

## **Novel Mutations in NSP1 and PLPro of SARS-CoV-2 NIB-1 Genome Mount for Effective Therapeutics**

Mohammad Uzzal Hossain<sup>1¶</sup>, Arittra Bhattacharjee<sup>1,2¶</sup>, Md. Tabassum Hossain Emon<sup>3¶</sup>, Zeshan Mahmud Chowdhury<sup>3</sup>, Md. Golam Mosaib<sup>4</sup>, Md. Moniruzzaman<sup>5</sup>, Md. Hadisur Rahman<sup>5</sup>, Md. Nazrul Islam<sup>6</sup>, Irfan Ahmed<sup>5</sup>, Md. Ruhul Amin<sup>7</sup>, Asif Rashed<sup>8</sup>, Keshob Chandra Das<sup>5</sup>, Chaman Ara Keya<sup>2</sup>, Md. Salimullah<sup>5\*</sup>

¶These Authors equally contributed to this manuscript and represents the first author

<sup>1</sup>*Bioinformatics Division, <sup>5</sup>Molecular Biotechnology Division, <sup>6</sup>Plant Biotechnology Division, National Institute of Biotechnology, Ganakbari, Ashulia, Savar, Dhaka-1349, Bangladesh*

<sup>2</sup>*Department of Biochemistry and Microbiology, North South University, Bashundhara, Dhaka-1229, Bangladesh*

<sup>3</sup>*Department of Biotechnology and Genetic Engineering, Life Science Faculty, Mawlana Bhashani Science and Technology University, Santosh, Tangail-1902, Bangladesh.*

<sup>4</sup>*Department of Biochemistry and Molecular Biology, Faculty of Health & Medical Sciences, Gono Bishwabidyalyo, Ashulia, Savar, Dhaka-1344, Bangladesh*

<sup>7</sup>*Center for Medical Biotechnology, MIS, Directorate General of Health Services, Dhaka, Bangladesh*

<sup>8</sup>*Department of Microbiology, Mugda Medical College, Dhaka, Bangladesh*

### **\*Correspondence:**

Dr. Md. Salimullah

Chief Scientific Officer

Molecular Biotechnology Division

National Institute of Biotechnology

Ganakbari, Ashulia, Savar, Dhaka-1349, Bangladesh

Tel: 880-2-7788443

E-mail: salim2969@gmail.com

## **Abstract**

Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2), the etiologic agent of Coronavirus Disease- 2019 (COVID-19), is rapidly accumulating new mutations. Analysis of these mutations is necessary for gaining knowledge regarding different aspects of therapeutic development. Recently, we have reported a Sanger method based genome sequence of a viral isolate named SARS-CoV-2 NIB-1, circulating in Bangladesh. The genome has four novel mutations in V121D, V843F, A889V and G1691C positions. V121D substitution has the potential to destabilize the Non-Structural Protein (NSP-1) which inactivates the type-1 Interferon-induced antiviral system hence this mutant could be the basis of attenuated vaccines against SARS-CoV-2. V843F, A889V and G1691C are all located in NSP3. G1691C can decrease the flexibility of the protein while V843F and A889V changed the binding pattern of SARS-CoV-2 Papain-Like protease (PLPro) inhibitor GRL0617. V843F PLPro showed reduced affinity for Interferon Stimulating Gene-15 (ISG-15) protein whereas V843F+A889V double mutants exhibited the same binding affinity as wild type PLPro. Here, V843F is a conserved position of PLPro that damaged the structure but A889V, a less conserved residue, most probably neutralized that damage. Mutants of NSP1 could provide attenuated vaccines against coronavirus. Also, these mutations of PLPro could be targeted to develop anti-SARS therapeutics.

**Keywords:** SARS-CoV-2, COVID-19, NSP1, Papain-Like protease, GRL0617, ISG-15

## 1. Introduction

Coronavirus Disease- 2019 (COVID-19) is now an irrepressible pandemic that continuously threatening humanity since its causative agent Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) is rapidly changing and adapting in different environments [1]. At the beginning of the COVID-19 outbreak, this viral infection caused flu-like symptoms such as cough, fever, and trouble breathing [2]. However, this acellular pathogen is constantly evolving with novel mutations thus viral diarrhea, cardiovascular injury, strokes, psychosis, dementia and Kawasaki like syndrome are becoming more frequent in COVID-19 cases [3-5]. According to the World Health Organization (WHO), this constant evolution in different selection pressures are letting the virus to be airborne [6]. Proper analysis of these genomic mutations can help to decipher issues regarding viral transmission, epidemiology, pathogenicity and therapeutics.

Coronaviruses (CoVs) are enveloped viruses that belong to the *Coronaviridae* family and *Orthocoronavirinae* subfamily [7, 8]. SARS-CoV-1 and Middle East Respiratory Syndrome (MERS-CoV) belong to the *Betacoronavirus* genus of *Orthocoronavirinae* subfamily [7, 8]. These pathogens were the causative agents of SARS and MERS epidemics. SARS-CoV-2, a member of *Betacoronavirus*, shares nearly 82.30% sequence similarity with SARS-CoV-1 while showing strikingly less similarity with MERS (28%) [9].

SARS-CoV-2 has a 29.9 kb single-stranded (+) positive RNA genome with 11 genes. The virus also contains two Untranslated Regions (UTRs) in 5' and 3' terminal. Nearly 67% of the genome consist of Orf1ab polyprotein gene. The rest of the genome constructs structural proteins e.g., Surface (S), Envelope (E), Membrane (M), and Nucleocapsid (N) proteins [10]. According to the

recent update of Global initiative on sharing all influenza data (GISAID) database, with particular prevalent mutations, the virus has made six distinct clades such as S, L, V, G, GR and GH [11]. Among them, GR is the largest and predominates in South America, Europe and Asia (gsaid.org). The members of this clade contain S-D614G and N-G204R mutations. Currently, GR is relatively less prevalent in North America. However, recently we have sequenced a GR type viral isolate SARS-CoV-2/human/BGD/NIB\_01/2020 (SARS-CoV-2 NIB-1) (NCBI GenBank Accession: MT509958, GISAID Accession ID: EPI\_ISL\_447904) from a young female COVID-19 patient of Bangladesh [12]. According to our report, the viral isolate shares a common ancestor with three North American SARS-CoV-2 isolates that were collected from Washington, United States of America (USA). Additionally, our isolate showed longer branch length or more genomic mutations than its closely related members [12]. When we aligned the genome with the reference genome (Accession: NC\_045512.2), we found 11 mutations of which 4 of them were new. Among them, three were found in Papain-Like Protease (PLPro) and C-terminal domain of Non-structural Protein (NSP)-3 and another one was located in NSP1 (Table: I).

Several studies have demonstrated that NSP1 and the PLPro of NSP3 from SARS-Cov-1 and 2 play a critical role to inhibit the type 1 interferon (IFN) response in the host cell [13]. Moreover, PLPro cleaves off Interferon Stimulating Gene-15 (ISG15) protein, therefore, the host cell cannot execute antiviral signals properly [14, 15].

In this study, we have analyzed the effect of these novel mutations on the respective protein structures. We have shown that some of these mutations have the potential to destabilize the structure of NSP1 and the C-terminal domain of NSP3. We also explored how the loss and gain of Valine in V843F and A889V mutations respectively affect the binding of protease inhibitor GRL0617 and ISG15 for their implication in new therapeutics development.

## **2. Methods:**

### **2.1. Genome retrieval and identification of the mutations:**

The genome of SARS-CoV-2 NIB-1 was collected from the National Center for Biotechnology Information (NCBI) (GenBank accession: MT509958.1). The genome was aligned with the reference genome by NCBI Nucleotide Basic Local Alignment Search Tools (BLASTN) to identify the mutations in the Untranslated Regions (UTRs) [16]. The Non-synonymous mutations were collected from GISAID CoVsurver ([gisaid.org](https://gisaid.org)).

### **2.2. Effect of the mutations on the proteins:**

Mutations that affect the structural stability of the proteins can behave differently than the wild type one [17, 18]. To assess the effect of the novel mutations, MUpro, Protein Variation Effect Analyzer (PROVEAN) and HOPE were employed [19-21]. Afterward, PLPro was taken for further analysis due to the availability of structural and molecular mechanisms on this protein.

### **2.3. *In silico* mutagenesis, molecular modeling and refinement of PLPro:**

To observe the mutational effect of the protein at three dimensional (3D) level, the crystal structure of PLPro was collected from Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB) ([www.rcsb.org](http://www.rcsb.org)) [22]. The chain A of PLPro (PDB ID: 6W9C) was preserved via BIOVIA Discovery Studio (<https://www.3ds.com/>). Heterogeneous atoms were also removed. Mutated protein sequences went under homology modeling by SWISS-MODEL using chain A of 6W9C as a template [23]. The generated structures were energetically minimized by 3D Refine and then again refined by Galaxy Refine [24, 25].

#### **2.4. Structural assessment of mutant proteins:**

The structural quality of the mutant proteins was evaluated. To execute this step, RAMPAGE and SWISS-MODEL structure assessment were employed [23, 26].

#### **2.5. Inhibitor binding analysis against the wild type and mutant PLPros**

GRL0617 or 5-Amino-2-methyl-N-[(R)-1-(1-naphthyl) ethyl] benzamide (PubChem CID: 24941262) is a PLPro inhibitor with a relatively low level of cytotoxicity [27]. Hence, this naphthalene based small molecule has the potential to be a prospective antiviral drug. Here, we analyzed the interactions between GRL0617 and PLPro by AutoDock Vina using our previously reported methodology for exploring drug-receptor interactions [28, 29]. In short, the GRL0616 was taken as ligand and the wild type and mutant proteins were used as receptors. The active site of the palm domain in PLPro was enclosed with Grid Box [30]. (Parameters are given in Supplementary file:1). The molecular interactions between GRL0617-PLPro were visualized by UCSF Chimera [31], BIOVIA Discovery Studio and PyMOL Molecular Graphics System, Version 2.3.3 Schrödinger, LLC.

#### **2.6. Analysis of ISG15-PLPro interactions:**

ISG15 C-terminal domain that interacts with PLPro was collected from PDB ID: 6XA9. As a ligand/ substrate, this protein was docked against the PLPro receptors by GalaxyTongDock\_A [32]. Model 1 with the highest cluster size and docking score was further analyzed by PRODIGY (PROtein binDing enerGY prediction) to determine the binding affinity of the protein complexes at 25° C temperature [33].

### **3. Results:**

The overall scheme of the present study has been described in **Figure 1**

#### **3.1. Transversion mutations are predominant in SARS-CoV-2 NIB-1:**

SARS-CoV-2 NIB-1 is a member of the GR clade (Table 1). The genome contains 11 mutations in total. Among them, 6 were transversion mutations and 5 were transition mutations. According to GISAID, 7 of these mutations were reported previously and 4 of them were first identified in SARS-CoV-2 NIB-1 isolate. In these 7 mutations, 4 were transition mutations; Three Purine →Purine transitions (Table 1; Serial # 4-6) and one Pyrimidine →Pyrimidine transition (Table 1; Serial # 1). The other three mutations were transversion mutations: Two Purine → Pyrimidine transversions (Table 1; Serial # 2, 3) and one Pyrimidine →Purine transversion (Table 1; Serial # 7). The 4 novel mutations were located in the Orf1b gene. Among them, only A889V substitution occurred due to C → T (Pyrimidine → Pyrimidine) transition. The other 3 mutations were transversion substitutions; Two Purine → Pyrimidine transversions (Table 1; Serial # 9, 11), one Pyrimidine →Purine transversion.

#### **3.2. Novel mutations have the potential to alter the functions and structures of protein:**

According to the MUPro, V121D, V843F and G1691C destabilize the proteins (Table 2). Moreover, PROVEAN determined that the V121D and G1691C substitutions are deleterious for its biological functions. HOPE demonstrated that most of the mutant amino acids are located in a conserved region and they are bigger than the wild type residues. They have the potential to abolish the protein functions. In contrast, only A889V mutant PLPro can increase protein stability (Table 2).

### **3.3. Interaction patterns of GRL0617-PLPro altered in mutant proteases**

The interactions between the inhibitor GRL0617 and PLPro were identified via molecular docking simulation. Before molecular docking, the quality of the receptors was evaluated by analyzing Ramachandran plots and MolProbity scores. Ramachandran plots showed that the mutant PLPros have 98.1-98.4% amino acids in favored regions (~98% expected). The MolProbity score of the wild type crystal structure was 2.55 whereas the mutants have MolProbity scores between 1.16-1.32 (Supplementary file 3). Therefore, after the homology modeling, the quality of the simulated structures were satisfactory for further analysis [26, 34].

V843F PLPro (-6.7 kcal/mol) and A889V PLPro (-6.9 kcal/mol) showed more binding affinity scores than the wild type one (-6.6 kcal/mol) (Table 3). However, V843F+ A889V PLPro showed slightly less affinity (-6.2 kcal/mol).

Although the binding affinity did not change significantly due to the mutations, the inhibitor binding site changed and hence the interacting amino acids (Table 3). The wild type receptor interacted with ASP164, ARG166, GLU167, SER170, PRO248, TYR264 whereas ASP76, ARG82 were common interacting residues for mutant type receptors. Moreover, residues that participated in hydrogen bonding in Wild type single mutants and double mutants were not the same. V843F and A889V single mutants have the same residue CYS155 which interacted with the inhibitor through Hydrogen bonding.

### **3.4. V843F mutation significantly reduced the affinity of PLPro toward ISG15**

SARS-CoV-2 PLPros cleaves off ISG15 that is essential for type 1 IFN mediated antiviral mechanisms [14]. The wild type, A889V and V843F+ A889V mutants showed nearly the same binding affinity (-15.4 to -15.5 kcal/mol) toward the C terminal domain of ISG15. However,

V843F mutant exhibited a significantly lower docking score and binding affinity (-11.1 kcal/mol) with higher dissociation constant (Table 3).

#### **4. Discussions:**

SARS-CoV-2 is a zoonotic pathogen that initiated the current global pandemic COVID-19 [2]. Since SARS-CoV-2 is a novel coronavirus and mostly resemble with SARS-CoV-1 (80% sequence similarity), we therefore focused on SARS-CoV-1 related reports to understand the results of the mutations [35]. However, we also explored other well studied *betacoronavirus* members such as Middle East Respiratory Syndrome (MERS-CoV), Human Coronavirus OC43, Mouse Hepatitis Coronavirus (MHV) to comprehend different genomic and proteomic functions of SARS-CoV-2.

Since the report of reference genome SARS-CoV-2 from Wuhan in December 2019, the deadly pathogen has spread throughout the world and produced numerous genetic variants with six predominant clades ([www.gisaid.org](http://www.gisaid.org)). The GISAID CoVsurver analysis revealed that our previously reported isolate ‘SARS-CoV-2 NIB-1’ is a member of GR clade which has been mutated 10 times in the coding regions [12]. The NCBI BLASTN unveiled another mutation in 5' UTR region. Interestingly, among these mutations, 54% occurred due to transversion mutations (Table 1) although species are usually biased to transition mutations [36]. Transversion mutations are detrimental for RNA viruses e.g. Influenza A and Human Immunodeficiency Viruses (HIV) since it can radically change the amino acids [37].

The novel mutations V121D, V843F and G1691C occurred in SARS-CoV-2 NIB-1 isolate due to transversion substitutions have the potential to destabilize or alter the protein structure and function (Table 2). Especially, V121D, V843F position are highly conserved in the protein and G1691C

might reduce the essential flexibility of NSP3 (Table 2). Amino acid substitution in V843F and G1691C took place due to G  $\rightarrow$  T transversion that was possibly introduced by Oxo-guanine generated from reactive oxygen species (ROS) [37, 38]. Conversely, C  $\rightarrow$  T transition, the most frequent transition of SARS-CoV-2, increased the structural stability of PLPro (Table 2) [38]. Besides, the C  $\rightarrow$  T transition in the 5' UTR region might interfere with the function of N and NSP1 protein [39].

SARS-CoV-2 NIB-1 has V121D and L122I amino acid substitutions in the NSP1. SARS-CoV-2 NSP1 has 84.4% similarity with SARS-CoV-1 NSP1 [40]. NSP1 is a potent virulence factor of SARS-CoV-1 that reduces the host mRNA expression by binding with the host 40S ribosome to inhibit the translation and specifically accelerate host mRNA degradation keeping the viral mRNA intact [40, 41]. NSP1 also disrupts the activation of IFN-dependent antiviral signaling pathways and represses the expression of the innate immune response genes such as type I IFN, ISG56 and ISG15 [13, 42]. That IFN and ISG15 gene-mediated antiviral pathways are crucial in host defense mechanisms against SARS-CoV-1 & 2 were shown by the suppression of their multiplication through *in vitro* type-1 IFN treatment [43, 44]. An attenuated mutant of NSP1 in SARS-CoV-1 can replicate as efficiently as wild-type strains with an intact IFN response [45]. Moreover, mutant NSP1 of MHV exhibited an adequate amount of cytotoxic T cells that protected the mice against further viral infections [46]. Hence, there is a higher possibility that a suitable NSP1 mutant of SARS-CoV-2 would work as an attenuated vaccine for COVID-19. Through computational analysis, we have observed that V121D and L122I mutant might destabilize the structure of NSP1 (Table 2 & Supplementary: 2) yet with these two mutations, the virus caused mild fever, cough and throat congestion in a young female patient [12]. Therefore, V121D and L122I mutations along with 93<sup>th</sup> C  $\rightarrow$  T in 5' UTR did not change the pathogenicity of SARS-CoV-2 significantly.

However, keeping in mind that the quick recovery of the patient within 10 days of the onset of symptoms without any notable care, it might be speculated that NSP1 with these mutations could be a potent candidate for the development of an attenuated vaccine against SARS-CoV-2.

Main protease and papain-like protease (PLPro) are two essential viral enzymes that are vital for polypeptide processing during viral maturation [40]. Therefore, inhibitors of these proteins could be a prospective antiviral drug [47]. SARS-CoV-1 PLPro has been targeted to develop inhibitors for decades [30]. SARS-CoV-1 & 2 PLPros have 4 domains (Figure 3) [27]. The active site of this enzyme resides in the thumb domain. PLPro interacts with the N and C terminal domain of ISG15 through Ubiquitin binding subsites 2 and 1 respectively [30]. This interaction finally cleaves off ISG15. Hence, type 1 IFN induced ISG15 antiviral response cannot function properly in host cells [14]. A recent study also demonstrated that when SARS-CoV-1 & 2 infected Vero and Calu3 2B4 cell lines are treated with type 1 INF, the SARS-CoV-2 showed more susceptibility to the treatment [43, 44]. Most possibly the IFN treatment enhanced the synthesis of ISG15 that ceased the viral replications in the cell lines. PLPro mediated pathogenesis can be halted with naphthalene based protease inhibitor GRL0617 since this ligand can make non-covalent bonds with SARS-CoV-2 PLPro [14, 27]. Here, V843F and A889V mutations were located in the PLPro domain of NSP3 protein. V843F, A889V single mutants and V843F+A889V double mutants changed the binding site of GRL0617. The inhibitor binds in the thumb domain of the wild type receptor whereas the mutant receptors interacted through the Ubiquitin-Like Domain (Figure 2). Whether these alterations in binding patterns reduce the efficacy of the ligand or not is yet to be explored.

The mutant PLPros were also docked against the C-terminal domain of ISG-15. To execute this step, the ISG15 portion was separated from the PLPro-ISG15 C-terminal domain complex (PDB ID: 6XA9). Only the V843F single mutant showed reduced affinity toward the Ubiquitin binding

subsite 1. Mutant A889V, V843F+A889V and Wild type PLPro did not show any difference when docked against ISG15. Therefore, V843F is a detrimental and destabilizing transversion mutation that might reduce the degradation of ISG15. It seems like A889V substitution protected the protease from the damage of V843F. Since A889V occurred due to the most common C → T transition, plausibly this mutation appeared first in the protein and then V843F took place due to ROS. Yet, further studies are needed to explore how this “yin and yang” incident affects the viral physiology. We hope our findings will give better insights during the development of attenuated vaccine and antiviral drugs.

## **5. Conclusions:**

This study focused on some novel mutations in the NSP1 and NSP3 proteins of SARS-CoV-2. NSP1 is a potent virulence factor of *Betacoronavirus* and hence a potential target for an attenuated vaccine. PLPro of SARS-like coronavirus has appeared as suitable targets for the development of anti-SARS therapeutics. In-depth, mutational studies on these proteins gave us better insights into the SARS-CoV-2 viral structure and pathogenicity for their fittingness as effective therapeutics candidates. In the future, wet lab research on the mutants will accelerate research and developments regarding attenuated vaccines and antiviral drugs.

## **6. Acknowledgments**

The authors are grateful to the Ministry of Science and Technology, Bangladesh for the extensive support during this work.

## **7. Competing Interest**

The authors declare that they don't have competing interests on this study.

## 8. Funding

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

## 9. References:

1. Sanche S, Lin YT, Xu C, Romero-Severson E, Hengartner N, Ke R. High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2. *Emerg Infect Dis.* 2020;26(7):1470-1477.
2. Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in China [published correction appears in *Nature*. 2020 Apr;580(7803):E7]. *Nature.* 2020;579(7798):265-269.
3. Varatharaj A, Thomas N, Ellul M et al. Neurological and neuropsychiatric complications of COVID-19 in 153 patients: a UK-wide surveillance study. *The Lancet Psychiatry.* 2020.
4. Bansal M. Cardiovascular disease and COVID-19. *Diabetes Metab Syndr.* 2020;14(3):247-250.
5. Verdoni L, Mazza A, Gervasoni A, et al. An outbreak of severe Kawasaki-like disease at the Italian epicentre of the SARS-CoV-2 epidemic: an observational cohort study. *Lancet.* 2020;395(10239):1771-1778.
6. Transmission of SARS-CoV-2: implications for infection prevention precautions. Who.int. <https://www.who.int/news-room/commentaries/detail/transmission-of-sars-cov-2-implications-for-infection-prevention-precautions>. Published 2020. Accessed August 19, 2020.
7. Fehr AR, Perlman S. Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol Biol.* 2015;1282:1-23.
8. Miłek J, Blicharz-Domańska K. Coronaviruses in Avian Species - Review with Focus on Epidemiology and Diagnosis in Wild Birds. *J Vet Res.* 2018;62(3):249-255. Published 2018 Dec 10.
9. Hossain M, Bhattacharjee A, Emon M et al. Recognition of Plausible Therapeutic Agents to Combat COVID-19: An Omics Data Based Combined Approach. 2020. doi:10.21203/rs.3.rs-25807/v1
10. Khailany RA, Safdar M, Ozaslan M. Genomic characterization of a novel SARS-CoV-2 [published online ahead of print, 2020 Apr 16]. *Gene Rep.* 2020;19:100682.
11. Shu Y, McCauley J. GISAID: Global initiative on sharing all influenza data - from vision to reality. *Euro Surveill.* 2017;22(13):30494.
12. Moniruzzaman M, Hossain MU, Islam MN, et al. Coding-Complete Genome Sequence of SARS-CoV-2 Isolate from Bangladesh by Sanger Sequencing. *Microbiol Resour Announc.* 2020;9(28):e00626-20. Published 2020 Jul 9.

13. Narayanan K, Ramirez SI, Lokugamage KG, Makino S. Coronavirus nonstructural protein 1: Common and distinct functions in the regulation of host and viral gene expression. *Virus Res.* 2015;202:89-100.
14. Shin, D., Mukherjee, R., Grewe, D. et al. Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature* (2020).
15. Morales DJ, Lenschow DJ. The antiviral activities of ISG15. *J Mol Biol.* 2013;425(24):4995-5008.
16. Zhang Z, Schwartz S, Wagner L, Miller W. A greedy algorithm for aligning DNA sequences. *J Comput Biol.* 2000;7(1-2):203-214.
17. Fleischmann WR Jr. Viral Genetics. In: Baron S, editor. *Medical Microbiology*. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston; 1996. Chapter 43. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK8439/>
18. Rahman, A., Bhattacharjee, A., Jaber, A., Hossain, M., Hasan, K., Islam, S., & Bhuyan, Z. (2020). High-Risk Non-synonymous SNPs of Human Bcl-2 Gene Alters Structural Stability and Small Molecule Binding. *Bioresearch Communications*, 6(1), 793-804. Retrieved from: <http://bioresearchcommunications.com/index.php/brc/article/view/20>
19. Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins.* 2006;62(4):1125-1132.
20. Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics.* 2015;31(16):2745-2747.
21. Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics.* 2010;11:548. Published 2010 Nov 8.
22. Berman HM, Westbrook J, Feng Z, et al. The Protein Data Bank. *Nucleic Acids Res.* 2000;28(1):235-242.
23. Waterhouse A, Bertoni M, Bienert S, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 2018;46(W1):W296-W303.
24. Bhattacharya D, Nowotny J, Cao R, Cheng J. 3Drefine: an interactive web server for efficient protein structure refinement. *Nucleic Acids Res.* 2016;44(W1):W406-W409.
25. Heo L, Park H, Seok C. GalaxyRefine: Protein structure refinement driven by side-chain repacking. *Nucleic Acids Res.* 2013;41(Web Server issue):W384-W388.
26. Lovell SC, Davis IW, Arendall WB 3rd, et al. Structure validation by C $\alpha$  geometry: phi,psi and C $\beta$  deviation. *Proteins.* 2003;50(3):437-450.
27. Freitas B, Durie I, Murray J et al. Characterization and Noncovalent Inhibition of the Deubiquitinase and deISGylase Activity of SARS-CoV-2 Papain-Like Protease. *ACS Infect Dis.* 2020;6(8):2099-2109.
28. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem.* 2010;31(2):455-461.
29. Bhattacharjee A, Hossain MU, Chowdhury ZM, et al. Insight of druggable cannabinoids against estrogen receptor  $\beta$  in breast cancer [published online ahead of print, 2020 Mar 11]. *J Biomol Struct Dyn.* 2020;1-10.

30. Báez-Santos YM, St John SE, Mesecar AD. The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds. *Antiviral Res.* 2015;115:21-38.
31. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.* 2004;25(13):1605-1612.
32. Park T, Baek M, Lee H, Seok C. GalaxyTongDock: Symmetric and asymmetric ab initio protein-protein docking web server with improved energy parameters. *J Comput Chem.* 2019;40(27):2413-2417.
33. Xue LC, Rodrigues JP, Kastiris PL, Bonvin AM, Vangone A. PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics.* 2016;32(23):3676-3678.
34. Chen VB, Arendall WB 3rd, Headd JJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr.* 2010;66(Pt 1):12-21.
35. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. *Nat Med.* 2020;26(4):450-452.
36. Gojobori, T., Li, W. & Graur, D. Patterns of nucleotide substitution in pseudogenes and functional genes. *J Mol Evol.* 18, 360–369 (1982).
37. Lyons DM, Lauring AS. Evidence for the Selective Basis of Transition-to-Transversion Substitution Bias in Two RNA Viruses. *Mol Biol Evol.* 2017;34(12):3205-3215.
38. Koyama, T., Platt, D., & Parida, L. (2020). Variant analysis of SARS-CoV-2 genomes. Retrieved 26 July 2020, from <https://www.who.int/bulletin/volumes/98/7/20-253591/en/>
39. Yang D, Leibowitz JL. The structure and functions of coronavirus genomic 3' and 5' ends. *Virus Res.* 2015;206:120-133.
40. Yoshimoto FK. The Proteins of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS CoV-2 or n-COV19), the Cause of COVID-19. *Protein J.* 2020;39(3):198-216.
41. Huang C, Lokugamage KG, Rozovics JM, Narayanan K, Semler BL, Makino S. SARS coronavirus nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs: viral mRNAs are resistant to nsp1-induced RNA cleavage. *PLoS Pathog.* 2011;7(12):e1002433.
42. Kamitani W, Narayanan K, Huang C, et al. Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. *Proc Natl Acad Sci U S A.* 2006;103(34):12885-12890.
43. Mantlo E, Bukreyeva N, Maruyama J, Paessler S, Huang C. Antiviral activities of type I interferons to SARS-CoV-2 infection. *Antiviral Res.* 2020;179:104811.
44. Lokugamage, K. G., Hage, A., Schindewolf, C., Rajsbaum, R., & Menachery, V. D. (2020). SARS-CoV-2 is sensitive to type I interferon pretreatment. *bioRxiv : the preprint server for biology*, 2020.03.07.982264.
45. Wathelet MG, Orr M, Frieman MB, Baric RS. Severe acute respiratory syndrome coronavirus evades antiviral signaling: role of nsp1 and rational design of an attenuated strain. *J Virol.* 2007;81(21):11620-11633.
46. Züst R, Cervantes-Barragán L, Kuri T, et al. Coronavirus non-structural protein 1 is a major pathogenicity factor: implications for the rational design of coronavirus vaccines. *PLoS Pathog.* 2007;3(8):e109. doi:10.1371/journal.ppat.0030109

47. Patick AK, Potts KE. Protease inhibitors as antiviral agents. *Clin Microbiol Rev.* 1998;11(4):614-627.

**Tables:**

**Table 1: Mutations in SARS-CoV-2 NIB-1.**

| <b>Previously reported mutations</b> |                           |                          |                          |                             |                            |
|--------------------------------------|---------------------------|--------------------------|--------------------------|-----------------------------|----------------------------|
| <b>SL No.</b>                        | <b>Reference Position</b> | <b>Position in NIB-1</b> | <b>Reference → NIB-1</b> | <b>Change of Amino Acid</b> | <b>Coding region/ Gene</b> |
| 1                                    | 241                       | 93                       | C → T                    | -                           | 5` UTR                     |
| 2                                    | 1163                      | 1015                     | A → T                    | I120F                       | Orf1ab                     |
| 3                                    | 629                       | 481                      | C → A                    | L122I                       | Orf1ab                     |
| 4                                    | 8171                      | 8023                     | G → A                    | A1818T                      | Orf1ab                     |
| 5                                    | 23403                     | 23255                    | A → G                    | <b>D614G*</b>               | S                          |
| 6                                    | 28882                     | 28734                    | G → A                    | R203K                       | N                          |
| 7                                    | 28883                     | 28735                    | G → C                    | <b>G204R*</b>               | N                          |
| <b>Novel mutations</b>               |                           |                          |                          |                             |                            |
| 8                                    | 627                       | 479                      | T → A                    | V121D                       | Orf1ab                     |
| 9                                    | 5246                      | 5098                     | G → T                    | V843F                       | Orf1ab                     |
| 10                                   | 5385                      | 5237                     | C → T                    | A889V                       | Orf1ab                     |
| 11                                   | 7790                      | 5642                     | G → T                    | G1691C                      | Orf1ab                     |

**Asterisk (\*) containing D614G and G204R made the strain a member of the GR clade.**

**Table 2: Effect of the novel mutations in protein stability, function and structure.**

| <b>SL No.</b> | <b>Change of Amino Acid</b> | <b>Protein Stability (MUPro)</b> | <b>Protein Function (PROVEAN; cutoff= -2.5)</b> | <b>Protein Structural Properties (HOPE)</b>   |
|---------------|-----------------------------|----------------------------------|---|---|
| 1             | V121D                       | DECREASE                         | Deleterious                                     | <p><b>Change of Amino Acid :</b><br/>NEUTRAL→ NEGATIVE</p> <p><b>Change in Structural Stability:</b> Slight destabilization of PLPro</p> <p><b>Level of Conservation:</b><br/>Highly Conserved</p>  |
| 2             | V843F                       | DECREASE                         | Neutral   | <p><b>Change of Amino Acid Size:</b><br/>Mutant residue size is bigger</p> <p><b>Change in structural stability:</b><br/>Disturbing the domain and might abolish its function</p> <p><b>Level of Conservation:</b><br/>Highly conserved</p> |
| 3             | A889V                       | INCREASE                         | Neutral   | <p><b>Change of Amino Acid Size:</b></p>  |

|   |        |          |             |   |
|---|--------|----------|-------------|---|
|   |        |          |             | <p>Mutant residue size is bigger.</p> <p><b>Change in Structural Stability:</b></p> <p>Disturbing the domain and might abolish its function</p> <p><b>Level of Conservation:</b></p> <p>Located near a highly conserved position, this mutation might occur without damaging the protein.</p>   |
| 4 | G1691C | DECREASE | Deleterious | <p><b>Change of Amino Acid Size:</b></p> <p>Mutant residue size is bigger and more hydrophobic.</p> <p><b>Change in Structural Stability:</b></p> <p>The wild-type residue is a glycine that might be necessary to give the protein essential flexibility. Mutation of this glycine might abolish the function of the protein.</p> <p><b>Level of Conservation:</b></p> |

|  |  |  |  |   |
|--|--|--|--|---|
|  |  |  |  | <p>The mutant residue is located near a highly conserved position. The mutant and wild-type residue are not very similar. Based on this conservation information this mutation is probably damaging to the protein.</p> |
|--|--|--|--|---|

**Table 3: Interaction between wild type and mutant SARS-CoV-2 Papain-Like viral protease with GRL0617 and ISG15.**

| <b>SARS-CoV-2 Papain-Like viral protease and GRL0617 interaction patterns</b> |  |  |  |  |
|---|--|--|--|--|
| <b>Receptor</b>   | <b>Binding affinity score (kcal/mol)</b> | <b>Participating amino acids in interaction</b>        |  |  |
| Wild type   | -6.6                                     | ASP164, ARG166, GLU167, <i>SER170</i> , PRO248, TYR264 |  |  |
| V843F   | -6.7                                     | ASP76, ARG82, <i>CYS155</i> , TYR171                   |  |  |
| A889V   | -6.9                                     | ASP76, ARG82, <i>CYS155</i> , TYR171                   |  |  |
| V843F+<br>A889V   | -6.2                                     | ASP76, ARG82, <i>THR75</i> , PRO77                     |  |  |
| <b>SARS-CoV-2 Papain-Like viral protease and ISG15 interaction patterns</b>   |  |  |  |  |
| <b>Receptor</b>   | <b>TongDock_A score</b>                  | <b>Cluster size</b>                                    | <b>Binding affinity score (kcal/mol)</b> | <b>Dissociation constant (K<sub>d</sub>) (M)</b> |
| Wild type   | 915.278                                  | 41   | -15.4                                    | 4.8E-12  |
| V843F   | 720.354                                  | 23   | -11.1                                    | 7.2E-09  |

|                 |          |    |       |         |
|-----------------|----------|----|-------|---------|
| A889V           | 1019.921 | 31 | -15.4 | 4.7E-12 |
| V843F+<br>A889V | 968.296  | 28 | -15.5 | 4.2E-12 |

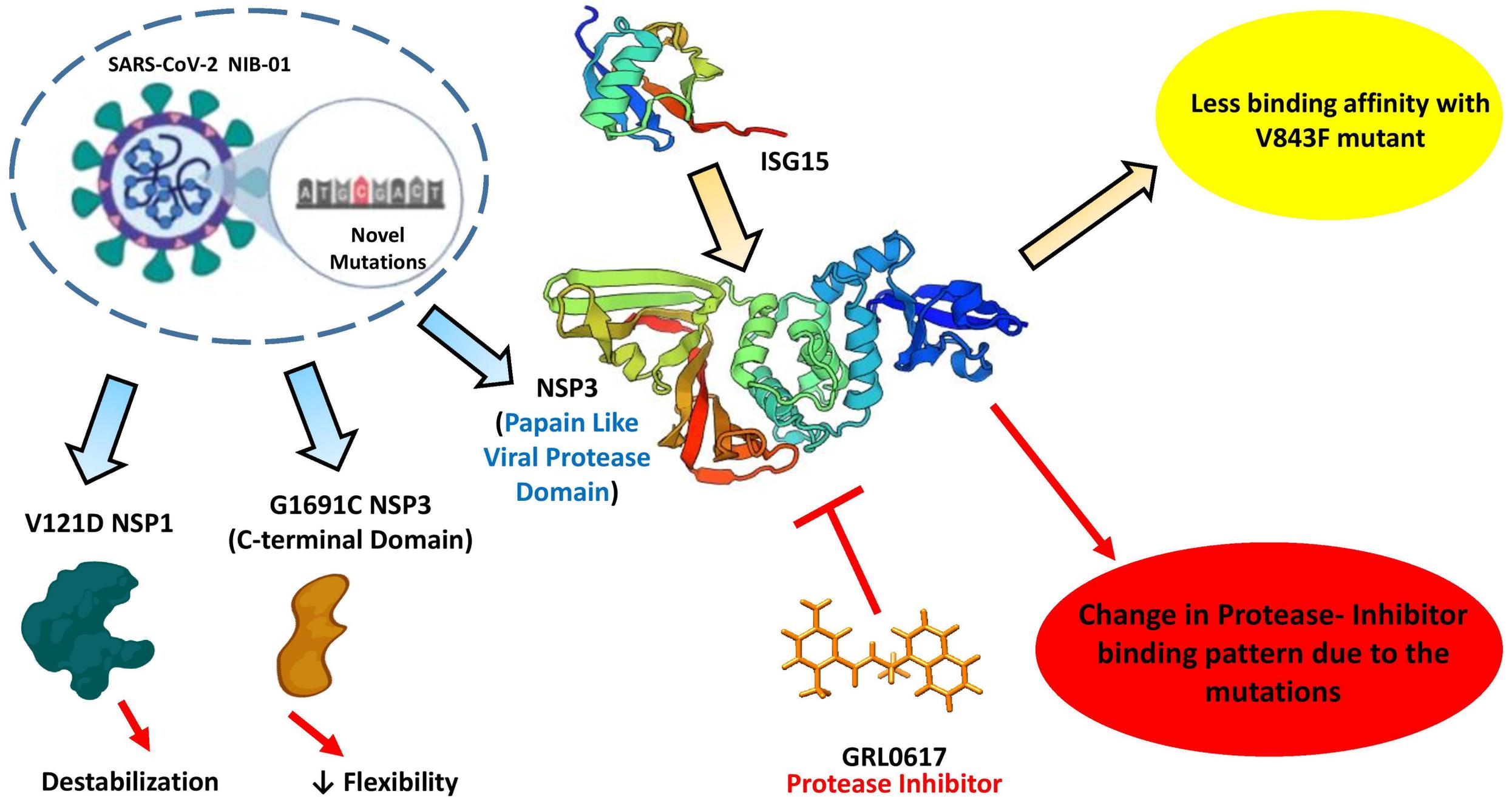
**Italicized residues interacted via hydrogen bonds.**

### Figure Legends:

**Figure 1:** Graphical representation of the whole work. Some of the parts of this figure were generated with the help of BioRender (<https://biorender.com/>).

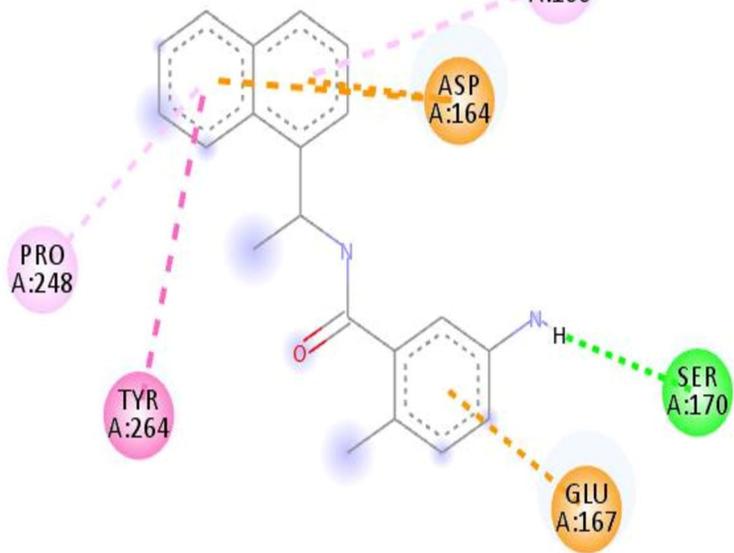
**Figure 2:** Interactions between GRL0617 inhibitor and superimposed SARS-CoV-2 wild type and mutant Papain-Like viral proteases (**a-d**). Interacting amino acid residues in wild type and PLPro mutant SARS-CoV-2. Inhibitor binding into the wild type and mutant protein. The red and blue boxes indicate the location and the interactions between ligand-receptors in wild and in mutant proteases respectively.

**Figure 3:** Interactions between the C-terminal domain of Interferon-Stimulated Gene 15 (ISG15) and SARS-CoV-2 wild type Papain-Like viral proteases. The red chain is the C-terminal domain of ISG 15. The blue chain represents the protease.



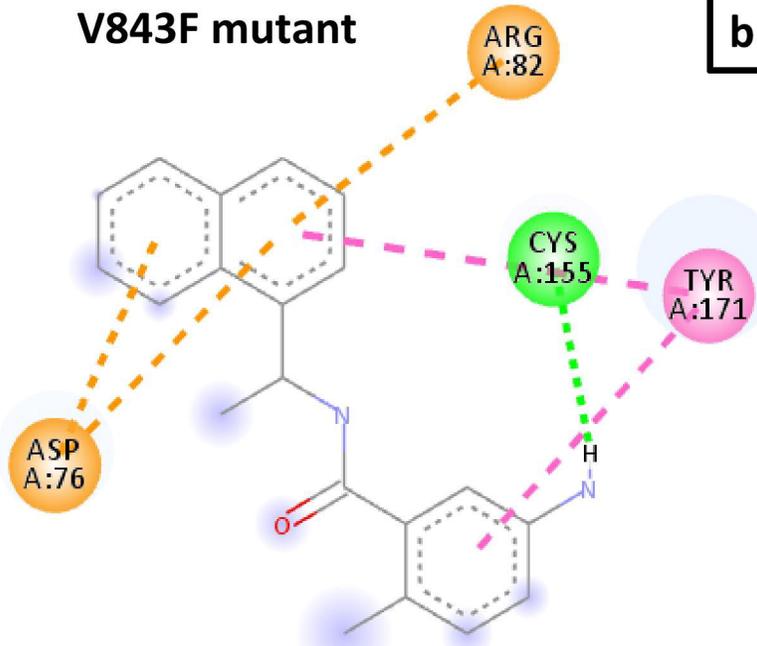
Wild type

a



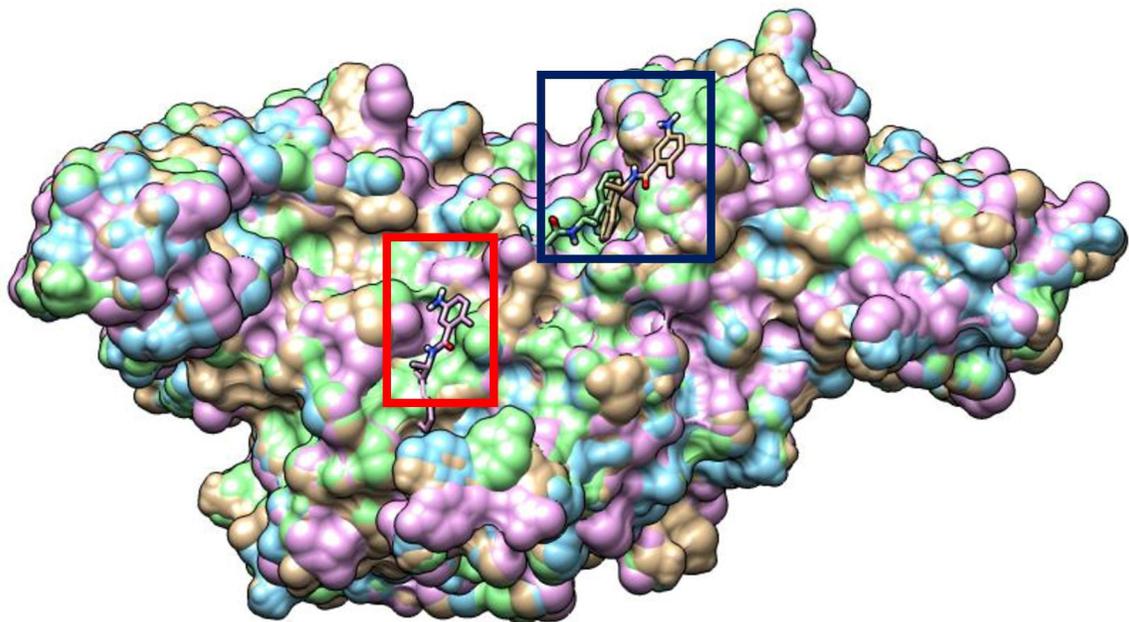
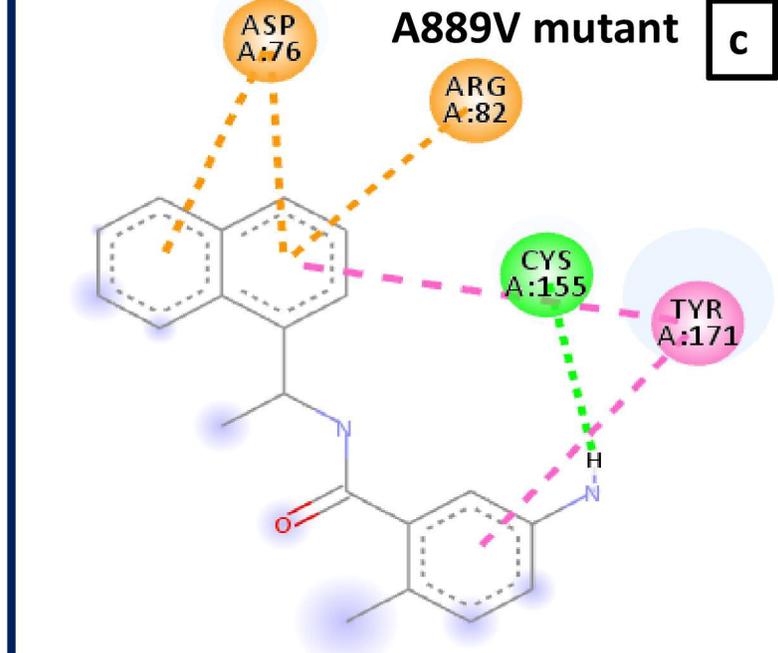
V843F mutant

b



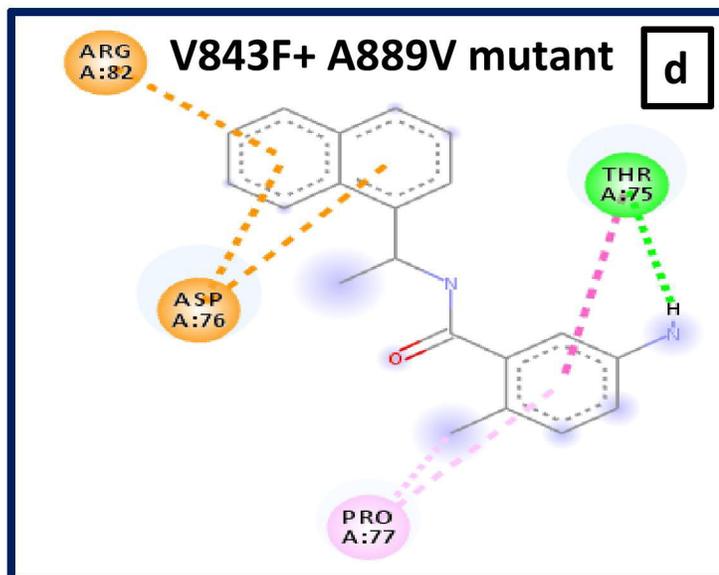
A889V mutant

c



V843F+ A889V mutant

d



Interactions

Conventional Hydrogen Bond

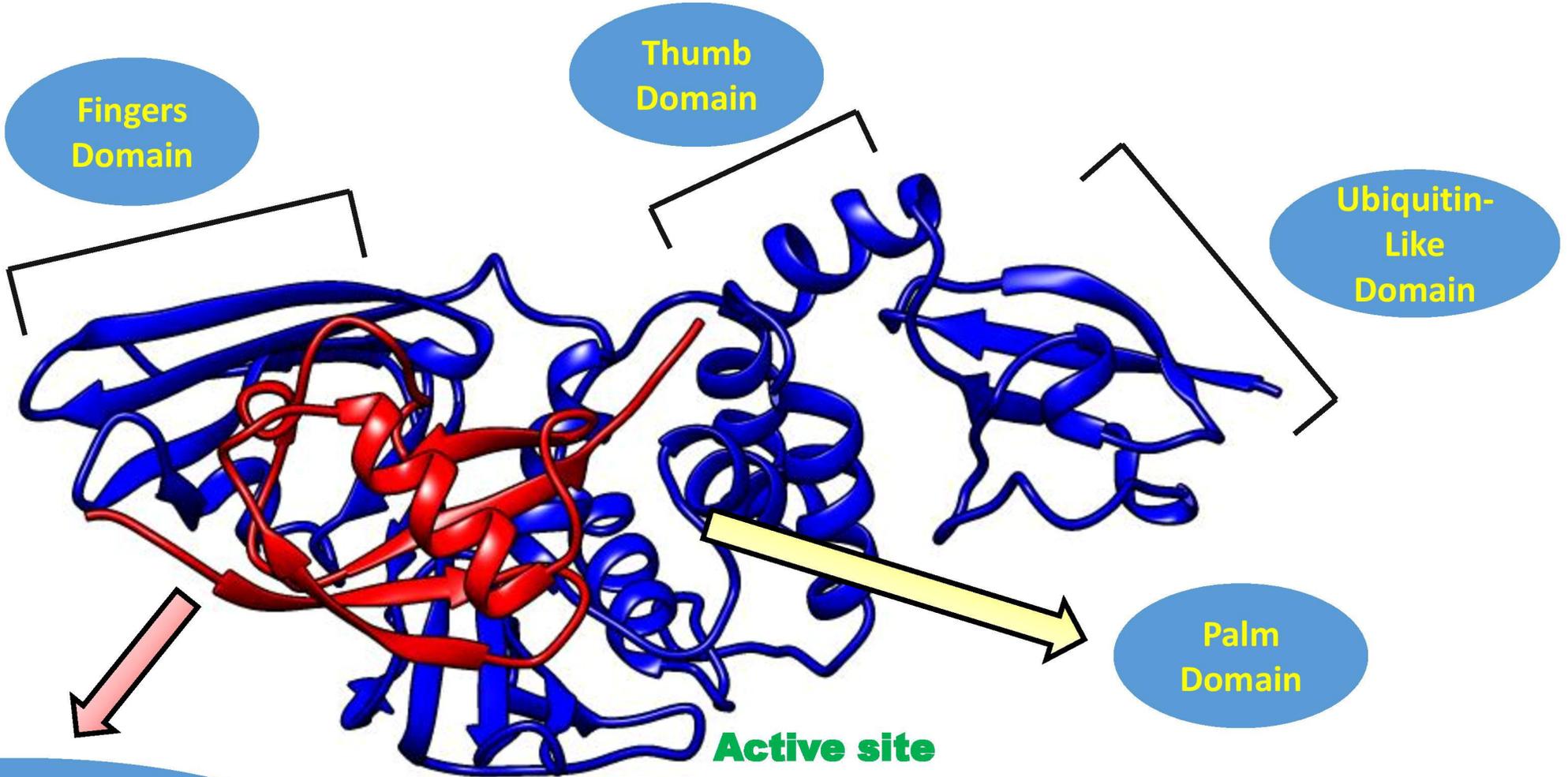
Pi-Cation

Pi-Anion

Amide-Pi Stacked

Alkyl

Pi-Alkyl



Fingers Domain

Thumb Domain

Ubiquitin-Like Domain

Palm Domain

Active site

ISG15 C-terminal Domain complex with Ubiquitin binding site 1