Repurposing nonnucleoside antivirals against SARS-CoV2 NSP12 (RNA dependent RNA polymerase) and identification of domain specific interactions

Feroza Begum[1,2], Upasana Ray[1,2]*

[1]CSIR-Indian Institute of Chemical Biology, 4, Raja S.C., Mullick Road, Jadavpur, Kolkata-700032, West Bengal, India.

[2]Academy of Scientific and Innovative Research (AcSIR), Ghaziabad- 201002, India

*Correspondence: ray.upasana@gmail.com

upasana.ray@iicb.res.in

Key words: SARS-CoV2, RdRp, NSP12, drug repurposing

Abstract

The pandemic of SARS-CoV-2 has necessitated expedited research efforts towards finding potential antiviral targets and drug development measures. While new drug discovery is time consuming, drug repurposing has been a promising area for elaborate virtual screening and identification of existing FDA approved drugs that could possibly be used for targeting against functions of various proteins of SARS-CoV-2 virus. RNA dependent RNA polymerase (RdRp) is an important enzyme for the virus that mediates replication of the viral RNA. Inhibition of RdRp could inhibit viral RNA replication and thus new virus particle production. Here, we screened non-nucleoside antivirals and found three out of them to be strongest in binding to RdRp. We propose these three drugs as potential RdRp inhibitors based on the site of binding.

Introduction

SARS-CoV2 RNA dependent RNA polymerase (NSP12) is a 932 amino acids long protein which is divided into two main functional domains: nidovirus RdRp associated nucleotidyl transferase (NIRAN) domain and RNA dependent RNA polymerase (RdRp) domain. NIRAN domain is important in nucleotidylation activity and helps in nucleotide transfer while RdRp domain is involved in the polymerisation activity (Fig 1). The N terminal region of the protein contains a beta hairpin structure (28th -50th a.a.) followed by the NIRAN domain (4th -28th, 69th -249th a.a.). The beta hairpin is important in stabilizing the overall structure of the enzyme. The NIRAN domain and the RdRp domain is linked via an Interface domain (250th -

365th). The RdRp domain is further divided into Finger, Palm and Thumb subdomains. The Finger subdomain (366th -581st a.a long termed as region 1 and 621st -679th a.a. long termed as region 2 in the study) and the Thumb subdomain (816th -919tha.a.) form a closed conformation which is supported by Nsp7 and Nsp8 and one molecule of Nsp8 binds to the Finger subdomain to interact with the Interface domain. Overall, the two subdomains create a particular conformation to help the Palm subdomain to function. The Palm subdomain (582nd -620th a.a. long termed as region 1 and 680th -815th a.a. long termed as region 2 in the study) possess different motifs (Motif A to E) each performing different functions and form the active site of the enzyme. Motif A Motif A (611th -626th) and Motif C (753rd -767th) form the main region of the active site that is important in RNA binding and polymerisation. Motif C possess catalytic residues (759th -767th a.a.) that is found to be conserved in other viral RdRps. The Finger subdomain possess two motifs (Motif F and Motif G) that works together and perform different functions. Motif F (538th -560th a.a.) and Motif G (500th -513th a.a.) forms a groove through which RNA template enters the active site composed of Motif A and C. Motif F also forms an NTP entry channel for the enzyme. Motif E (810th -821st a.a.) and the thumb subdomain support primer strand during RNA synthesis. The function of Motif B (680th -710th a.a.) and Motif D (774th -796th a.a.) is yet to be elucidated.

The RdRp Complex (NSP12) is composed of NSP12, NSP8 and NSP7 proteins. The NSP7-NSP8 heterodimer stabilise the closed conformation of NSP12 and is important in mediating NSP12-RNA interactions.

Materials and Methods

The different FDA approved nonnucleoside antivirals were downloaded from the ZINC database in SDF format and was converted to pdbqt file using Autodock tool. The crystal structure of SARS-CoV2 NSP12-NSP7-NSP8 complex was downloaded in PDB format from the Protein Data Bank (PDB ID: 6m71). Using the Pymol software the NSP12 (RNA dependent RNA polymerase) structure was extracted and colored as per domains. Initially docking was done using the Full protein to screen drugs based on the affinity cut off value of -9Kcal/mol. The positive control used in the study are Known approved antivirals Remdesivir monophosphate and Favipiravir ribose monophosphate. The structures were isolated using the Crystal structures available in PDB (PDB ID: 7bv2 and 4kn6 respectively). The binding affinity of the two drugs were found to be -6.9Kcal/mole and -7.8Kcal/mol respectively. As a negative

control Cinnamaldehyde and Thymoquinone structure was used due to its known lower affinity toward Polymerases of other viruses (ZINC ID: ZINC000001532777 and ZINC000000164367 respectively). The affinity was found to be -4.7 and -5.4 Kcal/mol respectively.

The ligand and the protein was prepared in autodock tool and saved in pdbqt format. Docking was performed by using Autodock vina tool and the docking result was analysed using Pymol software and Discovery Studio tool. The COACH tool was also used to identify the binding pockets of the protein. Domain specific analysis was performed by extracting different domains of NSP12 by Pymol software.

Results and Discussion

The drugs listed in Table 1 were screened based on their binding affinities towards the NSP12 protein. Keeping the cut off value of -9Kcal/mol, 3 drugs (Paritaprevir, Grazoprevir and Ledispavir) were shortlisted which showed affinities above -9Kcal/mol. These drugs were further used to perform domain specific interaction study to identify the interacting domains and their corresponding residues.

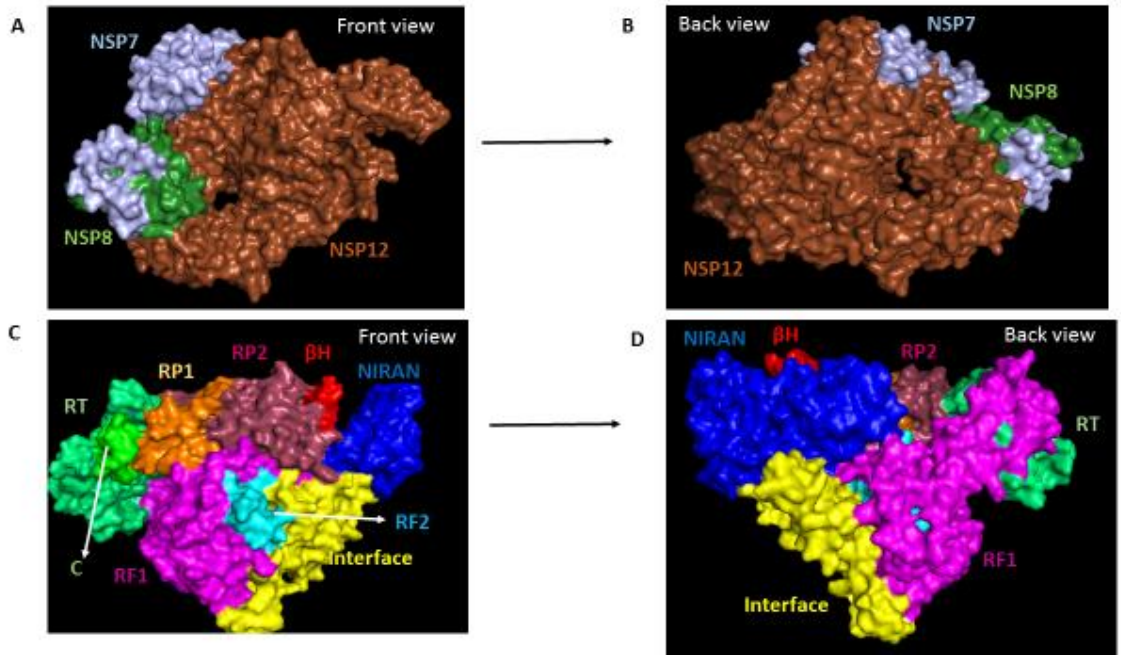

Figure1: Representation of the NSP12 and NSP12-NSP7-NSP8 complex. FigA and FigB. Represents the front and back surface view of the complex (PDB ID: 6m71). The different proteins are color coded. Figure C and D represents the different domains of the NSP12 with different color codes. βH- Beta hairpin region(V31-K50), NIRAN- NIRAN domain (Y69-R249), Interface- Interface region (A250-R365), RF1- NSP12 Finger subdomain region 1 (L366-A581), RF2- NSP12 Finger subdomain

region 2 (K621-G679), RP1-NSP12 Palm subdomain region 1 (T582-P620), RP2- NSP12 Palm subdomain region 2 (T680-Q815), RT- NSP12 Thumb subdomain (H816-E919), C- NSP12 C terminal region (F920-Q932).

Category	Drug name	ZINC ID	LogP value	Affinity(Kcal/mol)
HIV nonnucleoside Reverse Transcriptase Inhibitors	1. Rilpivirine 2. Delavirdine 3. Efavirenz 4. Etravirine 5. Nevirapine	1. ZINC000001554274 2. ZINC000018516586 3. ZINC000002020233 4. ZINC000000602632 5. ZINC000000004778	4.989 2.717 4.073 4.717 2.651	-7.5 -7.1 -6.7 -6.5 -6.5
HIV based Protease inhibitors	1. Saquinavir 2. Ritonavir 3. Indinavir 4. Nelfinavir 5. Amprenavir 6. Atazanavir 7. Tipranavir 8. Darunavir	1. ZINC000003914596 2. ZINC000003944422 3. ZINC000022448696 4. ZINC000003833846 5. ZINC000003809192 6. ZINC000003941496 7. ZINC000100016058 8. ZINC000003955219	3.092 5.905 2.867 4.748 2.403 4.212 7.326 2.375	-8.9 -6.1 -8.0 -7.6 -7.1 -7.5 -7.4 -7.2
HCV based Protease Inhibitors	1. Boceprevir 2. Asunaprevir 3. Simeprevir 4. Telaprevir 5. Vaniprevir 6. Paritaprevir 7. Grazoprevir	1. ZINC000014210455 2. ZINC000085540202 3. ZINC000169289453 4. ZINC000003992480 5. ZINC000096174226 6. ZINC000669678887 7. ZINC000095551509	1.711 4.694 5.254 2.446 4.511 3.637 4.142	-6.6 -6.6 -8.2 -6.3 -8.9 -9.2 -9.5
HIV based Integrase Inhibitors	1. Raltegravir 2. Dolutegravir 3. Elvitegravir	1. ZINC000013831130 2. ZINC000058581064 3. ZINC000013682481	1.486 1.353 4.281	-7.9 -7.7 -7.2
HCV NS5A and NS5B inhibitors	1. Daclatasvir 2. Dasabuvir 3. Elbasvir 4. Ledipasvir 5. Ombitasvir	1. ZINC000068204830 2. ZINC000095616937 3. ZINC000150588351 4. ZINC000150338819 5. ZINC000150601177	6.222 4.024 8.116 8.607 7.687	-6.6 -7.4 -8.0 -9.1 -7.3
Influenza virus Inhibitors	1. Amantidine 2. Oseltamivir 3. Rimantadine	1. ZINC000000968256 2. ZINC000003929508 3. ZINC000003831429	1.914 1.285 2.55	-5.2 -4.9 -5.9
Positive Control	1. Remdesivir monophosphate 2. Favipiravir monophosphate	1. Structure isolated from PDB ID: 7bv2 2. Structure isolated from PDB ID: 4kn6		-6.9 -7.8
Negative Control	1. Thymoquinone 2. Cinnamaldehyde	1. ZINC000000164367 2. ZINC000001532777	1.667 1.899	-5.4 -4.7

Table 1: List of the FDA approved antivirals used to screen for its binding affinity against NSARS-CoV2 NSP12 (RDRP) protein. The drugs showing affinity above -9Kcal/mol and used in the study is highlighted in red.

Paritaprevir, an inhibitor of HCV NS3-4A serine protease showed a high affinity of -9.2 Kcal/mol for NSP12 (RDRP) protein as a whole (Table 1). Also, as the protein associates with other proteins of SARS-CoV2 (NSP7 and NSP8) for its function, we tried to do docking with the whole complex (PDB ID: 6m71) and found it to show a high affinity of -9.5 Kcal/mol (Fig. 2). Hence, the drug might interact with RDRP both in single as well as complex state which is important as the protein in its functional state associates with other proteins/co-factors. Next, we tried docking the drug with different domains and subdomains of the NSP12 to identify important interacting residues of the individual domains. Domains that showed an affinity of -9 Kcal/mol and above were considered further to identify the interacting amino acids. The affinity of different domains have been listed in Table 2. Affinity of NIRAN (-8.6Kcal/mol) and Interface (-8.2 Kcal/mol) domains were found to be lesser than that of RDRP domain (-10.5 Kcal/mol) (Fig S1, S2, S3). Considering the docking results of different subdomains of the RDRP domain, all the three subdomains (Finger, Palm and Thumb) showed a higher affinity, further supporting our initial analysis (Fig S3-S5). Out of the three subdomains, the Thumb subdomain showed the highest affinity of -9.2 Kcal/mol among others (FigS5). Also, the RDRP domain based docking showed most of the interacting residues to be of the Thumb subdomain. The interacting residues have been listed in Table 2. Docking analysis of RDRP domain and Thumb subdomain revealed two common residues Ile864 and Tyr867, each showing multiple bonds with the drug (Table 2, Fig 3, S5). Ile864 bound with the drug via Pi Sigma, Pi Alkyl and conventional H bond and Tyr867 showed Pi Alkyl, conventional H bond and Pi Donor H bond. Multiple bond formation indicates the importance of the residue in binding with the drug and it strengthens the interaction. Moreover, H bond being the strongest of all non-covalent interactions, is one of the most important interactions in biology. These residues tend to show H bond interactions with drug which would strengthen the interaction further.

The RDRP Finger subdomain which showed an affinity of -8.8 Kcal/mol were further analysed using Region 1(L366-A581) and region 2(K621-G679). Residues1 showed a higher affinity of -10 Kcal/mol (Fig 4) as compared to the Residues 2 interaction (-6.9 Kcal/mol). Also, the docking analysis result of RDRP Finger subdomain as a whole picked up Region 1 amino acids

for the interaction. Hence, the drug might interact with the Finger subdomain region 1 along with the Thumb domain. The Palm subdomain analysis showed an affinity of -8.7 Kcal/mol and study of two of its regions (Region 1 T582-P620 and Region 2 T680-Q815) revealed a stronger interaction of -11 Kcal/mol with the region 2 than with region 1 (-7.5Kcal/mol)(Table 2 and Fig 5). Even the Palm subdomain docking showed interactions with the amino acid residues of Region 2 which further strengthens the observation. However, no common residues were identified in the docking results of Finger Subdomain and its respective regions as well as Palm subdomain and its respective regions. But the docking result of Palm subdomain and its region 2 identified Met755 and Ala 762 respectively that bound with the drug using Pi Sigma and Alkyl/Pi Alkyl interactions respectively. Both the residues lie in the Motif C of the Palm subdomain that forms the active site of the enzyme where polymerisation takes place. Also, Palm subdomain region 2 based docking showed Glu811 interaction with the drug. The residue lies in the Motif E of the Palm subdomain and is important to support the primer strand during polymerisation (1)

Hence, the drug interacts with RDRP domain as a whole, showing high affinity for the Thumb subdomain, Palm subdomain region 2 and Finger subdomain region 1. The Finger and Thumb subdomains help the palm domain to function and it is very important to further validate a drug that can target all the three subdomains for an efficient inhibition.

Domain/Complex used for docking	**Affinity (Kcal/mol)**	**Interacting amino acid residues**	**Bond length (Angstrom)**	**Types of bonds formed**
Interface domain	-8.2	Phe283 Leu280 Thr293 His295 Leu302 Asn300 Asn297	4.83 4.66 2.23 2.10,3.61,3.74 5.12 2.23, 2.38 2.52	Pi-Pi T shaped Pi Alkyl Conventional Hydrogen Bond Conventional Hydrogen Bond, Carbon Hydrogen bond, Pi Hydrogen bond Pi Alkyl Conventional Hydrogen Bond Conventional Hydrogen Bond
NIRAN domain	-8.6	Asp208 Leu240 Ala125 Val128 Tyr217	3.76,4.02 4.75 4.14,4.84 5.21 5.41	Pi Anion, Carbon Hydrogen bond Pi alkyl Pi alkyl Pi alkyl Pi alkyl

RDRP domain	-10.5	Ile864 Phe594 Tyr867 Pro868 Pro830 Leu829	2.82 5.19 4.03 4.26 5.28 3.58	Conventional Hydrogen Bond Pi-Pi T shaped Pi Donor Hydrogen bond Pi alkyl Pi alkyl Pi Sigma
RDRP Finger subdomain	-8.8	Leu372 Ala375 Met380 His381 Lys508 Trp509	5.07 5.24 5.12 2.59,3.95,4.90 2.44,2.63 4.36,4.53,5.04	Alkyl/Pi Alkyl Alkyl/Pi Alkyl Alkyl/Pi Alkyl Conventional Hydrogen bond, Pi Sigma, Pi Sulfur Conventional Hydrogen Bond Pi-Pi stacked
RDRP Finger subdomain Region 1	-10	Ala382 Ile539 Gln541 Arg392 Ala406	4.89 4.56 2.82 2.59 5.40	Alkyl/Pi Alkyl Alkyl/Pi Alkyl Conventional Hydrogen Bond Conventional Hydrogen Bond Alkyl/Pi Alkyl
RDRP Finger subdomain region 2	-6.9	----	----	----
RDRP Palm subdomain	-8.7	Val742 Tyr746 Leu749 Met755 Cys697	3.33 4.93,5.02 5.33 3.94 5.11	Conventional Hydrogen Bond Pi-Pi T shaped Pi Alkyl Pi Sigma Pi Alkyl
RDRP Palm subdomain region 1	-7.5	---	---	----
RDRP Palm subdomain region 2	-11	Met794 Phe782 Lys798 Glu811 Trp800 Ala762	5.56 5.37 3.36,4.68 3.54,3.62 2.78,3.74 4.15	Pi Sulfur Alkyl/Pi Alkyl Conventional hydrogen bond, Alkyl/Pi Alkyl Pi Sulfur Conventional hydrogen bond Alkyl/Pi Alkyl
RDRP Thumb subdomain	-9.2	Pro918 Val860 Ile864 Ala863	3.45 5.19,5.40 3.77,4.90 4.59	Carbon Hydrogen bond/Pi Donor Hydrogen Bond Pi Alkyl Pi Sigma, Pi Alkyl Pi Alkyl

		Tyr867	4.19,4.45	Conventional Hydrogen bond, Pi Alkyl
RDRP Complex (NSP12-NSP8-NSP7 complex)	-9.5	Lys500 Val557 Tyr546 Asp845 Gln408	4.66 4.15,4.81 3.02,3.06 3.34,3.79 3.07	Pi Alkyl Pi Alkyl Conventional Hydrogen bond Conventional Hydrogen bond, Pi Anion Conventional Hydrogen bond

Table 2: List of all the interactions between Paritaprevir and NSP12 protein and its complex (NSP12-NSP7-NSP8). The common residues between Thumb subdomain and RDRP domain is highlighted in red. The residues that lie in motif C of the Palm subdomain is highlighted in blue and the residue that lie in motif E is highlighted in orange.

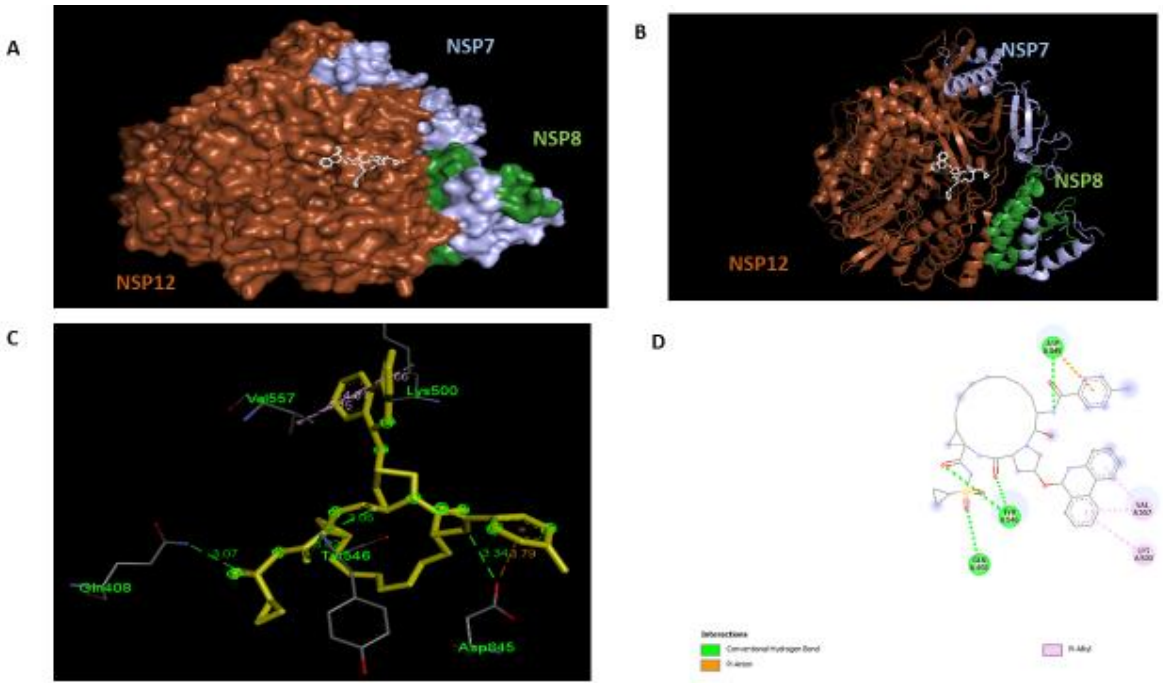

Figure2: Representation of the docking result of Paritaprevir with NSP12 complex: A and B. Surface and cartoon view of the ligand in its best docked position. C. The interaction between the ligand (Highlighted in yellow) with the interacting amino acids of the domain. D. The 2D image of the interactions formed between the ligand and the protein.

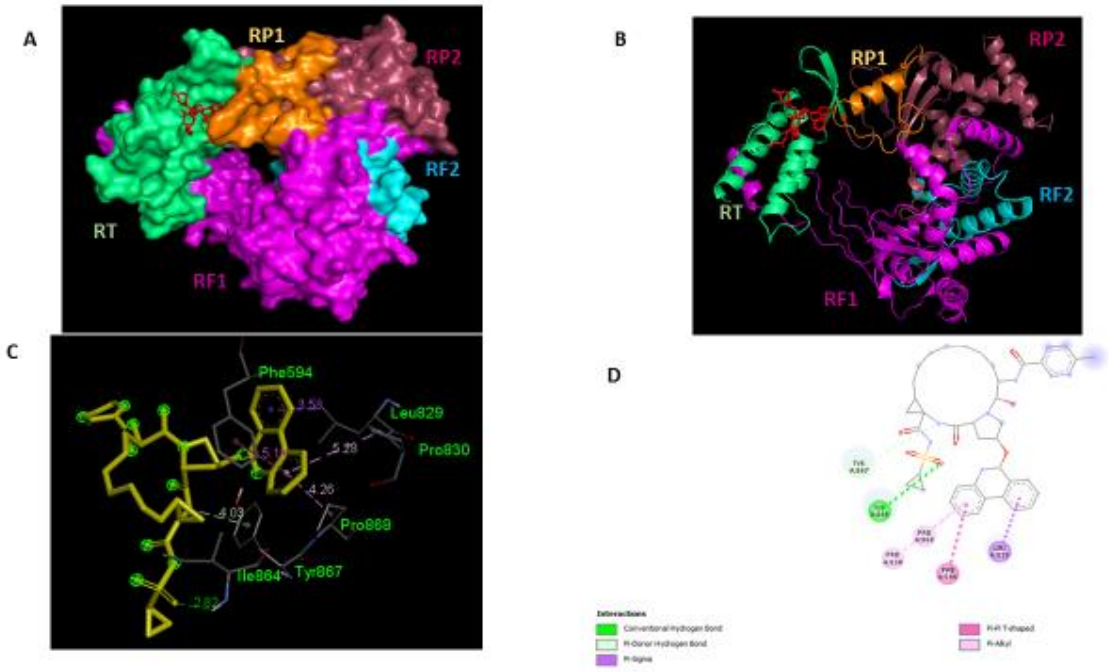

Figure3: Representation of the docking result of Paritaprevir with NSP12 RDRP domain: A and B. Surface and cartoon view of the ligand in its best docked position. C. The interaction between the ligand (Highlighted in yellow) with the interacting amino acids of the domain. D. The 2D image of the interactions formed between the ligand and the protein.

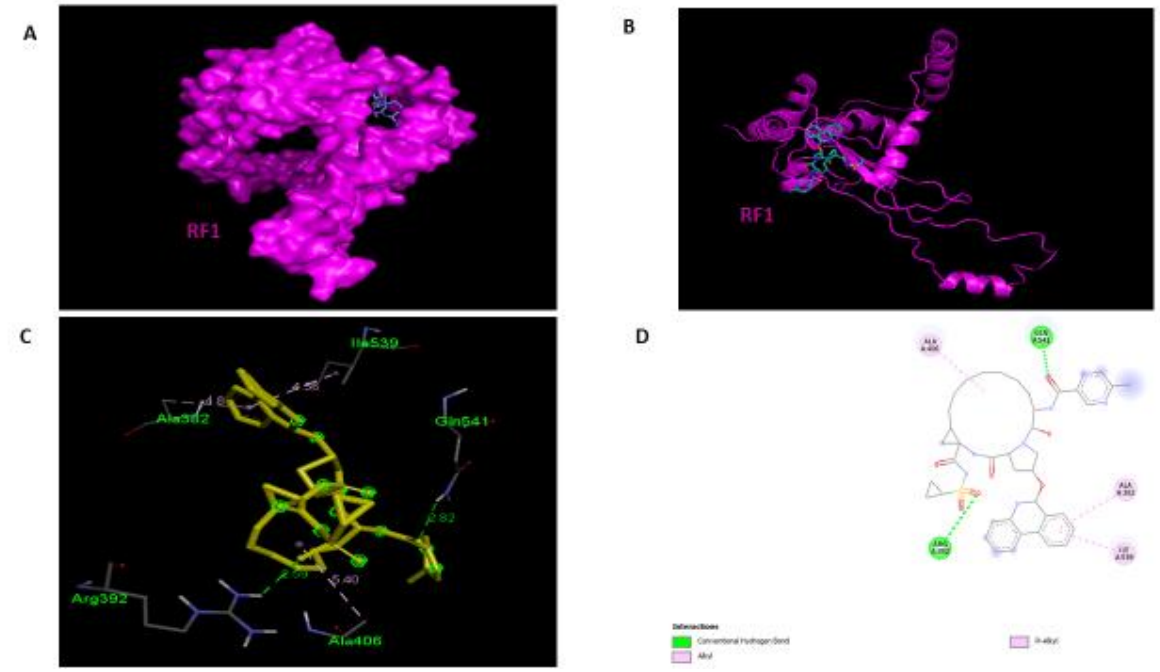

Figure4: Representation of the docking result of Paritaprevir with Finger subdomain region 1(RF1): A and B. Surface and cartoon view of the ligand in its best docked position. C. The interaction between the ligand (Highlighted in yellow) with the interacting amino acids of the domain. D. The 2D image of the interactions formed between the ligand and the protein.

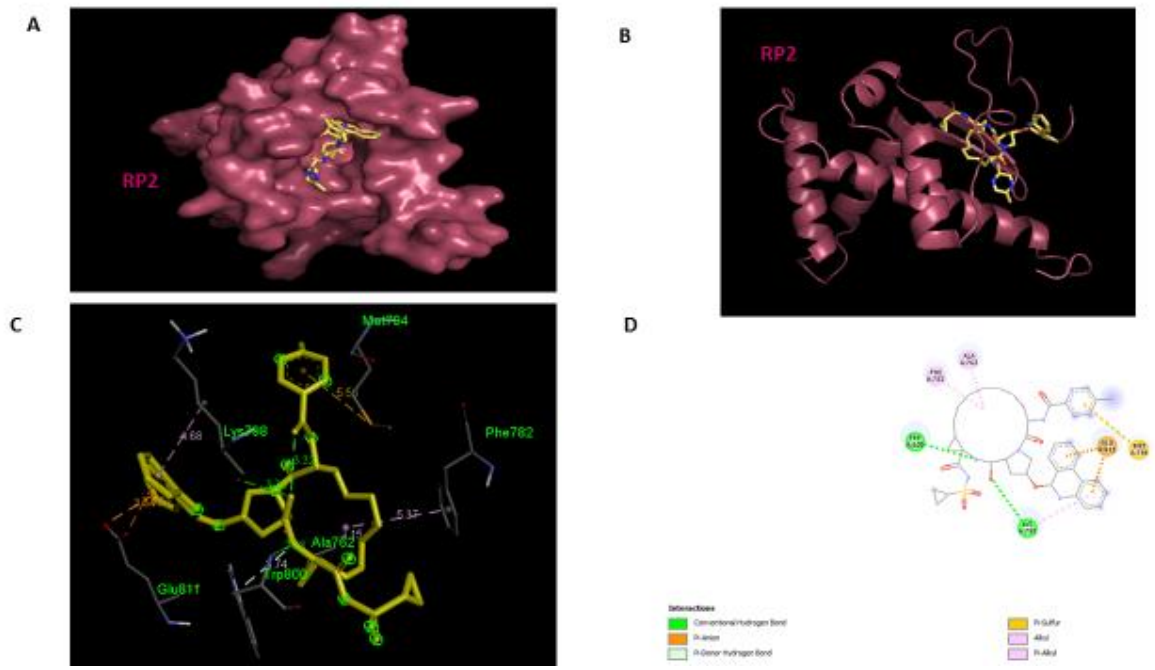

Figure5: Representation of the docking result of Paritaprevir with Palm subdomain region 2 (RP2): A and B. Surface and cartoon view of the ligand in its best docked position. C. The interaction between the ligand (Highlighted in yellow) with the interacting amino acids of the domain. D. The 2D image of the interactions formed between the ligand and the protein. The domains with affinity below the cut off value (-8 Kcal/mol) are not mentioned in terms of its interacting residues.

Grazoprevir is another HCV NS3/4A protease inhibitor that has showed a reasonable affinity of -9.5 Kcal/mol when interacting with the RDRP protein (NSP12) (Table 2). Also, the docking result of Nsp12-Nsp7-Nsp8 complex showed a high affinity of -8.3 Kcal/mol hence the drug is predicted to interact with RDRP in its functional state inside the cell (Fig 6). RDRP domain-based docking showed an affinity of -9.3 Kcal/mol which is higher than the affinity of the other domains, -6.8 Kcal/mol for NIRAN domain and -7.3 for the interface domain (Table 3 and Fig 7).

The docking result of the different subdomains of RDRP showed varying affinities with thumb subdomain showing the highest affinity of -8.2 Kcal/mol followed by Palm subdomain and Finger subdomain with an affinity of -7.7 Kcal/mol and -7.3 Kcal/mol respectively (Fig S6-8). The analysis of Thumb subdomain revealed Pro868 to be a common residue identified even in the docking result of RDRP domain showing Alkyl/ Pi Alkyl interaction (Table 3). Also, half of the interacting residues in the RDRP domain based docking was found to lie in the Thumb subdomain. Analysis of the Finger subdomain region 1 and 2 showed an affinity of -8.9Kcal/mol and-6.6 Kcal/mol respectively (Table 3 and Fig8). Hence, the drug is likely to interact with the Finger subdomain region 1 than region 2 as even the docking analysis of

RDRP Finger subdomain revealed Region 1 based amino acid residues to interact with the drug (Fig S7). Docking analysis of Palm subdomain region 1 and region 2 showed an affinity of -6.9 Kcal/mol and -10.4 Kcal/mol respectively indicating a much higher affinity for region 2 than 1 (Table 3 and Fig 9). Region 2 docking result revealed two Motif C residues Ala762 and Val764 to interact with the drug via Alkyl/Pi Alkyl and Conventional H bonds (Table 3). Motif C is one of the motifs that form the active site of the enzyme. Amino acid residues of Region 2 of Palm subdomain were found to interact with the ligand when the whole Palm subdomain was used for docking, hence supporting the previous results (Fig S7). A high affinity of Thumb domain as compared to the other two domains (Finger and Palm) may likely arise because of its small size as compared to the other two domains. But when the individual regions of the Finger and Palm subdomain were studied, it showed a higher binding affinity than the Thumb domain. Hence, it is possible that the drug interacts with the different domains of the protein with different affinities. Also, domain specific docking helps to identify the interacting amino acids more specifically than considering the whole protein or its domain. However, the affinity of the drug towards the protein is best determined when the whole protein is considered for docking as that is the natural state of the protein which us accessible to the drug.

Overall, this drug shows a strong affinity for the RDRP complex (NSP12-NSP7-NSP8), RDRP domain as a whole and also for its different domains. Most importantly, it has a high affinity for Palm subdomain region 2 and targets the active site of the enzyme along with its thumb and finger subdomain. Paritaprevir and Grazoprevir share similar structure and hence the docking results were also similar, showing a high affinity for Thumb subdomain, Finger subdomain region 1 and Palm subdomain region 2. It is important to identify drugs that have multiple targets to bind in the protein as it increases its efficiency for binding.

Domains/Complex used for docking	**Affinity (Kcal/mol)**	**Interaction amino acids**	**Bond length (in Angstrom)**	**Types of bonds formed**
RDRP Complex (NSP12-NSP8-NSP7 complex)	-8.3	Asp135 Tyr32 Lys47 Ser709 Asn781 Lys780 Try129	3.18 5.24 4.67,4.67,4.74 2.20,3.22 2.32 1.92,2,23 2.45	Conventional Hydrogen Bond Pi-Pi T shaped Pi Carbon bond, Alkyl bond Conventional Hydrogen Bond, Carbon hydrogen bond Conventional Hydrogen Bond Conventional Hydrogen Bond Conventional Hydrogen Bond
NIRAN domain	-6.8	----	-----	----

Interface domain	-7.3	------	-----	------
RDRP domain	-9.3	Ile864 Ser592 Phe594 Tyr595 Leu829 Pro868	4.12 2.17 5.22 2.55 4.64,4.98 4.93	Alkyl/Pi Alkyl Conventional Hydrogen Bond Pi-Sulfur Conventional Hydrogen Bond Alkyl/Pi Alkyl Alkyl/Pi Alkyl
RDRP Finger Subdomain	-7.3	Leu371 Ala375 Leu366 Trp509	5.0,5.03,5.06 4.52,4.91 5.44 2.58	Alkyl/Pi Alkyl Alkyl/Pi Alkyl Alkyl/Pi Alkyl Conventional Hydrogen Bond
RDRP Finger Subdomain Region 1	-8.9	Ala382 Leu401 Pro378 Val398 Ile539 Phe504	4.84 5.34 5.27 4.78,5.37 4.81,5.45 4.08,5.56	Pi Alkyl Pi Alkyl Pi Alkyl Pi Alkyl Pi Alkyl Pi Donor Hydrogen bond, Pi Sulfur bond
RDRP Finger Subdomain Region 2	-6.6	-----	-----	-------
RDRP Palm subdomain	-7.7	Leu708 Ser709 Gly774 Asn781 Lys780 Tyr788	3.91 3.20 3.70 2.31 4.24 5.34	Pi Sigma Carbon Hydrogen bond Carbon Hydrogen bond Conventional Hydrogen bond Alkyl/Pi Alkyl Alkyl/Pi Alkyl
RDRP Palm subdomain region 1	-6.9	-----	----	----
RDRP Palm subdomain region 2	-10.4	Asn790 **Ala762** **Val764** Phe782	2.47 **3.61** **2.30,4.37** 4.30	Conventional Hydrogen Bond **Alkyl/Pi Alkyl** **Conventional Hydrogen Bond,** **Alkyl/Pi Alkyl** Alkyl/Pi Alkyl
RDRP Thumb subdomain	-8.2	Tyr915 Leu885 Tyr867 Pro868 Pro830	3.30 4.37 4.71 5.37 5.15	Carbon Hydrogen bond Alkyl/Pi Alkyl Alkyl/Pi Alkyl Alkyl/Pi Alkyl Alkyl/Pi Alkyl

Table 3: List of all the interactions between Grazoprevir and NSP12 protein and its complex (NSP12-NSP7-NSP8). The common residues between Thumb subdomain and RDRP domain is highlighted in red. The residues that lie in motif C of the Palm subdomain is highlighted in blue. The domains with affinity below the cutoff value (-8 Kcal/mol) are not mentioned in terms of its interacting residues.

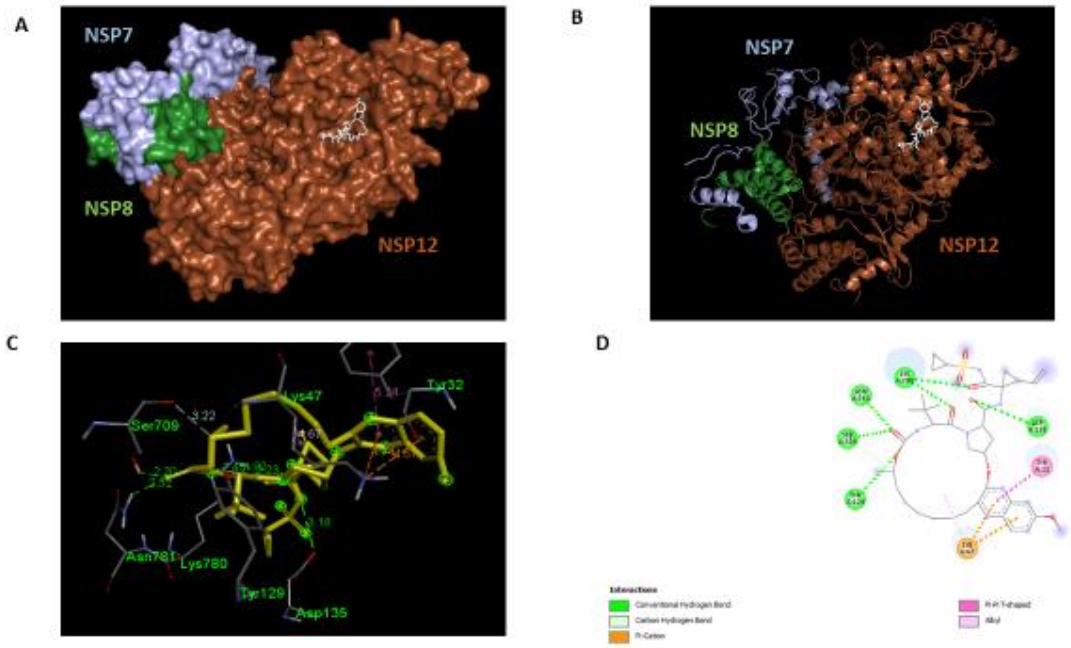

Figure 6: Representation of the docking result of Grazoprevir with NSP12 complex: A and B. Surface and cartoon view of the ligand in its best docked position. C. The interaction between the ligand (Highlighted in yellow) with the interacting amino acids of the domain. D. The 2D image of the interactions formed between the ligand and the protein.

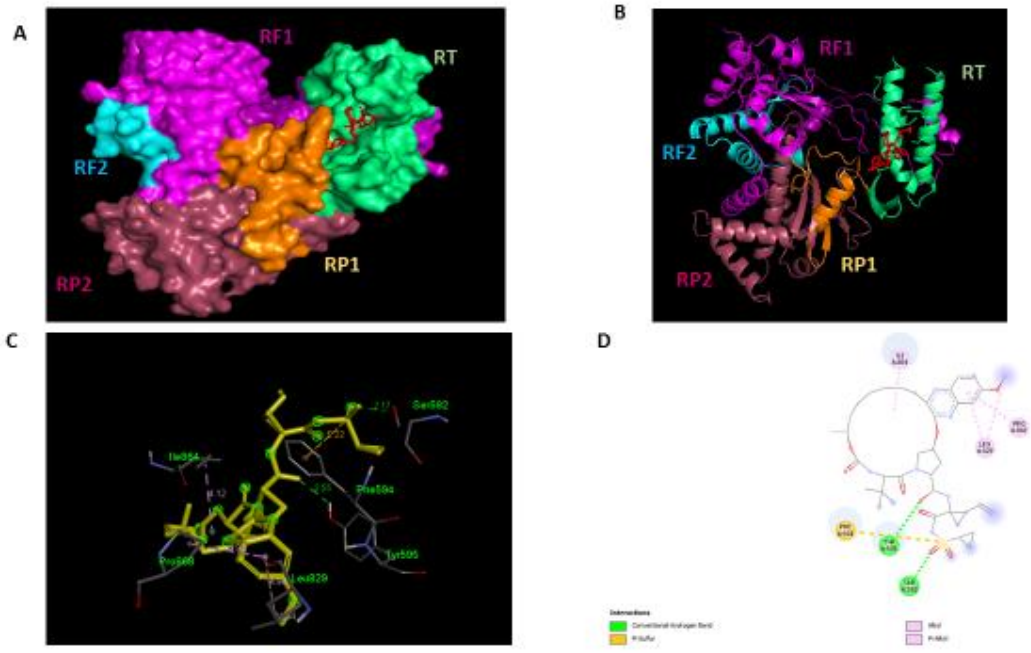

Figure 7: Representation of the docking result of Grazoprevir with RDRP domain: A and B. Surface and cartoon view of the ligand in its best docked position. C. The interaction between the ligand (Highlighted in yellow) with the interacting amino acids of the domain. D. The 2D image of the interactions formed between the ligand and the protein.

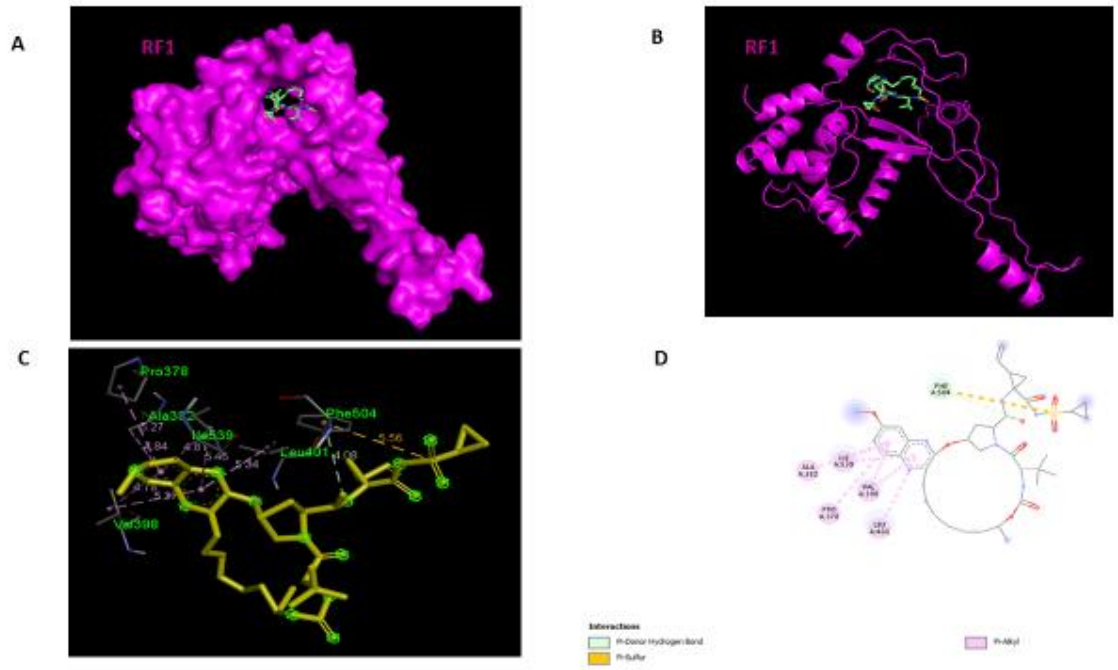

Figure 8: Representation of the docking result of Grazoprevir with RDRP Finger subdomain region 1 (RF1): A and B. Surface and cartoon view of the ligand in its best docked position. C. The interaction between the ligand (Highlighted in yellow) with the interacting amino acids of the domain. D. The 2D image of the interactions formed between the ligand and the protein.

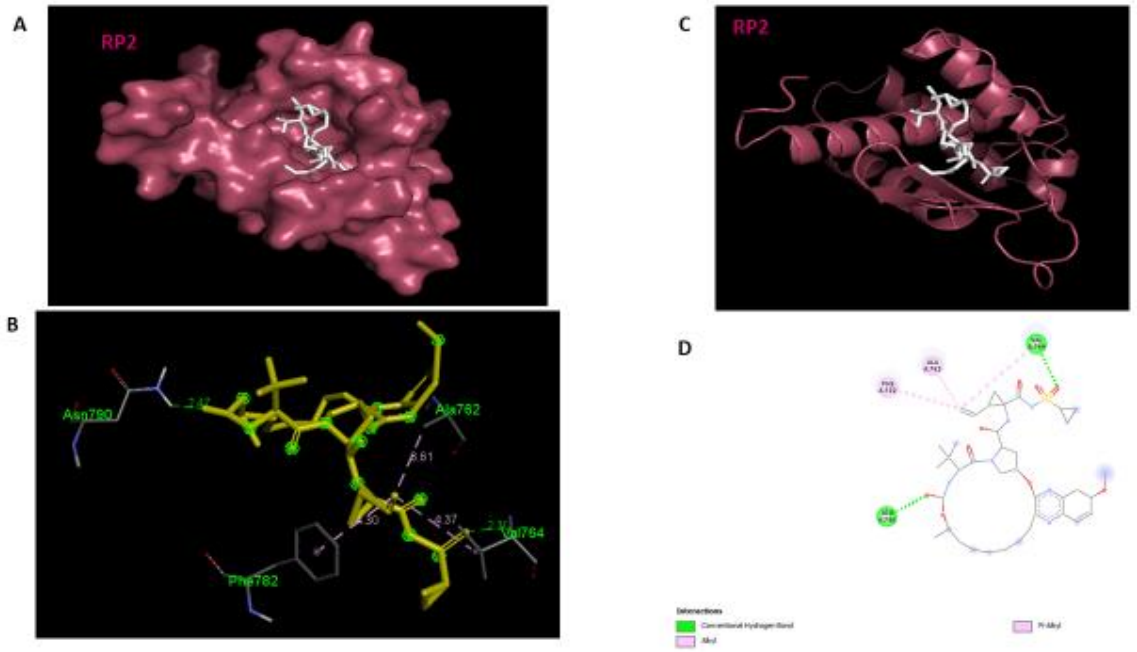

Figure 9: Representation of the docking result of Grazoprevir with RDRP Palm subdomain region 2 (RP2): A and B. Surface and cartoon view of the ligand in its best docked position. C. The interaction between the ligand (Highlighted in yellow) with the interacting amino acids of the domain. D. The 2D image of the interactions formed between the ligand and the protein.

Another drug identified in this study is Ledispavir which is an HCV NS5A inhibitor. The drug shows a high affinity for NSP12 protein (-9.1 Kcal/mol) (Table 1). It also shows a high affinity towards NSP12-NSP7-NSP8 complex, hence the binding sites of the drug is accessible in the complex state of the protein as well (Table4 and Fig10). The analysis of domain specific docking results stated a high affinity of the drug for NIRAN and RDRP domain (-8.4 Kcal/mol and -9.1 Kcal/mol respectively) (Fig S9 and 11). Considering the cut off value used for the study (-9Kcal/mol), the RDRP domain was used for further study and subdomain specific docking analysis revealed a higher affinity for Finger and Palm subdomain than with Thumb subdomain (-8.6,-8.6,-8.2 Kcal/mol respectively)(Fig S10-12). Further analysis of the Finger subdomain regions 1 and 2 indicate a higher affinity of Region 1 than region 2 for the ligand (Table 4). The docking results of RDRP domain and Finger subdomain (Fig 11,S10) identified most of the interacting residues to be of Finger subdomain Region 1 (Fig 12) and Leu401 was a common residue identified in the docking results of Finger subdomain Region 1 and Finger subdomain as a whole (Table 4). Hence, the Finger subdomain region 1 could potentially be the specific binding site for the drug as it is binding pocket used in the docking of the ligand with RDRP domain and Finger subdomain as a whole. Study of the Palm subdomain identified most of the bonding amino acid residues to belong to the region 2 of the subdomain. Region specific docking revealed Region 2 to show a higher affinity of -9.5 Kcal/mol than region 1 and Asp760 is one of the residues detected in case of Palm subdomain region 2 docking (Table 4 and Fig 13). Asp760 is the catalytic residue that lies in Motif C of the Palm subdomain and is an active site of the Polymerase.

Hence, the drug shows high affinity towards RDRP domain of the enzyme and has a mixed affinity towards different subdomains and regions. Finger subdomain region 1 and Palm subdomain region 2 is likely an important region for the ligand binding.

To predict the binding pockets of NSP12 COACH tool was used and it was found that the 1st binding pocket encompass the region surrounded by RDRP Finger subdomain and Palm subdomain region 2 with a C score of .09, the second detected binding pocket was formed of Palm subdomain region 2 and Thumb subdomain with a C score of .07. The docking analysis of our study identifies the binding pockets that are predicted by the tool, hence further confirming the reliability of the study. Further studies involving molecular dynamics and *in vitro/ ex vivo* studies will further validate the efficacy of the predicted drugs.

Protein Complex/Domains used for Docking	Affinity (Kcal/mol)	Interacting amino acids	Bond length (in Angstrom)	Types of Bonds
NSP12-NSP7-NSP8 complex	-8.8	Lys47	3.23,3.35,4.01	Conventional Hydrogen bond, Pi cation
		Val31	4.88,5.40	Alkyl
		Asp140	3.57,3.59	Carbon Hydrogen bond
		Ser709	3.76	Carbon Hydrogen bond
		Lys780	2.98	Conventional Hydrogen bond
		Gln773	3.24	Carbon Hydrogen bond
		Lys714	5.01	Alkyl
NIRAN domain	-8.4	Ala130	4.95	Alkyl/Pi alkyl
		Tyr129	2.11,5.74	Conventional Hydrogen bond, Pi Pi T shaped
		Asp126	3.26,3.88	Pi anion, Halogen (Fluorine)
		Ala125	4.82	Alkyl/Pi alkyl
		Val71	5.46	Alkyl/Pi alkyl
Interface domain	-7.6	----	----	-----
RDRP domain	-9.1	Phe368	3.91	Pi Sigma
		Leu371	3.21,4.35	Halogen (Fluorine),Alkyl/Pi Alkyl
		Ala375	4.38,4.63	Alkyl/Pi Alkyl
		Ser384	2.77	Conventional Hydrogen bond
		His381	3.06,3.38	Conventional Hydrogen bond, Carbon Hydrogen bond
		Trp509	4.52,5.37	Pi Pi stacked
		Leu514	4.86	Alkyl/Pi Alkyl
RDRP Finger subdomain	-8.6	Phe368	4.53	Alkyl/Pi Alkyl
		Leu371	3.11,3.53,4.41	Halogen (Fluorine), Alkyl/Pi Alkyl
		Ala375	4.47,4.53	Alkyl/Pi Alkyl
		His381	3.06	Conventional Hydrogen bond
		Ser384	2.71	Conventional Hydrogen bond
		Leu401	3.77	Carbon Hydrogen bond
		Trp509	4.58,5.49	Pi Pi stacked
		Leu514	4.88	Alkyl/Pi Alkyl
RDRP Finger subdomain region1	-9.9	Pro378	4.90,5.32	Alkyl/Pi alkyl
		Ala382	5.08	Alkyl/Pi alkyl
		Leu401	4.79	Alkyl/Pi alkyl
		Asp390	3.58	Carbon Hydrogen bond
		Val405	3.31	Carbon Hydrogen bond
		Ala406	2.34	Conventional Hydrogen bond
		Ile539	4.17	Alkyl/Pi alkyl
		Thr540	2.96	Halogen (Fluorine)
		Ala449	4.69	Alkyl/Pi alkyl

RDRP Finger subdomain region2	-6.7	----	-----	-----
RDRP Palm subdomain	-8.6	Val764 Val700 Ile696 Val587 Thr582 Val742 Leu731 Cys730	4.34 4.51 4.78 4.97 2.91 4.70,4.77 5.19 3.35	Alkyl/Pi alkyl Alkyl/Pi alkyl Alkyl/Pi alkyl Alkyl/Pi alkyl Conventional hydrogen bond Alkyl/Pi alkyl Alkyl/Pi alkyl Carbon Hydrogen bond
RDRP Palm subdomain Region1	-7.4	----	-----	-----
RDRP Palm subdomain Region2	-9.5	Ala690 Phe694 **Asp760** Asn695 Asn790	5.25 4.34,4.60 **3.10** 2.90 2.96,3.32	Pi alkyl Pi alkyl **Conventional Hydrogen bond** Conventional Hydrogen bond Conventional Hydrogen bond
RDRP Thumb subdomain	-8.2	Ala863 Tyr867 Ser913 Leu829 Tyr828 Val827	5.33 3.93,5.24 3.55 4.82,5.20 3.34 4.74,4.91	Alkyl/Pi alkyl Alkyl/Pi alkyl Carbon Hydrogen bond Carbon Hydrogen bond Carbon Hydrogen bond Alkyl/Pi alkyl

Table 4: List of all the interactions between Ledispavir and NSP12 protein and its complex (NSP12-NSP7-NSP8). The common residues between Finger subdomain and Finger subdomain region 1 is highlighted in red. The residues that lie in motif C of the Palm subdomain is highlighted in blue. The domains with affinity below the cut off value (-8 Kcal/mol) are not mentioned in terms of its interacting residues.

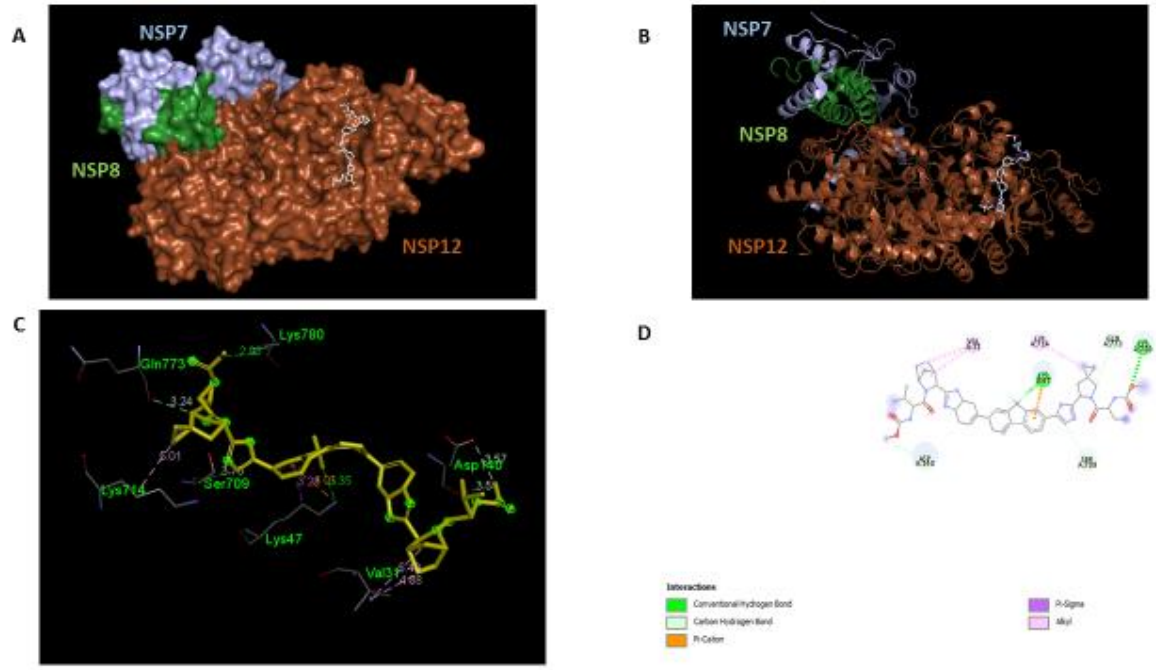

Figure 10: Representation of the docking result of Ledispavir with NSP12 complex: A and B. Surface and cartoon view of the ligand in its best docked position. C. The interaction between the ligand (Highlighted in yellow) with the interacting amino acids of the domain. D. The 2D image of the interactions formed between the ligand and the protein.

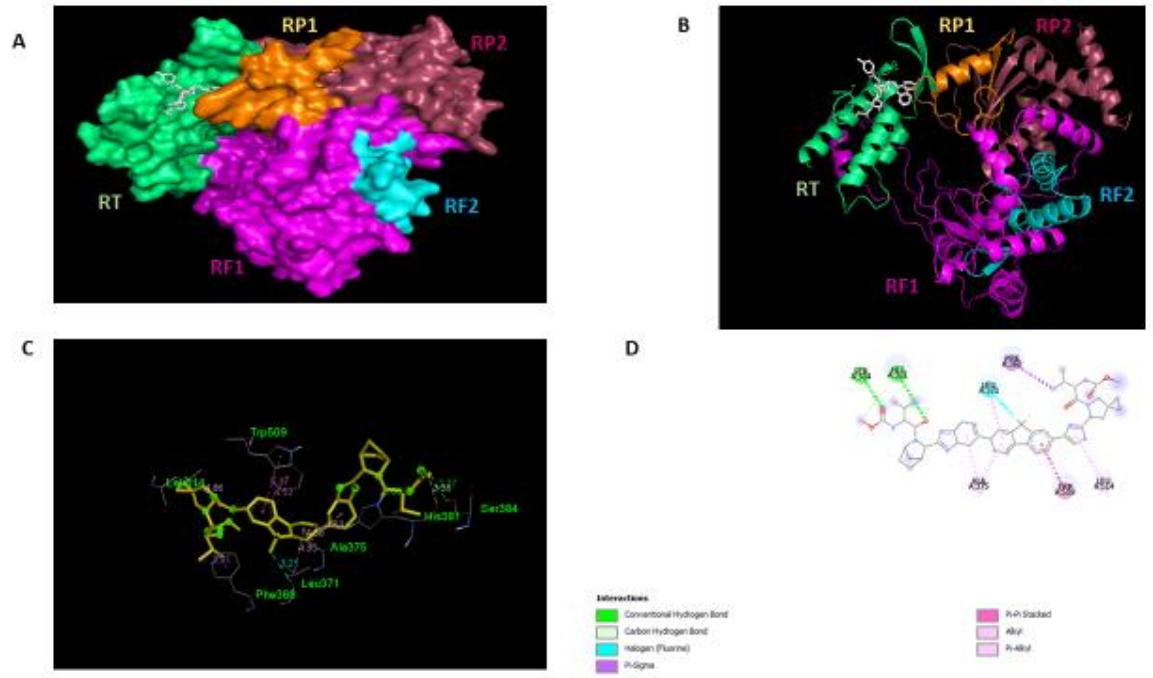

Figure 11: Representation of the docking result of Ledispavir with NSP12: A and B. Surface and cartoon view of the ligand in its best docked position. C. The interaction between the ligand (Highlighted in yellow) with the interacting amino acids of the domain. D. The 2D image of the interactions formed between the ligand and the protein.

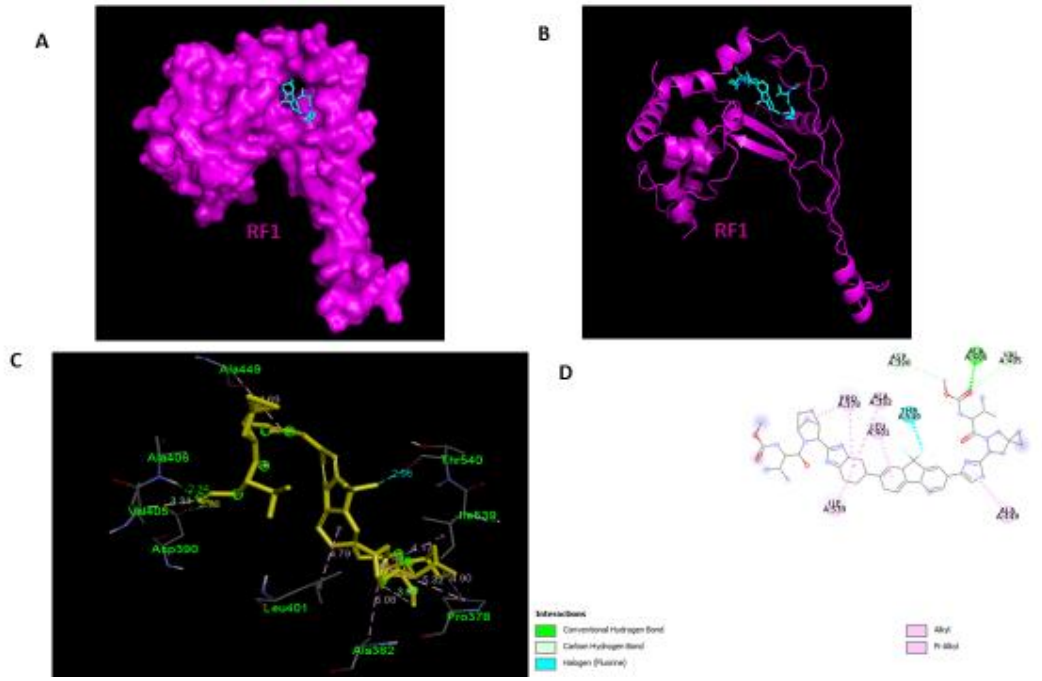

Figure 12: Representation of the docking result of Ledispavir with RDRP Finger subdomain region 1 (RF1): A and B. Surface and cartoon view of the ligand in its best docked position. C. The interaction between the ligand (Highlighted in yellow) with the interacting amino acids of the domain. D. The 2D image of the interactions formed between the ligand and the protein.

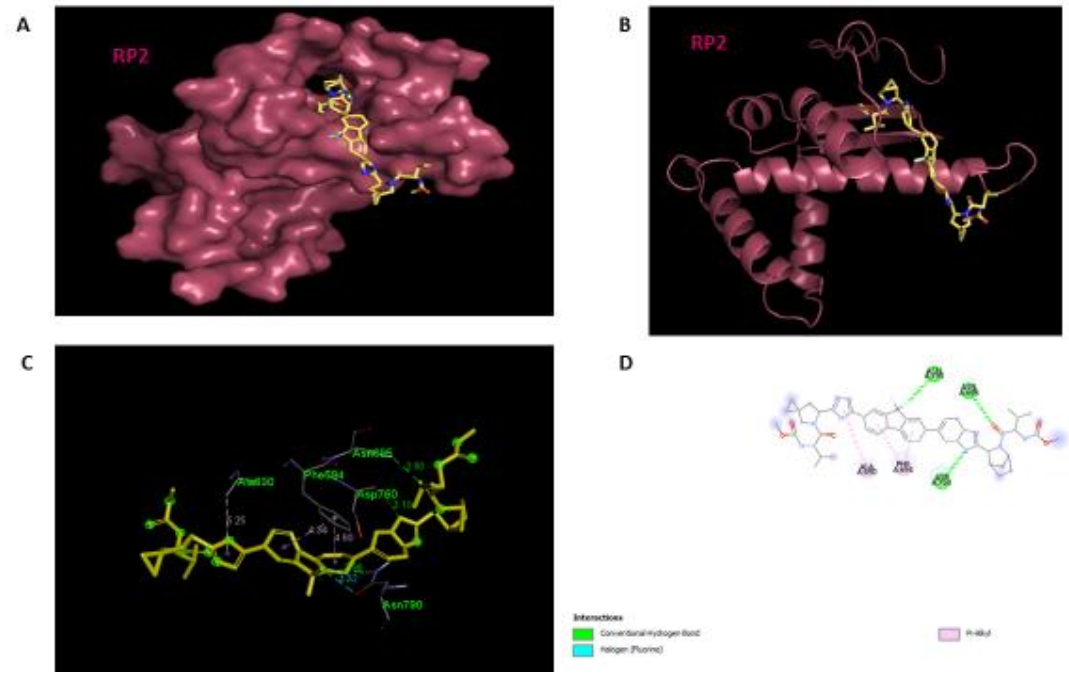

Figure 13: Representation of the docking result of Ledispavir with RDRP Palm subdomain region 2 (RP2): A and B. Surface and cartoon view of the ligand in its best docked position. C. The interaction between the ligand (Highlighted in yellow) with the interacting amino acids of the domain. D. The 2D image of the interactions formed between the ligand and the protein.

References:

1. Gao, Y. et al. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. (2020) Science 368: 779-782.
2. Yin, W. et al, Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. (2020) Science 368: 1499-1504.
3. Naesens, L. et al. Role of human hypoxanthine Guanine phosphoribosyltransferase in activation of the antiviral agent T-705 (favipiravir). (2013) Mol Pharmacol 84: 615-629.
4. Trott,O. et al. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. (2010) Journal of Computational Chemistry 31 455-461. DOI 10.1002/jcc.21334
5. Dassault Systèmes BIOVIA, Discovery Studio Modelling Environment, Release 2017, San Diego: Dassault Systèmes, 2016.
6. DeLano W L (2002). PyMOL. DeLano Scientific, San Carlos.
7. Sterling, T et al. ZINC 15 – Ligand Discovery for Everyone. (2015) *J. Chem. Inf. Model.* 55, 11, 2324–2337.
8. Yang,J. et al.Protein-ligand binding site recognition using complementary binding specific substructure comparison and sequence profile alignment.(2013). Bioinformatics. 29:2588-2595
9. Clercq,E. et al.Approved antiviral drugs over the past 50 years.(2016). Clin Microbiol Rev 29:695–747.