MT007544.1 (Australia-VIC01) and LC528232.1 (Kng-Hu)]; Human SARS coronavirus [AY278741.1 (Urbani/2016), AY463059.1 (ShanghaiQXC1) and AY291315.1 (Frankfurt)]; Other human coronavirus [KY684760.1 (229E-USA/2017), KF514433.1 (229E-USA/1993), KP209312.1 (MERS-UAE/2013), KF53.0112.1 (NL63-USA/2001), DQ415914.1 (HKU1) and KF530092.1 (OC43-USA/2000); Bat coronavirus [MN996532.1 (RaTG13/2020), KY417146.1 (RS4231/2017) and MK211376.1 (YN2018B/2019); Pangolin coronavirus [MT040333.1 (GX-P4L)]; and finally Camel alphacoronavirus [MF593473.1 (Abu Dhabi)] and Alpaca coronavirus [JQ410000.1 (CA08-2008)].

The alignment was performed based on the amino acid using MultAlin software (27). The genetic relationships between the virus isolates were examined using phylogenetic analysis, which was carried out by MEGA7 (28). The neighbour-joining method assessed the tree to estimate the genetic evolution of the generated datasets. The tree distance was computed under a Poisson correction model and was in the units of the number of amino acid substitutions per site. Statistical analysis was estimated using the bootstrap approach via 100 replications. The tree was generated for each replicate, the method was repeated 100 times, and a percentage score was given for the same grouping.

Hypothetical spike binding to other receptors

This study followed some computational work using various software to address the binding affinity of the inserted proteins in the receptor-binding domain (RBD) of the spike region with two suspected human receptors, CCR3 and ICAM-1, for these purposes. The amino acid sequences were retrieved from NCBI/GenBank, and some protein structures were parsed from the PDB database. Secondary structure alignment was performed by using the ESpript server. The 3D structure, alignment, superposition and distance calculations were built using PyMOL software. Python 3.7 was used to script the residue interface calculation, and molecular docking of protein pairs was conducted utilizing HEX_Cuda 8.0 to predict the binding affinity. Default parameters were selected in HEX_Cuda, except for the following. The number of solutions was set to 1000, and the receptor step size was set to 7.5. The translation substep was set to 2. A steric scan of 20 and a final search of 30 were selected to optimize the output.

The RBD of the spike region was split into three areas (each region consists of the Cysteine residues in the beginning and the end connected by a disulfide bond), and they were separately fed to Clustal Omega for pairwise alignment with CCR3 and ICAM-1 amino acid sequence. The SARS-CoV2 ectodomain open-state (6VYB) and closed-state (6VXX) were used as 3D structure templates to correctly identify the secondary structure and annotation on the alignment using the ESpript server. These

peptides then aligned with 2MPM (the structural basis of receptor sulfotyrosine recognition) in the N-terminal region of CCR3 and 1IAM peptide in the ICAM-1 receptor. Moreover, the structural superposition and distance calculation of the SARS-CoV2 Spike with CCR3 and ICAM-1 were built using PyMOL software.

Results and Discussion

1. Classification of the SARS-CoV2 based on the Spike region

The phylogenetic tree analysis of the full length of SARS-CoV2 showed a close relation and high similarity with SARS-CoV, Bat and Pangolin coronavirus (data not shown). Therefore, we repeated the analysis by comparing different coronavirus strains based on the spike glycoprotein genes [Fig. 1]. The tree showed that all the novel coronavirus strains are approximately 100% similar and closely related to the Bat coronavirus (RaTG13). Moreover, the 2019 novel coronavirus's close genetic relationship is observed with another strain of Bat coronavirus and Pangolin coronavirus with 46% similarity. However, the diversity was found among the other coronavirus strains separated from the novel SARS-CoV2 and grouped in a different clade.

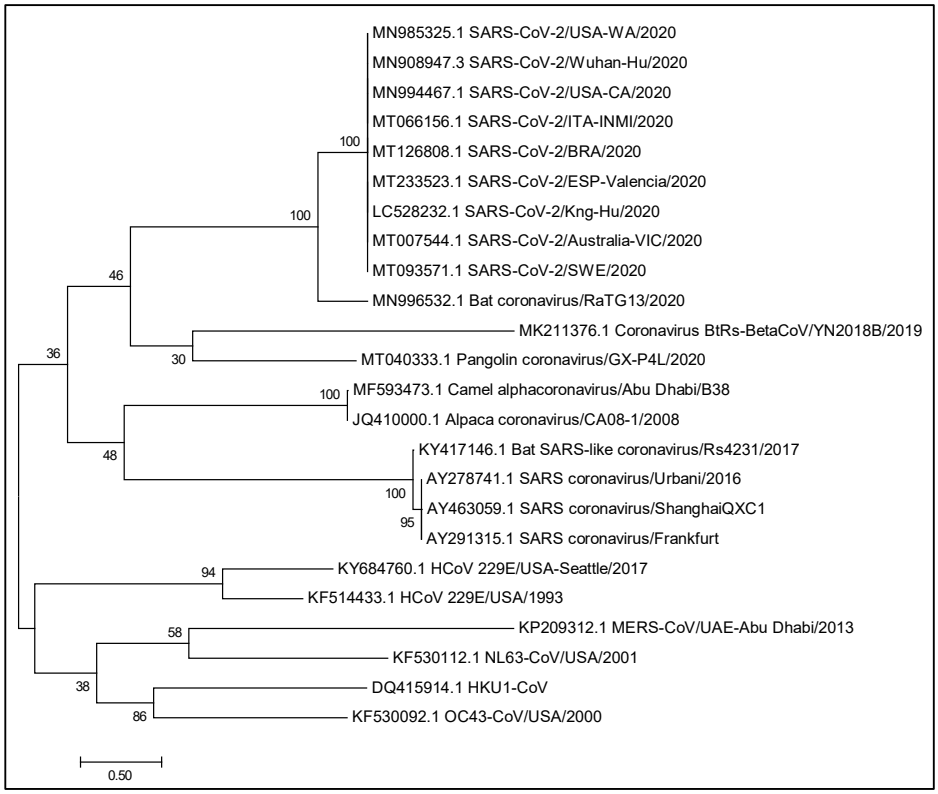


Fig. 1: Phylogenetic analysis based on the spike glycoprotein. The analysis is based on the Maximum Likelihood method and the Poisson correction model. The tree with the highest log likelihood (44231.1897) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model and then selecting the topology superior log-likelihood value.

2. Most variation is located in the S1 subunit.

The analysis of the spike region's multiple sequence alignment (the alignment presented in the supplementary data) between different coronavirus strains from humans (SARS-CoV2 and SARS-CoV strains), Bat and Pangolin showed a total of 873 amino acid variations among these strains (supplementary data). Among this total, 78% of the variants are located in the S1 region and 22% in the S2 subunit. There are only two amino acid variations in the spike glycoprotein among the novel SARS-CoV2 at position 251 and 801 in both Australian and Swedish strains. Interestingly, the lowest degree of difference between the SARS-CoV2 spike region and Bat is less than 3%, and with Pangolin, it is less than 10%. In contrast, the highest percentage of amino acid variation, approximately 29%, is seen between SARS-CoV2 and SARS-CoV strains in their spike region.

On the other hand, the alignment outputs also showed four amino acids deleted in the former part of the N-terminal domain and 23 amino acids inserted in five different N-terminal and Receptor-binding positions domain (Fig. 2). Surprisingly, all the inserted amino acids are found in all the novel coronavirus strains. Still, none of these inserts are found in the SARS-CoV strains. Moreover, the same amino acid inserts are detected in the Bat and Pangolin coronavirus except for 4 and 6 amino acids, respectively. Our analysis showed that the four deleted amino acids and the five insertion groups of the amino acid are unique to the novel SARS-CoV2 and could not be detected in any previously published coronavirus strains.

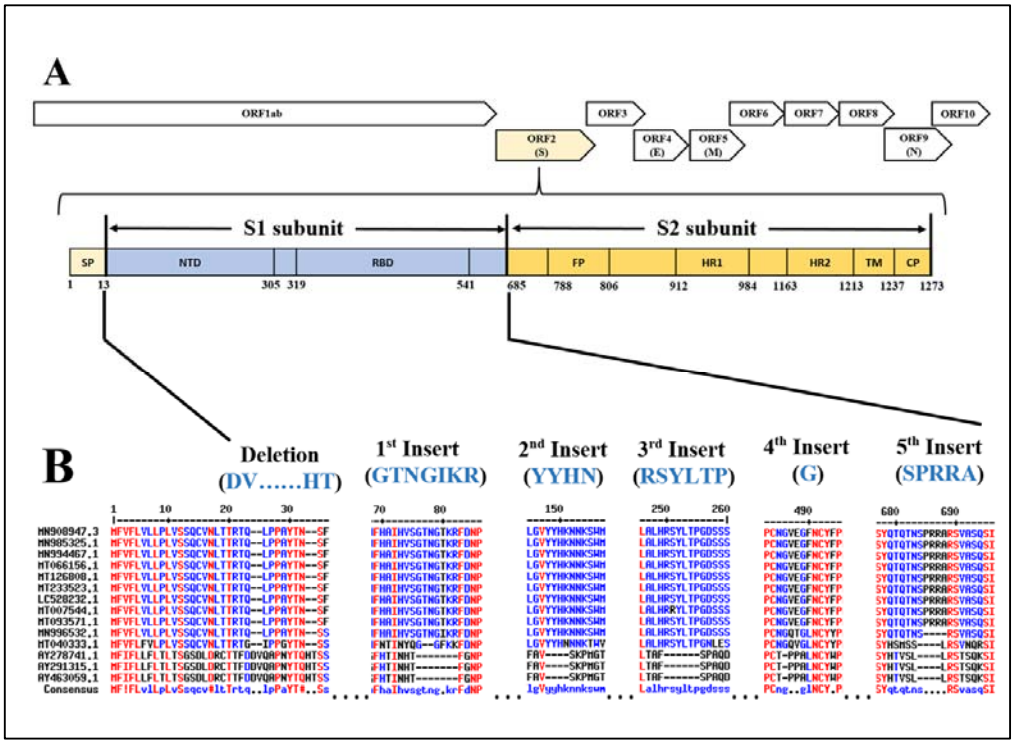
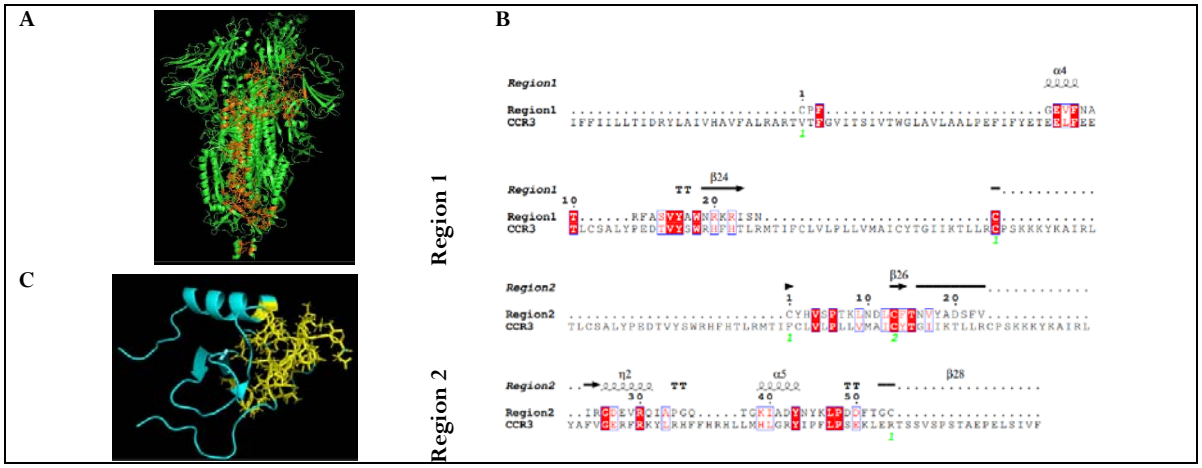


Fig. 2: A) Diagrammatic representation of the SARS-CoV2 spike glycoprotein and its genes. The scheme shows the functional domains in the S protein. The Signal Peptide (SP) is followed by the S1 subunit, which includes the N-terminal domain (NTD) and Receptor-Binding Domain (RBD). In contrast, the S2 subunit consists of Fusion Peptide (FP), Heptad Repeat 1 (HR1), Heptad Repeat 2 (HR2), Transmembrane Domain (TM), Cytoplasmic Domain (CP) (8). B) Part of the spike region's multiple sequence alignment of some selected coronavirus, including SARS-CoV2, SARS-CoV, Bat coronavirus and Pangolin coronavirus. The alignment shows the four deleted proteins and the insertion of 23 proteins in the S1 subunit.

3. Spike glycoprotein structural modelling and hypothetical receptors binding

The 3D structure of the spike glycoprotein showed many residue interfaces, which can produce polar binding with the aid of the hydrophilic bond in their surface [Fig. 3A]. For this reason, the selected regions were located in these positions to show the binding affinity with other receptors. In this study, we picked two cell receptors, CCR3 and ICAM-1, to demonstrate their ability to bind to the spike glycoproteins. The amino acid alignment of the three selected regions in the RBD of the spike glycoprotein to the CCR3 receptor showed a similarity of approximately 8% in some alpha-helix and more Beta pleated sheets. The identity and overlap amino acid ratio are different among the regions, and according to the Smith-Waterman Algorithm, it shows 30% identity when comparing 32 amino acids. Moreover, the homo-trimeric SARS-CoV2 Spike and CCR3 cartoon model's superposition showed these interactions in 35 proteins. However, these interactions are not entirely promising because of the weak bonding interaction between these proteins.



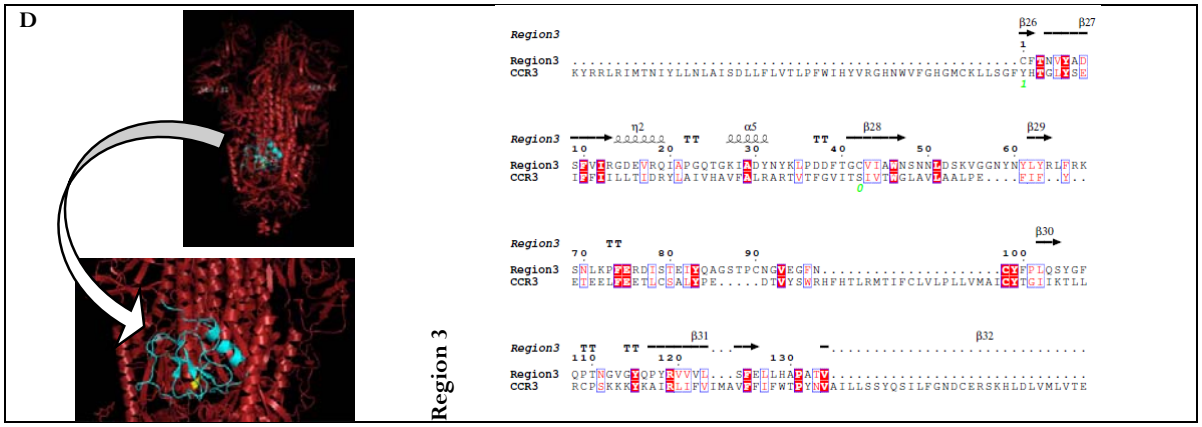


Fig. 3: Hypothetical binding between spike glycoprotein with CCR3 receptor. A) The interface of SARS-CoV2 Spike protein (orange colour). B) Multiple sequence alignment of some selected regions showing these regions' alpha-helix, Beta and gamma sheets. C) The interface of CCR3 protein (the yellow highlighted sticks are the interface residues for possible interaction). D) Superposition of CCR3 on COVID-19 spike protein (the yellow stick is CCR3 superposed on COVID-19 spike protein).

The super command alignment in PyMOL software shows a good alignment in a different region of the spike glycoproteins with the ICAM-1 receptor [Fig. 4]. As evident from the structural alignment, the ligand-binding area of ICAM-1 has a degree of affinity to bind to various regions in the S1 and S2 subunits of the SARS-CoV2 spike protein. The similarities are all in the beta-pleated sheets except for one alpha-helix part. Based on the Smith-Waterman Algorithm and comparing 120 amino acids in the spike region and ICAM-1 protein, over 50% similarity was detected, consistent with the superimposition's 3D structure.

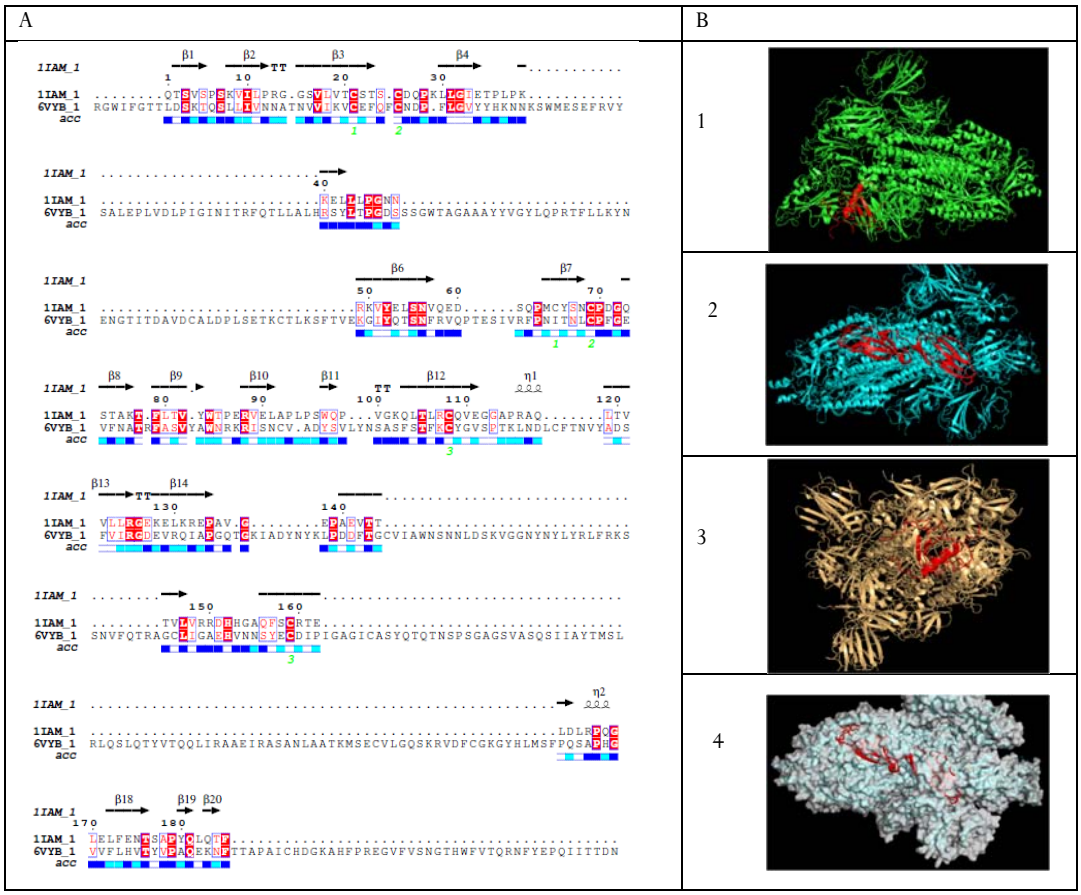


Fig. 4: Hypothetical binding between spike glycoprotein with ICAM-1 receptor. A) Multiple sequence alignment of some regions showing beta and gamma sheets of these regions. B) 3D structure superimposition showing 1- Interface 2MPM (Red) superimposed on 6VYB (Green) open state; 2- superposition of ICAM-1 on Spike protein 6VYB (Green) open state; 3- 2MPM (Red) superimposed on 6VXX (Yellow) closed state and 4- Structure and surface accessibility of 6VYB (Opened state) and 1IAM.

The RBD structure of SARS-CoV2 was extracted as an object from the 6VW1 and 6VYB, and with the aid of PyMOL software, the affinity of binding observed with 1IAM from ICAM-1 receptor as a 3D structure [Fig. 5A]. The ICAM-1 protein was superimposed with the 6VW1, and their interactions are in the Beta-sheet structure. The closest amino acids in this interaction are Ser494 on RBD and Val17 on ICAM-1. The distance measurement between the two sheets is 4.3 angstroms apart. Moreover, an extra interaction was observed outside the interface residue of the 1Iam, Cys108 (purple color). Another possible (polar) interaction region is between Asn440 in the RBD of the spike region and Arg116 of the ICAM-1 (red color in Fig. 5).

On the other hand, the interaction tested in the open state (6VYB) of the spike protein (purple color) with the ICAM-1 receptor. The crystal structure showed an interaction between the spike protein (colored white) with the 1IAM of the ICAM-1 receptor (red color) in a Proximity R of 28.6 Å. This was followed by fetching the spike protein into the software and tweaking the position with the receptor [Fig. 5B] so that they get into proximity. Because the HEX_Cuda software docks proteins best at distances no longer than 30 Å, the Interface residue calculation was first performed and adapted to the proteins using a Python script developed by (29) including these parameters, a complex of the first and second chain, and Cutoff value of the different area over the residues. The docking calculation was performed at 54.66 intermolecular separations in +/- steps of 0.75 Å. The applied rotational searching was then used in several angular increments of about 7.5 degrees in each of the two proteins (i.e. the receptor and the spike protein). The internal step sizes were set to fit the sample size and correspond to the FFT grid.

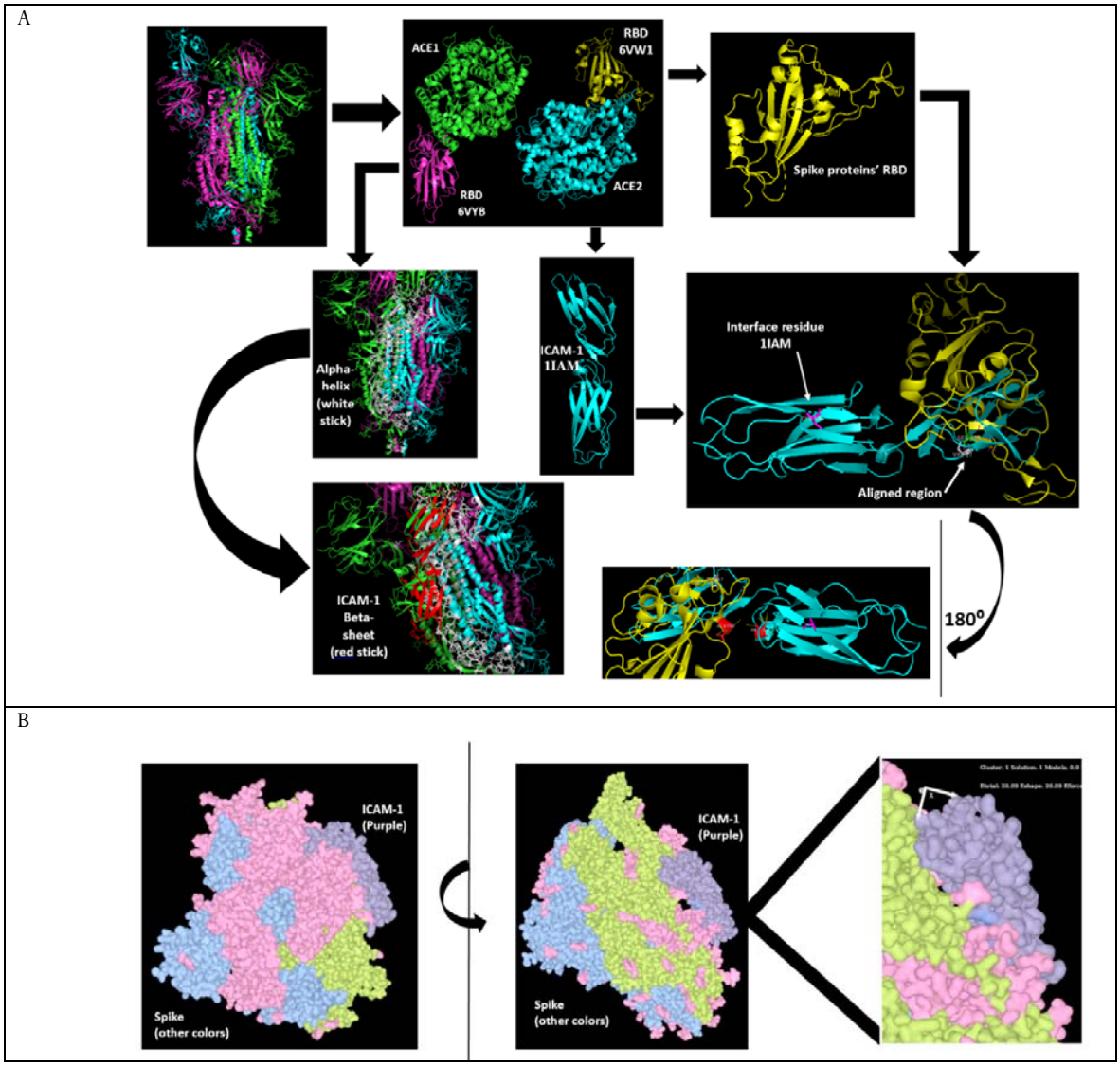


Fig. 5: Crystal Structure Presenting the Affinity and Binding Position. A) Both 6VW1 and 6VYB of the Spike were tested for their affinity binding against the ICAM-1 receptor. The analysis was performed by PyMOL software. B) Protein docking presented on both sites between the Spike glycoproteins (pink, blue and green color) with the ICAM-1 proteins (purple color). The analysis was performed by HEX_Cuda software.

The spike glycoprotein in the SARS-CoV2 is an essential element for the viral attachment and entry into the host cell, and any alteration of this protein may impact the host specifity. Therefore, we aimed to analyze this vital region using computational methods in this study. The phylogenetic analysis of the published sequences from NCBI showed that novel coronavirus is derived from Bat and Pangolin coronavirus, and they have a massive diversity with the nearest human coronavirus relatives. Moreover, 23 new proteins inserted in five regions of the spike S1 subunit are conserved among all the SARS-CoV2, which may impact the degree of infectivity and increase the number of host cells the virus can infect. The hypothetical binding of some of these inserted proteins and others in the RBD of the Spike to the ICAM-1 protein opens the possibility of addressing these proteins for specific detection of SARS-CoV2 and the development of treatment against COVID-19.

The unique amino acid insertion and deletion are only found in all the published SARS-CoV2 and are not present in other human coronaviruses. We found the presence of 3 insertion groups [first insert (seven amino acids), second insert (four amino acids), and third insert (six amino acids)] and deletion of 4 amino acids in the N-terminal domain of the RBD of the S1-subunit. Furthermore, another two insertion groups [fourth insert (one amino acid) and fifth insert (five amino acids)] in the C-terminal domain of the S1-subunit. Other research groups detected some of these inserts in the spike glycoprotein sequences of coronaviruses. A group from China demonstrated three insertions, and a group from India mentioned four insertions (30) due to comparing fewer sequences (4). However, the number and amino acids are different from these studies. The multiple sequence alignment of the proteins is only aligned with varying coronavirus strains, in contrast to Pradhan, Pandey (30), which showed the similarity between these inserts of the spike glycoproteins and HIV glycoproteins.

On cell entry, the coronavirus needs a cleavage of the spike region's S1 and S2 subunit, which is most probably done by furin. In all the human coronavirus, this site located in the S2 region (31). Interestingly, the 5th insert is the new furin-like cleavage site responsible for cleavage of S1 and S2 subunits before exiting with no need for further cleavage on entry (32). As evident in the multiple sequence alignment, the new cleavage site also exists in the Bat strain RaTG13 [Fig. 2]. We thought that the upstream inserted amino acids enhance the new furin-like cleavage site as a potential "gain-of-function" for the new virus. This potential for an infectivity gain seems to be an important idea because it needs cleaving of S1/S2 on exit instead of the SARS necessity for cleaving on entry.

On the other hand, further analysis showed the close relation of the SARS-CoV2 to the Bat coronavirus and Pangolin coronavirus (33). Other human coronavirus members were also derived from Bats as they represent a natural reservoir for this virus and others (3). Our analysis based on the spike region also showed the same similarity with Bat coronavirus RaTG13 [MN996532.1] and Pangolin coronavirus GX-P4L [MT040333.1. Moreover, the considerable diversity observed between SARS-CoV2 and SARS-CoV is based on the spike glycoproteins.

The 3D structure of the spike glycoproteins' 3D structure showed that the inserted amino acids provide a hydrophilic loop for the protein structure, which gives flexibility and increases the binding affinity of the Spike glycoprotein to different cell receptors (34, 35). As presented from our theoretical analysis, the inserted proteins have a crucial role in a high binding to the ICAM-1 and, to a lesser extent, to CCR3 receptors (36, 37). Our hypothesis may agree with Liu, Jiang (38), who mentioned a decrease in T-lymphocytes due to COVID-19. As these receptors can be detected in many immune and epithelial cells, it is crucial to be addressed in more detail to enhance the discovery of treatment against this virus.

Conclusion

In summary, this study shows approaches to utilizing data and analyzing the spike glycoprotein of SARS-CoV2. The unique inserts not detected in any other coronavirus may enhance the rapid molecular detection of this virus as the primary stage to control the infection. Moreover, more analysis of the spike region may help understand the route of infection. Further, the 3D modelling shows the binding of the SARS-CoV2 with other cell receptors like ICAM-1, which may be the reason for the new coronavirus's decrease in immune cells during infection. Altogether, our findings may have a crucial outcome for understanding the pathogenesis, disease severity, discovery of the treatment and vaccine development.

Acknowledgments: None

Conflict of interest: All the authors of this work have no conflict of interest

Financial support: None

Ethics statement: This article does not contain any studies with animals performed by any of the authors.

References

1. WHO. Situation report-1-2019-ncov. 2020.
2. Greco F, Dimonte S, Babakir-Mina M, Hama-Soor T, Ali S. Is It Scaly Anteater or Bat A Real Origin of The 2019-Novel CoV: A Probable Hypothesis? Kurdistan Journal of Applied Research. 2020;5(3):1-12. DOI https://doi.org/0.24017/covid.1.
3. Balboni A, Battilani M, Prosperi S. The SARS-like coronaviruses the role of bats and evolutionary relationships with SARS coronavirus. NEW MICROBIOLOGICA. 2012;35:1-16. PMID: 22378548.
4. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature.

2020;579(7798):270-3. https://doi.org/10.1038/s41586-020-2012-7.

5. Hama-Soor T, Jalal PJ, Ali S, Babakir-Mina M, Dimonte S, Greco F, et al. Is The SARS-CoV2 evolved in Human Being: A prospective Genetic Analysis. Kurdistan Journal of Applied Research. 2020:169-77.

6. Ji W, Wang W, Zhao X, Zai J, Li X. Cross-species transmission of the newly identified coronavirus 2019-nCoV. J Med Virol. 2020;92(4):433-40. https://doi.org/10.1002/jmv.25682.

7. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. N Engl J Med. 2020;382(13):1199-207. DOI: 10.056/NEJMoa2001316.

8. Jiang S, Du L, Shi Z. An emerging coronavirus causing pneumonia outbreak in Wuhan, China: calling for developing therapeutic and prophylactic strategies. Emerg Microbes Infect. 2020;9(1):275-7. https://doi.org/10.1080/22221751.2020.1723441.

9. Chan JF-W, Yuan S, Kok K-H, To KK-W, Chu H, Yang J, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. The Lancet. 2020;395(10223):514-23. DOI:https://doi.org/10.1016/S0140-6736(20)30154-9.

10. Cai J, Sun W, Huang J, Gamber M, Wu J, He G. Indirect Virus Transmission in Cluster of COVID-19 Cases, Wenzhou, China, 2020. Emerging Infectious Diseases journal. 2020;26(6):DOI: 10.3201/eid2606.200412.

11. Yi Y, Lagniton PNP, Ye S, Li E, Xu RH. COVID-19: what has been learned and to be learned about the novel coronavirus disease. Int J Biol Sci. 2020;16(10):1753-66. doi: 10.7150/ijbs.45134.

12. Maulud SQ, Majed SO, Ali BA, Jalal PJ, Azeez SH, Mohammad KA. Epidemiological Approach of SARS-CoV2 in the First Month of Appearance in the Kurdistan Region of Iraq. European Journal of Molecular & Clinical Medicine. 2020;7(11).

13. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet. 2020;395(10223):497-506. DOI:https://doi.org/10.1016/S0140-6736(20)30183-5.

14. WHO. Situation report-89-covid-19. 2020.

15. Mojarradgandoukmolla S, Akan H. Physiological activity and GC-Mass analysis of Trigonella strangulata, Trigonella filipes and Trigonella uncinata against Ethanol-Induced Hepatorenotoxicity in rats. Pakistan Journal of Zoology. 2022;55(2):513-24.

16. Mojarrad S. Responses of liver and renal function markers against arjuna tree extract in induced Hyperlipidemia Rats. Issue S. 2020;10(6).

17. Corman VM, Muth D, Niemeyer D, Drosten C. Hosts and Sources of Endemic Human Coronaviruses. Adv Virus Res. 2018;100:163-88. DOI: 10.1016/bs.aivir.2018.01.001.

18. Weiss SR, Navas-Martin S. Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. Microbiol Mol Biol Rev. 2005;69(4):635-64. doi:10.1128/MMBR.69.4.

19. Chan JF, Lau SK, To KK, Cheng VC, Woo PC, Yuen KY. Middle East respiratory syndrome coronavirus: another zoonotic betacoronavirus causing SARS-like disease. Clin Microbiol Rev. 2015;28(2):465-522. DOI: 10.1128/CMR.00102-14.

20. Mousavizadeh L, Ghasemi S. Genotype and phenotype of COVID-19: Their roles in pathogenesis. J Microbiol Immunol Infect. 2020;S1684-1182:30082-30087:http://doi.org/10.1016/j.jmii.2020.3.022.

21. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. Nature Medicine. 2020;26:450-2. DOI: 10.1038/s41591-020-0820-9.

22. Bosch BJ, van der Zee R, de Haan CA, Rottier PJ. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. J Virol. 2003;77(16):8801-11. DOI: 10.1128/JVI.77.16.

23. Li F. Structure, Function, and Evolution of Coronavirus Spike Proteins. Annu Rev Virol. 2016;3(1):237-61. DOI: 10.1146/annurev-virology-110615-042301.

24. 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;https://doi.org/10.1016/j.cell.2020.02.058

25. Backer JA, Klinkenberg D, Wallinga J. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20-28 January 2020. Euro Surveill. 2020;25(5):DOI: 10.2807/1560-7917.ES.2020.25.5.2000062.

26. Ibrahim IM, Abdelmalek DH, Elshahat ME, Elfiky AA. COVID-19 spike-host cell receptor GRP78 binding site prediction. J Infect. 2020;80(5):554-62. doi: 10.1016/j.jinf.2020.02.026.

27. Corpet F. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Research. 1988;16(22):DOI: 10.1093/nar/16.22.10881.

28. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016;33(7):1870-4. DOI: 10.093/molbev/msw054.

29. Vertrees J, Wrabl JO, Hilser VJ. An energetic representation of protein architecture that is independent of primary and secondary structure. Biophys J. 2009;97(5):1461-70. DOI: 10.016/j.bpj.2009.06.020.

30. Pradhan P, Pandey AK, Mishra A, Gupta P, Tripathi PK, Menon MB, et al. Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag. bioRxiv preprint. 2020;https://doi.org/10.1101/2020.01.30.927871 (retracted).

31. Hasan A, Paray BA, Hussain A, Qadir FA, Attar F, Aziz FM, et al. A review on the cleavage priming of the spike protein on coronavirus by angiotensin-converting enzyme-2 and furin. J Biomol Struct Dyn. 2020:1-13. DOI: 0.1080/07391102.2020.1754293.

32. Coutard B, Valle C, de Lamballerie X, Canard B, Seidah NG, Decroly E. The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. Antiviral Res. 2020;176:104742. DOI: 10.1016/j.antiviral.2020.

33. Zhang T, Wu Q, Zhang Z. Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID-19 Outbreak. Current Biology. 2020;30 1346-51. DOI: 10.016/j.cub.2020.03.022.

34. Kirchdoerfer RN, Cottrell CA, Wang N, Pallesen J, Yassine HM, Turner HL, et al. Pre-fusion structure of a human coronavirus spike protein. Nature. 2016;531(7592):118-21. DOI:10.1038/nature17200.

35. Ou X, Liu Y, Lei X, Li P, Mi D, Ren L, et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. Nat Commun. 2020;11(1):1620. https://doi.org/10.038/s41467-020-15562-9.

36. Barnes PJ. Distribution of receptor targets in the lung. Proc Am Thorac Soc. 2004;1(4):345-51. DOI: 10.1513/pats.200409-045MS.

37. Hanel K, Stangler T, Stoldt M, Willbold D. Solution structure of the X4 protein coded by the SARS related coronavirus reveals an immunoglobulin like fold and suggests a binding activity to integrin I domains. J Biomed Sci. 2006;13(3):281-93. DOI: 10.1007/s11373-005-9043-9.

38. Liu X, Jiang W, Zeng Z, Fan D, Zhang Y, Chen Z, et al. Decreased Counts of T Lymphocyte Subsets Predict Prognosis in SARS-CoV-2-Infected Pneumonia in Wuhan, China: a Retrospective Study. Research square. 2020:DOI: 10.21203/rs.3.rs-18694/v1.