Current Trends in Biotechnology and Pharmacy
Vol. 18 (Supplementry Issue 4A) 41 - 52, November 2024, ISSN 0973-8916 (Print), 2230-7303 (Online)
10.5530/ctbp.2024.4s.4

Mutational Stability Profiling and Functional Analysis of Spike Protein In Indian Sars Cov-2 Delta Variants: An *In Silico* Analysis

Prisho Mariam Paul[1,2], Krupakar Parthasarathy[1]*, Sudhanarayani S Rao[1], Vignesh Sounderrajan[1], Sakthivel Jayaraj[1] and Swetha Sunkar[3]

[1]Centre for Drug Discovery and Development, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu
[2]Biotechnology Department, CMS College Kottayam, Kerala
[3]Department of Bioinformatics, School of Bio and chemical Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu
*Corresponding author: pkrupakar.cddd@sathyabama.ac.in

Abstract

Amidst the global pandemic caused by the SARS-CoV-2 virus, researchers are actively studying its evolution and impact. This research delves into the spike protein of the SARS-CoV-2 Delta variant, using advanced computational methods to analyze mutations and their effects. By examining mutations and their evolutionary connections in the spike protein using data from databases, the study aims to gain insights into the virus's behavior.To gauge the stability of the spike protein variants, we used various methods. We predicted protein sequence disorders using PONDR. We also used the I-Mutant v2.0 bioinformatics tool to assess how mutations affected stability. We then performed functional studies using Protparam and NetPhos to analyze the consequences of these mutations. These analyses identified several key mutations in the spike protein, especially the D614G substitution. This polar-to-nonpolar amino acid change was particularly noteworthy.Our research identified significant mutations in the spike protein of the Delta Indian variant of SARS-CoV-2, including: T19R, E156G, L452R, T478K, P681R, and D950N. Deletions occurred at positions Y145, F157, and R158. Analysis showed increased phosphorylation of the spike protein, hinting at potential regulation by these mutations. Using structural modeling by Spdbv tool, we found that mutations within the Receptor-Binding Domain (RBD) region reduced protein stability. This study enhances our understanding of the mutations and their impact on the function of the Delta Indian variant's spike protein. Our research explores how mutations in the SARS-CoV-2 virus affect how its proteins work and stay stable. By understanding these effects, we can lay the foundation for future studies that design treatments specifically for new SARS-CoV-2 variants. In summary, our work shows that SARS-CoV-2 is constantly evolving, making it crucial to keep track of changes and keep researching to help guide health policies.

Keywords: SARS CoV-2, Spike protein, Mutation, Coronaviruses.

Introduction

The spread of COVID-19 around the world has shown that the severity of the disease can vary a lot. This variation depends on things like age, sex, and where the person lives. Even though there have been many ideas proposed, none of them can fully explain why there are such big differences in COVID-19 symptoms. The virus was first found in Wuhan, China in December 2019 and has since spread all over the world. Scientists are trying to understand how SARS-CoV-2 has changed by comparing the genetic material of the virus from different countries to the original material from Wuhan (1). The Coronaviridae family has two main groups: Letovirinae and Orthocoronavirinae. Orthocoronavirinae has

Vol. 18 (Supplementry Issue 4A) 41 - 52, November 2024, ISSN 0973-8916 (Print), 2230-7303 (Online)
10.5530/ctbp.2024.4s.4

several types, including alpha, beta, gamma, and delta. Beta-coronaviruses cause respiratory diseases in humans, including SARS-CoV, MERS-CoV, and SARS-CoV-2. Scientists classify variants of these viruses based on how easily they spread, how severe the illness they cause is, how well they can evade immunity, how well they respond to treatment, and how easy they are to detect with tests.Different types of SARS-CoV-2 virus, called variants, are categorized into two groups: Variants of Concern (VOC) and Variants of Interest (VOI). Currently, eleven variants have been identified (Alpha to Lambda). It's likely that more variants will continue to appear in the future(2).

Coronaviruses are viruses that have an outer layer made of fat. Inside, they carry a short strand of RNA that can infect both animals and birds. The RNA in the SARS-CoV-2 virus, which causes COVID-19, is about 30, 000 units long. It contains instructions to make 29 different proteins. These proteins include 16 that help the virus spread and survive, 9 that play a supporting role, and 4 that form the virus's structure. The structural proteins are: Spike (S) glycoprotein: Helps the virus attach to cells. Envelope (E) protein: Forms part of the virus's outer layer. Membrane (M) protein: Forms part of the virus's outer layer and helps assemble new viruses. Nucleocapsid (N) protein: Protects the virus's RNA (3).Coronaviruses have prominent glycoproteins called Spike proteins that extend outside their envelope. These proteins assist in the early stages of viral infection by binding to receptors on host cells. SARS-CoV and SARS-CoV-2 share about 75% similarity in their Spike proteins. These proteins have a molecular weight of about 141, 178 Daltons and are made up of 1, 273 amino acids. The Spike protein has three major parts: an outer ectodomain, a membrane-spanning part, and a brief section within the cell.The Spike protein has two subunits, S1 and S2, which help the virus bind to cells and enter them. It also allows the virus to spread from cell to cell, contributing to the infection's progression(4).

Scientists have identified a specific type of COVID-19 virus (SARS-CoV-2) by looking at the most common changes in its genetic material, especially those in the Spike protein. Viruses like SARS-CoV-2 change their genetic makeup in various ways to increase their chances of staying alive.Despite potential mutations that could weaken the virus, it often regains its strength due to the emergence of "compensatory mutations." To better grasp this process, we analyzed the frequency of mutations in all genes of the SARS-CoV-2 virus.(2) The WHO announced the start of the third wave of the COVID-19 pandemic. The dominant strain is the Delta variant, first seen in India. This variant is better at dodging antibodies, making it more easily spread and possibly causing more severe illness. Experts have also identified a more contagious subvariant of Delta, called Delta plus (B.1.617.2). (5). This research employs bioinformatics tools to assess the impact of mutations on the stability and function of the spike (S) protein found in the SARS-CoV-2 delta variant from India. The analysis is based on the genetic sequence of the virus.

Methodology

Sequence Retrieval

According to the NCBI protein database (http://www.ncbi.nlm.nih.gov/), we obtained 128 full protein sequences from India. Among these, we selected 15 spike protein sequences of the Delta variant strain from different regions for analysis. The reference sequence for comparison is the ancestral strain from Wuhan, China. Table 1

Table 1: Accession id's of all SARS CoV-2 sequences that were used in this study
NCBI Accession id(Spike Protein)
YP_009724390.1 QYM90002.1QYM89991.1QYM89980.1 QYM89969.1QYM89958.1QYM89445.1QYM89434.1QYM89423.1QYM89412.1QYM89401. 1QYM88739.1QYM88727.1QYM88716.1QYM88705.1QYM88694.1

Vol. 18 (Supplementry Issue 4A) 41 - 52, November 2024, ISSN 0973-8916 (Print), 2230-7303 (Online)
10.5530/ctbp.2024.4s.4

contains the accession numbers for all spike protein sequences.

Multiple Sequence Analysis and Phylogenetic analysis

We used BioEdit to align the amino acid sequences of spike protein variants. Conserved sites were verified using ClustalW. Sequence comparisons were conducted with MEGA(Molecular Evolutionary Genetics Analysis). The Neighbour-Joining method was utilized to study evolutionary history (6), with 1000 replicates used to create the bootstrap consensus tree (7).

Mutational analysis

PONDR, a predictor of natural disordered regions available at (http://www.pondr.com/), forecasts regions of disorder within protein sequences. These disordered regions can impact the specificity or affinity of protein binding (8). PONDR relies on five different algorithms, including VLXT, XL1_XT, and VSL2. The VLXT predictor combines three feedforward neural networks: VL1, N-terminus (XN), and C-terminus (XC) predictors. XL1 focuses on disorder regions of more than 39 amino acids (9), while VSL2 predicts both long and short disordered regions.

The iMutant 2.0 server is used to compare the stability of wild type and mutant proteins by calculating the effect of mutations. It predicts the outcome based on changes in Gibbs free energy (10). ΔG is calculated by subtracting the energy of the unfolded state from the energy of the folded state. The difference between the ΔG of the wild type and mutant type is referred to as ΔΔG. If both values are negative, the protein's energy and stability decrease. A positive ΔΔG value indicates that the mutant sequence has a higher ΔG and stabilizes the protein. All mutations were conducted at 25 °C, close to the global average temperature, and pH 7, which does not impact any type of mutation. Learn more about the iMutant 2.0 server at (https://folding.biofold.org/imutant/i-mutant2.0.html). Changing these values has an impact on protein stability.

Physicochemical properties of S-proteins

The S-proteins' physical and chemical characteristics were analyzed using Expasy'sProtParam tool (http://expasy.org/tools/protparam.html). This included predicting the theoretical isoelectric point (pI), extinction coefficient, instability index, molecular weight, aliphatic index, and grand average hydropathy (GRAVY) of the spike protein.

Post-Translational Modification

The NetPhos tool available at (http://www.cbs.dtu.dk/services/NetPhos/) can predict Serine, Threonine, and Tyrosine phosphorylation sites. Similarly, the NetNGlyc tool found at (www.cbs.dtu.dk/services/NetNGlyc/) is used to predict glycosylation sites on spike proteins.

Structural Analysis

The ACE2 bound to the spike protein of SARS CoV-2 was obtained from the PDB database (ID: 6M0J). The RBD region of the spike protein was extracted using the Autodock tool, and a specific mutation was introduced using Spdbv. The stability of the structure was then assessed by performing energy calculations after the mutation insertion.

Results and Discussions

Multiple Sequence Analysis of Spike protein delta variants

We compared 15 spike protein sequences of Indian delta variant strain from different regions of India with the reference sequence from Wuhan, China. Total 29 mutations are identified in sequences. Result of multiple sequence alignment is shown in Figure 1. Each variation and its proportion is displayed in Figure 2. Substitution of amino acid occurs in non polar amino acids and polar amino acids. Also a polar to non polar and non polar to amino acid substitution is identified in multiple sequence alignment. The D614G mutation is the most notable change from a polar to a non-polar amino acid, and it is a significant mutation. Additionally, there are other important mutations observed, such

Vol. 18 (Supplementry Issue 4A) 41 - 52, November 2024, ISSN 0973-8916 (Print), 2230-7303 (Online)
10.5530/ctbp.2024.4s.4

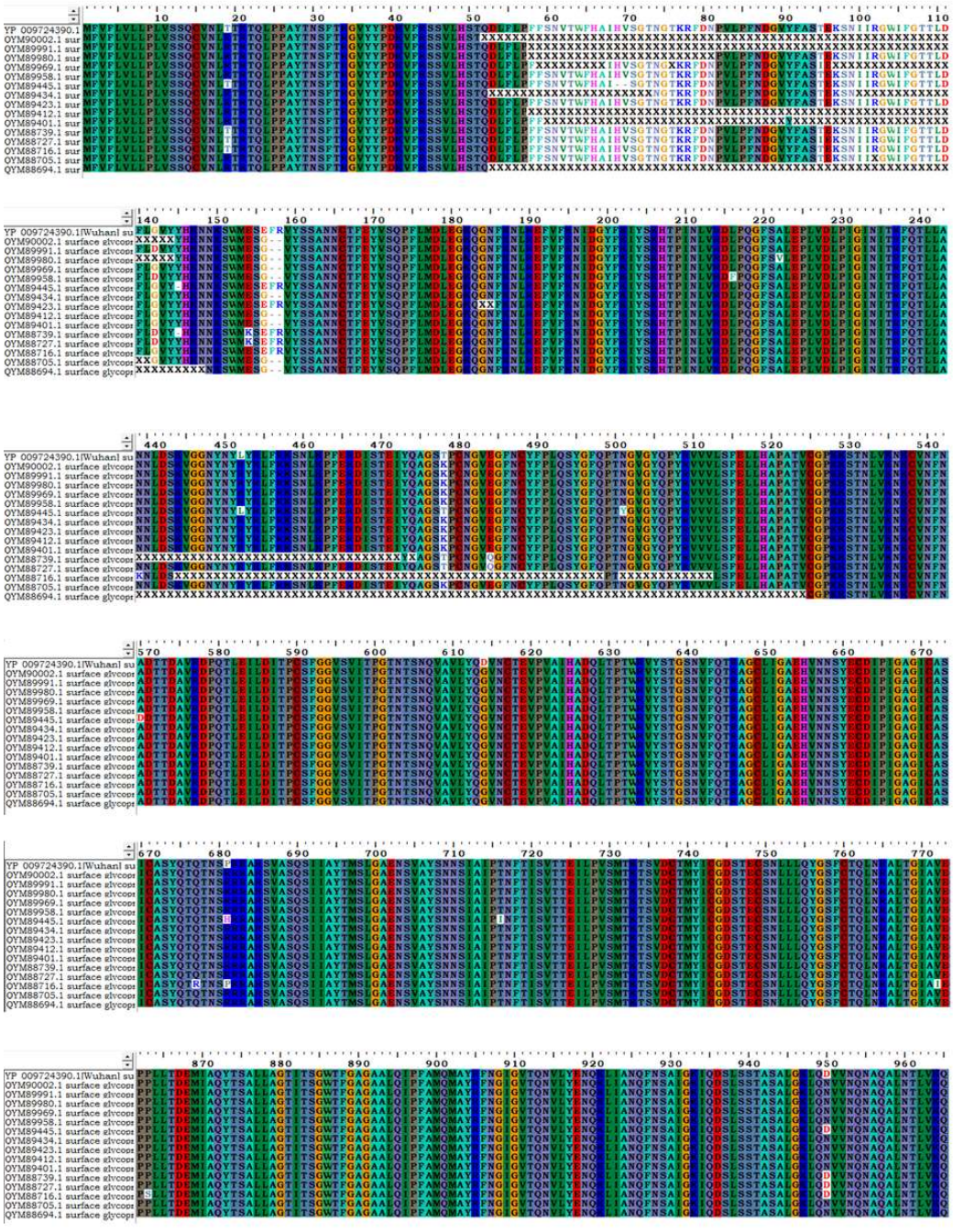

Figure 1: Multiple Sequence Analysis of SARSCoV- 2 Spike protein

Mutational Stability Profiling and Functional Analysis

Vol. 18 (Supplementry Issue 4A) 41 - 52, November 2024, ISSN 0973-8916 (Print), 2230-7303 (Online)
10.5530/ctbp.2024.4s.4

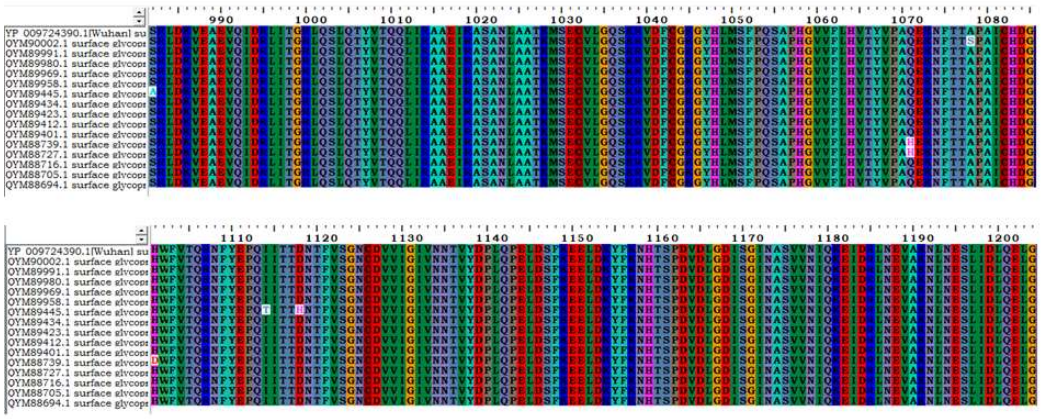

Figure 1: Multiple Sequence Analysis of SARSCoV- 2 Spike protein (*Contd.*)

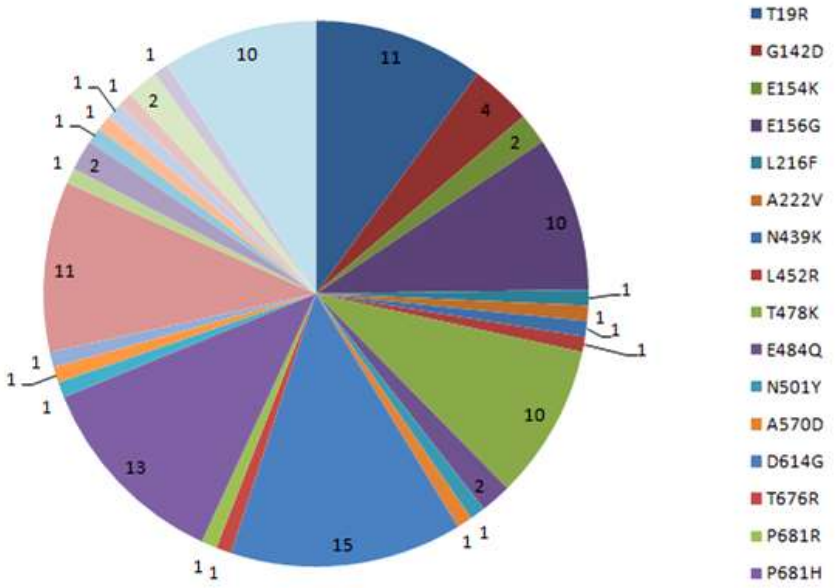

Figure 2: Proportion of variation in Sequences of SARS CoV-2

as the T19R switch from polar to polar, E156G switch from polar to non-polar, L452R switch from non-polar to polar, T478K switch from polar to polar, P681R switch from non-polar to polar, and D950N switch from polar to polar. Furthermore, there are deletions of Y145 in two sequences, as well as deletions of F157 and R158 in ten sequences.

Phylogenetic Analysis

Phylogenetic analyses of spike proteins were done with MegaX tool. Neighbor Joining method was used to interpret the evolutionary history (6). The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa

Vol. 18 (Supplementry Issue 4A) 41 - 52, November 2024, ISSN 0973-8916 (Print), 2230-7303 (Online)
10.5530/ctbp.2024.4s.4

analyzed(7). Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. A Poisson correction method is used to calculate the evolutionary distances and is represented as units of the number of amino acid substitutions per site (11). 16 amino acid sequences are involved in this analysis. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There were a total of 1273 positions in the final dataset. Resulting tree is shown in Figure 3. Result demonstrated that only one sequence is positioned in same clade of Wuhan strain. All other sequences are showed in different clade based on similarity.

Mutational analysis

PONDR (http://www.pondr.com/) Predictor of Natural Disordered Regions predicts the disordered region upon sequences. For predicting disordered region in this study, we took three algorithms XL1-XT, VLXT, VSL2. We consider the region which has been predicted by 3 algorithms as a disordered region of protein. The resulting disordered region is shown in Table 2.

To calculate the effect of mutation on protein stability between wild type and mutant protein the iMutant 2.0 server (https: //folding.biofold.org/imutant/i-mutant2.0. html) is used. Each mutation was considered separately for prediction. Table 3 showed the mutation stability of each mutation which identified after multiple sequence analysis. For analyzing mutational stability the individual mutation took the Wuhan protein as the reference.

Physicochemical properties of S-proteins

Protparamanalysis resulted that spike protein is an acidic peptide due to the theoretical pI range from 6.2 to 6.8 (percentage of acidic amino acid). The instability indexes of all sequences are below 40 and it showed that all protein was a stable peptide.Positive factors for the increased thermo stability of proteins are represented by aliphatic index and this factor is range from 69.33 to 84.67 which showed that all proteins are thermostable. GRAVY is ahydropathy index which augmented with theincrease in the positive score. GRAVY value of all protein is approximately near to -0.079. So this interpret that all proteins were hydrophilic in nature. Protparam result of all spike protein is given in Table 4.Due to some ambiguous sequence Protparam tooldid not calculates the pI of some sequencesand it was shown as undefined. The mutation causes very less variation in physiochemical properties. But further studies on this will reveal more details of physiochemical parameters.

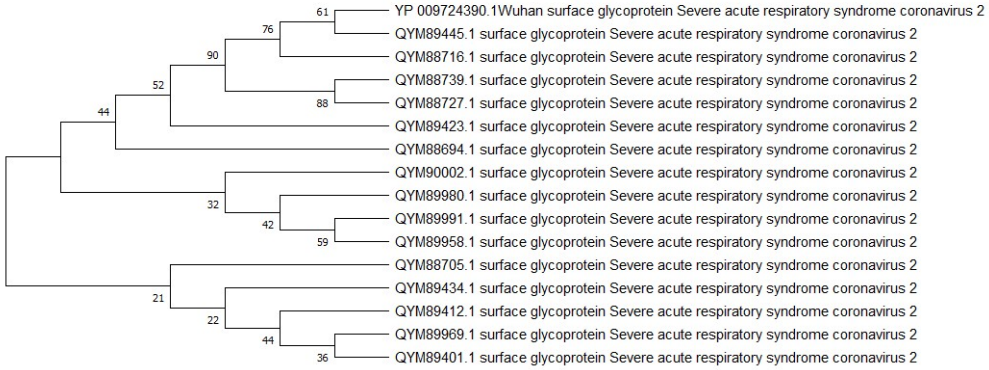

Figure 3: Phygenetic Analysis of SARS CoV-2 Delta variant Spike protein

Mutational Stability Profiling and Functional Analysis

Vol. 18 (Supplementry Issue 4A) 41 - 52, November 2024, ISSN 0973-8916 (Print), 2230-7303 (Online)
10.5530/ctbp.2024.4s.4

Table 2: PONDR – Disordered regions in the sequences

Sl.No.	Protein Name (total length)	XL1-XT ,VLXT, VSL2
1	YP_009724390.1(Wuhan)	679-689, 945-950
2	QYM90002.1	677-687, 700, 938-953
3	QYM89991.1	677-687, 700-702, 938-953
4	QYM89980.1	677-687, 700-702, 938-953
5	QYM89969.1	677-687, 700-702, 938-953
6	QYM89958.1	677-687, 700-702, 938-953
7	QYM89445.1	681-685, 942-946
8	QYM89434.1	93-96, 677-687, 700-702, 938-953
9	QYM89423.1	679-689, 940-955
10	QYM89412.1	677-687, 700-702, 938-953
11	QYM89401.1	677-687, 700-702, 938-953
12	QYM88739.1	677-687, 700-703, 944-948
13	QYM88727.1	679-689, 702-704, 945-950
14	QYM88716.1	678-689, 702-704, 945-950
15	QYM88705.1	677-687, 700-702, 938-953
16	QYM88694.1	677-687, 700-702, 938-953

Table 3: Stability of the induced Mutant with reference to wild type SARS CoV-2 Wuhan strain

Mutation	Stability				
T19R	Decrease ↓	E484Q	Decrease ↓	P863S	Decrease ↓
G142D	Decrease ↓	N501Y	Increase ↑	D950N	Decrease ↓
E154K	Decrease ↓	A570D	Decrease ↓	S982A	Decrease ↓
E156G	Decrease ↓	D614G	Decrease ↓	A1070H	Decrease ↓
L216F	Decrease ↓	T676R	Decrease ↓	A1078S	Decrease ↓
A222V	Increase ↑	P681R	Decrease ↓	H1101D	Decrease ↓
N439K	Decrease ↓	P681H	Decrease ↓	I1114T	Decrease ↓
L452R	Decrease ↓	T716I	Decrease ↓	D1118H	Decrease ↓
T478K	Decrease ↓	V772I	Decrease ↓		

Post-Translational Modification

NetPhos and NetNGlyc are used for predicting modification sites. Our result showed that spike protein is highly phosphorylated. All variant strains also have the same result. So, mutation is not affected in post translational modification.

Structural Analysis

After mutational analysis, study identified five mutations in RBD region of spike protein. They are N439K, L452R, T478K, E484Q, and N501Y. Using Spdbv inserted these mutations in RBD region and calculated the energy.

Vol. 18 (Supplementry Issue 4A) 41 - 52, November 2024, ISSN 0973-8916 (Print), 2230-7303 (Online)
10.5530/ctbp.2024.4s.4

Table 4: Protparam result for the SARS CoV-2 Spike protein sequences

Protein id	Theoretical Isoelectric Point(pI)	Instability index	Aliphatic index	Grand average hydropathy(GRAVY)
YP_009724390.1(Wuhan)	6.24	33.01	84.67	-0.079
QYM90002.1	Undefined	30.19	74.38	-0.086
QYM89991.1	Undefined	30.46	76.14	-0.085
QYM89980.1	Undefined	31.48	78.44	-0.069
QYM89969.1	Undefined	32.66	80.06	-0.091
QYM89958.1	6.78	32.75	84.19	-0.091
QYM89445.1	6.35	32.82	84.65	-0.078
QYM89434.1	Undefined	30.87	76.6	-0.097
QYM89423.1	Undefined	32.24	80.84	-0.076
QYM89412.1	Undefined	30.78	76.68	-0.091
QYM89401.1	Undefined	30.75	76.37	-0.087
QYM88739.1	Undefined	31.41	78.61	-0.045
QYM88727.1	6.8	32.81	84.67	-0.084
QYM88716.1	Undefined	29.16	72.05	-0.031
QYM88705.1	Undefined	32.79	83.19	-0.094
QYM88694.1	Undefined	28.09	69.33	-0.047

Table 5: Phosphorylation and Glycosylation sites in the SARS CoV-2 Sequences

NetPhos (Phosphorylation sites) and NetNGlyc(Glycosylation sites)
Phosphorylation sites 13, 19, 29, 46, 50, 51, 71, 73, 76, 91, 95, 108, 109, 112, 116, 151, 160, 161, 162, 170, 172, 200, 204, 208, 221, 240, 247, 250, 255, 256, 279, 286, 302, 305, 313, 316, 349, 351, 359, 366, 375, 376, 380, 383, 396, 415, 449, 469, 470, 473, 477, 494, 514, 523, 531, 547, 553, 555, 572, 573, 588, 591, 596, 599, 602, 604, 612, 630, 632, 637, 640, 659, 660, 673, 676, 678, 680, 686, 689, 691, 698, 707, 721, 734, 735, 741, 746, 758, 768, 778, 789, 791, 803, 813, 875, 881, 917, 929, 937, 939, 940, 961, 968, 974, 975, 982, 998, 1003, 1009, 1021, 1027, 1030, 1037, 1100, 1138, 1147, 1155, 1161, 1170, 1196, 1209, 1215, 1242, 1252, 1261
Glycosylation sites 17, 61, 74, 122, 149, 165, 234, 282, 331, 343, 603, 616, 657, 709, 717, 801, 1074, 1098, 1134, 1158, 1173, 1194

For SARS CoV-2 RBD region showed the energy -10337.443 kcal/mol and after inserting mutation it showed -9755.156 kcal/mol. So it revealed that mutation decreases the stability of the structure Figures 4 and 5.

Due to mutations, the SARS CoV-2 virus has spread more easily. The virus's low genetic diversity suggests a recent origin. Mutations' impact increases over time, as seen in HIV-1. The virus's novelty is indicated by the presence of acommon phenotypic

Vol. 18 (Supplementry Issue 4A) 41 - 52, November 2024, ISSN 0973-8916 (Print), 2230-7303 (Online)
10.5530/ctbp.2024.4s.4

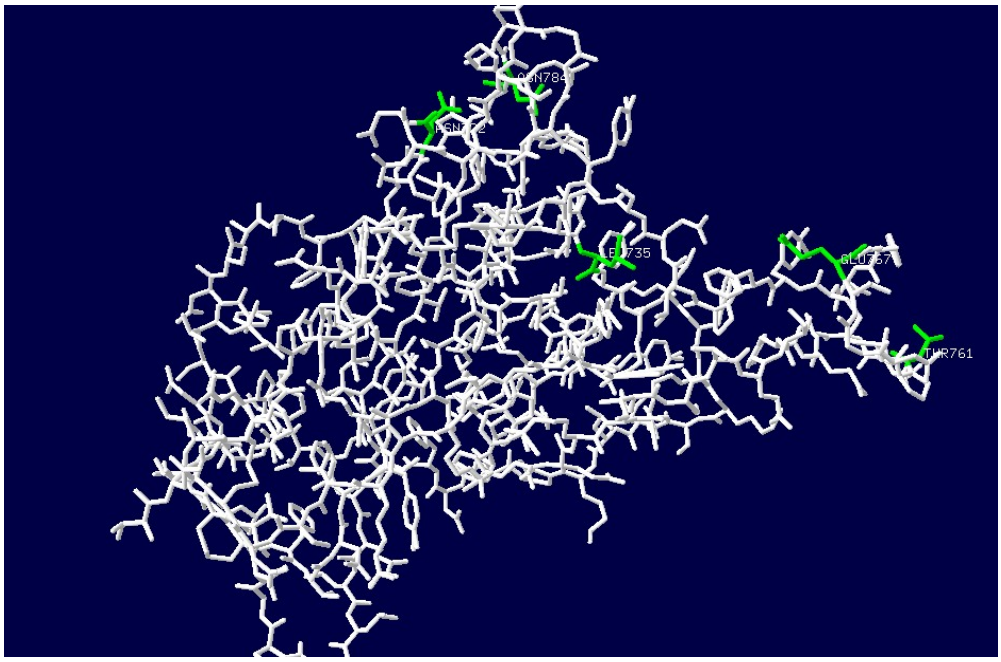

Figure 4: RBD Structure of CoV-2 before mutation

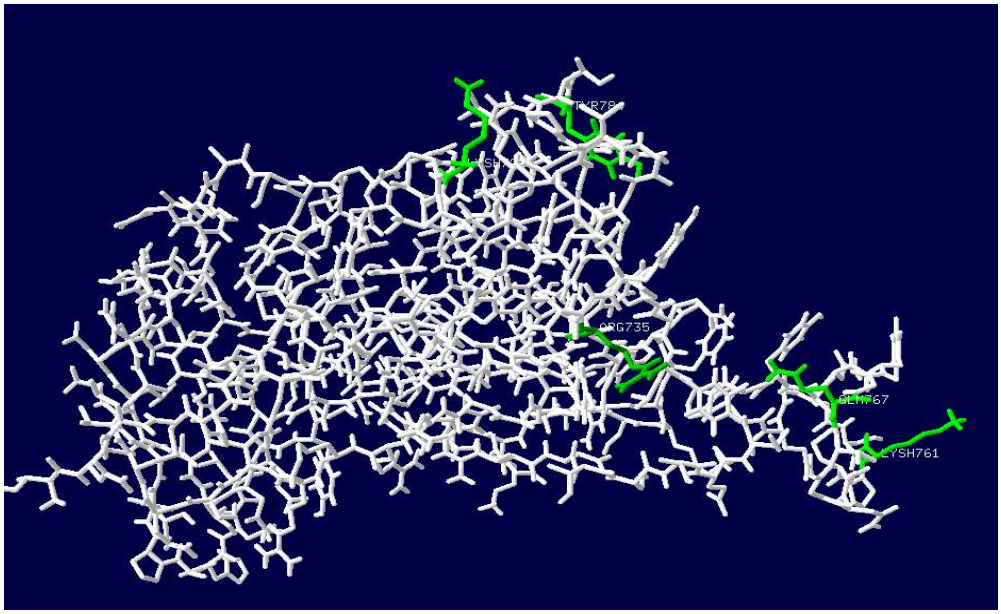

Figure 5: RBD Structure of CoV-2 after induced Mutation

Vol. 18 (Supplementry Issue 4A) 41 - 52, November 2024, ISSN 0973-8916 (Print), 2230-7303 (Online)
10.5530/ctbp.2024.4s.4

mutation(12). If a virus has a single unique mutation that often means, it's new. The pattern of mutations in spike protein shows how all the possible changes in these proteins have built up over time. Generally, most mutations make the proteins less stable, which drives the virus to change further to become more stable and avoid being killed off by the immune system (13).

Protein mutations weaken their stability, both functionally and structurally. Research indicates that proteins with smaller cores are more susceptible to mutations. Mutations can destabilize proteins, causing structural changes that alter their conformation. These changes can disrupt protein function and stability (14). Similarly, in the case of HIV-1, drug resistance is a significant concern in the use of combination therapies. Through sequence analysis of drug-resistant strains, it has been discovered that the majority of mutations destabilize the virus (15).The destabilizing mutations in SARS CoV-2 are the main reason for strain diversification and huge variation of the virus. Some mutations affect the phenotype of protein and viral transmissibility (16). Protein stability change due to mutation leads to evolvability of virus and increases genetic diversity (17).To predict intrinsic disorder regions of S protein in this study 3 prediction methods are used. VSL2, XL1-XT and VLXT (18). All three predictive models used characteristics derived from amino acid compositions. Since the S protein's primary role is to bind to host proteins, these disordered areas may impact how the virus spreads. Mutations in the receptor-binding area include L452R, T478K, E484Q, and N501Y.

Phosphorylation sites have prominent role in coronavirus. Phosphorylation plays a vital role in the retention of spike protein at cell surface. For assembling the trimer phosphorylation sites on S protein is needed (19). Glycosylation has a vital role in antigenicity, fusogenic, and immunomodulatory activities of the spike protein (20).Previous studies show that these positions are conserved and these can be considered as inhibitory sites.

The SARS-CoV-2 delta variant, which originated in India, has spread to multiple countries and is known for having three unique mutations: E156del/R158G in the N-terminal domain and T478K in an important receptor binding domain. The T478K mutation in the S gene's functional domain may be linked to the transmissibility and severity of this variant. However, despite efforts, researchers have not been able to precisely locate this specific mutation site in the δ variant through detailed protein structure analyses, possibly due to limitations in laboratory conditions as compared to the complex viral infection process in hosts. Additionally, these mutations could impact the way the virus interacts with other molecules, possibly resulting in stronger interactions with different amino acids (21).

Conclusion

Since December 2019, COVID-19, caused by the SARS CoV-2 virus, has become a global pandemic. A major challenge in developing specific treatments is that the virus evolves over time, creating new strains with different protein sequences. This study used various computer-based methods to examine and analyze mutations in ten SARS CoV-2 proteins. By grouping and comparing protein sequences, researchers identified variations in the virus. They also assessed the impact of these mutations on protein structure and function using mutational profiling and analysis.Understanding the nature of specific viral residues can guide the development of antiviral drugs to combat a broader spectrum of mutations. This knowledge is crucial because mutations can lead to drug resistance, requiring combinatorial therapies like those used for HIV-1. These therapies combine multiple drugs to prevent resistance and increase efficacy(22).Other respiratory viruses have also been treated successfully through therapies that combine multiple drugs (23).

Vol. 18 (Supplementry Issue 4A) 41 - 52, November 2024, ISSN 0973-8916 (Print), 2230-7303 (Online)
10.5530/ctbp.2024.4s.4

The COVID-19 pandemic, caused by the SARS-CoV-2 virus, has posed a major public health challenge globally. The virus's tendency to mutate rapidly makes it essential to develop treatment strategies that can adapt to these changes. This study used advanced computer simulations to analyze mutations in key viral proteins. This analysis revealed important variations in the virus's genetic makeup, which can affect protein structure and function. By identifying specific viral "weak spots" and their role in drug resistance, this research provides valuable information for designing more effective antiviral treatments.Like successful strategies against viruses like HIV-1, proactively anticipating and countering emerging mutations is crucial for fighting COVID-19. This helps healthcare professionals adapt to the virus's changing nature, preventing its spread and minimizing its global impact. Researchers can customize vaccines to target specific mutations, such as those found in the Delta Indian variant, in order to enhance their effectiveness against new variants. By honingin on crucial mutations, vaccine makers could enhance the range and strength of vaccines, potentially expanding protection against different strains of SARS-CoV-2. Studying how mutations affect the stability and function of the spike protein can help develop treatment options like monoclonal antibodies or antiviral drugs. These therapies could target spike protein areas made susceptible by mutations, potentially blocking the virus from entering host cells or improving virus elimination.

Acknowledgment

Authors acknowledge ICMR, Govt.of.India for funding by the project ICMR-VIR/COVID19/33/2021/ECD-I

Funding

This work was supported by ICMR, Govt.of.India for funding by the project ICMR-VIR/COVID19/33/2021/ECD-I

Competing interest

All authors declare that there is not any actual or potential conflict of interest including anyfinancial, personal, or other relationships with other people or organizations within five years ofbeginning the submitted work.

Author Contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by [Prisho Mariam paul], [Vignesh Sounderrajan], [Krupakar Parthasarathy], [Sudhanarayani S Rao], [Sakthivel Jayaraj] and [Shwetha Sunkar] The first draft of the manuscript was written by [Prisho Mariam paul] and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data Availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. Biswas, S. K., & Mudi, S. R. (2020). Genetic variation in SARS-CoV-2 may explain variable severity of COVID-19. Medical Hypotheses, 143, 109877.
2. Kannan, S. R., Spratt, A. N., Cohen, A. R., Naqvi, S. H., Chand, H. S., Quinn, T. P., et al. (2021). Evolutionary analysis of the Delta and Delta Plus variants of the SARS-CoV-2 viruses. Journal of Autoimmunity, 124, 102715.
3. Gao, T., Gao, Y., Liu, X., Nie, Z., Sun, H., Lin, K., et al. (2021). Identification and functional analysis of the SARS-COV-2 nucleocapsid protein. BMC Microbiology, 21(1), 58.
4. Satarker, S., &Nampoothiri, M. (2020). Structural Proteins in Severe Acute Respiratory Syndrome Coronavirus-2. Archives of Medical Research, 51(6), 482–491.
5. Arora, P., Kempf, A., Nehlmeier, I., Graichen, L., Sidarovich, A., Winkler, M. S., et al. (2021). Delta variant (B.1.617.2) sublineages do not show increased neutralization resistance. Cellular & Molecular Immunology, 18(11), 2557–2559.

Vol. 18 (Supplementry Issue 4A) 41 - 52, November 2024, ISSN 0973-8916 (Print), 2230-7303 (Online)
10.5530/ctbp.2024.4s.4

6. Saitou, N., & Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution, 4(4), 406–425.
7. Felsenstein, J., & Kishino, H. (1993). Is there something wrong with the bootstrap on phylogenies? A reply to Hillis and Bull. Systematic Biology.
8. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., & Obradović, Z. (2002). Intrinsic Disorder and Protein Function. Biochemistry, 41(21), 6573–6582. https://doi.org/10.1021/bi012159.
9. Romero, P., Obradovic, Z., & Dunker, A. K. (1997). Sequence data analysis for long disordered regions prediction in the calcineurin family. Genome Informatics, 8, 110–124.
10. Capriotti, E., Fariselli, P., &Casadio, R. (2005). I-Mutant2.0: Predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Research, 33(Web Server issue), W306-W310.
11. Zuckerkandl, E., & Pauling, L. (1965). Evolutionary divergence and convergence in proteins. Evolving Genes and Proteins, 97–166.
12. Storz, J. F. (2018). Compensatory mutations and epistasis for protein function. Current Opinion in Structural Biology, 50, 18–25.
13. Serohijos, A. W. R., Rimas, Z., &Shakhnovich, E. I. (2012). Protein biophysics explains why highly abundant proteins evolve slowly. Cell Reports, 2(2), 249–256.
14. Hadfield, J., Megill, C., Bell, S. M., Huddleston, J., Potter, B., Callender, C., et al. (2018). Nextstrain: Real-time tracking of pathogen evolution. Bioinformatics, 34(23), 4121–4123.
15. Olabode, A. S., Avino, M., Ng, G. T., Abu-Sardanah, F., Dick, D. W., & Poon, A. F. Y. (2019). Evidence for a recombinant origin of HIV-1 Group M from genomic variation. Virus Evolution, 5(1), 039.
16. Chen, G., Wu, D., Guo, W., Cao, Y., Huang, D., Wang, H., et al. (2020). Clinical and immunological features of severe and moderate coronavirus disease 2019. Journal of Clinical Investigation, 130(5), 2620–2629.
17. Petit, C. M., Melancon, J. M., Chouljenko, V. N., Colgrove, R., Farzan, M., Knipe, D. M., et al. (2005). Genetic analysis of the SARS-coronavirus spike glycoprotein functional domains involved in cell-surface expression and cell-to-cell fusion. Virology, 341(2), 215–230.
18. Peng, K., Radivojac, P., Vucetic, S., Dunker, A. K., & Obradovic, Z. (2006). Length-dependent prediction of protein intrinsic disorder. BMC Bioinformatics, 7, 208.
19. Davidson, A. D., Williamson, M. K., Lewis, S., Shoemark, D., Carroll, M. W., Heesom, K. J., et al. (2020). Characterisation of the transcriptome and proteome of SARS-CoV-2 reveals a cell passage induced in-frame deletion of the furin-like cleavage site from the spike glycoprotein. Genome Medicine, 12(1), 68.
20. Fung, T. S., & Liu, D. X. (2018). Post-translational modifications of coronavirus proteins: Roles and function. Future Virology, 13(6), 405–430.
21. Jhun, H., Park, H. Y., Hisham, Y., Song, C. S., & Kim, S. (2021). SARS-CoV-2 Delta (B.1.617.2) Variant: A Unique T478K Mutation in Receptor Binding Motif (RBM) of Spike Gene. Immune Network, 21(5), e32. https://doi.org/10.4110/in.2021.21.e32.
22. McFadden, K., Fletcher, P., Rossi, F., Kantharaju, Umashankara, M., Pirrone, V., et al. (2012). Antiviral breadth and combination potential of peptide triazole HIV-1 entry inhibitors. Antimicrobial Agents and Chemotherapy, 56(2), 1073–1080.
23. Hayden, F. G. (1996). Combination antiviral therapy for respiratory virus infections. Antiviral Research, 29(1), 45–48.