3D Modeling of SARS-Cov-2 RDRP Mutant Proteins in Drug Resistance and Viral Evolution

(Chemistry Award)

Helen H. Zheng

Student Researcher

Watchung Hills Regional High School, Warren, NJ 07059 USA

Stephen K. Burley

Professor and Faculty Mentor

Rutgers Institute for Quantitative Biomedicine (IQB), Piscataway, NJ 08854 USA

TITLE: 3D Modeling of SARS-Cov-2 RDRP Mutant Proteins in Drug Resistance and Viral Evolution

ABSTRACT

SARS-CoV-2 is mutated over time, making it difficult to contain. Central to viral replication is RDRP, a heterotetramer RNA polymerase composed of Nsp12, Nsp7 and Nsp8. RDRP is targeted by remdesivir, an FDA-approved COVID19 drug. Currently, how mutations affect RDRP is poorly understood. In this study, I tested the hypothesis that certain mutations alter RDRP structure and function. To this end, I performed comprehensive computational 3D modeling of 1351 unique sequence variants (USVs) of RDRP from over 90,000 sequenced viral genomes. Most changes are found on protein surfaces and boundary layers, rather than hydrophobic cores. Only a handful substitutions occur in the active sites or protein-protein interfaces. It is noteworthy that D684 of Nsp12 bind to the -1 position of RNA. The D684G substitution could weaken interaction with RNA and alter shape of the active site, lowering replication fidelity or drug affinity. The impact of other notable substitutions (e.g. P323L) on RDRP and viral evolution to more contagious strains is also discussed. Collectively, my analyses also provide possible explanations for how certain USVs impact structure and function of RDRP, promoting viral evolution and drug resistance. These 3D models create a useful framework for future research on antiviral drug development.

KEY WORDS:

SARS-CoV-2, 3D Structure Modeling, Variants, Antiviral Drug Resistance, Viral Evolution

TABLE OF CONTENT:

HIGHLIGHTS OF RESEARCH

- Generated a comprehensive catalog of 3D structure models for SARS-CoV-2 RDRP mutant proteins, a resource potentially useful for future research on SARS-CoV-2 RDRP and antiviral drug development.

- Provided a possible explanation for how certain mutations might impact the 3D structure of SARS-CoV-2 RDRP, causing it to lose replication fidelity and accelerating viral evolution.

- Provided a possible explanation for how certain mutations might impact the 3D structure of SARS-CoV-2 RDRP, causing viral variants to become refractory to the US FDA-approved antiviral drug remdesivir.

INTRODUCTION

Background

SARS-CoV-2 is the causative virus for the current COVID-19 pandemic [1]. Its viral genome contains a single stranded RNA of 29.9kb in size, which expresses four structural virion proteins, seven open reading frames (Orfs) and eleven non-structural proteins (Nsps) [2, 3]. Coronaviruses are the longest single-stranded RNA viruses. Central to viral replication is the RNA-dependent RNA polymerase (RDRP) protomer consists of Nsp7, and 2 subunits each of Nsp8 and Nsp12. Together with the RNA helicase Nsp13 and exonuclease Nsp14, RDRP replicates the RNA genome [4]. Nsp14 provides proof-reading and ensures some accuracy in genomic RNA replication. Over the course of the COVID-19 pandemic since its first reported in late 2019, SARS-CoV-2 viruses have evolved by accumulation of mutations as they passage from one host to the next around the world [5, 6]. Viral genome sequencing provides the genetic basis for analyzing how sequence variants affect viral proteins involved in infectivity [7].

Nsp12 has 3 major domains: nucleotidyltransferase domain, RDRP domain and interface domain [8]. Active site residues of Nsp12 includes Motif A (T611 to M626) [9]. Nsp12 recognizes dsRNA, and the bonding of D760 and D761 with the RNA 3'-end that is involved in replication of the growing RNA strand [10]. dsRNA is bound by alpha helical extensions from the N-termini of two Nsp8 proteins, which stabilize RDRP association with dsRNA. Nsp8 binds to RNA through a series of positively-charged residues, including K58 that is essential for producing replication-competent viruses [10]. RDRP is the molecular target of the US FDA-approved anti-SARS-CoV-2 small molecule drug remdesivir. The high resolution Cryo-EM structure of SARS-CoV-2 RDRP bound to remdesivir has been determined (PDB ID 7bv2) [11], which provided the three-dimensional (3D) structure basis for the current research project.

Research question

Viruses are known to undergo constantly changes through mutations, causing new variants of a virus to occur. New variants will emerge and persist if they adopt well in host populations, while others will disappear if lose competitive advantages. SARS-CoV-2 is known to frequently acquire new mutations that have led to mumerous SARS-CoV-2 variants documented globally since the start of this pandemic. Several virus variants have been documented to be much more contagious (e.g. the delta strain), responsible for resurgence of Covid-19 pandemic. Moreover, some new strains are potentially more evasive to vaccines or drug therapies. Because RDRP is responsible for faithfully replicating viral genomes, certain mutations in RDRP are likely to compromise

4

accuracy of genomic replication, resulting in rapid viral evolution and emergence of new variant strains. However, RDRP mutations have not been systematically surveyed for their impact on the 3D protein structure and hence function of this key viral enzyme. To address this question, I carried out a comprehensive analysis of how unique sequence variants (USVs or unique mutations) (arose during the nine months of the pandemic between December 2019 and August 30th 2020) impact SARS-CoV-2 RDRP structure in 3D and function, by combination of virus genome data (https://www.gisaid.org), the experimental 3D structures of SARS-CoV-2 RDRP and RDRP-remdesvir [11] available from RCSB PDB, and 3D structure modeling using the Rosetta Computational Modeling Suite [12].

METHODS

SARS-CoV-2 polypeptide sequences and Cryo-EM structures

Sequencing of over 90,000 viral genomes revealed polypeptide sequence variations from 12/ 2019 to 9/2020 available at GISAID (https://www.gisaid.org). Pre-aligned protein sequences in FASTA format were downloaded from GISAID (gisaid.org) [13, 14] on September 1st, 2020. Sequence alignments for Nsp7, Nsp8 and Nsp12 proteins of SARS-CoV-2 are made by removing nonviral sequences, and truncated, incomplete and duplicated viral sequences. Reference sequences were those released on January 10th 2020 (GenBank accession code MN908947.3) [15] for comparing amino acid substitutions of individual study proteins. The observed USVs were presumed to have preserved biochemical functions necessary for viral replication, infectivity and transmission. Atomic coordinates for the experimental structure of RDRP protomer complexed with RNA and remdesvir were downloaded from Protein Danta Bank (PDB) (PDB IDs 7BV1 and 7BV2).

Molecular visualization and graphics

Mol* was used for analysis of reference and mutated SARS-CoV-2 viral protein structures. Space-filling graphic figure was produced by Illustrate. Ribbon/atomic stick images were drawn by Mol* and PyMOL.

Computational assessment of variant site(s), similarity and energetics

The machine learning software tool PyRosetta was applied to mapping analysis. Amino acid (AA) pairs with C_α-C_α distance of less than 5.5Å were defined as neighbors, and those with C_α-C_α distance of over 11Å were also defined as neighbors as long as the angle their C_α-$C\beta$ vectors was less than 75°. The Layer Determination Factor (LDF) was calculated by the following formula: $(\cos(\theta)+0.5)/1.5)^2/(1 + \exp(d - 9))$. Θ, angle between the C_α-C_β vector; d, the C_α-C_α distance. AA with LDF < 2 was located on the surface; AA with LDF > 5.2 was located within the protein core; AA with a score in between is located in the boundary. Variant protein structural models were generated by computation with replacement of reference atomic coordinates, which were optimized by Monte Carlo optimization and gradient-based energy minimization. Structural models were further optimized were performed by combinations of scoring functions. Energetics of AA variants were calculated by side chain optimization and energy minimization.

RESULTS

Mapping spatial localization of sequence variants of RDRP

Reference structures or 3D models were constructed for reference and mutant RDRP subunits. For each RDRP subunit, AA variants were mapped to three major types of locations: the hydrophobic core, the protein surface, and the boundary layer. Most mutations for Nsp7, Nsp8 and Nsp12 take place on the surface or boundary, rather than the core (Fig. 1-3). The substitutions were more likely to be conservative in the core and less likely to be conservative within the boundary layer or on the protein surface (Fig. 4), reflecting the residues within the core tend to be more essential than those on the periphery. Majority of nonconservative variants can be a result of single base substitutions. For example, A to D and E changes can be generated through changes of the 2nd base of corresponding codons. Changes involving two base changes (*e.g.*, proline to aspartic acid) were less common.
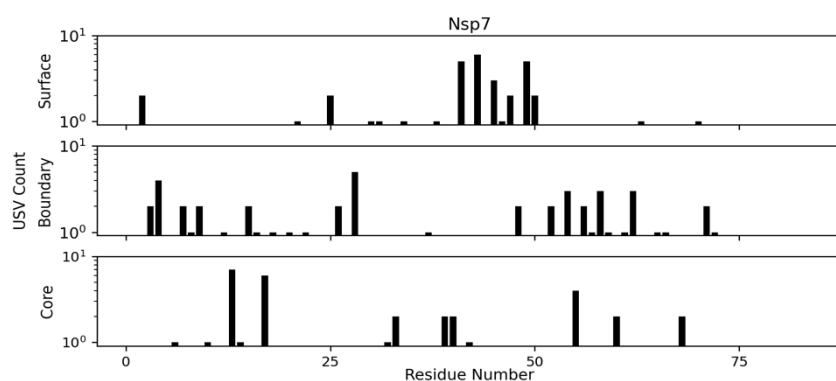

Figure 1. Distribution of USVs in the 3D structure of Nsp7 protein.

Shown are distribution of USVs on the surface, in the boundary or core of Nsp7 protein. The position and abundance are as indicated.
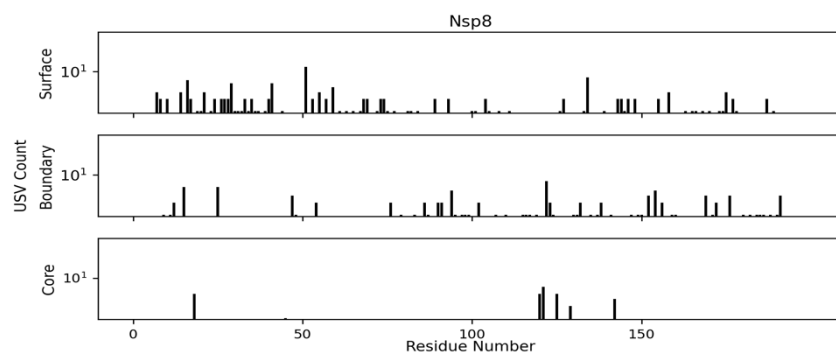

Figure 2. Distribution of USVs in the 3D structure of Nsp8 protein.

Shown are distribution of USVs on the surface, in the boundary or core of Nsp8 protein. The position and abundance are as indicated.
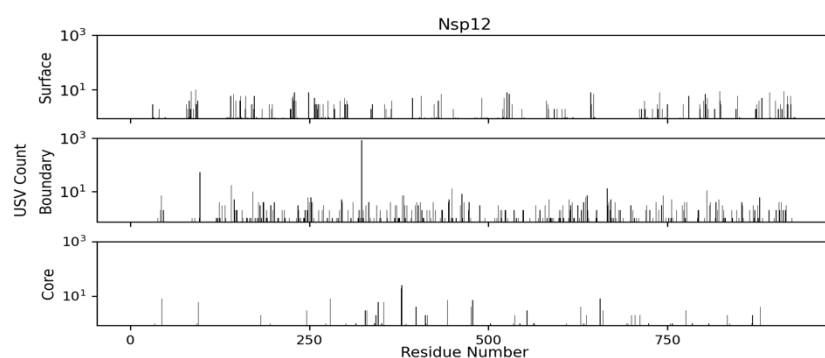
Figure 3. Distribution of USVs in the 3D structure of Nsp12 protein.
Shown are distribution of USVs on the surface, in the boundary or core of Nsp12 protein. The position and abundance are as indicated.
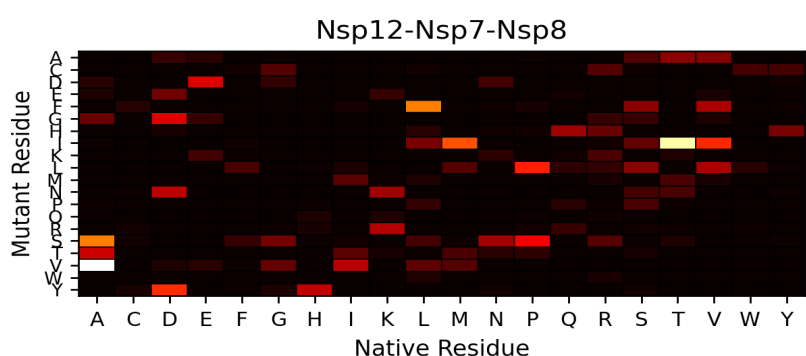

Figure 4. Patterns of unique sequence variants (USVs) compared with native residues in Nsp7, Nsp8 and Nsp12 proteins.
Heat map indicates the frequency of different native amino acids changed to other amino acids in USVs.

Overall unique sequence variant (USV) profiles for Nsp7, Nsp8 and Nsp12

The cryo-EM 3D structure of the RDRP-dsRNA complex (PDB ID 6yyt) was used as the reference for mutant structures of RDRP subunits Nsp7, Nsp8, and Nsp12. For Nsp7, there were 86 USVs out of 90,864 viral genomes from the GISAID dataset compared with the reference, with 74 single substitutions, 7 doubles, 1 triple, 2 quadruples, 1 quintuple, and 2 multi-point substitutions. Vast majority of mutations occurred in only one unique variant (USV). The most frequent USV (1473 times) for Nsp7 is S25L, a nonconservative substitution on the protein surface. For Nsp8, there were 212 USVs compared with the reference sequence, with 192 single substitutions, 11 doubles, 6 triple, and 3 multi-point substitutions. Substitutions were mainly presented in a single USV. The most frequently observed USV for Nsp8 is Y138H (nonconservative, boundary, 283 times). Nsp8 contains nine positively-charged residues (K36, 37, 39, 40, 46, and R51, 57, and K58, 61) that engage the binding of the two Nsp8 subunits to RNA. No substitutions were found in K58, the essential RNA-binding residue. However, substitutions were observed with K37, K40, R51, R57, and K61. There were distinct substitutions in R51, R51L, R51C, R51H, of which R51L and R57L are predicted to disrupt salt bridge formation. The weakened binding to another USV exhibits

apparent synergy between individual AA substitutions: M90S;L91F. The two residues were mapped to the interface between Nsp12 and a Nsp8 subunit, which was close to the shared interface of the other Nsp7 and Nsp12 subunit.

For Nsp12, there were 1070 USVs of Nsp12 with 274 single substitutions, 645 doubles, 124 triples, 15 quadruples, and 3 quintuples, and 9 multi-point substitutions. Majority of mutations were present in a single USV. The most frequent USV (60,523 times) for Nsp12 is P323L, a nonconservative substitution in the boundary layer. P323L represents a distinct SARS-CoV-2 clade that has since spread worldwide. Approximately 70% of AA substitutions among all proteins were nonconservative (1659 nonconservative *versus* 705 conservative), with most of the nonconservative variants present in the boundary and surface of proteins. Most USVs were slightly less stable than the reference.
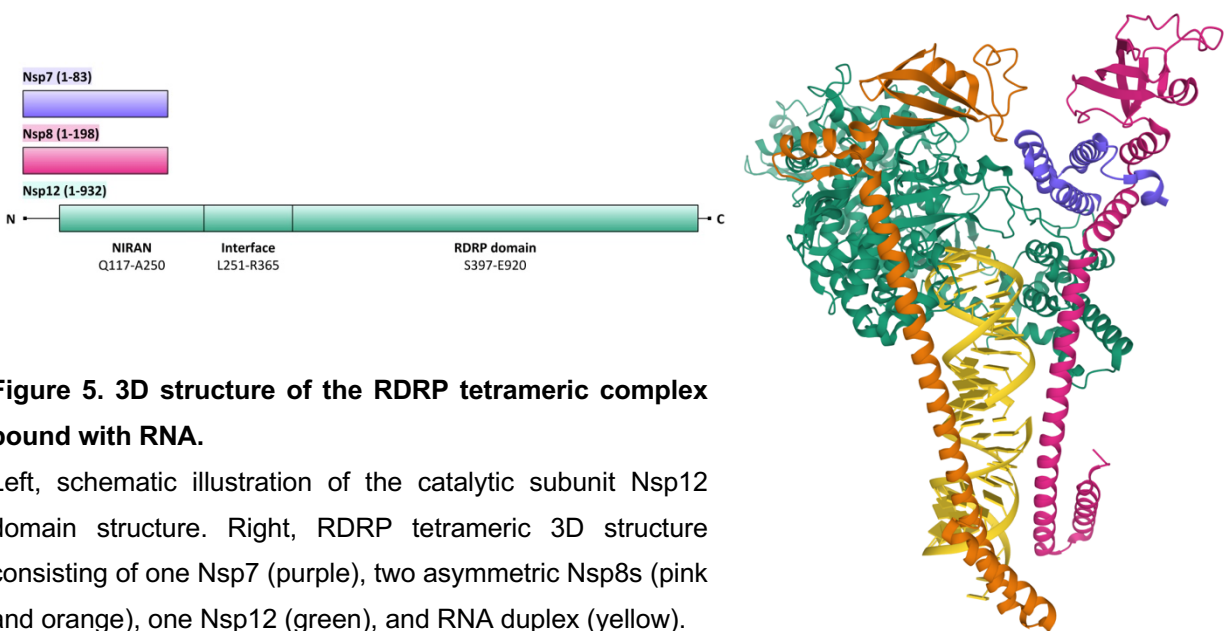

Figure 5. 3D structure of the RDRP tetrameric complex bound with RNA.

Left, schematic illustration of the catalytic subunit Nsp12 domain structure. Right, RDRP tetrameric 3D structure consisting of one Nsp7 (purple), two asymmetric Nsp8s (pink and orange), one Nsp12 (green), and RNA duplex (yellow).

Nsp12 active site mutations

Nsp12 is composed of three distinct domains: the amino-terminal nucleotidyltransferase domain, the interface domain, and the carboxyl-terminal RDRP domain (Fig. 5). Nsp12 interacts with dsRNA. D760 and D761 bind to RNA 3'-end, which is needed for RNA synthesis. Of the residues comprising active site Motif A in the catalytic pocket (Fig. 6), mutations were observed in P612, H613, L614, M615, G616, W617, D618, Y619, P620, and A625 (Fig. 6). These mutations were mostly single or low count substitutions, except H613Y with 104 occurrences that are often

accompanied by the P323L substitution. It is noteworthy that they are pointed toward the protein's hydrophobic core opposite the active site, which should remain catalytically active.
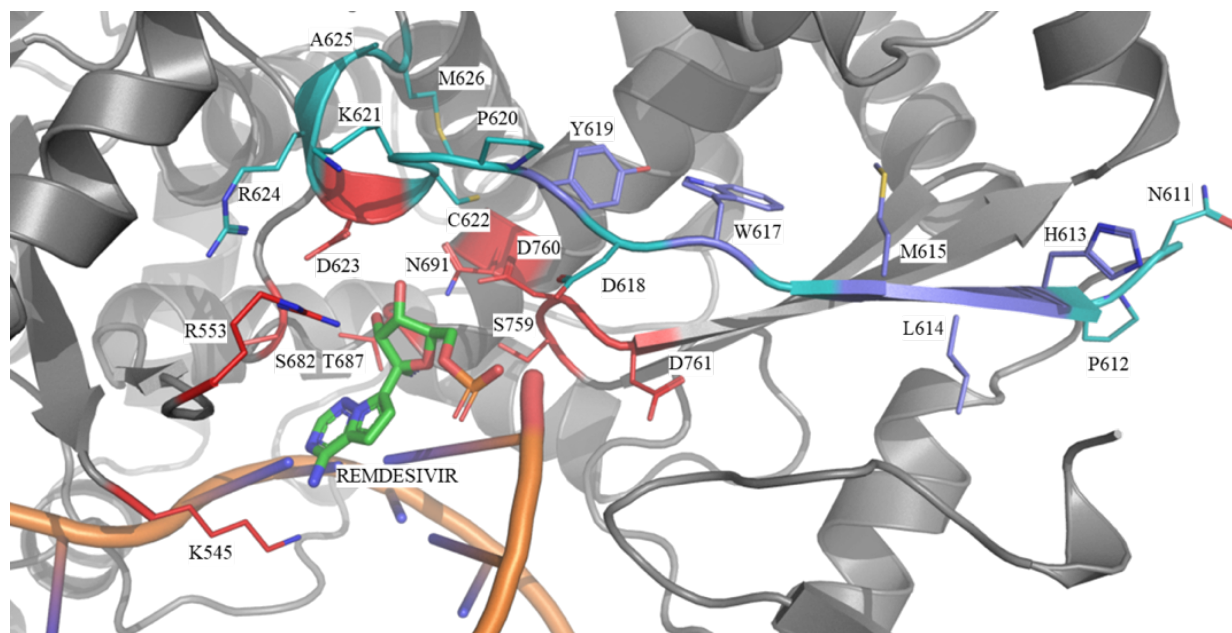

Figure 6. Close-up view of the catalytic site of Nsp12 bound to dsRNA and remdesivir.
The active site of Nsp12 in complex with remdesvir shown in ribbon-atomic stick (PDB ID 7bv2). RNA is shown in orange, remdesivir in atomic stick figure and active site Motif A in magenta and purple. AAs in contacts with remdesivir is shown in light red.

Remdesivir is a ribonucleotide prodrug that can be transformed into an active triphosphate metabolite, which is incorporated into the RNA molecule [16]. Remdesivir monophosphate (RMP) is positioned in the middle of the active site, thereby blocking access by nucleotide triphosphate into the active site [11] (Fig. 6). RMP interacts with K545 and R555. Two Mg^{2+} and a pyrophosphate facilitate the binding of RMP, with both Mg^{2+} interact with the phosphate group. The pyrophosphate (an opening around this may provide resistance) in the nucleotide entry site to the catalytic center. It may precent NTP from getting into the catalytic site [11].
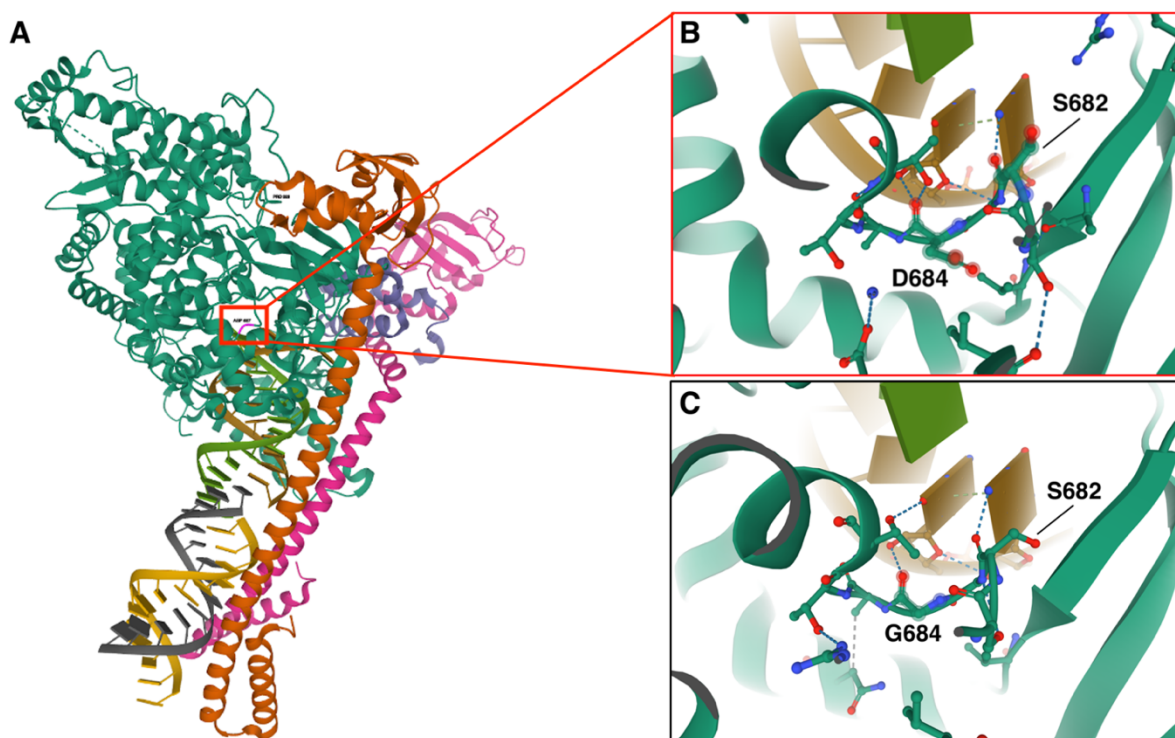
Figure 7. Impact of D684G substitution on RNA binding by RDRP.

(A) 3D structure of RDRP-RNA-RMP complex. Nsp12 is in green; the two copies of Nsp8 are in brown and pink; RNA is in grey; the active site is boxed in red. (B) Close-up view of RDRP active site from the reference 3D structure showing binding of S682 and D684 to RNA. (C) Close-up view of RDRP catalytic center from the D682G variant 3D structure model show alteration in binding of G684 to RNA and the overall more opening of the active site.

No mutations were detected for residues bonding with remdesivir (K545, R553, D623, N691, D760, S759, D760) (Fig. 6), suggesting that any substitutions are not tolerated as these residues are essential for RDRP's catalytic activity. Intriguingly, non-conservative substitutions were observed with two residues, S682 and D684 that form hydrogen bonding with -1 and +1 nucleotide, respectively, of the template RNA (Fig. 7). The S684G substitution would likely weaken the grip of the RNA helix, which may reduce replication accuracy. Moreover, the lost RNA binding at the -1 position, when combined with the smaller glycine substitution, appeared to open up more space in the catalytic pocket (Fig. 7B and 7C), which could open up the active site sufficiently to allow entry of nucleotide triphosphate even in the presence of RMP, rendering drug resistance.

Protein-protein interaction interface mutations

RDRP protomer has a tetrameric structure composed of 1 Nsp7, 2 Nsp8 and 1 Nsp12 subunits. It is anticipated that substitutions at the various protein-protein interfaces, while preserving the basic catalytic activity of the viral RNA polymerase. 11 (5 non-conservative) substitutions were found in 6 Nsp7 residues that influences binding to Nsp12 (K7, L14, S15, S26, L40, and L41). Conversely, substitutions of Nsp12 at T409, P412, F415, Y420, E436, A443 and D445 could also affect the Nsp7:Nsp12 interaction. In particularly, Y420S is likely to break hydrogen bonding to with D5 of Nsp7. E436G/K could interfere with the salt bridge to K43 of Nsp7. Substitutions in Nsp7 and Nsp12 tend to present within their contact interface were highly destabilizing with $\Delta\Delta G^{App}$>+10 REU. Nsp7 only makes substantial contact with one Nsp8 subunit, with mutations within this interaction interface in V6, T9, S15, V16, L20, L28, Q31, F49, E50, M52, S54, L56, S57, V58, L60, S61, V66, I68, and L71. It is noteworthy that the S54P mutation is potentially disruptive to an interfacial helix as it is positioned in the middle of this helix. Conversely, substitutions of residues occurred in V83, T84, S85, T89, M90, L91, M94, L95, N100, A102, I107, V115, P116, I119, L122, V131, and A150 in Nsp8 on the Nsp7:Nsp8 interface, which may affect their interaction. Two Nsp8 monomers are asymmetric. Therefore, a substitution could have differential effect on the interaction of individual chain with Nsp12. Substitutions at 23 sites were found that may affect the interaction of one or both Nsp8 with Nsp12. While mutation at T68, K72, R75, S76, and K79 appears to only impact the Nsp8-Nsp12 interface that wraps around Nsp7, mutations at V83, M90, and M94 would be predicted to disturb both interfaces. Of these 38 mutations across 31 positions, 19 were conservative. A P121S substitution in Nsp8 is likely to form a hydrogen bond to V398 of Nsp12. W154C and W182C in Nsp8 may induce backbone rearrangement due to the change of the bulkier tryptophan.

Substitutions were observed in the following residues of Nsp12 at the interface with Nsp8 chain, On the other hand, substitutions of these 10 residues of Nsp12 occurred at the interface with the second chain of Nsp8. About half of Nsp12 mutations were non-conservative. It is particularly noteworthy that the P323L mutation is located at the carboxy-terminus of helix 10 of Nsp12 on the interface between Nsp12 and Nsp8 (Fig. 8). Leucine substitution of P323 is not only more flexible, but also appears to create a new beta-sheet that could interact with the corresponding beta-sheet of Nsp8, making a more stable Nsp12:Nsp8 interaction. Because Nsp8 binding affinity has been shown to increase RNA polymerase activity [17], this variant may be more efficient in viral replication, and hence could be more transmissible.
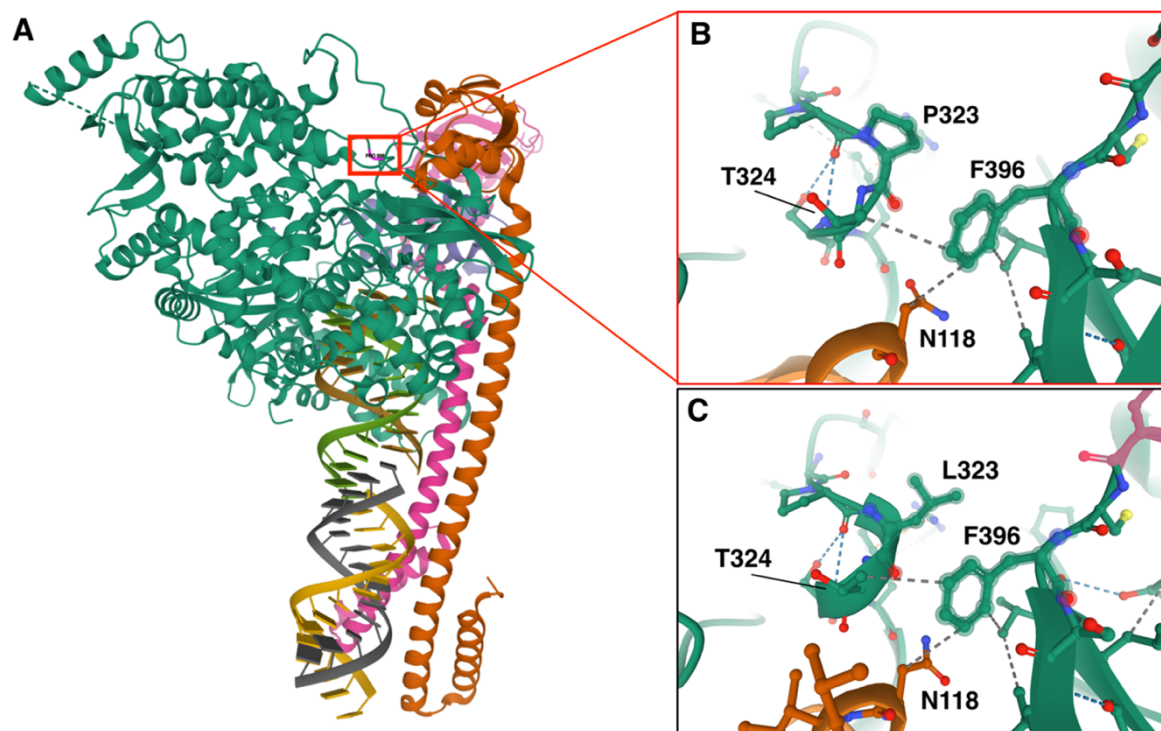
Figure 8. Impact of P323L substitution on Nsp12:Nsp8 interaction.

(A) 3D structure of SARS-CoV-2 RDRP-RNA-RMP complex. Nsp12 is in green; the two copies of Nsp8 are in brown and pink; RNA is in grey; the active site is boxed in red. (B) Close-up view of the Nsp12:Nsp8 interface of the reference 3D structure. (C) Close-up view of the Nsp12:Nsp8 interface of the P323L variant 3D structure model. A new beta-sheet is created as a result of the P323L substitution that is in proximity to the beta-sheet of Nsp8 at the Nsp12:Nsp8 interface.

DISCUSSION

Implications and significance of this work

In this study, the 3D impact of amino acid substitutions in RDRP subunits Nsp7, Nsp8 and Nsp12 were analyzed *versus* the original reference sequence for over 90,000 sequenced SARS-CoV-2 genomes during the first nine months of the COVID19 pandemic. The computed 3D structure models of the protein variants show that mutations primarily occur in the boundary and surface of RDRP complex. Most AA changes appeared to be moderately destabilizing, as judged by the energetics ($\Delta\Delta G^{App}$) values. SARS-CoV-2 isolates were obtained from infected individuals and hence should be infectious. As such, RDRP variants were also presumed to be functional and capable of viral reproduction due the essential role of RDRP in viral genome replication. The 3D structure models of RDRD variants generated in this study from over 90,000 SARS-CoV-2 genomes provide a potentially useful framework for future research on biochemical functions of RDRP, viral genome replication and development of antiviral therapeutics for SARS-CoV-2 and related coronavirues.

SARS-CoV-2 is known to frequently acquire new mutations. Because RDRP is responsible for faithfully replicating viral genomes, some USVs are likely to alter replication efficiency and accuracy, and hence contribute to evolution of the virus. P323 is located at the interaction surface between Nsp12 and the first Nsp8 unit. The P323L substitution is predicted to form a new beta-sheet in Nsp12 at the Nsp12:Nsp8 inter-surface, which could interact with a nearby beta-sheet in Nsp8, strengthening the affinity between Nsp12 and Nsp8. It has been shown that the level of Nsp8 binding to Nsp12 is an important determinant for SARS-CoV-1 RNA polymerase activity [17]. Hence the P323L substitution could improve RDRP's replication efficiency through higher level of Nsp8 association, providing a possible explanation for the dominant presence of this viral clade that includes the most contagious delta strain. Another noteworthy non-conservative substitution occurred at D684 that form hydrogen bonding with the -1 position of RNA. Such substitution could weaken the grip of template RNA, compromising RNA replication accuracy and accelerating viral evolution.

RDRP, like other viral polymerases, is a major anti-viral therapeutic target. It is inhibited by remdesivir, the only small molecule drug fully approved by US FDA for treatment of COVID19 patients [18, 19]. This small molecule was originally discovered in a screen for anti-Ebola drugs. Although it did not show sufficient efficacy against Ebola infections, it did exhibit excellent human safety profile [16]. Subsequent clinical studies led to its repurposing for treating CoV2 patients

[16]. A major challenge for successful anti-viral therapies is the emergence of acquired resistant mutations at the drug-binding sites [20]. Interestingly, no USVs were observed with key residues of RDRP in direct contact with remdesivir, suggesting that substitution of these residues is unfavorable, which would make such viral variants 'less fit'. It is conceivable that such 'less fit', drug-resistant strains may emerge under drug treatment pressure, which otherwise would not compete effectively in untreated populations. Strikingly, substitutions of two residues that do not have direct contact with remdesivir, S682L/P and D684G, were observed. The D684G mutation in particular, appears to open up the catalytic pocket due to weakened RNA binding, which could temper drug-binding affinity or restore nucleotide triphosphate's entry into the active site even in the presence of RMP. Hence, these substitutions are potential cause for drug resistance. The computed 3D structural models of RDRP USVs generated in this study should provide useful information for future drug developers to anticipate sources of drug resistance and design next generation of more potent antiviral drug candidates.

Future work

Several interesting findings are made based on analysis of the 3D computational models of RDRP variants. First, the P323L and S684G substitutions may alter RNA replication efficiency or accuracy, accelerating viral evolution. This hypothesis can be tested by genetically engineering these RDRP mutants and examining their replication efficiency and accuracy in cell model systems or cell-free reconstituted reactions [11]. Second, the D684G substitution may open up the catalytic pocket and render resistance to the antiviral drug remdesivir. This hypothesis can also be tested using the previously established in vitro assay for RDRP inhibition [11]. Third, if S684G is proven to be remdesivir-resistant in subsequent in vitro studies, the 3D RDRP variant models would be useful for designing or computationally screening novel, more potent antiviral drugs to overcome such drug resistant variants.

REFERENCES

[1]   Wiersinga WJ, Rhodes A, Cheng AC, Peacock SJ, Prescott HC. Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19): A Review. JAMA 2020;324:782-793.

[2]   Harrison AG, Lin T, Wang P. Mechanisms of SARS-CoV-2 Transmission and Pathogenesis. Trends in Immunology 2020;41:1100-1115.

[3]   Hu B, Guo H, Zhou P, Shi Z-L. Characteristics of SARS-CoV-2 and COVID-19. Nature Reviews Microbiology 2020.

[4]   Denison MR, Graham RL, Donaldson EF, Eckerle LD, Baric RS. Coronaviruses: an RNA proofreading machine regulates replication fidelity and diversity. RNA Biol 2011;8:270-279.

[5]   Hartley P, Tillett RL, Xu Y, AuCoin DP, Sevinsky JR, Gorzalski A, et al. Genomic surveillance revealed prevalence of unique SARS-CoV-2 variants bearing mutation in the RdRp gene among Nevada patients. medRxiv 2020.

[6]   Wang R, Hozumi Y, Yin C, Wei G-W. Decoding SARS-CoV-2 Transmission and Evolution and Ramifications for COVID-19 Diagnosis, Vaccine, and Medicine. Journal of Chemical Information and Modeling 2020;60:5853-5865.

[7]   Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell 2020;182:812-827.e819.

[8]   Kim D, Lee JY, Yang JS, Kim JW, Kim VN, Chang H. The Architecture of SARS-CoV-2 Transcriptome. Cell 2020;181:914-921.e910.

[9]   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 2020;368:779-782.

[10]   Hillen HS, Kokic G, Farnung L, Dienemann C, Tegunov D, Cramer P. Structure of replicating SARS-CoV-2 polymerase. Nature 2020;584:154-156.

[11]   Yin W, Mao C, Luan X, Shen DD, Shen Q, Su H, et al. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. Science 2020;368:1499-1504.

[12]   Chaudhury S, Lyskov S, Gray JJ. PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. Bioinformatics 2010;26:689-691.

[13]   Elbe S, Buckland-Merrett G. Data, disease and diplomacy: GISAID's innovative contribution to global health. Glob Chall 2017;1:33-46.

[14]   Shu Y, McCauley J. GISAID: Global initiative on sharing all influenza data - from vision to reality. Euro Surveill 2017;22.

[15]   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;579:265-269.

[16]   Eastman RT, Roth JS, Brimacombe KR, Simeonov A, Shen M, Patnaik S, et al. Correction to Remdesivir: A Review of Its Discovery and Development Leading to Human Clinical Trials for Treatment of COVID-19. ACS Cent Sci 2020;6:1009-1009.

[17]   Subissi L, Posthuma CC, Collet A, Zevenhoven-Dobbe JC, Gorbalenya AE, Decroly E, et al. One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. Proceedings of the National Academy of Sciences 2014;111:E3900-E3909.

[18]   Beigel JH, Tomashek KM, Dodd LE, Mehta AK, Zingman BS, Kalil AC, et al. Remdesivir for the Treatment of Covid-19 — Final Report. New England Journal of Medicine 2020;383:1813-1826.

[19]   Rubin D, Chan-Tack K, Farley J, Sherwat A. FDA Approval of Remdesivir — A Step in the Right Direction. New England Journal of Medicine 2020;383:2598-2600.

[20]   Strasfeld L, Chou S. Antiviral drug resistance: mechanisms and clinical implications. Infect Dis Clin North Am 2010;24:413-437.
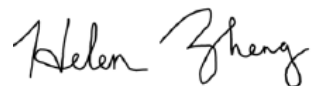
ACKNOWLEDGEMENT

Commitments on Academic Honesty and Integrity

We hereby declare that we

1.   are fully committed to the principle of honesty, integrity and fair play throughout the competition.
2.   actually perform the research work ourselves and thus truly understand the content of the work.
3.   observe the common standard of academic integrity adopted by most journals and degree theses.
4.   have declared all the assistance and contribution we have received from any personnel, agency, institution, etc. for the research work.
5.   undertake to avoid getting in touch with assessment panel members in a way that may lead to direct or indirect conflict of interest.
6.   undertake to avoid any interaction with assessment panel members that would undermine the neutrality of the panel member and fairness of the assessment process.
7.   observe the safety regulations of the laboratory(ies) where we conduct the experiment(s), if applicable.
8.   observe all rules and regulations of the competition.
9.   agree that the decision of YHSA is final in all matters related to the competition.

We understand and agree that failure to honour the above commitments may lead to disqualification from the competition and/or removal of reward, if applicable; that any unethical deeds, if found, will be disclosed to the school principal of team member(s) and relevant parties if deemed necessary; and that the decision of YHSA is final and no appeal will be accepted.

(Signatures of full team below)

X _____

Name of team member: Helen H Zheng

X _____

Name of team member:

X _____

Name of team member:

X _____

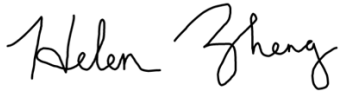Name of supervising teacher: Stephen K Burley

Declaration of Academic Integrity

The participating team declares that the paper submitted is comprised of original research and results obtained under the guidance of the instructor. To the team's best knowledge, the paper does not contain research results, published or not, from a person who is not a team member, except for the content listed in the references and the acknowledgment. If there is any misinformation, we are willing to take all the related responsibilities.

Names of team members

Helen H. Zheng

Signatures of team members

Name of the instructor

Stephen K. Burley

Signature of the instructor

Date: September 14, 2021