Heliyon 7 (2021) e06167

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

The molecular assessment of SARS-CoV-2 Nucleocapsid Phosphoprotein variants among Indian isolates

Gajendra Kumar Azad *

Department of Zoology, Patna University, Patna, 800005, Bihar, India

ARTICLE INFO

Keywords:
COVID-19
SARS-CoV-2
Mutations
Nucleocapsid Phosphoprotein (N)
Infectious diseases
India

ABSTRACT

Coronavirus disease- 2019 (COVID-19) has rapidly become a major threat to humans due to its high infection rate and deaths caused worldwide. This disease is caused by an RNA virus, Severe Acquired Respiratory Syndrome –Corona Virus-2 (SARS-CoV-2). This class of viruses have a high rate of mutation than DNA viruses that enables them to adapt and also evade host immune system. Here, we compared the first known Nucleocapsid Phosphoprotein (N protein) sequence of SARS-CoV-2 from China with the sequences from Indian COVID-19 patients to understand, if this virus is also mutating, as it is spreading to new locations. Our data revealed twenty mutations present among Indian isolates. Out of these, mutation at six positions led to changes in the secondary structure of N protein. Further, we also show that these mutations are primarily destabilising the protein structure. The candidate mutations identified in this study may help to speed up the understanding of variations occurring in SARS-CoV-2.

1. Introduction

In the Wuhan City of China (Hubei province), multiple cases of severe pneumonia like symptoms were reported for the first time in December 2019 [1]. Initially, it was thought to be a viral disease and after the complete genome sequencing, it was revealed that it belongs to a novel coronavirus and later renamed as SARS-CoV-2 because of its similarity with SARS-CoV [2]. The disease caused by SARS-CoV-2 has been termed Coronavirus disease-2019 (COVID-19). The COVID-19 has been declared by the World Health Organization (WHO) as a global pandemic in March 2020. As of 6th June 2020, there were more than 7 million confirmed cases of COVID-19 with 0.4 million confirmed deaths worldwide.

The SARS-CoV-2 is rapidly spreading to new locations across the globe and infecting human populations. The viruses have the tendency to mutate which helps them to survive better in the host. Most of the RNA viruses including coronaviruses have been reported to possess an inherent high rate of mutation [3]. Reports suggest that the Spike protein and RNA dependent RNA polymerase (RdRp) are two mutational hotspot proteins of SARS-CoV-2 that is frequently mutated as revealed by the sequencing data from different countries [4, 5, 6, 7].

One of the important proteins that protect RNA genome of SARS-CoV-2 is N protein [8]. The N protein comprises of three distinct conserved regions; the N-terminal domain (NTD), C-terminal domain (CTD) and intrinsically disordered regions (IDRs). There are three IDRs, namely at N-terminal end, C-terminal end and between the NTD and CTD [8]. All of these regions make contact with the viral genomic RNA. Further, evidences suggest that the NTD have various residues that are critical for its interactions with the RNA genome of SARS-CoV-2. The CTD is also known as dimerization domain because two molecules of N protein dimerise with the interactions of their CTD. The NTD and CTD are separated by serine and arginine rich, IDR also known as the linker region (LKR). The main function of N protein is to protect viral RNA genome of SARS-CoV-2 by proper packaging. It binds with the viral genomic RNA and generates long, relatively flexible and somewhat helical ribonucleoprotein complex also known as viral capsid [9].

In this study, we analysed the N protein sequences of SARS-CoV-2 from Indian COVID-19 patients and compared with the Wuhan virus sequence. Our study identified twenty mutations in N protein among Indian isolates, and discussed their possible consequences.

2. Material and methods

2.1. N protein sequence retrieval

As of 7th June 2020, there were 175 protein sequences of SARS-CoV-2 deposited in NCBI- virus database from India. We downloaded these

* Corresponding author.
E-mail address: gkazad@patnauniversity.ac.in.

https://doi.org/10.1016/j.heliyon.2021.e06167
Received 11 June 2020; Received in revised form 5 September 2020; Accepted 28 January 2021

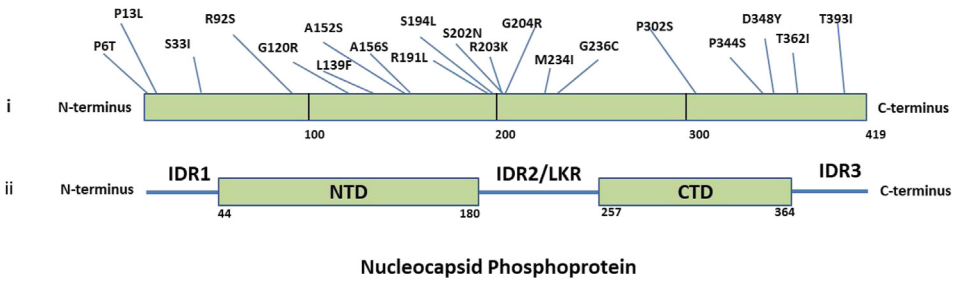

Figure 1. Mutational analysis of SARS-CoV-2 sequences reported from India. The sequence of SARS-CoV-2 N protein obtained from Indian COVID-19 patients was aligned with the sequence from Wuhan (wet seafood market) SARS-CoV-2. The mutations were recognised by amino acid sequence alignment by Clustal Omega. i) The mutant residues positions and mutations are highlighted in the schematic of N protein. The N protein sequence of Wuhan SARS-CoV-2 was used as reference for this analysis. ii) The schematic of N protein showing different domains. N-terminal domain (NTD), C-terminal domain (CTD), intrinsically disordered regions 1, 2, 3 (IDR1, IDR2 and IDR3), IRD2 also known as linker region (LKR) because it connects NTD and CTD. NTD have RNA binding motif while CTD helps in protein dimerization.

sequences and their protein identification accession number are mentioned in supplementary table 1. The N protein sequence from the earliest sequenced virus from Wuhan wet seafood market area was used as a reference (accession number: YP_009724397).

2.2. Multiple sequence alignments

We used Clustal Omega online tool to perform multiple sequence alignment (MSA) to identify mutations present among Indian isolates as described earlier [10]. The Indian sequences were compared with the Wuhan N protein (accession number: YP_009724397). The mutant residues were carefully identified and marked. Clustal Omega tool offers fast and reliable algorithm to perform sequence alignments to efficiently compare thousands of input sequences [11].

2.3. Secondary structure prediction

For secondary structure prediction, we used Chou and Fasman secondary structure prediction (CFSSP) online tool [12]. This tool provides the secondary structure pattern from the input primary amino acid sequence.

Table 1. Mutations across N proteins of SARS-CoV-2 among Indian isolates and their frequencies. The frequency of each mutation was calculated by counting the number of samples that exhibited the variation among 175 samples analysed in this study. Similarly, the percentage frequency was also calculated.

Mutation	Frequency	% Frequency
P6T	2	1.142857
P13L	16	9.142857
S33I	2	1.142857
R92S	1	0.571429
G120R	1	0.571429
L139F	1	0.571429
A152S	1	0.571429
A156S	1	0.571429
R191L	1	0.571429
S194L	49	28.0
S202N	5	2.857143
R203K	5	2.857143
G204R	5	2.857143
M234I	1	0.571429
G236C	1	0.571429
P302S	1	0.571429
P344S	1	0.571429
D348Y	1	0.571429
T362I	1	0.571429
T393I	1	0.571429

2.4. Calculations of difference in free energy (ΔΔG) between wild-type and mutant

The ΔΔG was calculated using online mCSM webserver [13]. It uses an algorithm to calculate the differences in free energy between the wild-type and mutant protein. The positive ΔΔG indicates stabilising mutation while negative ΔΔG represents destabilisation. mCSM tool also shows the impact of a mutation on the atomic-distance patterns surrounding an amino acid residue. For this analysis RCSB protein ID: 6VYO [14] was used for NTD molecular modelling and RCSB protein ID: 6WJI (Centre for Structural Genomics of Infectious Diseases-CSGID, unpublished) was used for CTD molecular modelling of N protein.

2.5. Analysis of protein dynamicity

We analysed the impact of mutation on dynamicity and structural variations using DynaMut webserver [15]. This webserver provides the location of a selected amino acid in the 3D-structure of the protein and also show the kind of interactions made by the residues with surrounding amino acids. For this analysis, the crystal structure of N protein was first uploaded on DynaMut webserver. The RCSB protein ID: 6VYO [14] and 6WJI were used for NTD and CTD molecular modelling of N protein respectively. It also provides the differences in the vibrational entropy between wild-type and mutant protein and predicts the impact of mutation on structural stability and dynamicity.

3. Results

3.1. N protein of SARS-CoV-2 is rapidly mutating in India

We downloaded all N protein sequences from India deposited in NCBI virus database till 7th June 2020. The accession numbers of the N protein used in this study are mentioned in supplementary table 1. The multiple sequence alignments were performed using Clustal Omega tool and the Wuhan virus N protein sequences was used as a reference (accession number: YP_009724397). Our data revealed twenty mutations present in the SARS-CoV-2 N protein from Indian COVID-19 patients (Figure 1). The complete details of mutations identified in this study are listed in supplementary table 2. Further, these mutations are spreading all over the N protein; however, a cluster is observed in the intrinsically disordered region 2 (IDR2) or linker region (LKR) present between NTD and CTD (Figure 1). There are seven mutations present at this location suggesting that this might be a mutational hotspot area of this protein. The N-terminal unstructured region harbours three mutations at 6, 13 and 33 positions. The NTD and CTD have various mutations as shown in

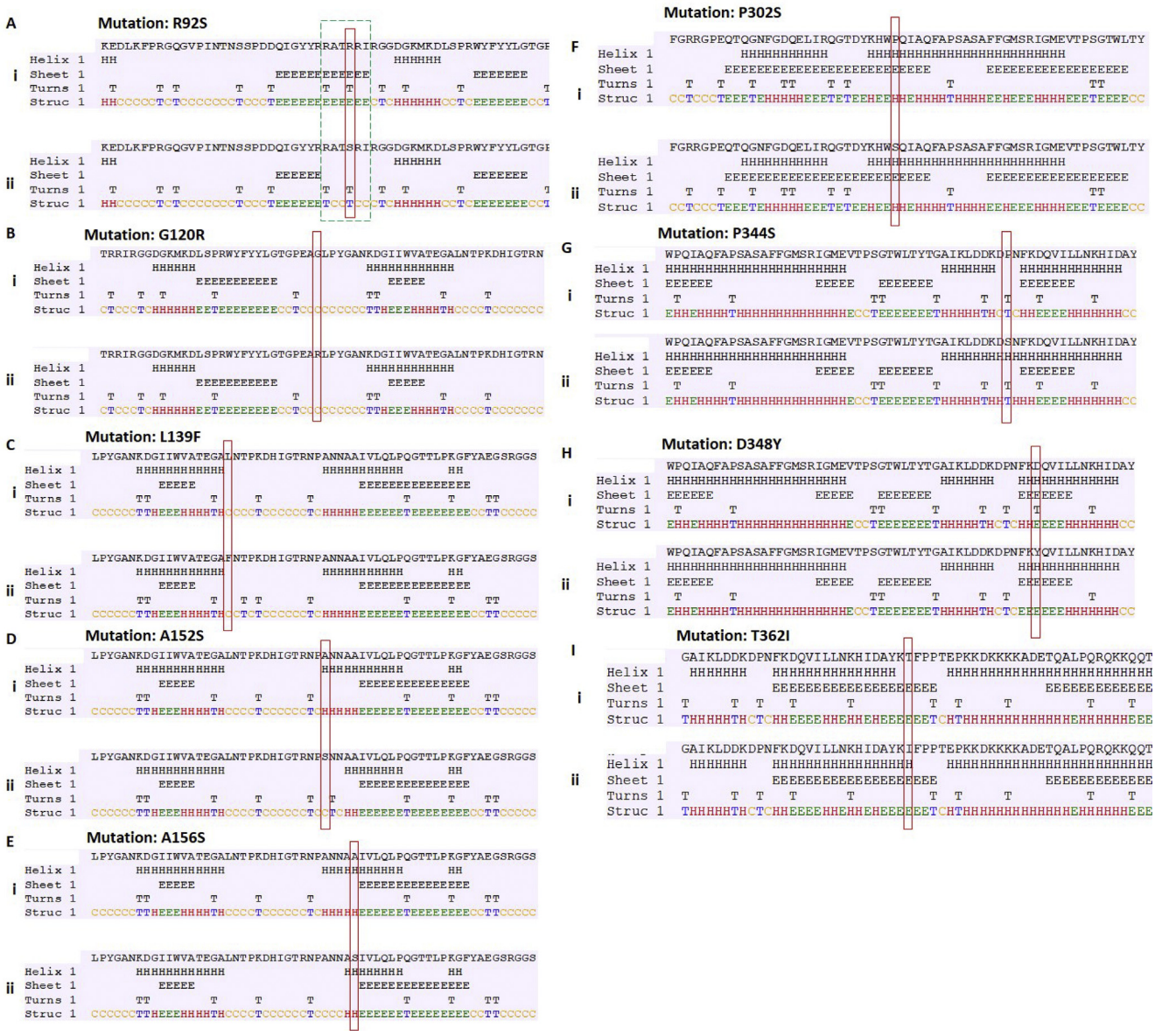

Figure 2. Prediction of the secondary structure of N protein. The panels A, B, C, D, E, F, G, H and I indicate nine mutations observed among Indian isolates that are localised in NTD and CTD of N protein. Panel A (R92S), B (G120R), C (L139F), D (A152S), E (A156S) and F (P302S) represent NTD mutations while, G (P344S), H (D348Y) and I (T362I) represent mutations in the CTD of N protein. In each panel, (i) represents the sequence of Wuhan isolate and (ii) represents sequence of Indian isolates. Both part (i) and (ii) shows the contribution of individual amino acid to the secondary structure of N protein (H- helix, T-turn, E-beta sheet). The location of mutant residue is highlighted in red box for better comparisons. We observed change in the secondary structure between Wuhan, and Indian isolate at six mutation sites (Panel A, D, E, G, H and I). The figure of each panel (A–I) was obtained from CFSSP webserver.

Figure 1. Subsequently, we also analysed the frequency of twenty mutations among Indian isolates (Table 1). Our data revealed S194L and P13L are most frequently mutated while rest of the mutants have low frequency (Table 1). Altogether, we have identified twenty mutations in N protein present among Indian isolates.

3.2. Mutations residing in NTD and CTD of N protein are leading to change in the secondary structure

Subsequently, to understand the possible implication of the identified mutations on N protein structure, we evaluated the secondary structure of mutant residues and compared with the wild-type sequence. Since, NTD and CTD only have defined secondary structure; therefore, we analysed only those mutations which reside in these two domains of N protein. Out of the 20 mutations identified in this study, nine mutations are present in NTD and CTD. Our data revealed that mutation at six positions alter the secondary structure as shown in Figure 2. We have used CFSSP webserver to computationally analyse the probable effect of mutation on secondary structure. Each panel of Figure 2 comprises of two parts (i and ii), part i represents Wuhan isolate (wild type sequence) and part ii represents Indian Isolate (mutant sequence). Both part i and ii shows the contribution of individual amino acid to the secondary structure of N protein (H- helix, T-turn, E-beta sheet). The location of mutant residue is highlighted in red box for better comparisons. Overall, Figure 2 suggests the effect of identified mutation on the secondary structure of the N protein where the mutant residue resides. Our data suggest that mutation at 92, 152 and 156 (Figure 2A, D and E; compare panel i with ii) positions in NTD led to alteration in the secondary structure. Similarly, mutation at position 344, 348 and 362 (Figure 2G, H and I; compare panel i with ii) alters the secondary structure in CTD. Among these, there

Table 2. Calculations of ΔΔG between wild-type and mutant Nucleocapsid Phosphoprotein. mCSM webserver was used to calculate the predicted ΔΔG. The negative values indicate the destabilization of protein upon mutation.

S. No.	PDB File	Wild-type Residue	Residue Position	Mutant Residue	Predicted ΔΔG (kcal.mol^{-1})	Outcome
1	6vyo.pdb	R	92	S	-1.146	Destabilizing
2	6vyo.pdb	G	120	R	-0.432	Destabilizing
3	6vyo.pdb	L	139	F	-0.94	Destabilizing
4	6vyo.pdb	A	152	S	-0.3	Destabilizing
5	6vyo.pdb	A	156	S	-1.422	Destabilizing
6	6wji.pdb	P	302	S	-1.475	Destabilizing
7	6wji.pdb	P	344	S	-0.346	Destabilizing
8	6wji.pdb	D	348	Y	-0.006	Destabilizing
9	6wji.pdb	T	362	I	-0.1	Destabilizing

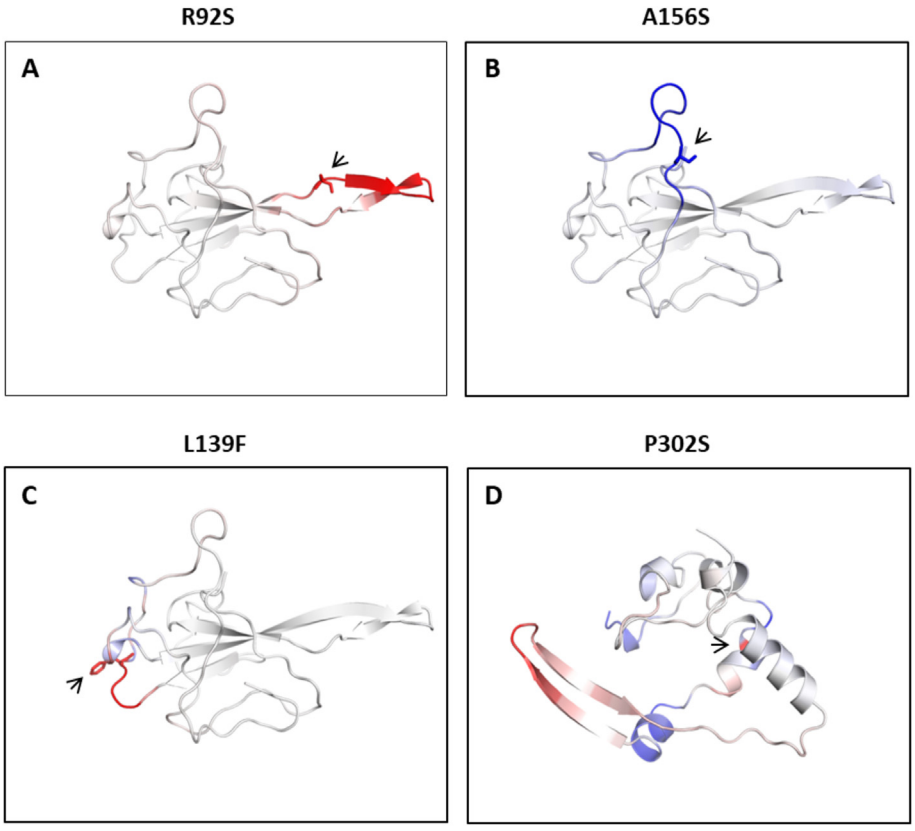

Figure 3. Prediction of the impact of mutations on structural integrity of N protein. Panel A, B, C and D represent the site of respective mutations. The panel A (R92S), B (A156S) and C (L139F) shows the mutations that are localised in the NTD while panel D (P302S) represents CTD mutant. In each panel, the blue color signifies a rigidification in the protein structure, and red represents the gain in flexibility upon mutation. The molecular structures of N protein shown in each panel were downloaded from DynaMut webserver. The arrowhead indicates the location of mutant amino acid in the protein structure.

is a major shift in the secondary structure at 92 position which harbours arginine to serine mutation. Due to this mutation the beta-sheet structure is completely replaced by coiled-coil secondary structure (Figure 2A, compare panel i with ii). Altogether, our data strongly suggest that the mutations identified in this study led to alteration in secondary structure that might affect the functioning of N protein.

3.3. Mutant N proteins have altered protein stability and dynamicity

Since the secondary structures were changed due to the mutations; therefore, we decided to study the effect of these mutations on the stability of the N protein. We used mCSM tool that provide the impact of mutation on the dynamic structure by calculating the difference in free energy (ΔΔG) between wild-type and mutant protein. The values more than zero indicates stabilising mutation and values below zero represents destabilising mutation. The analyses were done with those mutations that are present in the NTD and CTD of N protein. As shown in Table 2, the mutations are leading to the destabilisation of protein structure, and the effect was considerably higher at four positions namely, R92S, L139F, A156S and P302S. The ΔΔG for these four positions are negative in the range of 1.0–2.0 kcal.mol^{-1} (Table 2). Subsequently, we used DynaMut webserver to perform protein modelling to analyse the overall protein structural flexibility upon mutation at four positions that exhibited the maximum change in ΔΔG. Our data revealed that R92S led to gain in flexibility (Figure 3A) and A156S increased the rigidification (Figure 3B). However, L39F and P320S mutations caused both rigidity and flexibility in some areas of the protein structure (Figure 3C and D).

Since, our data show that mutations are altering ΔΔG as well as the protein dynamicity (rigidity and flexibility). To substantiate these data, we sought to analyse the impact of the mutations on intramolecular interactions made by mutant amino acids in the N protein. Our analysis revealed that the mutations were notably affecting the intramolecular bonds in the pockets where these amino acids reside (Figure 4). Specifically, the R92S mutant lost almost all interactions with its neighbouring amino acids (Figure 4). Minor changes in intramolecular interactions were also visible at other mutant positions (Figure 4). Altogether, our

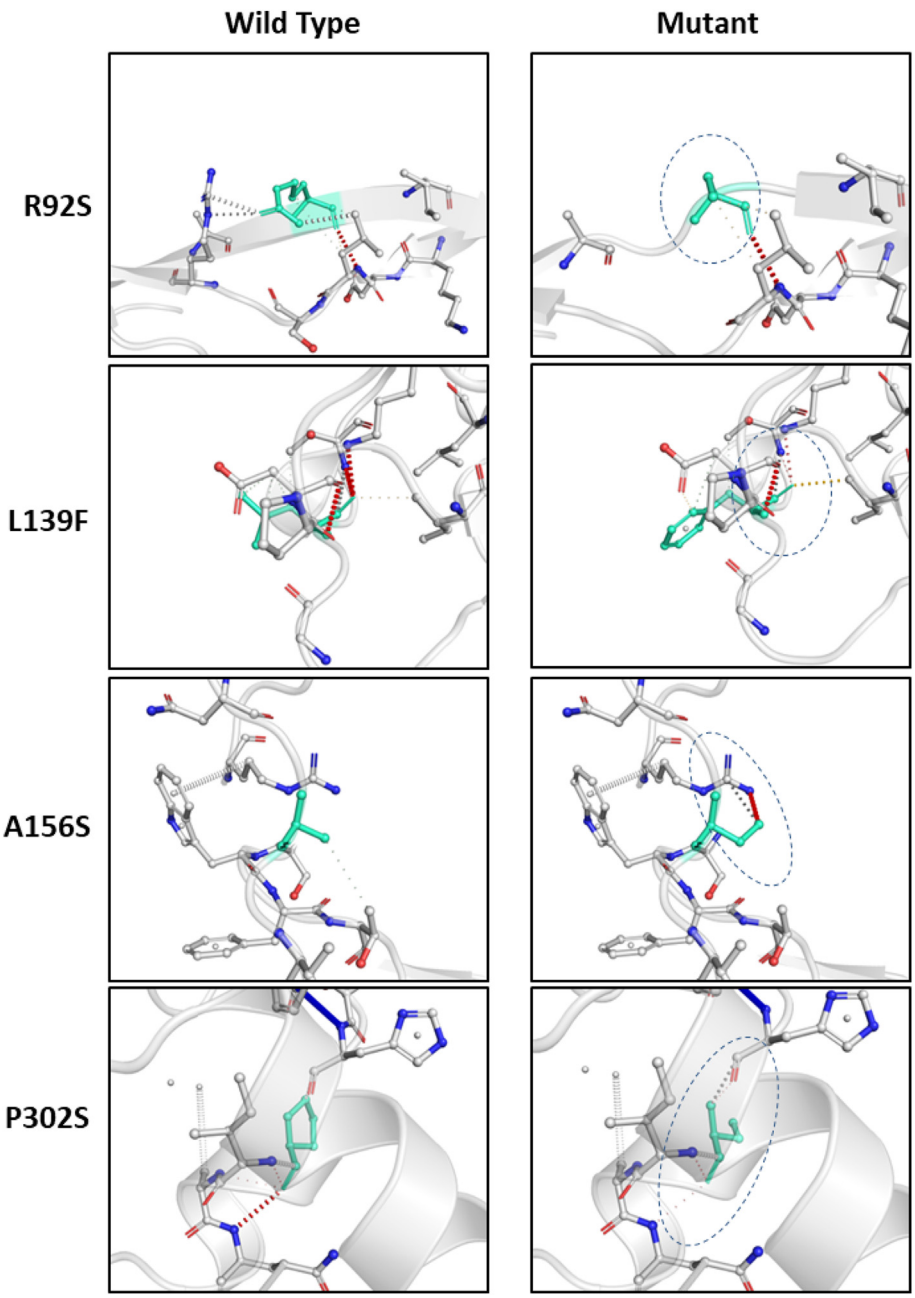

Figure 4. Visual representation of close range intramolecular interactions contributed by the highlighted residues in their respective three-dimensional positions. The molecular structures of wild-type and mutant amino acids of N protein are highlighted in the light-green color. The surrounding residues which are making close contacts with the wild-type and mutant residues are also depicted. The localisation of four mutants is shown from top to bottom that includes R92S, A156S, L139F and P302S respectively. Panel in the left side represents wild-type while the right side represents mutant version. The structures were downloaded from DynaMut webserver. The dashed circle highlights the variation in intramolecular bonds between the wild-type and mutant structure.

data strongly suggest that N protein mutations led to alteration in its structure and dynamicity.

4. Discussions

The SARS-CoV-2 is an RNA virus and its genome is replicated by RdRp that have weak proofreading activity. Therefore, the RNA viruses have relatively higher level of mutations than DNA viruses as they replicate [3]. These mutations enable them to survive in the host cells and also adapt them to new locations and environmental conditions. The SARS-CoV-2 pandemic rapidly spread after the first reported case from Wuhan, China and within 3 months it reached to almost all countries worldwide. As the virus migrated to new locations they also attained mutations. In order to understand the mutations occurring in India, we compared the N protein sequences of SARS-CoV-2 from Indian patients with the earliest known sequence from China. Our data revealed twenty mutations among Indian isolates that are distributed all over the protein. However, we found a distinct cluster in the LKR which is an IDR between the NTD and CTD (Figure 1). IDRs do not have a well-defined tertiary structure in the native form; however, IDRs of coronavirus N proteins play important role in binding with the viral genomic RNA [16]. LKR also plays primary role in the intracellular signalling [17, 18, 19]. Further, we report multiple mutations that reside in the NTD which are known to make contact with the RNA genome of this virus. Previous study shows that the Arg-76 in the infectious bronchitis virus (IBV) N protein is a well conserved across the coronavirus family and required for efficient RNA binding, and may structurally resemble the Arg-94 in the SARS-CoV N protein [20], and possibly Arg-92 of SARS-CoV-2 N protein. Our study revealed that Arg-92 is mutated to Serine and therefore, it might affect N protein and RNA interactions.

We also observed mutations in CTD that could possibly affect its dimerization potential. Earlier report shows that 302, 347, 352, 358 and other residues are involved in dimerization in SARS-CoV N protein [21]. This study identified mutations at 302, 344, 348 and 362 residues which

are in the close vicinity of these critical residues. Overall, our data predicts the effect of mutation on the stability of the N protein. Although, we lack 'experimental validation' of our computational data; however, based on the computational methods, we have identified twenty mutations and concluded that among those twenty, four mutations (R92S, L139F, A156S and P302S) are most likely to affect the structure and stability of the N protein. The secondary structure prediction tool shows that mutation at several locations (92, 152 and 156) will either favour helix/beta sheet or turn structure. This means that these mutations might cause shift in the secondary structure that can affect the stability of the mutant protein. Subsequently, the stability of protein was computationally analysed by DynaMut protein stability prediction tool.

Reports suggest that coronavirus N proteins also acts as an RNA chaperones [22, 23] and helps in the maintenance and proper folding of the RNA genome. Therefore, drugs that can interfere with N protein function are of great pharmacological interests. The drugs that inhibit oligomerisation of N protein via its CTD are promising candidates to inhibit virus assembly [24]. A recent study with SARS-CoV-2 N protein identified ten promising drugs that can inhibit its function including Conivaptan, Ergotamine, Venetoclax, Rifapentine and others [25]. Alanine at 156th position is among the critical residues that interacts with these promising drug candidates [25]. Our study shows mutation at 156th position; thereby, it could possibly cause alteration in drug binding pocket of N protein. Therefore, we predict that few of these mutants might have differential interaction with the drugs, and that can possibly contribute to the drug resistance.

Declarations

Author contribution statement

Gajendra Kumar Azad: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

Data included in article/supplementary material/referenced in article.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2021.e06167.

Acknowledgements

We would like to acknowledge Patna University, Patna for infrastructural support.

References

[1] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G.F. Gao, W. Tan, A novel coronavirus from patients with pneumonia in China, 2019, N. Engl. J. Med. (2020).
[2] R. Lu, X. Zhao, J. Li, P. Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu, Y. Bi, X. Ma, F. Zhan, L. Wang, T. Hu, H. Zhou, Z. Hu, W. Zhou, L. Zhao, J. Chen, Y. Meng, J. Wang, Y. Lin, J. Yuan, Z. Xie, J. Ma, W.J. Liu, D. Wang, W. Xu, E.C. Holmes, G.F. Gao, G. Wu, W. Chen, W. Shi, W. Tan, Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding, Lancet (2020).
[3] S. Duffy, Why are RNA virus mutation rates so damn high? PLoS Biol. (2018).
[4] M. Laamarti Jr., R. Medical, Large scale genomic analysis of 3067 SARS-CoV-2 genomes reveals a clonal geodistribution and a rich genetic variations of hotspots mutations, Biorxiv (2020).
[5] M. Pachetti, B. Marini, F. Benedetti, F. Giudici, E. Mauro, P. Storici, C. Masciovecchio, S. Angeletti, M. Ciccozzi, R.C. Gallo, D. Zella, R. Ippodrino, Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant, J. Transl. Med. (2020).
[6] G.B. Chand, A. Banerjee, G.K. Azad, Identification of novel mutations in RNA-dependent RNA polymerases of SARS-CoV-2 and their implications on its protein structure, PeerJ 8 (2020), e9492.
[7] G.B. Chand, A. Banerjee, G.K. Azad, Identification of twenty-five mutations in surface glycoprotein (Spike) of SARS-CoV-2 among Indian isolates and their impact on protein dynamics, Gene Rep. 21 (2020) 100891.
[8] C.K. Chang, M.H. Hou, C.F. Chang, C.D. Hsiao, T.H. Huang, The SARS coronavirus nucleocapsid protein - Forms and functions, Antivir. Res. (2014).
[9] M. Surjit, S.K. Lal, The SARS-CoV nucleocapsid protein: a protein with multifarious activities, Infect. Genet. Evol. (2008).
[10] G.K. Azad, Identification of novel mutations in the methyltransferase complex (Nsp10-Nsp16) of SARS-CoV-2, Biochem. Biophys. Rep. (2020).
[11] F. Madeira, Y.M. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A.R.N. Tivey, S.C. Potter, R.D. Finn, R. Lopez, The EMBL-EBI search and sequence analysis tools APIs in 2019, Nucleic Acids Res. (2019).
[12] T. Ashok Kumar, CFSSP: Chou and Fasman Secondary Structure Prediction Server, Wide Spectr, 2013.
[13] D.E.V. Pires, D.B. Ascher, T.L. Blundell, MCSM: predicting the effects of mutations in proteins using graph-based signatures, Bioinformatics (2014).
[14] S. Kang, M. Yang, Z. Hong, L. Zhang, Z. Huang, X. Chen, S. He, Z. Zhou, Z. Zhou, Q. Chen, Y. Yan, C. Zhang, H. Shan, S. Chen, Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites, Acta Pharm. Sin. B (2020).
[15] C.H.M. Rodrigues, D.E.V. Pires, D.B. Ascher, DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability, Nucleic Acids Res. (2018).
[16] C.-K. Chang, Y.-L. Hsu, Y.-H. Chang, F.-A. Chao, M.-C. Wu, Y.-S. Huang, C.-K. Hu, T.-H. Huang, Multiple nucleic acid binding sites and intrinsic disorder of severe acute respiratory Syndrome coronavirus nucleocapsid protein: implications for ribonucleocapsid protein packaging, J. Virol. (2009).
[17] S.A. Stohlman, R.S. Baric, G.N. Nelson, L.H. Soe, L.M. Welter, R.J. Deans, Specific interaction between coronavirus leader RNA and nucleocapsid protein, J. Virol. (1988).
[18] M.M. Parker, P.S. Masters, Sequence comparison of the N genes of five strains of the coronavirus mouse hepatitis virus suggests a three domain structure for the nucleocapsid protein, Virology (1990).
[19] R. McBride, M. van Zyl, B.C. Fielding, The coronavirus nucleocapsid is a multifunctional protein, Viruses (2014).
[20] Y.W. Tan, S. Fang, H. Fan, J. Lescar, D.X. Liu, Amino acid residues critical for RNA-binding in the N-terminal domain of the nucleocapsid protein are essential determinants for the infectivity of coronavirus in cultured cells, Nucleic Acids Res. (2006).
[21] C.Y. Chen, C. ke Chang, Y.W. Chang, S.C. Sue, H.I. Bai, L. Riang, C.D. Hsiao, T. huang Huang, Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA, J. Mol. Biol. (2007).
[22] S. Zúñiga, J.L.G. Cruz, I. Sola, P.A. Mateos-Gómez, L. Palacio, L. Enjuanes, Coronavirus nucleocapsid protein facilitates template switching and is required for efficient transcription, J. Virol. (2010).
[23] S. Zúñiga, I. Sola, J.L. Moreno, P. Sabella, J. Plana-Durán, L. Enjuanes, Coronavirus nucleocapsid protein is an RNA chaperone, Virology (2007).
[24] Y.S. Lo, S.Y. Lin, S.M. Wang, C.T. Wang, Y.L. Chiu, T.H. Huang, M.H. Hou, Oligomerization of the carboxyl terminal domain of the human coronavirus 229E nucleocapsid protein, FEBS Lett. (2013).
[25] O. Kadioglu, M. Saeed, H.J. Greten, T. Efferth, Identification of novel compounds against three targets of SARS CoV2 coronavirus by combined virtual screening and supervised machine learning, Bull. World Health Organ. (2020).