

Setayesh Yazdani

M.Sc. Student at Structural Genomics Consortium (SGC) Toronto

Pharmacology and Toxicology Department, University of Toronto

“Genetic Variability at the Catalytic Site of the Papain-like Protease and the Predicted Effects of Mutations on the binding of the Inhibitor VIR251”

Background:

SARS-CoV-2 Papain-like protease (PLPro) is a cysteine protease. PLPro recognizes the tetrapeptide LXGG motif found in-between the fused viral proteins nsp1 and nsp2, nsp2 and nsp3, and nsp3 and nsp4 (nsp1/2, nsp2/3, nsp3/4). It then hydrolyzes the peptide bonds on the carboxyl side of glycine at the P1 position. Consequently, nsp1, nsp2, and nsp3 proteins get released, which is necessary for viral replication. SARS-CoV-2 PLPro also helps virus evasion from the host innate immune responses through its effects on the deubiquitination and deISGylation process. Hence PLPro is essential for SARS-CoV-2 replication and is a promising target for antiviral drugs development.¹

Wioletta Rut and their colleagues have crystalized structures of two inhibitors (VIR250 and VIR251) in complex with SARS-CoV-2-PLpro which uncovers their inhibitory mechanisms (PDB: 6wx4).¹ The Catalytic cysteine 111 of CoV-2-PLPro participates in Michael Addition to the beta-carbon of the vinyl group of the VME warhead of VIR251 as shown in figure 1. This reaction results in the formation of a covalent thioether linkage between VIR251 and Cys111 of PLPro, as shown in figure 2.

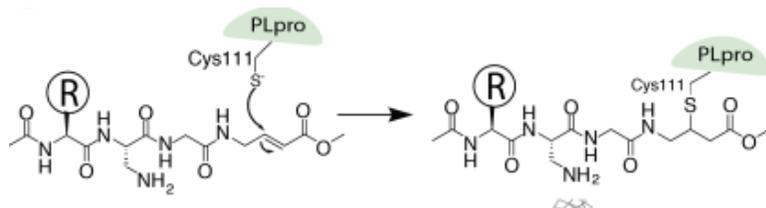


Figure 1. The Michael Addition to the beta-carbon of the vinyl group of the VME warhead of VIR251 by catalytic cysteine 111.¹

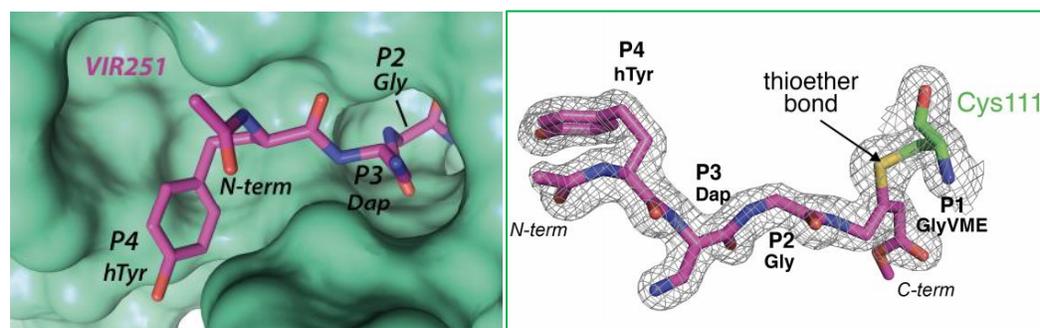


Figure 2. Formation of thioether bond between the non-natural amino acid-containing inhibitor, VIR251 and Cys111.¹

In the following section of this document, I will outline all the steps which I have taken to analyze the catalytic site of SARS-CoV-2 PLPro. In the context of emergence of future MERS-like or SARS-like coronaviruses from the bat strains circulating in bat reservoir species, it is imperative to do a broad survey of viral proteins to identify the best strategies for the development of broad-spectrum viral inhibitors.² Hence we map the genetic variations of coronaviruses onto PLPro crystal structure (PDB: 6wx4) to elucidate how those mutations might affect VIR251 binding to its active site.

Method:

Step 1: Finding drug-binding sites on PLPro and assessing their druggability (PDB: 6wx4).

Using ICM function PocketFinder (Molsoft, SD) we found the potential drug-binding sites using the PLPro x-ray crystal structure (PDB: 6wx4). Then using SiteMap (Schrodinger, NY) we assessed the druggability of the binding sites. The catalytic site of PLPro has a druggability score (Dscore) of 0.81 indicative of a difficult druggable site. It is important to note that PLPro catalytic site has shown to be druggable using covalent inhibitors such as VIR251.¹

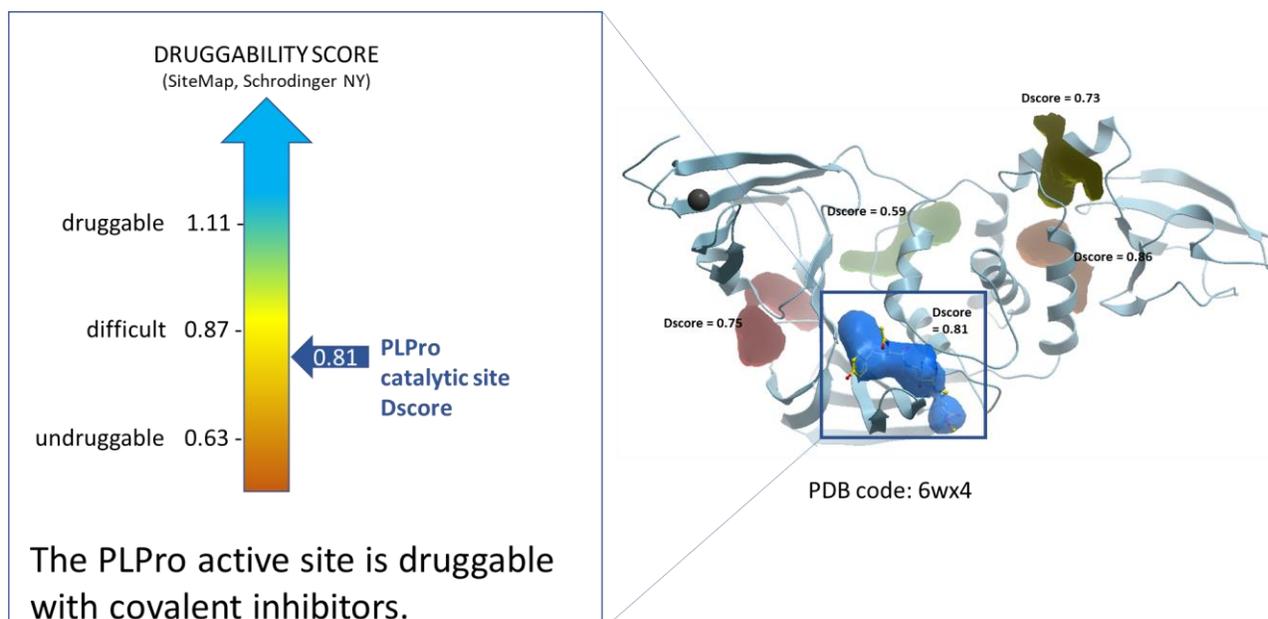


Figure 3. Druggability analysis of PLPro catalytic site. The blue pocket shows the catalytic site of the PLPro protein which is occupied by the non-natural amino acid-containing inhibitor, VIR251.

Step 2. Determine the residues that line the catalytic site of PLPro (PDB: 6wx4).

Using ICM, we found the amino acid sidechains that face the catalytic pocket within 2.8 Å vicinity of the catalytic pocket. We identified twenty-four sidechains in total which include: W106, N109, C111, Y112, L162, G163, D164, V165, R166, M208, S245, A246, P247, P248, Y264, N267, Y268, Q269, C270, G271, H272, Y273, T301, D302. Figure 4 shows these residues in sticks representation around the inhibitor-bound pocket of PLPro.

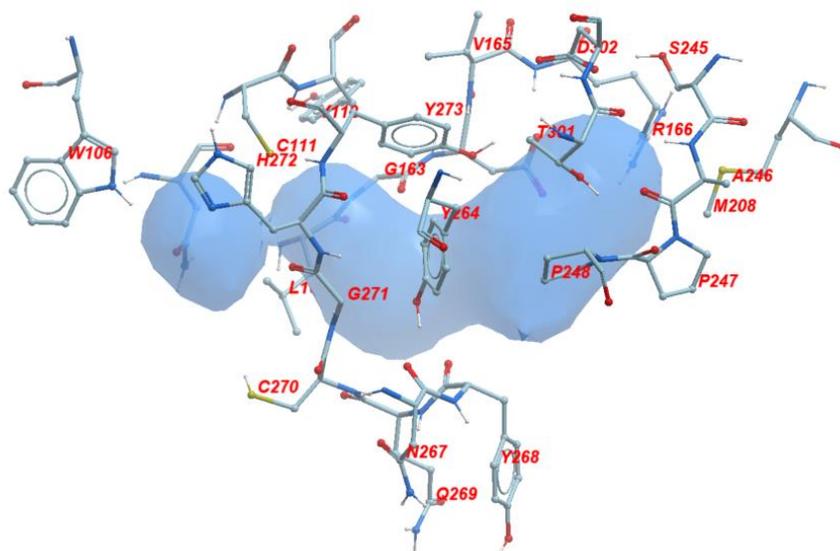


Figure 4. The neighboring residues lining the catalytic pocket of PLPro.

Step 3. Make diversity dendrograms for sequences of the Alpha- and Betacoronavirus genera entries in UniProt database.

Using 28 reviewed sequences from Uniprot, where 6 belong to entries of Alphacoronavirus genus and 21 belong to Betacoronavirus genus, I made a diversity dendrogram focusing on the amino acid residues of the PLPro catalytic site. Shown below in figure 5 is the diversity dendrogram, where the diversity profile is shown. Strikingly, only 5 out of 24 residues at PLPro catalytic site are conserved across these 28 entries: C111 (the catalytic residue which binds covalently to the inhibitor VIR251), D164, G271, H272, Y273.

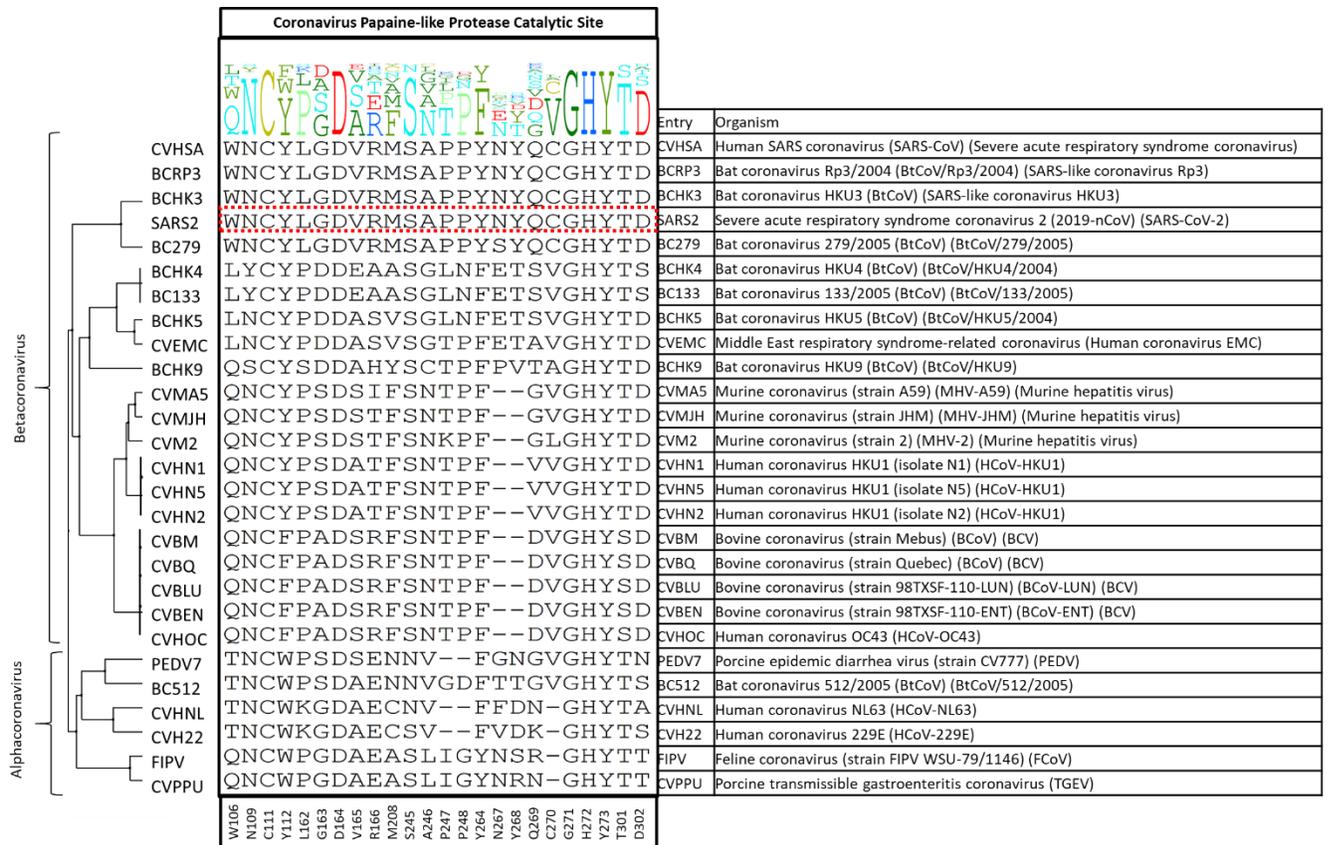


Figure 5. The diversity dendrogram of the catalytic site of PLPro for the entries of Alpha- and Betacoronavirus genera

Step 4: Make diversity dendrograms for reviewed sequences of the Betacoronavirus genus entries in UniProt database.

We made a diversity dendrogram of the 21 entries of the Betacoronavirus genus. As shown in figure 6, we still observe high variability among these entries where only 6 out of 24 residues are conserved at PLPro catalytic site: C111 (the catalytic residue which binds covalently to the inhibitor VIR251), D164, S245, G271, H272, Y273.

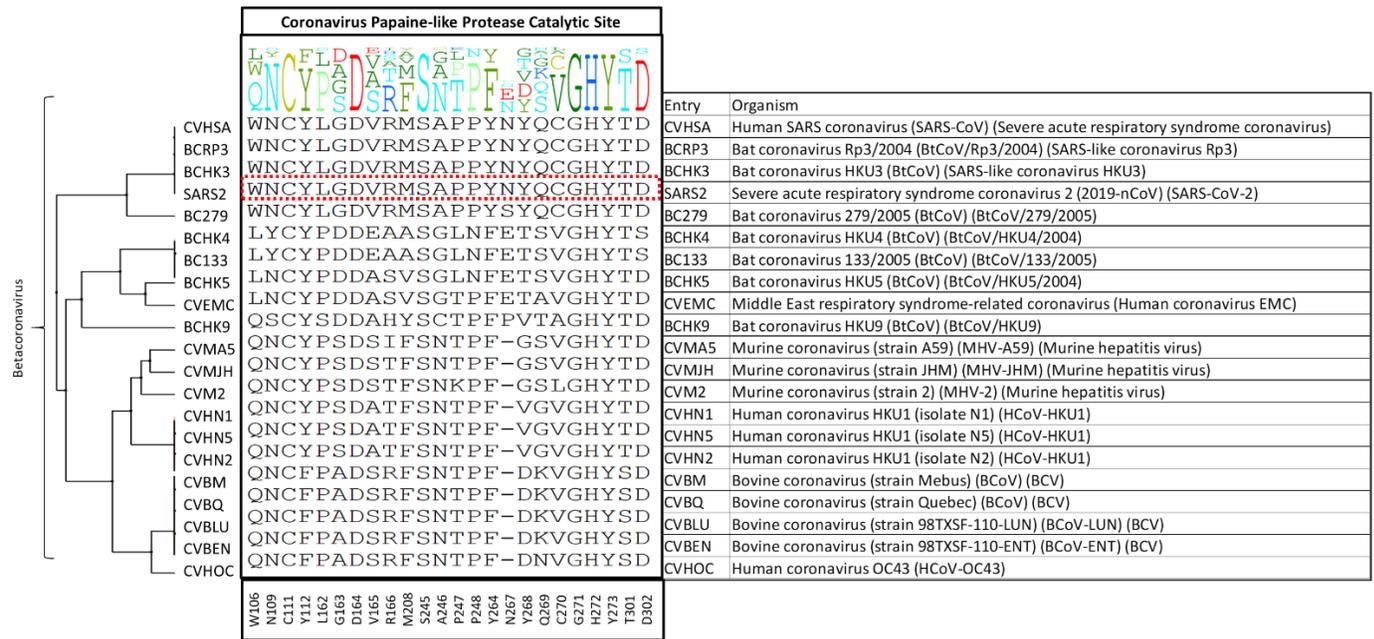


Figure 6. The diversity dendrogram of PLPro catalytic site residues among the UniProt entries of the Betacoronavirus genus.

Step 5: Map variations observed among Alpha- and Betacoronavirus genera entries and the variations within only the Betacoronavirus entries onto SARS-CoV-2 PLPro crystal structure (PDB: 6wx4).

We mapped the variations stated in step 3 and 4 onto SARS-CoV-2 PLPro crystal structure. In figure 7, you can see the color-coded structure of SARS-CoV-2 and the sidechains that are variable among the entries considered in step 3 and 4 highlighted in green. Later in step 8, the predicted effect of each variation on ligand binding is provided.

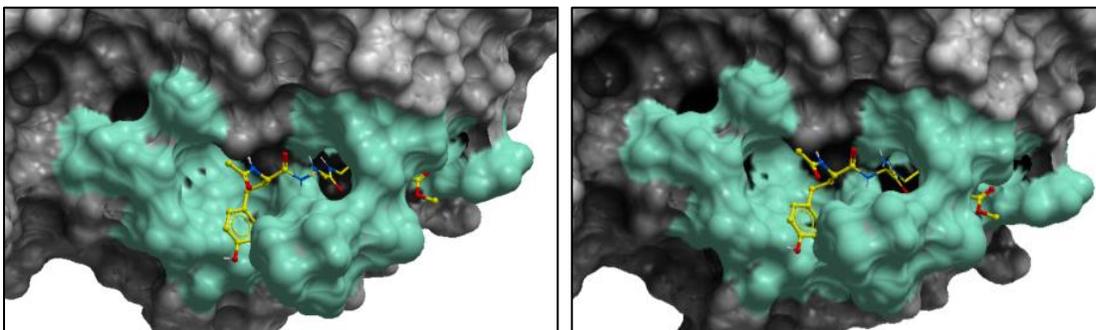


Figure 7. Mapping the genetic variation of entries in the Alpha- and Betacoronavirus genera onto SARS-CoV-2 Crystal structure. The green color shows the non-conserved residues among the entries from the Alpha- and Betacoronavirus genera. On the left, all twenty residues that were non-conserved among the twenty-eight entries of Alpha- and Betacoronavirus genera are shown (W106, N109, N109, Y112, L162, G163, V165, R166, M208, S245, A246, P247, P248, Y264, N267, Y268, Q269, C270, T301, D302). On the right, all nineteen residues that were non-conserved among the twenty-one entries of Betacoronavirus genus are shown (W106, N109, N109, Y112, L162, G163, V165, R166, M208, A246, P247, P248, Y264, N267, Y268, Q269, C270, T301, D302).

Step 6: Assessing the genetic variability of PLPro catalytic site among SARS-CoV-2 samples.

Nicola De Maio, our collaborator from Nick Goldman's lab at EBI, looked at more than 15000 SARS-CoV-2 samples and identified all mutations at the PLPro catalytic site. All the five non-synonymous variants are shown in table 1 below. For example, at residue number 247, the wild-type sidechain in SARS-CoV-2 PLPro is a proline. In the sample batch Nicola has looked at, he saw that 15868 had proline at that position while there were 3 other variants at this position: 1) serine 2) leucine 3) glutamine.

Residue	Non-synonymous variations at PLPro catalytic site
P247	P (15868) S (1), L (2), Q (1)
P248	P (15871), S (1)
T301	T (15864), A (5)

Table 1. The non-synonymous variants at PLPro catalytic site across more than 15000 SARS-CoV-2 samples.

Step 7. Energy calculations for the SARS-CoV-2 variants at PLPro catalytic site and predicting their effects on VIR251 binding.

In ICM, we modeled the specific mutations from step 6 and used the change in Gibbs free energy (ddGbind) to predict the effect of those mutations on ligand binding. The ligand in this case is the viral inhibitor, VIR251 (PDB: 6wx4). In Table 2, the ddGbind values for all five non-synonymous mutations are shown.

Residue Number	Wild Type	Mutant	ddGbind	dGbind Wild type	dGbind Mutant
247	pro	leu	5.2	-34.2	-29.0
247	pro	ser	-0.5	-34.2	-34.7
247	pro	gln	0.7	-34.2	-33.5
248	pro	ser	3.5	-34.2	-30.7
301	thr	ala	0.0	-34.2	-34.2

Table 2. Energy calculations of non-synonymous mutations at SARS-CoV-2 catalytic site to predict their effect on the binding of VIR251. The energy unit for ddGbind values is Kcal/mol.

Looking at table 2, we see that ddGbind values for P247L, P248S are above 2Kcal/mol which is considered a significant change in binding according to a previous study.³ P247S, P247Q and T301A were predicted to have minor effects on ligand binding (-2kcal/mol < ddGbind < 2kcal/mol).

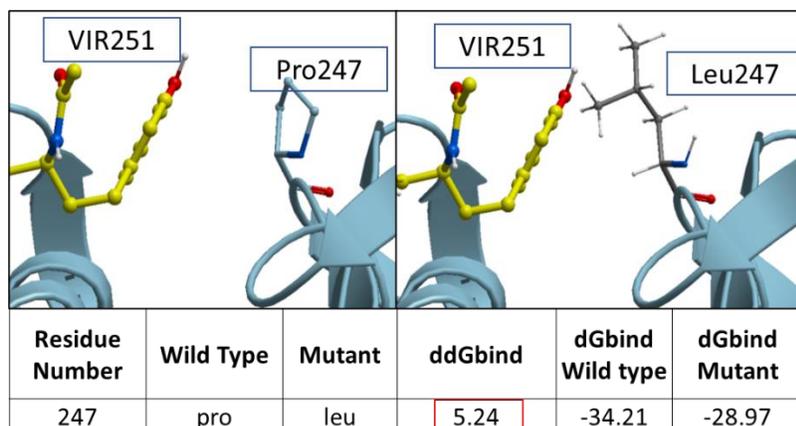


Figure 8. The ribbon diagram of SARS-CoV-2 PLPro's variant: P247L.

We think that the predicted destabilizing effect of P247L on ligand binding could be due to bulkiness of leucine compared to proline in proximity to the aromatic ring of VIR251.

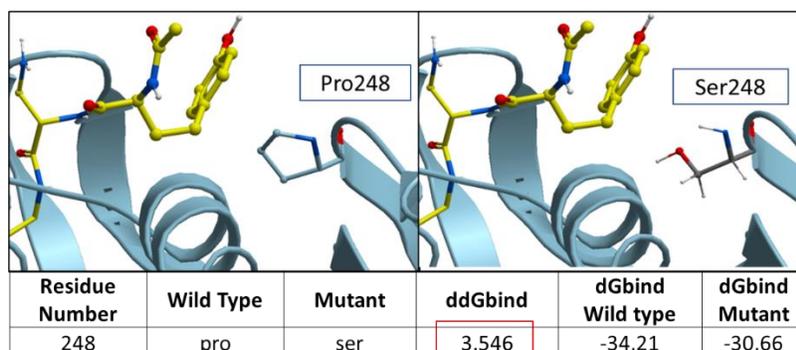


Figure 9. The ribbon diagram of SARS-CoV-2 PLPro's variant: P248S.

The ddGbind value corresponding to P248S is predicted to be at 3.547 Kcal/mol. Proline is a non-polar aliphatic residue which makes hydrophobic interaction with the aromatic ring of Tyrosine on VIR251. The change to serine which is a polar non-charged residue may negatively impact hydrophobic interactions with the ligand. In figure 10, the 2D ligand-protein diagram outlines the chemical interaction between VIR251 and SARS-CoV-2 PLPro. This 2D diagram also highlight P248's hydrophobic interaction with the aromatic ring of Tyrosine of the non-natural amino acid-containing inhibitor, VIR251.

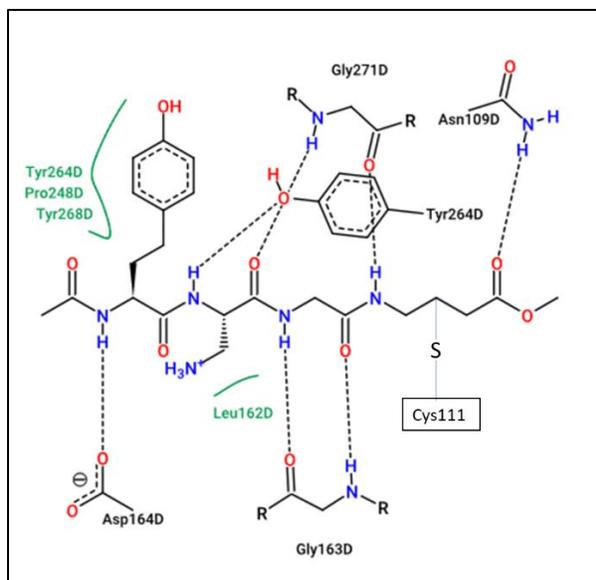


Figure 10. 2D ligand-protein diagram showing the chemical interactions between VIR251 and PLPro (PDB: 6wx4). This diagram was made using PoseView in Proteins.Plus.

The color-coded structure in figure 11 highlights the variants among the SARS-CoV-2 samples at PLPro catalytic site.

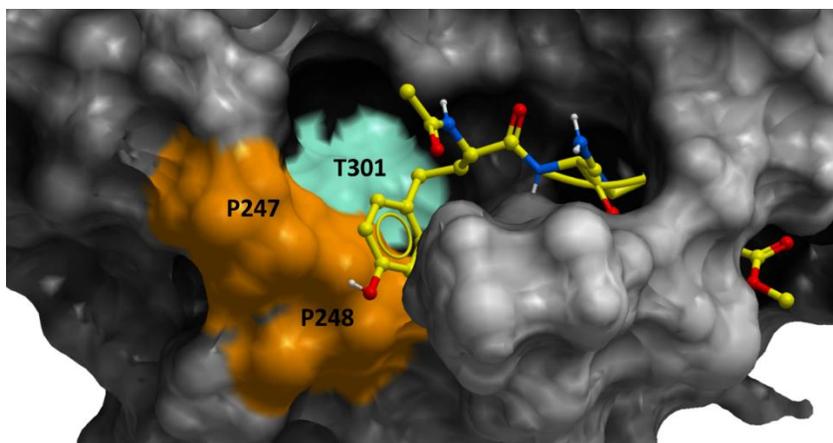


Figure 9. Color-coded surface representation of SARS-CoV-2 PLPro. The colored areas on the mesh show the non-conserved residues across more than 15000 SARS-CoV-2 samples from COVID-19 patients. The orange-colored residues correspond to variants, P247L and P248S, which penalizes VIR251-binding significantly, while the green-colored residues mutations T301A does not penalize ligand binding significantly.

Step 8: Predicting the effect of all possible mutations at the PLPro catalytic site on VIR251 binding.

We created an energy matrix table where we consider the effect of every possible mutation on the binding of VIR251 with PLPro. Table 3 includes a list of 23 residues expect Cys111. Cys111 is not included as it makes a covalent interaction with VIR251: any mutation at this position will be highly detrimental to ligand binding.

It is likely that the residues of the rows with green color are not making optimal interactions with VIR251. This is reflected in the very slight and non-significant lowering in ddGbind values of these

residues (mutations are stabilizing) and hence the green color. On the other hand, the sidechains whose mutations are shown in orange are likely the positions where PLPro would better interact with VIR251. As a result, the mutations of those sidechains are predicted to penalize inhibitor-binding significantly and hence the orange color.

Residue Number	WT	Mutant Residue																			
		gly	ala	val	leu	ile	pro	cys	met	ser	thr	phe	tyr	trp	asn	gln	asp	glu	his	lys	arg
106	W	0.1	0.1	0.2	0.2	0.2	0.0	0.5	-0.1	0.8	0.6	3.3	4.6	0.0	1.5	0.4	2.0	0.9	1.8	0.8	1.2
109	N	1.0	0.2	3.4	3.2	2.9	50.1	-0.4	-0.1	0.1	4.2	0.8	1.1	0.3	0.0	-1.1	1.0	84.6	-0.5	-0.4	0.4
112	Y	4.5	2.0	-1.8	-0.5	3.8	-0.3	1.5	0.9	2.2	0.3	-0.4	0.0	-1.2	0.7	0.4	2.7	1.8	0.2	1.0	1.3
162	L	6.8	3.7	5.0	0.0	41.7	5.5	2.5	1.6	3.7	4.9	0.1	0.7	18.4	2.8	3.8	3.2	2.4	2.3	-0.1	0.9
163	G	0.0	0.1	0.2	0.1	-0.1	201.5	-0.2	0.0	0.7	0.5	0.0	0.0	0.2	0.0	0.0	0.6	0.5	-0.3	-0.2	-0.6
164	D	2.6	2.6	3.7	1.3	8.4	42.8	2.7	11.3	2.7	4.6	3.4	2.6	79.0	5.5	25.2	0.0	-1.1	2.6	4.2	2.7
165	V	10.0	-0.3	0.0	65.5	-0.1	-0.3	-0.2	20.1	-0.1	-0.3	-0.5	-0.5	-0.3	-0.6	-0.1	0.3	0.4	-0.4	-0.7	-0.7
166	R	-1.3	-1.0	28.7	-1.5	0.7	-0.2	-1.1	-1.9	-1.0	-0.9	-1.1	-1.0	-0.8	-1.6	-1.4	-1.4	-1.3	-1.1	-1.1	0.0
208	M	0.1	0.3	0.1	27.3	0.1	54.4	-0.1	0.0	0.2	0.1	0.1	0.1	0.1	0.6	0.0	0.6	-0.3	-0.2	0.4	1.1
245	S	-0.8	-0.2	0.5	-0.1	0.3	41.4	0.0	-0.1	0.0	0.3	-0.3	-0.3	-0.4	-0.2	0.0	-0.1	-0.1	0.0	0.1	0.2
246	A	-0.2	0.0	0.0	-0.1	0.1	41.0	-0.1	0.0	-0.1	0.1	-0.1	-0.2	0.0	-0.1	0.1	-0.1	0.1	-0.1	-0.1	-0.2
247	P	0.3	21.4	1.9	5.2	1.9	0.0	0.5	0.8	-0.5	3.3	4.4	3.6	48.0	1.5	0.7	0.4	0.7	3.5	1.6	0.6
248	P	3.6	3.0	2.2	2.1	1.9	0.0	2.4	3.9	3.5	3.1	3.5	4.7	1.9	2.1	2.6	5.8	3.9	2.1	2.1	0.6
264	Y	9.1	8.9	8.0	10.8	7.4	69.9	9.4	11.9	10.5	8.1	3.5	0.0	9.2	8.6	12.4	11.8	16.0	7.3	88.7	49.6
267	N	-0.1	-0.3	-0.6	-0.4	-0.5	59.4	-0.2	-0.3	-0.3	30.5	-0.3	-0.6	-0.3	0.0	-0.5	-0.3	-0.6	-0.7	-0.7	-0.3
268	Y	0.8	0.2	6.0	-0.3	9.0	0.1	-0.5	-0.2	-0.1	2.5	-0.2	0.0	0.0	-0.5	0.0	-0.1	0.5	-0.4	-0.3	0.2
269	Q	0.1	-0.3	0.0	0.0	0.0	69.2	-0.3	-0.3	0.0	0.0	-0.3	-0.1	0.3	-0.3	0.0	0.1	-0.2	-0.3	9.9	-0.4
270	C	-0.3	-0.4	-0.5	-0.6	-0.5	16.3	0.0	-0.5	-0.2	-0.3	-0.8	-0.6	-0.5	13.2	-0.1	0.6	-0.1	-0.4	-0.4	1.6
271	G	0.0	-2.2	-2.9	-1.0	-3.3	45.8	-2.1	11.1	0.0	-1.9	-6.7	-6.7	-4.5	-2.1	2.6	3.0	8.6	-4.1	59.2	46.3
272	H	0.3	0.2	-1.0	-0.2	-0.9	106.1	-0.1	0.2	0.4	-0.2	-0.7	-0.8	-0.6	-0.5	0.2	0.4	-0.8	-0.2	0.1	-0.4
273	Y	0.6	0.3	0.2	-0.2	-0.8	0.2	0.3	-1.3	1.0	0.8	-0.3	0.0	-0.5	0.0	-0.5	1.2	0.6	1.0	1.7	3.3
301	T	-0.2	0.0	-0.6	0.2	-0.3	-0.6	0.1	-0.3	0.3	0.0	-1.9	-1.9	-3.4	-1.3	1.1	3.7	5.4	-1.5	9.6	15.4
302	D	0.1	-0.2	0.0	11.8	-0.1	-0.4	-0.2	0.0	-0.2	-0.1	-0.1	0.0	0.1	-0.4	-0.2	0.0	0.2	0.2	0.0	-0.6

Table 3. The energy matrix showing the effect of all possible mutations at SARS-CoV-2 PLPro catalytic site and the predicted effect on VIR251 binding. The energy unit is kcal/mol. WT = wild type residue.

Color-coding based on ddGbind value (X) in Kcal/mol
X > 10.0
4.0 > X > 10.0
2.0 > X > 4.0
-2.0 > X > 2.0
-4.0 > X > -2.0
-10.0 > X > -4.
-10.0 > X

Altogether, SARS-CoV-2 variants are predicted to penalize VIR251-binding only slightly with PLPro. From the diversity dendrograms, it became evident that the catalytic site residues of PLPro are mainly non-conserved and that would limit its application as a target-site for wide-spectrum anti-coronavirus inhibitors.

References:

1. Rut, W. et al. Activity profiling and structures of inhibitor-bound SARS-CoV-2-PLpro protease provides a framework for anti-COVID-19 drug design. <https://www.biorxiv.org/content/10.1101/2020.04.29.068890v2>
2. Sheahan, T. P. An orally bioavailable broad-spectrum antiviral inhibits SARS-CoV-2 in human airway epithelial cell cultures and multiple coronaviruses in mice. *Science Translational Medicine* 12 (541). 2020. <https://doi.org/10.1126/scitranslmed.abb5883>
3. Schapira, M., Totrov, M., Abagyan, R. Prediction of the Binding Energy for Small Molecules, Peptides and Proteins. *Journal of Molecular Recognition* 12(3), 177-90 (1999). [https://doi.org/10.1002/\(SICI\)1099-1352\(199905/06\)12:3<177::AID-JMR451>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1099-1352(199905/06)12:3<177::AID-JMR451>3.0.CO;2-Z)