TITLE

Identification of novel mutations in RNA-dependent RNA polymerases of SARS-CoV-2 and their implications

AUTHORS

Gyanendra Bahadur Chand[1] and Gajendra Kumar Azad[1#]

AFFILIATION

[1]Department of Zoology, Patna University, Patna-800005, Bihar, India

#To whom correspondence should be addressed: Gajendra Kumar Azad, PhD, Assistant Professor, Department of Zoology, Patna University, Patna-800005, Bihar, India. Email: gkazadpatnauniversity@gmail.com

Keywords: COVID-19, SARS-CoV-2, RNA-dependent RNA polymerases (RdRp), Nsp12, mutation

ABSTRACT

The rapid development of SARS-CoV-2 mediated COVID-19 pandemic has been the cause of significant health concern, highlighting the immediate need for the effective antivirals. SARS-CoV-2 is a RNA virus that has an inherent high mutation rate. These mutations drive viral evolution and genome variability, thereby, facilitating viruses to have rapid antigenic shifting to evade host immunity and to develop drug resistance. Viral RNA-dependent RNA polymerases (RdRp) perform viral genome duplication and RNA synthesis. Therefore, we compared the available RdRp sequences of SARS-CoV-2 from Indian isolates and 'Wuhan wet sea food market virus' sequence to identify, if any variation is present between them. We report three mutations observed in Indian SARS-CoV-2 isolates and one unique mutation at position 323 that showed changes in the secondary structure of the RdRp protein at region of mutation. We also studied molecular dynamics using normal mode analyses and found that this mutation increases the stability of RdRp protein. We propose this unique mutation leading to structural alteration in RdRp protein might have functional consequences that can interferes with RdRp targeting pharmacological agents.

INTRODUCTION

SARS-CoV-2 (a member of Corona viruses) outbreak occurred in Wuhan, China in December 2019 and it became a pandemic in March 2020 that has affected almost all countries worldwide. The SARS-CoV-2 causes COVID-19 disease that has created a global public health problem. As of May 5, 2020, more than 3.6 million confirmed COVID-19 cases were reported worldwide with 0.26 million confirmed deaths. To design antiviral therapeutics/ vaccines it is important to understand the genetic sequence, structure and function of the viral proteins. When a virus tries to adapt to a new host, in a new environment at a distinct geographical location, it make necessary changes in its genetic make-up to bring slight modifications in the nature of its proteins(1). Such variations are known to help viruses to exploit the host's cellular machinery that promotes its survival and proliferation. The β lymphocytes of host's adaptive immune system eventually identify the specific epitopes of the pathogenic antigen and starts producing protective antibodies, which in turn results in agglutination and clearance of the pathogen(2–4). Being an efficient unique pathogen, a virus, often mutate its proteins in a manner that it can still infect

the host cells, evading the host's immune system. Even when fruitful strategies are discovered and engaged, the high rate of genetic change displayed by viruses frequently leads to drug resistance or vaccine escape(5).

The SARS-CoV-2 has a single stranded RNA genome of approximately 29.8Kb in length and possesses 14 ORFs encoding 29 proteins that includes four structural proteins: Envelope (E), Membrane (M), Nucleocapsid (N) and Spike (S) proteins, 16 non-structural proteins (nsp) and 9 accessory proteins(6),(7) including the RNA dependent RNA polymerase (RdRp) (also named as nsp12). RdRp is comprised of multiple distinct domains that catalyze RNA-template dependent synthesis of phosphodiester bonds between ribonucleotides. The SARS-CoV-2 RdRp is an important constituent of the replication/transcription machinery. The structure of the SARS-CoV-2 RdRp has been recently solved(8). This protein contains a "right hand" RdRp domain (residues S367-F920), and a nidovirus-unique N-terminal extension domain (residues D60-R249). The polymerase domain and N-terminal domain are connected by an interface domain (residues A250-R365).

For RNA viruses, the RNA-dependent RNA polymerase (RdRp) presents an ideal target because of its vital role in RNA synthesis and absence of host homolog. RdRp is therefore, a primary target for antiviral inhibitors such as remdesivir(9) that is being considered as a potential drug for the treatment of COVID-19. Corona viruses have been known to undergo rapid mutations in its RNA genome. Since, RNA viruses constantly evolve, therefore, we decided to analyse RdRp protein sequence of SARS-CoV-2 from different geographical region to see if RdRp also mutate. Here, in the present study, we identified and characterised three mutations in RdRp protein isolated from India against the 'Wuhan wet sea food market'(7) SARS-CoV-2. Our data show that one mutation, out of three, causes considerable change in the RdRp structural dynamics. Therefore, the prevalence of mutations in the genome of SARS-CoV-2 needs to be considered to develop new approaches for targeting this virus.

RESULTS:

Identification of mutations in RdRp protein present in Indian isolates

The SARS-COV-2 sequencing data was downloaded from NCBI (NCBI-Virus-SARS-CoV-2 data hub). There are 14 SARS-CoV-2 sequences available from India. We downloaded all Indian SARS-CoV-2 sequences and Wuhan SARS-CoV-2 sequence, which was deposited for the first time after COVID-19 cases started to appear in

Wuhan province, China. We downloaded protein sequence of NSP12/ RDRP from the database. All these RDRP sequences were first aligned in CLUSTAL Omega to check for similarities or differences. The Clustal Omega algorithm produces a multiple sequence alignment by producing pairwise alignments. Using this algorithm, we identified three mutations in isolates from Indian SARS-CoV-2 samples compared to Wuhan SARS-CoV-2 RDRP sequence as shown in figure 1. We considered the original Wuhan sequence as the wild type for this comparison. These mutations on RdRp are A97V, P323L and A466V respectively.

P323L mutation causes the alteration in secondary structure of RdRp

Next, we studied the effect of these mutations on RdRp. At two positions (97 and 466) the alanine amino acid is substituted by valine. The valine side chain is larger than alanine, and substitution of valine at position 97 and 466 might impair packing of the protein; however, our secondary structure predications fail to predict any alteration in secondary structure in and around these two mutations (figure 2A and 2C, compare panel i and ii). Interestingly, the secondary structure prediction showed changes in secondary structure when proline is substituted by leucine at 323 position (Figure 2B, compare panel i and ii). The detailed analysis revealed that the mutant RdRp (P323L) have attained considerable changes in secondary structure at the mutation site. There is a loss of turn structures from position 323 and 324 and addition of five sheets at positions 321, 322, 323, 324 and 327 (Figure 2B, highlighted in green box). Proline possesses a unique property, its side chain cyclizes back on to the backbone amide position. Structurally, proline amino acid contributes to the secondary structure formation because of its bulky pyrrolidine ring places steric constraints on the conformation of the preceding residue in the helix. The substitution of proline to leucine in the mutant RdRp might losses the structural integrity provided by proline. Leucine is a hydrophobic amino acid and generally buried in the folded protein structure. Therefore, the substitution of leucine in mutant RdRp is changing the secondary structure of the protein as show in figure 2B that might have functional consequences.

P323L alters the stability dynamics of tertiary structure of RdRp.

To correlate the reflection of changes in secondary structure, if any, in the dynamics of the protein in its tertiary structure, we used DynaMut software(10) and performed

normal mode analyses and studied protein stability and flexibility. Our data revealed that the change in vibrational entropy energy (ΔΔSVib ENCoM) between the wild type (Wuhan isolate) and the mutant Indian isolate was -0.277 $kcal.mol^{-1}.K^{-1}$ (Figure 3A). Vibration entropy represents the major configurational-entropy contribution of proteins. By definition, it is an average of the configurational entropies of the protein within single minima of the energy landscape, weighted by their occupation probabilities(11). The negative ΔΔSVib ENCoM of mutant RdRp represents the rigidification of the protein structure. Further, we also calculated the free energy differences, ΔΔ*G,* between wild-type and mutant as 0.908 kcal/mol. The free energy differences, ΔΔ*G*, caused by mutation have been correlated with the structural changes, such as changes in packing density, cavity volume and accessible surface area and therefore, it measures effect of mutation on protein stability(12). In general, a ΔΔG below zero means that the mutation causes destabilization and above zero represents the protein stabilization. Here, our analysis show the positive ΔΔ*G* suggesting that P323L mutation is stabilising the protein structure (Figure 3A).

We further closely analysed the changes in the intramolecular interactions due to P323L mutation. Our data showed that it was also affecting the interactions of the residues which were present in the close vicinity of proline. The substitution of proline to valine alters the side chain leading to change of intramolecular bonds in the pocket, where this amino acid resides as shown in figure 3B and 3C.

Next, we analysed the atomic fluctuations and deformation energies of wild type and mutant RdRp (Figure 4). Atomic fluctuations calculate the measure of absolute atomic motion whereas the deformation energies detect the measure of flexibility of a protein. Figure 4 shows the visual representations of the atomic fluctuation and deformation energies and the specific positions that could be visibly detected to be different have been marked. Consistent with the above results the mutant RdRp have at least three positions which shows alteration in atomic fluctuations and deforming energies (Figure 4, highlighted by dashed circle) compared to the wild type protein. Therefore, it can be conclusively stated that the P323L mutation is in large changing the stability and intramolecular interactions in the protein and that might have functional consequences.

DISCUSSION

RNA viruses exploit all known mechanisms of genetic variation to ensure their survival. Distinctive features of RNA virus replication include high mutation rates, high yields, and short replication times (13). Mutation rates determine the amount of genetic variation generated in a population, which is the material upon which natural selection that can act (14). For this reason, a higher mutation rate correlates with a higher evolutionary rate. The mutation rates also determine the probability that a mutation conferring to drug resistance, antibody escape, or expanded host range will arise (15). Additionally, mutation rates can determine whether a virus population will be susceptible to drug-induced lethal mutagenesis.

The SARS-CoV-2 infected humans from Wuhan province of China and quickly spread to almost all countries worldwide. Since, this virus has already spread to different demographic areas having different climatic condition, temperature, humidity and seasonal variation; therefore, we can predict that this virus might be mutating to adapt to new environments. Towards this approach, we investigated the muations in RdRp of SARS-CoV-2. We focused primarily on RdRp because it is an essential protein of this virus and helps in its replication and transcription. More importantly, there are many drugs which specifically targets RdRp and are potent antivirals. Our study is the first to report the RdRp mutations present in the Indian isolates of SARS-CoV-2 and it will help to understand the effect of variations on RdRp protein. Our data demonstrate that P323L mutation leads to significant change in the protein secondary structure (Figure 2). This mutation lies in the interface domain (residues A250-R365) of the RdRp protein. This domain helps in the coordination of N and C terminal domains of RdRp; therefore, a mutation in the interface domain might have drastic impact on the function of RdRp. A recent virtual molecular docking studies, screened approximately 7500 drugs to identify SARS-CoV-2 RdRp inhibitor revealed several potential compounds (DOI: 10.20944/preprints202003.0024.v1) namely, Simeprevir, Filibuvir and Tegobuvir, etc. Same study also predicted that these drugs bind RdRp at a putative docking site (a hydrophobic cleft) that includes phenylalanine at 326th position. Interestingly, the mutation identified in our study is very close to the docking site (P323L). Therefore, it is reasonable that the substitution of P323L might interfere with the interaction of these drugs with RdRp. Further, our data also revealed that this mutation is causing stabilisation of the protein structure. The mutations in RdRp have already been linked to drug resistance in different viruses. Such as a study on

RdRp of influenza A virus demonstrate that a K229R mutation confers resistance to favipiravir(16). Similarly, mutation in hepatitis C virus RdRp at P495, P496 or V499 & T389 and have been linked to drug resistance(17). Altogether, it is conceivable that many RNA viruses acquire drug resistance through mutations in RdRp. However, the functional characterisation of RdRp mutation investigated in our study needs to be carried out to understand the exact role of P323L mutation. These variations might lead to virus diversification and eventual emergence of variants or antibody escape mutants. Therefore, elaborate studies on SARS-CoV-2 protein sequence variations needs to be carried out that will help scientific community in better therapeutic targeting.

METHODS

Sequence retrieval:

We downloaded all SARS-CoV2 sequences from the NCBI virus database as shown in table 1. As of now, there are 16 SARS-CoV-2 sequences from India have been deposited in this database. Out of 16, two sequences are not complete; therefore, we retrieved 14 datasets of SARS-CoV-2 (samples are from Indian). As a reference, we downloaded the sequence of SARS-CoV-2 that was first reported genome sequence deposited in the NCBI virus database from the ‘wet sea food market area’ from the early days of COVID-19 pandemic(7). This virus was formerly called 'Wuhan seafood market pneumonia virus and the accession number is NC_045512.

Sequence alignments and structure:

All the RDRP protein sequences were aligned by multiple sequence alignment platform of CLUSTAL Omega(18). Clustal Omega is a multiple sequence alignment program that uses seeded guide trees and HMM profile-profile techniques to generate alignments between three or more sequences. The alignment file was carefully studied and differences in the amino acid changes were recorded.

Secondary structure predictions:

We used CFSSP(19) (Chou and Fasman secondary structure prediction), an online server, to predict secondary structures of SARS-CoV-2 RdRp protein. This server predicts the possibility of secondary structure such as alpha helix, beta sheet, and turns from the amino acid sequence. CFSSP uses Chou-Fasman algorithm, which is

based on analyses of the relative frequencies of each amino acid in secondary structures of proteins solved with X-ray crystallography.

RdRp dynamics study:

To investigate the effect of mutation on the RdRp protein structural conformation, its molecular stability and flexibility, we used DynaMut software(10) (University of Melbourne, Australia). To run this software we first downloaded the known protein structure of SARS-CoV-2 RdRp from RCSB (PDB ID: 6M71(8)) and used it for analysis. Next, the 6M71 structure was uploaded on DynaMut software and effect of mutation in various protein structure stability parameters such as vibrational entropy; the atomic fluctuations and deformation energies were determined. DynaMut, a web server implements well established normal mode approaches that is used to analyse and visualise protein dynamics. This software samples conformations and measures the impact of mutations on protein dynamics and stability resulting from vibrational entropy changes and also predicts the impact of a mutation on protein stability. DynaMut provides a comprehensive suite for protein motion and flexibility analysis and visualisation. Here, we used DynaMut for atomic fluctuation and deformation energy calculations of RdRp protein. The calculations were performed over first ten non-trivial modes of the molecule.

ACKNOWLEDGEMENTS

We would like to thank Patna University, Patna for providing necessary infrastructural support. No funding was used to carry out this research.

CONFLICT OF INTEREST

Authors declare no conflict of interests.

FIGURE LEGENDS

Figure 1: Multiple sequence alignment of Wuhan SARS-CoV-2 RDRP protein with sequences obtained from India. The mutations are highlighted in red box. Only those sequences are shown that have variations, rest of the sequences are identical among all samples.

Figure 2: Effect of mutations on secondary structure of RDRP. Figure A, B and C demonstrate three mutations observed in Indian isolates. Panel (i) represents sequence of Wuhan isolate and panel (ii) represents sequence of Indian isolates. We observed difference of secondary structure between Wuhan and Indian isolate at P323L position only, as highlighted in green box.

Figure 3: A) Δ Vibrational Entropy Energy between Wild-Type and Mutant (P323L) RdRp, amino acids are colored according to the vibrational entropy change as a consequence of mutation (P323L) of RDRP protein. BLUE represents a rigidification of the structure. B) Interatomic interactions- Wild-type and mutant (P323L) residues are colored in light-green and are also represented as sticks alongside with the surrounding residues which are involved on any type of interactions.

Figure 4 (A and B): Visual analysis of Atomic Fluctuation, the calculations performed over the first 10 non-trivial modes of the molecule. The magnitude of the fluctuation is represented by thin to thick tube colored blue (low) and red (high). The difference between wild-type and mutant (P323L) are highlighted in dashed circle. (C and D): Visual analysis of Deformation Energies, the calculations performed over the first 10 non-trivial modes of the molecule. The magnitude of the deformation is represented by thin to thick tube colored blue (low) and red (high). The difference between wild-type and mutant (P323L) are highlighted in dashed circle.

REFERENCES

1. Sackman AM, McGee LW, Morrison AJ, Pierce J, Anisman J, Hamilton H, et al. Mutation-driven parallel evolution during viral adaptation. Mol Biol Evol. 2017;
2. Alcami A, Koszinowski UH. Viral mechanisms of immune evasion. Trends in Microbiology. 2000.
3. Jensen S, Thomsen AR. Sensing of RNA Viruses: a Review of Innate Immune Receptors Involved in Recognizing RNA Virus Invasion. J Virol. 2012;
4. Chaplin DD. Overview of the immune response. J Allergy Clin Immunol. 2010;
5. McKeegan KS, Borges-Walmsley MI, Walmsley AR. Microbial and viral drug resistance mechanisms. Trends in Microbiology. 2002.
6. Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, O'Meara MJ, et

al. A SARS-CoV-2-Human Protein-Protein Interaction Map Reveals Drug Targets and Potential Drug-Repurposing. bioRxiv. 2020;

7. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. Nature. 2020;
8. Gao Y, Yan L, Huang Y, Liu F, Zhao Y, Cao L, et al. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. Science (80-). 2020;
9. Gordon CJ, Tchesnokov EP, Woolner E, Perry JK, Feng JY, Porter DP, et al. Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency. J Biol Chem. 2020;
10. Rodrigues CHM, Pires DEV, Ascher DB. DynaMut: Predicting the impact of mutations on protein conformation, flexibility and stability. Nucleic Acids Res. 2018;
11. Goethe M, Fita I, Rubi JM. Vibrational entropy of a protein: Large differences between distinct conformations. J Chem Theory Comput. 2015;
12. Eriksson AE, Baase WA, Zhang XJ, Heinz DW, Blaber M, Baldwin EP, et al. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. Science (80-). 1992;
13. Domingo E, Holland JJ. RNA VIRUS MUTATIONS AND FITNESS FOR SURVIVAL. Annu Rev Microbiol. 1997;
14. Duffy S. Why are RNA virus mutation rates so damn high? PLoS Biol. 2018;
15. Irwin KK, Renzette N, Kowalik TF, Jensen JD. Antiviral drug resistance as an adaptive process. Virus Evol. 2016;
16. Goldhill DH, Te Velthuis AJW, Fletcher RA, Langat P, Zambon M, Lackenby A, et al. The mechanism of resistance to favipiravir in influenza. Proc Natl Acad Sci U S A. 2018;
17. Delang L, Froeyen M, Herdewijn P, Neyts J. Identification of a novel resistance mutation for benzimidazole inhibitors of the HCV RNA-dependent RNA polymerase. Antiviral Res. 2012;
18. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 2019;
19. Ashok Kumar T. CFSSP: Chou and Fasman Secondary Structure Prediction server. Wide Spectr. 2013;

Table 1: Details of SARS-CoV-2 sequences used in the analysis

S.No	Accession Number	authors	Institute/ University
1	NC_045512	Wu,F., Zhao,S., et al., 2020	Fudan University, Shanghai, China
2	MT012098	Yadav,P.D., et al., 2020	NIV, Pashan, Pune, Maharashtra 411021, India
3	MT050493	Potdar,V., et al., 2020	NIV, Pashan, Pune, Maharashtra 411021, India
4	MT358637	Pandit,R.,Shah,T., et al., 2020	GBRC, Gandhinagar, Gujarat 382010, India
5	MT396242	Pattabiraman,C., et al., 2020	NIMHANS, Bangalore, Karnataka 560029,India
6	MT396243	Pattabiraman,C., et al., 2020	NIMHANS, Bangalore, Karnataka 560029,India
7	MT396244	Pattabiraman,C., et al., 2020	NIMHANS, Bangalore, Karnataka 560029,India
8	MT396245	Pattabiraman,C., et al., 2020	NIMHANS, Bangalore, Karnataka 560029,India
9	MT396246	Pattabiraman,C., et al., 2020	NIMHANS, Bangalore, Karnataka 560029,India
10	MT396247	Pattabiraman,C., et al., 2020	NIMHANS, Bangalore, Karnataka 560029,India
11	MT396248	Pattabiraman,C., et al., 2020	NIMHANS, Bangalore, Karnataka 560029,India
12	MT415320	Muttineni,R., et al., 2020	VRL, Dept. of Zoology, Osmania University, Hyderabad-500007,India
13	MT415321	Muttineni,R., et al., 2020	VRL, Dept. of Zoology, Osmania University,

			Hyderabad-500007,India
14	MT415322	Muttineni,R., et al., 2020	VRL, Dept. of Zoology, Osmania University, Hyderabad-500007,India
15	MT415323	Muttineni,R., et al., 2020	VRL, Dept. of Zoology, Osmania University, Hyderabad-500007,India

Figure 1:

A

Mutation 1: A97V

```
MT415322    EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVVKHDFFKFRIDGDMVPHISRQRLT  120
MT358637    EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQRLT  120
MT415321    EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQRLT  120
MT415323    EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQRLT  120
Wuhan       EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQRLT  120
MT050493    EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQRLT  120
MT415320    EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQRLT  120
MT012098    EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQRLT  120
MT396244    EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVVKHDFFKFRIDGDMVPHISRQRLT  120
MT396245    EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVVKHDFFKFRIDGDMVPHISRQRLT  120
MT396242    EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQRLT  120
MT396243    EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQRLT  120
MT396247    EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQRLT  120
MT396246    EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVVKHDFFKFRIDGDMVPHISRQRLT  120
MT396248    EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVVKHDFFKFRIDGDMVPHISRQRLT  120
            ************************************.***********************
```

B

Mutation 2: P323L

```
MT415322    CLDDRCILHCANFNVLFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN  360
MT358637    CLDDRCILHCANFNVLFSTVFPLTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN  360
MT415321    CLDDRCILHCANFNVLFSTVFPLTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN  360
MT415323    CLDDRCILHCANFNVLFSTVFPLTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN  360
Wuhan       CLDDRCILHCANFNVLFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN  360
MT050493    CLDDRCILHCANFNVLFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN  360
MT415320    CLDDRCILHCANFNVLFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN  360
MT012098    CLDDRCILHCANFNVLFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN  360
MT396244    CLDDRCILHCANFNVLFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN  360
MT396245    CLDDRCILHCANFNVLFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN  360
MT396242    CLDDRCILHCANFNVLFSTVFPLTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN  360
MT396243    CLDDRCILHCANFNVLFSTVFPLTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN  360
MT396247    CLDDRCILHCANFNVLFSTVFPLTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN  360
MT396246    CLDDRCILHCANFNVLFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN  360
MT396248    CLDDRCILHCANFNVLFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN  360
            ********************** ************************************
```

C

Mutation 3: A466V

```
MT415322    LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFY  420
MT358637    LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFY  420
MT415321    LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFY  420
MT415323    LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFY  420
Wuhan       LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFY  420
MT050493    LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFY  420
MT415320    LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFY  420
MT012098    LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVVFQTVKPGNFNKDFY  420
MT396244    LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFY  420
MT396245    LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFY  420
MT396242    LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFY  420
MT396243    LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFY  420
MT396247    LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFY  420
MT396246    LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFY  420
MT396248    LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFY  420
            ********************************************.**************
```

Figure 2

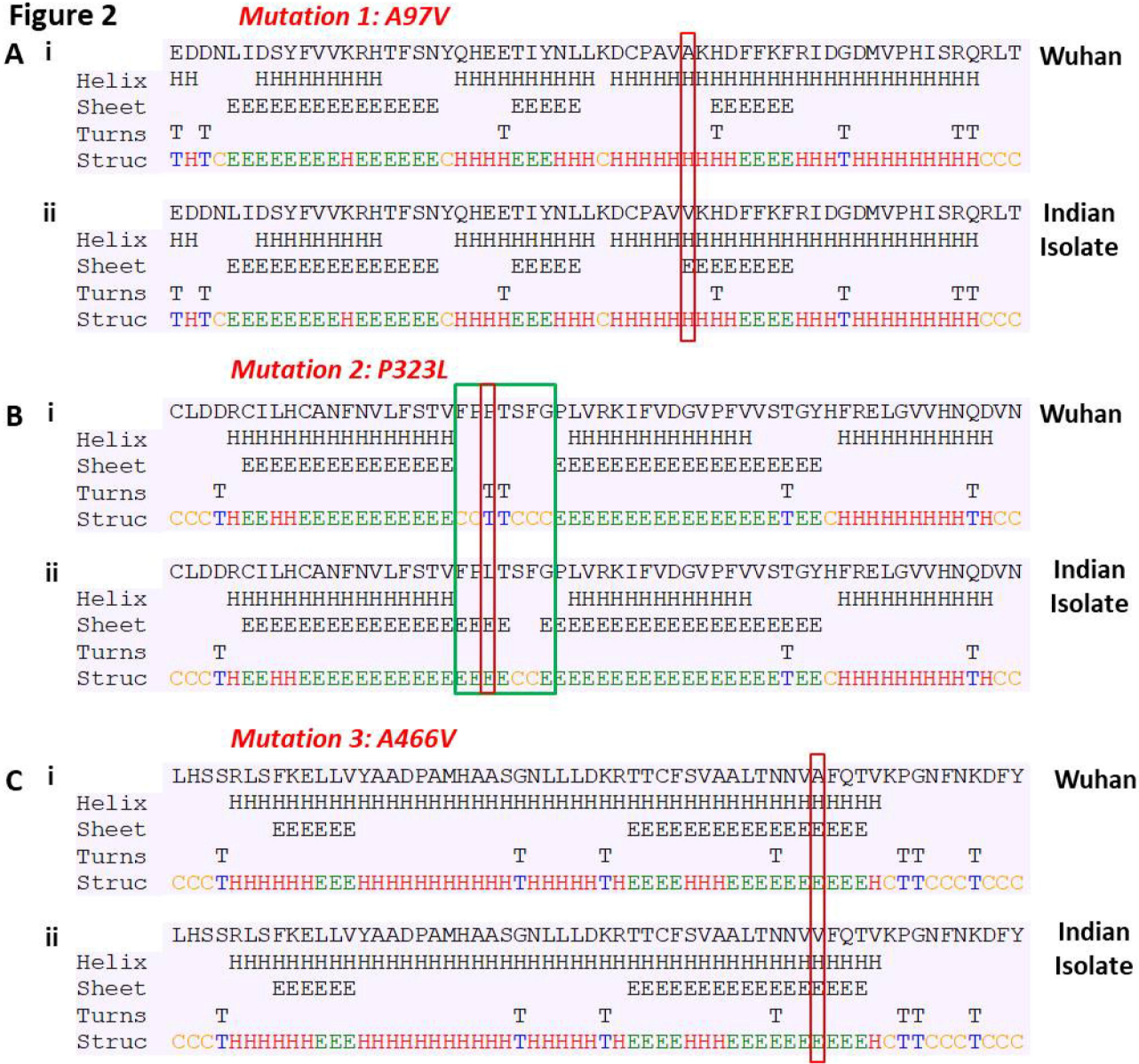
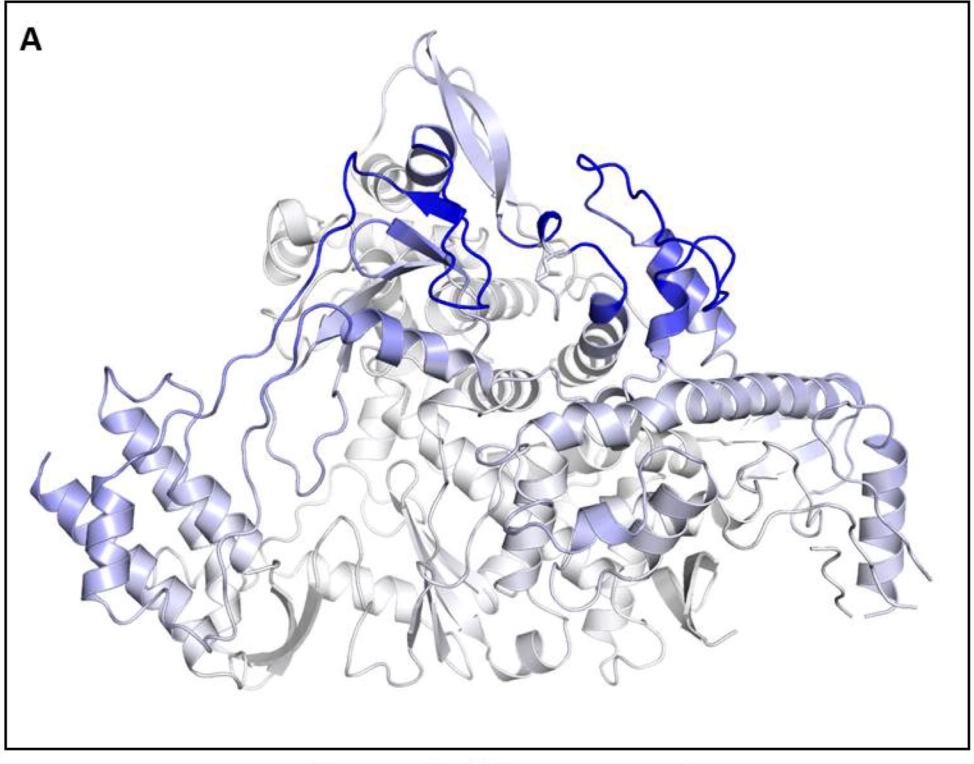
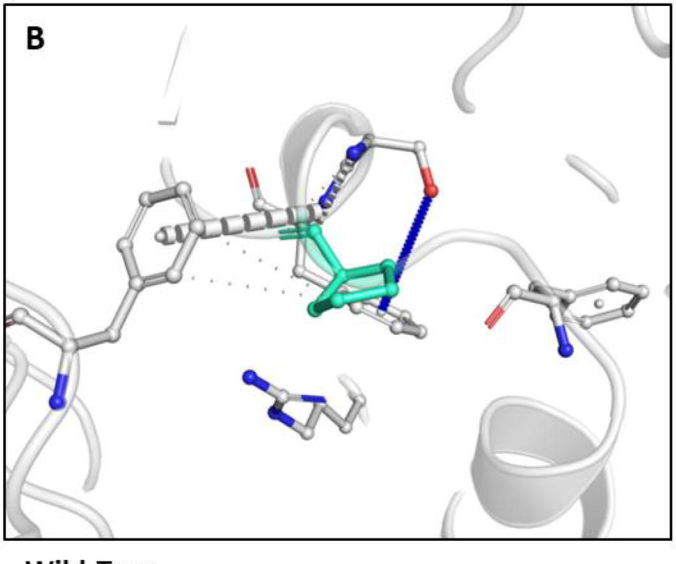
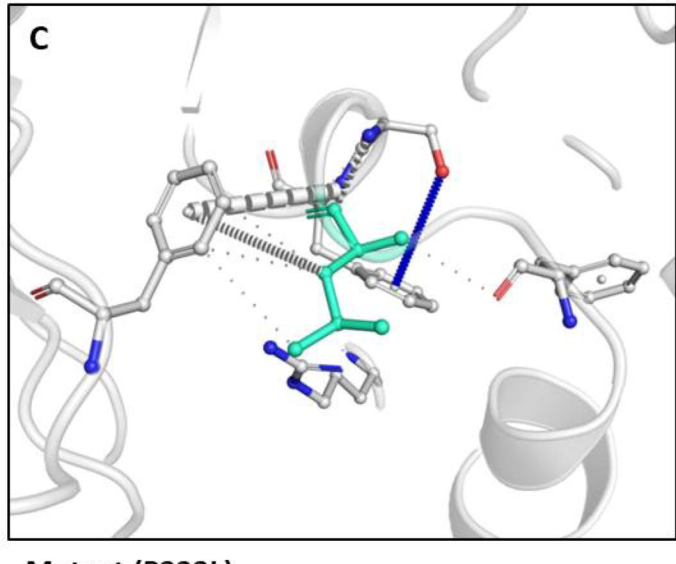

Mutation 1: A97V
A
i
EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQRLT
Helix HH HHHHHHHHH HHHHHHHHHH HHHHHHHHHHHHHHHHHHHHHHHH
Sheet EEEEEEEEEEEEEE EEEEE EEEEEE
Turns T T T T T TT
Struc THTCEEEEEEEEHEEEEEECHHHHEEEHHHCHHHHHHHHHEEEEHHHTHHHHHHHHHCCC
Wuhan
ii
EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVVKHDFFKFRIDGDMVPHISRQRLT
Helix HH HHHHHHHHH HHHHHHHHHH HHHHHHHHHHHHHHHHHHHHHHHH
Sheet EEEEEEEEEEEEEE EEEEE EEEEEEEE
Turns T T T T T TT
Struc THTCEEEEEEEEHEEEEEECHHHHEEEHHHCHHHHHHHHHEEEEHHHTHHHHHHHHHCCC
Indian Isolate
Mutation 2: P323L
B
i
CLDDRCILHCANFNVLFSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN
Helix HHHHHHHHHHHHHHHH HHHHHHHHHHHHH HHHHHHHHHHH
Sheet EEEEEEEEEEEEEE EEEEEEEEEEEEEEEEEE
Turns T TT T T
Struc CCCTHEEHHEEEEEEEEEEECCTTCCCEEEEEEEEEEEEEEEETEECHHHHHHHHHTHCC
Wuhan
ii
CLDDRCILHCANFNVLFSTVFPLTSFGPLVRKIFVDGVPFVVSTGYHFRELGVVHNQDVN
Helix HHHHHHHHHHHHHHHH HHHHHHHHHHHHH HHHHHHHHHHH
Sheet EEEEEEEEEEEEEEEEEE EEEEEEEEEEEEEEEEEE
Turns T T T
Struc CCCTHEEHHEEEEEEEEEEEEEEECCEEEEEEEEEEEEEEEETEECHHHHHHHHHTHCC
Indian Isolate
Mutation 3: A466V
C
i
LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKPGNFNKDFY
Helix HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
Sheet EEEEEE EEEEEEEEEEEEEEEE
Turns T T T T TT T
Struc CCCTHHHHHHEEEHHHHHHHHHHHTHHHHHTHEEEEHHHEEEEEEEEEEHCTTCCCTCCC
Wuhan
ii
LHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVVFQTVKPGNFNKDFY
Helix HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
Sheet EEEEEE EEEEEEEEEEEEEEEE
Turns T T T T TT T
Struc CCCTHHHHHHEEEHHHHHHHHHHHTHHHHHTHEEEEHHHEEEEEEEEEEHCTTCCCTCCC
Indian Isolate

Figure 3







Wild-Type

Mutant (P323L)

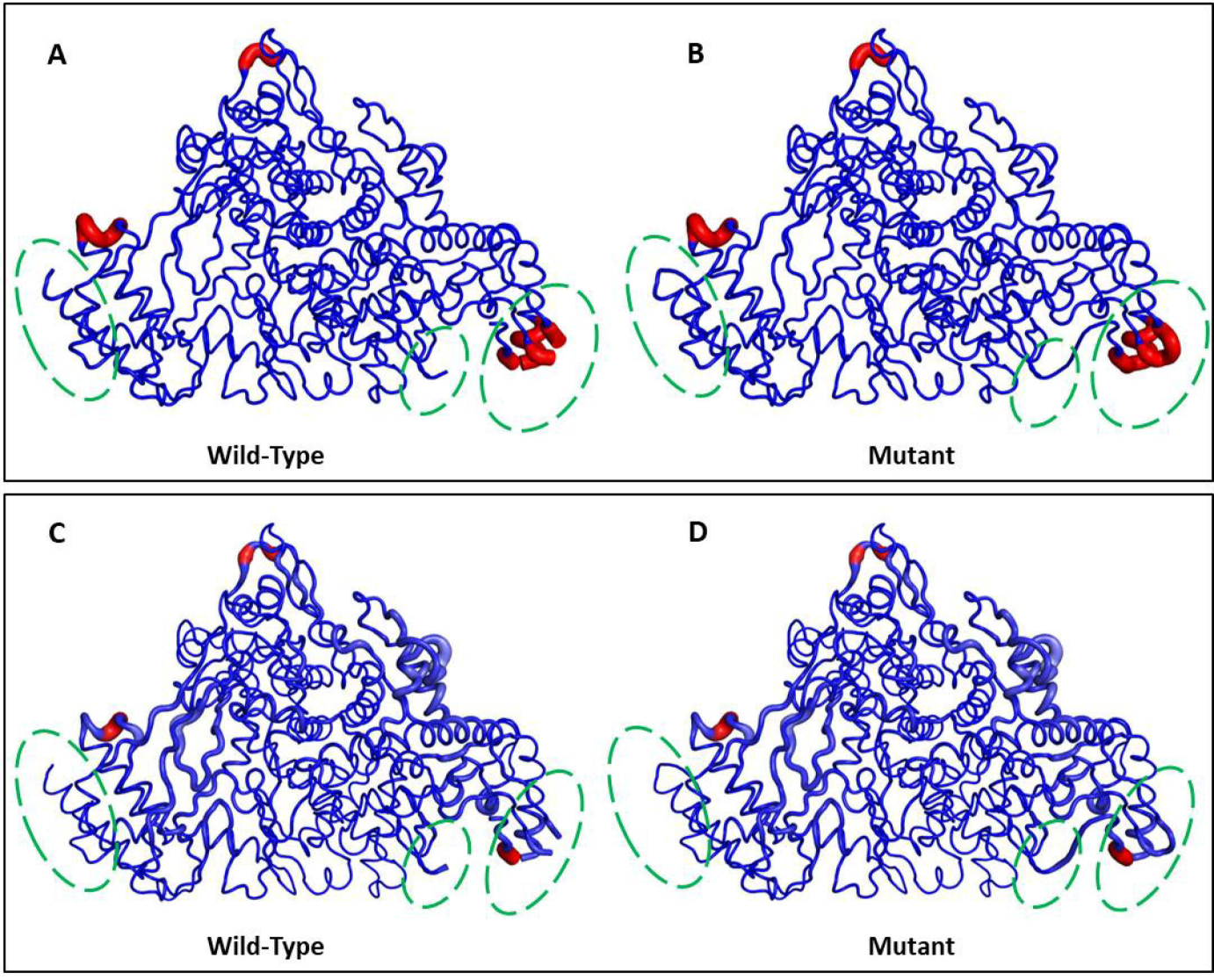



Figure 4