

# Targeting allosteric pockets of SARS-CoV-2 main protease M<sup>pro</sup>

Zahoor Ahmad Bhat<sup>#1</sup>, Dheeraj Chitara<sup>#2</sup>, Jawed Iqbal<sup>1</sup>, Sanjeev. B.S.<sup>2</sup> and Arumugam Madhumalar<sup>1\*</sup>

<sup>1</sup> Multidisciplinary Centre for Advanced Research and Studies, Jamia Millia Islamia, New Delhi.

<sup>2</sup> Department of Applied Sciences, Indian Institute of Information Technology, Allahabad.

# authors contributed equally to this work

## \*Corresponding author

Dr. Arumugam Madhumalar,  
Email: [amadhumar@jmi.ac.in](mailto:amadhumar@jmi.ac.in)  
Contact: +91 8800733110.

1. Multidisciplinary Centre for Advanced Research and Studies,

Jamia Millia Islamia, Jamia Nagar, New Delhi-110025, India.

2. Department of Applied Sciences, Indian Institute of Information Technology, Allahabad, Devghat, Jhalwa, Prayagraj, -211012, India

**Keywords:** SARS-CoV-2, Molecular Dynamics Simulations, M<sup>pro</sup>protease, Elbasvir, Glecaprevir, Ritonavir, Allostery, Combination therapy, COVID-19

## ABSTRACT

Repurposing of antivirals is an attractive therapeutic option for the treatment of COVID-19. M<sup>pro</sup> (also called 3CL<sup>pro</sup>) is a key protease of SARS-CoV-2 involved in viral replication, and is a promising drug target for testing the existing antivirals. A major challenge to test the efficacy of antivirals is the conformational plasticity of M<sup>pro</sup> and its future mutation prone flexibility. To address this, we hereby propose combination therapy by drugging two specific additional pockets of M<sup>pro</sup> probed in our studies. Long scale Molecular Dynamics (MD) simulations provide evidence of these additional sites being allosteric. Suitable choice of drugs in catalytic and allosteric pockets appear to be essential for combination therapy. Current study, based on docking and extensive set of MD simulations, finds the combination of Elbasvir, Glecaprevir, Ritonavir to be a viable candidate for further experimental drug testing/pharmacophore design for M<sup>pro</sup>.

## INTRODUCTION

Coronavirus (CoV) is an enveloped, non-segmented, and positive-sense RNA virus with 30 kb genome. Among six open reading frames (ORFs), ORF1a/1b directly translate into polyprotein 1a/1ab (pp1a/pp1ab) which is processed by viral main protease (M<sup>pro</sup>), also known as chymotrypsin-like protease (3CL<sup>pro</sup>), and papain-like protease into 16 non-structural proteins (nsps). The other ORFs on the one-third of the genome near 3'-terminus encode four main structural proteins: spike (S), membrane (M), envelope (E), and nucleocapsid (N) proteins (1). A newly identified Severe Acute Respiratory Syndrome coronavirus 2 (SARS-CoV-2), that causes the disease COVID-19, purportedly originated sometime during December of 2019 in a wet market of Wuhan, China. It spread rapidly across the globe within a short time and was declared a pandemic in March, 2020 (2). SARS-CoV-2 is so named as its RNA genome is about 82% identical to the previous SARS coronavirus (SARS-CoV) from the same genus, *Betacoronavirus* (3). Of the seven-known human CoVs, MERS, SARS-CoV, and SARS-CoV-2 cause severe lower respiratory disease, while HKU1, NL63, OC43 and 229E cause only mild upper respiratory disease (4). It was observed that SARS-CoV-2 genome sequence has 96.2% identity to the bat RaTG13-CoV. On the basis of virus genome sequencing and evolutionary analysis, it is suspected that bat could be the natural host of SARS-CoV-2, and it might have transmitted from bats through unknown intermediate hosts to infect humans. Recent studies have shown Angiotensin-Converting Enzyme 2 (ACE2) as a host receptor for SARS-CoV-2 (5) to infect humans. This virus mainly spreads through the respiratory

secretions, especially by droplets, and direct human to human contacts (6). Currently there are no viable vaccine or drug(s) available for COVID-19 patients. Hence there is an urgent need for repurposing drugs to control this disease. The main protease ( $M^{pro}$ ) of SARS-CoV-2 plays a major role in polyprotein cleavage, which is a crucial step to drive viral transcription and replication. Absence of a homologous human protease makes  $M^{pro}$  an attractive target for the development of COVID-19 therapeutics. Previous studies, both *in vitro* and clinical, reported success when SARS and Middle East Respiratory Syndrome (MERS) patients were treated with Human Immunodeficiency Virus (HIV) and Hepatitis C Virus (HCV) antivirals. Hence, given the structural similarity between SARS-CoV-2 main protease with HCV and HIV proteases, it was to be expected that drugs for the treatment of HCV and HIV to have beneficial therapeutic outcome for SARS-CoV-2(7-11). Two crystal structures of  $M^{pro}$  (PDB IDs 6Y2G and 6LU7), co-crystallized with potential inhibitors  $\alpha$ -ketoamide 13b (1.75 Å) and peptide like inhibitor N3 (2.16 Å), were recently solved (12,13). This opened up immense possibilities of structure-based drug discovery to check existing antivirals as possible therapeutics. The crystal structures revealed  $M^{pro}$  to be a crystallographic dimer, formed by two monomers with each monomer consisting of N-terminal catalytic domain (residues from 8 to 101, forming domain I, residues 102-184, forming domain II) made of anti-parallel  $\beta$ -barrel and C-terminal all  $\alpha$ -helical domain (residues from 201-303) [Figure 1, 12].  $M^{pro}$  proteases of SARS-CoV and SARS-CoV-2 differ only by 12 residues/sites (highlighted in green in Figure 1 and A), spread all over the structure. Structurally,  $M^{pro}$  belongs to the family of cysteine proteases with catalytic dyad at His41 and Cys145. The third catalytic member is a water molecule that makes key hydrogen bond interactions, including the ones with catalytic Histidine. Interestingly, the crystal structure solved by Prof. Rolf Hilgenfeld *et al.* of  $M^{pro}$  inhibitor  $\alpha$ -ketoamide 13b is optimised for its pharmacophore. It has been successfully tested in mice through inhalative route with no adverse effects (12). This further enhanced the confidence for designing and repurposing known viral drugs (13). However, to the best of our knowledge, in all the reported studies so far on virtual screening of  $M^{pro}$  (14-17), the conformational flexibility of  $M^{pro}$  has not been taken into account, which is essential part of drug design (18). More so, for SARS-CoV-2  $M^{pro}$ , the conformational plasticity associated with its flexibility is evident from recent structural studies based on room temperature crystal structure, network model (19,20) and Molecular Dynamics (MD) simulation studies (21) and this poses a major challenge for drug design.

To address the above challenges, including conformational flexibility, we carried out extensive MD simulations on M<sup>Pro</sup> bound to various drugs and their combinations. As the subsequent sections elucidate, combination therapy (*via* drugging catalytic and two newly identified allosteric pockets) may enable the efficient inhibition of M<sup>Pro</sup>. We believe that this may be the key to designing mutation-resilient drugs targeting M<sup>Pro</sup>, hence COVID-19. And, these newly identified pockets, are also conserved in SARS-CoV, suggesting the possibility of broad spectrum of combination of antivirals targeting M<sup>Pro</sup> of SARS-CoVs.

## Methods

### Molecular Docking

Virtual screening by molecular docking of 130 known antivirals was carried out using AutoDock Vina 1.1.2 (22,). The structural information of the target protein, SARS-CoV-2 main protease M<sup>Pro</sup> complete monomer (hereafter referred as M<sup>Pro</sup> unless explicitly stated otherwise) consisting of all the three domains with the catalytic pocket located in between domain I and II, was considered for all the following studies. The monomer was obtained from the protein databank; PDB ID 6Y2G and it was converted to PDBQT format using AutoDock Tools (ADT 1.5.6) (23).

The 130 approved antiviral drugs were selected for molecular docking/ virtual screening [Table S1] based on the search in DrugBank database (24) (<https://www.drugbank.ca/>), retrieved from Chempider database (25) (<http://www.chemspider.com/>) in SDF (3D) format. These compounds were further converted to the PDBQT format using open source software Open Babel (26) wherein the polar hydrogens and atom types suitable for AutoDock Vina, were added. Before running the virtual screening, these molecules were analysed using PyMOL (27) to check their structural integrity.

Two sets of molecular docking, viz., targeted and blind dockings, were carried out. In the former, the catalytic pocket of the protease monomer was targeted while in the latter the whole protease monomer was searched for other possible potential druggable pockets. For targeted docking, a grid box with the dimensions 30Å x 30Å x 30Å around the catalytic site, encompassing the residues interacting with co-crystallized ligand ( $\alpha$ -ketoamide inhibitor), was generated using ADT. The centre of the grid was set to the centroid of the binding pocket

residues. For the blind docking, the grid box dimensions were set as 126Å x 126Å x 126Å, effectively enclosing the entire receptor with the centre of grid being the receptor centroid. AutoDock Vina (version 1.1.2) (22) was used to calculate the predicted docking poses and their respective binding energies (Kcal/mol). Docking results were visualized and analysed using PyMOL and the respective images of the docked poses were generated using PyMOL. The performance of AutoDock Vina 1.1.2 for this target protein and the docking protocol was validated by redocking the  $\alpha$ -ketoamide into the catalytic site of the protease (6Y2G).

## **Molecular Dynamics Simulation**

Molecular Dynamics simulations of M<sup>pro</sup>-drug complexes were carried out using AMBER18 package (28 D.A. Case et.al). Force field parameters for the top hits (drugs) were generated using Antechamber (29) wherein the AM1BCC charges (30) were used. Each system was solvated with a box of TIP3P water molecules such that the boundary of the box was at least 14 Å from any protein atom. The net charges in the system were balanced by adding counter ions. The force field ff14SB (31) was used for intermolecular interactions and Generalized Amber Force Field (GAFF) (32) was used to treat the drugs. Particle Mesh Ewald method (33) was used for computing long-range electrostatics. All bonds involving hydrogen were constrained by SHAKE (34). An integration time step of 2 fs was used for propagating the dynamics. Each system was initially minimized for 10,000 steps to remove any unfavourable interactions between the protein and the solvent, followed by heating to 300 K over 150 ps under normal pressure/temperature (NPT) conditions. Subsequently, each system was simulated for 250ns at constant temperature (300 K) and a pressure of 1atm NPT conditions and the structures were stored every 1ps for analysis. All the analysis was carried out using AMBER18 software and trajectories were visualised using VMD (35). The simulation of M<sup>pro</sup>-G, M<sup>pro</sup>-DG were terminated at 100ns as the drugs (Glecaprevir and Danoprevir) moved out of the binding pocket early in the simulation. The total simulation time for the reported M<sup>pro</sup>-drug complexes was 2.95 $\mu$ s.

## **Results**

### **Virtual Screening of antiviral drugs for the catalytic pocket of M<sup>pro</sup>**

To identify the potential inhibitors for M<sup>pro</sup>, virtual screening by molecular docking of the selected 130 antivirals was carried out using AutoDock Vina 1.1.2 (22) To begin with, benchmark docking was carried out to evaluate the performance of AutoDock Vina for the M<sup>pro</sup>

inhibitor/ligand  $\alpha$ -ketoamide 13b of the crystal structure. The top two poses of the docked ligand had Root Mean Square Deviations (RMSD) of 1.4 Å and 0.7 Å with respect to the crystal structure pose, corresponding to the binding scores of -8.5 Kcal/mol and -7.9 Kcal/mol [Figure S1]. Then based on the confidence obtained from the benchmark, as well as other successful docking studies reported earlier using AutoDock Vina (36), two sets of docking were performed, one targeting catalytic pocket of M<sup>pro</sup>, which can block the catalysis and the second wherein the whole protein was selected (blind docking) to check the possibility of any other preferable binding site(s). The top 10 hits of targeted docking are given in Table 1 along with their molecular targets and known toxicity data, obtained from DrugBank (<https://www.drugbank.ca>). Out of top 10 hits, 8 were anti-HCV drugs and 2 were anti-HIV drugs. Also, 7 out of top 10 hits were protease inhibitors and 2 were HCV-NS5A inhibitors. Interestingly, HCV-NS5A protein is also involved in viral replication, host-cell interactions and viral pathogenesis (37). Based on the binding scores from the targeted docking, top 4 hits with binding scores better than that of inhibitor  $\alpha$ -ketoamide 13b, namely Glecaprevir, Paritaprevir, Danoprevir, Elbasvir were identified as potential inhibitors for M<sup>pro</sup> of SARS-CoV-2. Interestingly, all three are anti-HCV drugs. The docked poses of the top 4 hits are shown in Figure S2.

### **Virtual screening of antiviral drugs for the entire M<sup>pro</sup> protease**

In the second docking, the entire surface of M<sup>pro</sup> was scanned for any preferential site potentially binding the antiviral drugs. The top 30 hits of favourable compounds, along with their preferential sites and known toxicity are given in Table 2A, 2B and 2C. The first inference from this docking is the identification of two other potential pockets [Figure 2]. Of the two sites, the first is near the dimeric interface and other is near the intra domain interface [Figure 2]. Nearly 47% of drugs docked near the dimeric interface, 37% at the catalytic site and 17% at the intra-domain interface. Since there is always a limitation in the exhaustiveness of the search space in the docking calculations, potential pockets of M<sup>pro</sup> were also independently calculated using Computed Atlas of Surface Topography of proteins (CASTp) [<http://cast.engr.uic.edu>, 38,]. The top three pockets identified by CASTp are shown in Figure 3; the catalytic pocket having a large accessible surface area (ASA) of 353 Å<sup>2</sup>, the second near the dimeric interface with ASA 78 Å<sup>2</sup> (pocket 2) and the third at the intra domain interface with ASA 148 Å<sup>2</sup> (pocket 3) and these two sites are close to the probable sites obtained from blind docking [Figure 3]. It is important to note that, unbiased/blind docking and CASTp

employ different methods, however, the predicted pocket residues by CASTp, overlap with the residues constituting the docked drugs obtained from blind docking experiment. These additional pockets, which are far from the catalytic pocket, could be a crucial factor for inhibiting M<sup>pro</sup> given its plasticity and the potential mutability. In addition to the inhibitors/drugs at the catalytic pocket, the combination of drugs at the other pockets could offer viable options for effective inhibition of M<sup>pro</sup>. To test this idea, based on the docking score of the drugs for three different pockets from targeted and blind dockings, 10 promising M<sup>pro</sup>-drug complexes, along with M<sup>pro</sup>-apo (drug free form), were chosen for further MD studies.

### **Molecular Dynamics simulations of M<sup>pro</sup>-drug complexes**

All-atom MD simulations were carried out for the various drug combinations with M<sup>pro</sup>. The drugs used in these simulations (along with their single letter representations) were Danoprevir (D), Glecaprevir (G), Elbasvir (E), Ritonavir (R), and Velpatasvir (V). For brevity, in the following discussion, only single letter representations are used.

10 simulations were performed for various combinations in which (monomer) M<sup>pro</sup> is complexed with 1 to 3 drugs. They were as follows:

**Single drug complexes:** Three simulations of M<sup>pro</sup> were carried out with single drug bound at catalytic site. They were, G – the top hit from the targeted docking. E, and D were the top hits for catalytic pocket from the blind docking.

**Double drug combinations:** Four 2-drug-M<sup>pro</sup> complex simulations were carried out. They were, M<sup>pro</sup> with D at the catalytic site and G at the dimeric interface (DG) (as two of them were top hits for the respective pockets from blind docking), E at catalytic site and G at the dimeric interface (M<sup>pro</sup>-EG), E at catalytic site and V at the intra-domain interface, pocket3 (EV) (V being top hit for pocket3), and E at the catalytic site and R at, pocket3 (ER). The reason for choosing the Ritonavir at the third pocket, next to Velpatasvir is, based on the success of reported clinical studies (7). Also, docking scores did not differ much among the other choices.

**Triple drug combinations:** Three 3-drug combinations were simulated. They were D, G and V (DGV), E, G, and R (EGR), and E, G, and V (EGV) at catalytic pocket (pocket1), dimeric interface (pocket2) and intra domain interface (pocket3) respectively.

To check the stability of the drugs as well as to understand the dynamics of M<sup>pro</sup> in the presence of various combinations of drugs at the three pockets, very long all-atom MD simulations of 250ns were carried out. In the case of M<sup>pro</sup>-G and M<sup>pro</sup>-DG, the drugs moved out of the respective pockets and simulations were terminated at 100 ns.

All the 250 ns simulations (except two 100ns of M<sup>pro</sup>-G, M<sup>pro</sup>-DG simulations) were stable as inferred from their C<sup>α</sup>-RMSDs with respect to their starting structures, which were within the range of 1.6 Å to 2.4 Å with standard deviations in the range of 0.17 to 0.72 [ Figure S3]. The stability of drugs in the respective simulations were analyzed for each system.

### **Stability of the drugs during the course of simulation**

In the single drug M<sup>pro</sup>-G and M<sup>pro</sup>-D simulations, Glecaprevir and Danoprevir were mobile inside the catalytic pocket due to their small cyclic structures unable to pack well into the wider dynamic groove of the catalytic pocket [Movie S1]. In fact, in M<sup>pro</sup>-D simulation, Danoprevir completely moved out of the catalytic pocket. In another single drug complex, M<sup>pro</sup>-E, though Elbasvir was stable for a long time, its central benzene ring flipped out of the pocket after 200 ns due to the dynamics of the loop forming the catalytic pocket (mainly from residues 131-143) [Movie 1].

In the two-drug combination, M<sup>pro</sup>-DG, Danoprevir at the catalytic pocket was again not stable due to its cyclic ring structure and was unable to fit well in the catalytic pocket. Glecaprevir was also not stable in its pocket and moved towards the interface of all  $\alpha$ -helical domain early in the simulation. Among other combinations, Elbasvir in the catalytic pocket was more stable in only in ER [Movie 2] combination than in EG, and EV [Movie S2 and S3]. This is probably due to the relatively stable Ritonavir in the pocket 3 than Glecaprevir and Velpatasvir at the pocket 2, pocket 3 respectively. Whereas the central benzene ring of Elbasvir flipped out of catalytic pocket in both EG, and EV combination as early as 650 ps and started moving out of the pocket [Movie S2 and S3]. M<sup>pro</sup>-ER was the most stable combination among all two drug combinations.

In the case of M<sup>pro</sup>-three drug combination of M<sup>pro</sup>-DGV, though Danoprevir and Glecaprevir were not stable similar to the two-drug combination, Velpatasvir was quite stable at the pocket3 (intradomain interface) and it was seen to be stable in the EGV combination as well [Movie

S4]. In the next three drug combination, M<sup>pro</sup>-EGR simulation, Elbasvir was quite stable in the catalytic pocket amidst the dynamic catalytic loop and linker, where Glecaprevir at pocket2, was more mobile as seen from the other simulations. However, after 150 ns, it became stable near the  $\alpha$ -helical domain and Ritonavir at pocket3 (intradomain interface) was seen to be quite stable [Figure 4A and Movie 3]. The stability of Elbasvir at the catalytic site is due to the following favorable interactions; its central benzene ring is nestled very well within the relatively hydrophobic groove with residues like Thr24, Leu27, His41, Phe140, Cys141, Cys145, His163, Met165, Pro168, and His172, and its methoxycarbonyl and carbamate on either side making polar contacts with Thr28, Thr29, Thr30 and Arg198, spanning the entire pocket, thus stabilizing it [Figure 4B]. At the pocket2, Glecaprevir was unstable and it was mobile throughout the simulation and had only one interaction with Asn277 of  $\alpha$ -helical domain [Figure 4D]. At pocket3, (the intradomain interface), Ritonavir was observed to well packed among the residues Asn151, Ser158, Cys160, Ile106, Gln107, Pro108, Gly109, Gln110, while its peptide backbone was seen to interact with Thr292, Pro293, backbone of Phe294, Gln110, Gly109 and its benzyl ring was packed in a hydrophobic pocket formed by residues Pro293, Phe294, Asn151, Val104, Ile106 [Figure 4C]. Another stable M<sup>pro</sup>-three drug complex was EGV, where Elbasvir, Glecaprevir and Velpatasvir were very stable at their respective pockets [Figures 5A and 5B, Movie 4]. The behaviour of Elbasvir and Glecaprevir was similar to that of EGR combination, and in-fact Glecaprevir in this combination was more stable due to the favourable packing and interactions from Asn277 and Tyr329. Velpatasvir was quite stable in the ; intradomain interface pocket 3; with well packed hydrophobic structure formed by Pro 108, Ile249, Phe294, Val227, Ile249, Gly109, Gln100. [Figure 5B]. In summary, M<sup>pro</sup>-EGR and M<sup>pro</sup>-EGV complexes were stable in the bound form for a longer time scale of simulation. The residues constituting the newly identified pockets along with the summary of all the M<sup>pro</sup>-monomer-drug simulations are given in the Table 3.

Further, to check whether similar to the most stable monomer M<sup>pro</sup>-EGR, EGV drug complexes are stable in the dimeric form, 250 ns simulation of M<sup>pro</sup>-dimer-EGR and EGV were carried out and the stability of the drugs were checked. While the drugs EGR were stable in the respective pockets throughout the simulation [Movie 5], EGV was stable in only one of the protomers [Movie 6]. This observation is of significance based on the experimental studies reported earlier on SARS-CoV M<sup>pro</sup>, where it was concluded that dimerization was a requirement for the enzymatic activity of the protease, though only one monomer in the

protease dimer displayed catalytic activity (39). The details of all the other dimer simulations will be discussed elsewhere. Interestingly, E is also reported to bind to three other targets of Cov-2 replication machinery such as RdRP, papain-like proteinase, and helicase, based on virtual screening of 54 FDA- approved antiviral drugs and ~3300 investigational drugs, emphasis its potential as promising viral replication inhibitor (40).

Overall, all the M<sup>pro</sup>-drug complexes exhibited different stability/mobility when compared to M<sup>pro</sup>-apo monomer (drug-free); the conformational flexibility as inferred by Root Mean Square fluctuations (RMSF) of C<sup>α</sup> atoms of M<sup>pro</sup> were also different [Figure 6]

### **Differential drug response of M<sup>pro</sup>**

The Root Mean Square Fluctuations (RMSF) of C<sup>α</sup> atoms, indicator of rigid and flexible region of M<sup>pro</sup>, in the presence of single, two and three drugs, differed mainly in the catalytic loop (residues Cys44 to Pro52), loop connecting strands which forms the catalytic pocket (especially the loop 136 to 146, where 3/10 residues are Gly, loop constituting the residues 166-170, 153-156, linker region (Phe185-Thr201) and α-helical domain III (Figures 6A, 6B and 6C). Compared to M<sup>pro</sup>-apo, the presence of drugs showed reduction in RMSF values, reduced dynamics; whereas three drug combinations showed the most. Also, the presence of drug and stability of drugs in the catalytic pocket, showed variation in the RMSF values of α-helical domain. The more stable drug in the catalytic pocket showed reduced RMSF values in the α-helical domain (eg., presence of E vs M<sup>pro</sup>-apo RMSF values, *see* Figure 6B). However, the variations in C<sup>α</sup> fluctuations of α-helical domain in all simulation could be due to, a) the absence of drug at the pocket3/pocket2 (G, E, DG, EG simulations), b) the instability of drug at the pockets (like Glecaprevir of EG, Velpatasvir of EV simulations), or c) unstable drug in the catalytic pocket (like Danoprevir of DG simulation, Elbasvir of EV simulation). The RMSF values of M<sup>pro</sup>-G (top hit of targeted docking) didn't show significant difference with that of M<sup>pro</sup> -apo [Figure S4]. Next, in order, to check for any correlation between the observed fluctuations in the three-dimensional space vis-à-vis any cross talk between the three pockets, dynamic cross correlation matrix (DCCM) of C<sup>α</sup> atoms were calculated for the snapshots taken from all M<sup>pro</sup>-drug simulations.

### **Correlated motions and allostery**

The ( $C^\alpha$ ) Dynamic Cross Correlation Matrices (DCCM) were computed of the structures taken from all  $M^{pro}$ -drug complexes and compared with that of apo- $M^{pro}$ . DCCM shows the correlated movements in proteins that occurred during the simulation as well as their direction. The correlated motions range between +1 or -1 and the correlations which are greater than 0.5 (colored in red) and less than -0.5 (colored in blue) are shown in the Figures 7A-E. The calculated DCCM were strikingly different for the mentioned systems, indicating that differential response of  $M^{pro}$  in the presence of various drugs. The apo- $M^{pro}$ -apo DCCM shows, positive correlated motions between residues of catalytic region and the  $\alpha$ -helical domain. [Figure 7A]. Since  $M^{pro}$ -EGR is observed to be the most stable complex, comparison of DCCM of  $M^{pro}$ -E, EG, ER, EGR was carried out. Compared to apo- $M^{pro}$ , the presence of single drug E, shows striking difference in the DCCM, where the positive correlated motions of the catalytic pocket and the helical domain is lost (boxed region in Figure 7C). The DCCM of  $M^{pro}$ -EG, ER complexes also showed reduced correlated motion compared to the  $M^{pro}$ -apo, and the three-drug combination EGR showed less correlated motion (both positive and negative correlated motions), indicating the effect of three drugs in reducing the conformational dynamics/flexibility and stronger binding efficiency. Strikingly all of them showed a presence of correlated motion between the catalytic pocket and the  $\alpha$ -helical domain III: an indication of long- range communication between the distant sites.

Thus, the druggable pockets at the dimeric interface (pocket2) and intra-domain interface (pocket3) are allosteric pockets, and appear to be coupled with the conformational dynamics of the catalytic pocket.

### **Binding energetics of $M^{pro}$ -drug complexes**

In order to check whether there is any cooperativity among the binding of drugs in the allosteric pockets, binding energetics of  $M^{pro}$ -drugs were calculated.

The binding free energy for  $M^{pro}$ -drug complexes were calculated using MM/GBSA method of AMBER18. Thee last 50 structures (considering them as most stable states) were taken for calculating energetics. MM/GBSA essentially combines the Molecular Mechanic (MM)/Generalised Born Surface Area (MM-GBSA). The MM constitutes of potential energy, GBSA constitutes polar and non-polar solvation energies. The binding free energy ( $\Delta G_{binding}$ ) can be computed following the theory as explained in (41,42) and is expressed in the following set of equations:

$$\Delta G_{\text{binding}} = \Delta G_{\text{water}}(\text{complex}) - (\Delta G_{\text{water}}(\text{protein}) + \Delta G_{\text{water}}(\text{ligand}))$$

$$\Delta G_{\text{water}} = E_{\text{MM}} + \Delta G_{\text{solvation}} - TS$$

$$G_{\text{solvation}} = G_{\text{solvation} - \text{electrostatic}} + G_{\text{nonpolar}}$$

$$E_{\text{MM}} = E_{\text{internal}} + E_{\text{electrostatic}} + E_{\text{vdW}}$$

$$E_{\text{internal}} = E_{\text{bond}} + E_{\text{angle}} - E_{\text{torsion}}$$

Here,  $\Delta G_{\text{binding}}$ ,  $\Delta G_{\text{water}}(\text{complex})$ ,  $\Delta G_{\text{water}}(\text{protein})$  and  $\Delta G_{\text{water}}(\text{ligand})$  are free energies of binding, complex, free receptor and unbound ligand respectively. The binding free energies of each components is sum of the absolute molecular mechanical energies ( $E_{\text{MM}}$ ) and the solvation free energies ( $G_{\text{solvation}}$ ). The entropy contribution is from vibrational, rotational and translational entropies obtained by normal modes. From the energetics calculations, it is evident that in single drug combination, E is relatively favourable due to its favourable packing and electrostatic interactions [Table 4]. The binding energetics of two drug combinations, both enthalpic and the entropic didn't differ much, though ER, EG, EV appear better than the DG combination. It is pertinent to note the enhanced binding enthalpies of M<sup>PRO</sup>-ER, EV, EG compared to M<sup>PRO</sup>-E, indicative of the cooperativity of binding in M<sup>PRO</sup>-drug complexes. It is further evident from the energetics of M<sup>PRO</sup>-EGR, which is clearly more favourable than the M<sup>PRO</sup>-E, ER, EV and EG combinations; indicating the cooperative association of drug complexes with M<sup>PRO</sup>.

### **Residue-wise contribution for M<sup>PRO</sup>-EGR interactions.**

Further, to understand the role of key residues/hotspots contributing to the drugs binding in the three sites, residue wise decomposition analysis was carried out for M<sup>PRO</sup>-EGR complex, the most stable combination. The present residual decomposition is obtained for the 50 snapshots of monomer M<sup>PRO</sup>-EGR simulations. At the catalytic site, the stability of Elbasvir was due to the favourable interactions mainly from Gln189 (where NH- $\pi$ ) interactions, just located at the entrance of the catalytic pocket, locking the pocket, and favourable packing of its central benzene ring surrounded by Met165, Glu166, His164, Asn142, Gly143, Cys145. And the favourable hydrogen bonds between Thr24, Thr25, Thr26, Leu 27 (backbone) of catalytic loop on one side, Gln192, Ala191, Thr190, Arg192 backbone with methoxycarbonyl and carbamate

on either side [Figures 8A and 4B]. And all of them emerged as per-residue contributors of favourable binding of E. At the pocket2, Glecaprevir quinoxaline ring was stabilised mostly by the hydrophobic residues Leu286,Leu287,Leu272, Met276,Ala285,Gly275,Gly278 and favourable polar contact from side chain of Asn277 of  $\alpha$ -helical domain [Figures 8B and 4D]. Which has emerged as major per-residue contributors. At pocket3, (the intradomain interface), Ritonavir was nicely packed between the residues Asn151, Ser158, Cys160, Ile106, Gln107, Pro108, Gly109, Gln110, while its peptide backbone was seen to interact with Thr 292, Pro293, backbone of Phe294, Gln110, Gly109 and its benzyl ring was packed in a hydrophobic pocket formed by residues Pro293,Phe294, Asn151, Val104, Ile106 [Figures 8C and 4C].

To conclude, more number of per-residue contributors were seen for Elbasvir and Ritonavir and which was also reflected in binding energetics [Table 4].

## Discussion

Repurposing available drugs is a promising strategy for treating (COVID-19) which has become a global pandemic at unprecedented pace. There is no specific treatment or drug that is readily available. Main protease ( $M^{pro}$ ) being one of the main proteases, essential for assembling viral transcription and replication complex, in addition to the absence of a similar protease in humans, makes it an attractive target for designing/repurposing antivirals as COVID therapeutics. All the efforts in this direction based on virtual screening, of massive library of compounds, natural products and repurposing of antiviral drugs, has taken into account the conformational flexibility and the plasticity of  $M^{pro}$ . Now more evidences suggest that flexibility of  $M^{pro}$  being a bottleneck for structure-based drug design/repurposing efforts (19-21). This arises mainly due to the dynamic interplay of charged residues, especially R131 and K137, which are strategically located on the loop connecting  $\beta$ 10-11 of catalytic pocket, and Asp289 and Glu290 of  $\alpha$ -helical domain. And, major dynamics of pocket is due to interplay of Arg40 (catalytic pocket)-Glu55 of helix 2 and Asp187 of linker catalytic loop, contiguous Glu47 and Asp48 of catalytic loop) [Figure 9 and Movie 7]. All this dynamic electrostatic interplay contributes to the major characteristic of  $M^{pro}$  catalytic pocket -- its flexibility and associated plasticity that results it in being a major challenge to structure-based drug design. In addition, another layer of complication may arise due to the potential mutability of  $M^{pro}$  at any site including the catalytic pocket, and its associated flexibility; and this has been already shown by the increase in catalytic activity by Thr285Ala mutations of  $\alpha$ -helical

domain, which is far from the catalytic pocket (12). To overcome the above mentioned challenges, a more plausible solution could be combination therapy, involving drugging the pockets other than the catalytic pocket of M<sup>pro</sup> with a viable combination of drugs.

For the reasons elucidated in preceding sections, select combinations of the drugs were examined in detail through long-scale MD simulations. Based on the results we conclude that, Elbasvir as a single drug could stably occupied the dynamic catalytic pocket for a fairly long timescale (~200 ns) owing to well-packed central heterotetracyclic ring and favorable polar interactions. However, it started to move out towards the end of the simulation. When the allosteric pockets were occupied with drugs Glecaprevir and Ritonavir, the stability of Elbasvir at the catalytic pocket was enhanced as seen from the simulations of double, triple drug combinations (ER, EGR). Reduced conformational flexibility of M<sup>pro</sup>, as inferred from RMSF, DCCM, and binding energetics, indicates more potent inhibition.

It is also pertinent to note that combination therapy of protease inhibitors is a well-established treatment regimen for HIV patients; where protease inhibitor Ritonavir has been shown to boost the effect of other protease inhibitors like Lopinavir (43).

## **Conclusion**

Key finding of this study is the identification of potential druggable allosteric pockets at the intra-domain and dimeric interface. Further, observed dynamics of catalytic pocket in the presence of various drugs may be fine-tuned for better affinity with the newly identified allosteric pockets. This could be important to circumvent the resistance mutations that may arise (even in the catalytic pocket of M<sup>pro</sup>). The modulation of dynamics of M<sup>pro</sup> by the presence of drug at the allosteric pockets can modulate drug binding at the catalytic pocket, which necessitates the need for the right choice of drugs for combination therapy. This study also opens up the new route for drug design; i.e., fine tuning of conformational dynamics of M<sup>pro</sup> by chemical groups of drugs/inhibitors to modulate the transition states of M<sup>pro</sup> for binding, hence its binding affinity/drug efficacy. Importantly, these newly identified pockets, are also conserved in SARS-CoV, suggesting the possibility of broad spectrum of combination of antivirals targeting M<sup>pro</sup> of CoVs. Collectively, the combination of anti-HCV drug Elbasvir, Glecaprevir and anti-HIV drug Ritonavir appear to be a viable candidate for further experimental testing and development of anti-COVID-19 drugs.

## **Competing interest**

Authors have declared no conflict of interest.

## **Acknowledgements**

AM would like to acknowledge University Grants Commission's UGC-Faculty recharge Programme (No.F.4-5(234-FRP)/2015(BSR)). We thank the Central Computing Facility of the Indian Institute of Information Technology-Allahabad for providing computational facilities for work related MD simulations. The support of Ramalingaswami Fellowship (BT/RLF/Re-entry/09/2015), Government of India, Ministry of Science and Technology, Department of Biotechnology (DBT) and support of Early Career Research Award (File No. ECR/2018/002114) from Science & Engineering Research Board (SERB), Department of Science & Technology, Government of India to author Jawed Iqbal are acknowledged. AM would like to thank Dr. Beena Krishnan, Dr. Tanveer Ahmad for valuable discussions and suggestions.

## **Authors contribution**

AM and JI conceived the idea. ZAB carried out the molecular docking. DC and BS carried out the simulations and analysis was carried out by AM and BS. AM, BS and JI wrote the manuscript.

## **References**

1. Anthony, .R.; Perlman, S., Coronaviruses: An Overview of Their Replication and Pathogenesis, *Methods Mol. Biol.* 2015, 1282, 1–23.
2. Ghebreyesus, T. A., WHO Director-General's Opening Remarks at The Media Briefing on COVID-19, WHO, 2020.
3. Wu, F.; Zhao, S.; Yu, B.; Chen, Y. M.; Wang, W.; Song, Z.G.; Hu, Y.; Tao, Z.W.; Tian, J.H.; Pei, Y.Y.; Yuan, M.L.; Zhang, Y.L.; Dai, F.H.; Liu, Y.; Wang, Q.M.; Zheng, J.J.; Xu, L.; Holmes, E.C.; Zhang, Y.Z., A New Coronavirus Associated with Human Respiratory Disease in China, *Nature*, 2020, 579, 265–269.
4. Chen, Y.; Liu, Q.; Guo, D., Emerging Coronaviruses: Genome Structure, Replication, and Pathogenesis, *J. Med. Virol.* 2020, 92, 418-423.
5. Zhou, P.; Yang, X.L.; Wang, X.G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.R.; Zhu, Y.; Li, B.; Huang, C.L.; Chen, H.D.; Chen, J.; Luo, Y.; Guo, H.; Jiang, R.D.; Liu, M.Q.;

Chen, Y.; Shen, X.R.; Wang, X.; Zheng, X.S.; Zhao, K.; Chen, Q.J.; Deng, F.; Liu, L.L.; Yan, B.; Zhan, F.X.; Wang, Y.Y.; Xiao, G.F.; Shi, Z.L., A Pneumonia Outbreak Associated with a New Coronavirus of Probable Bat Origin, *Nature*, 2020, 579, 270-3.

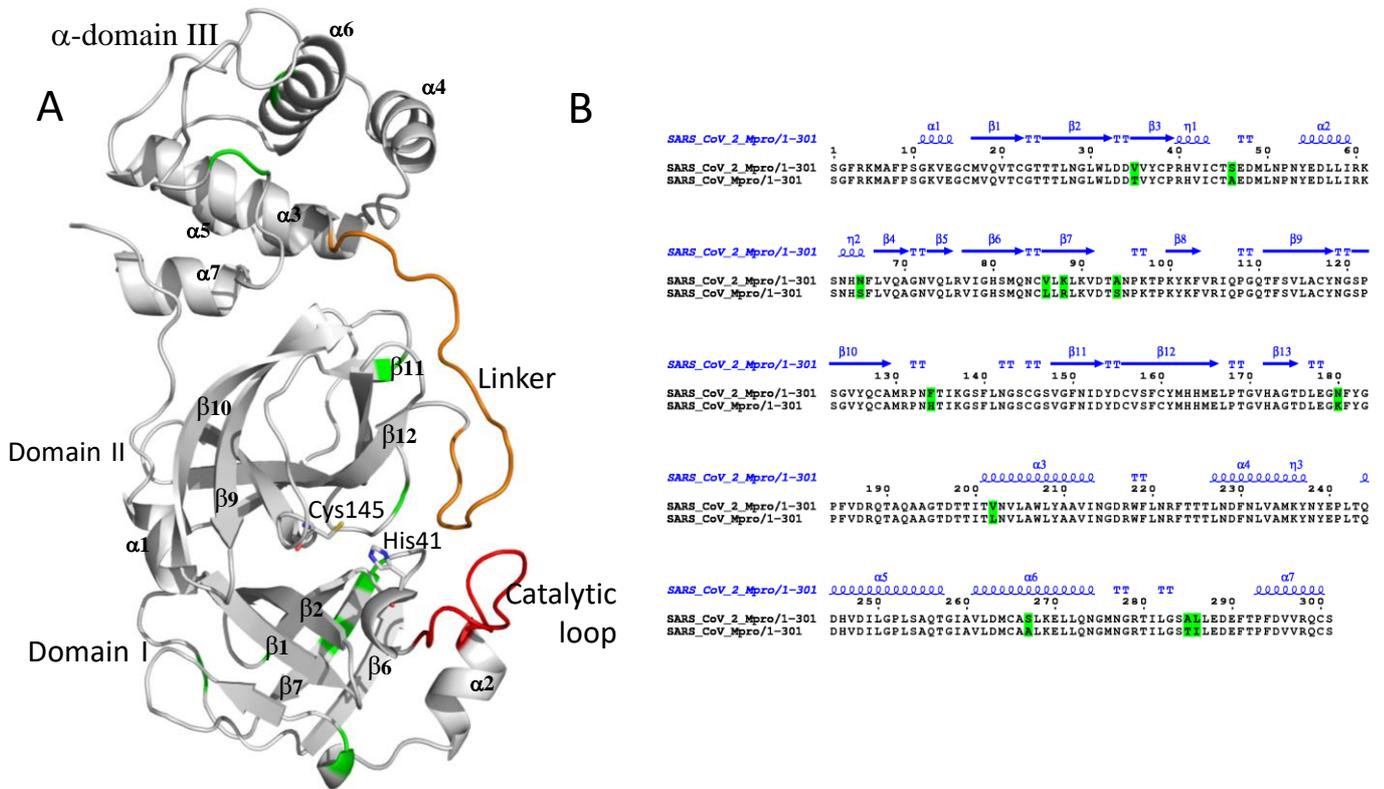
6. Li, Q.; Guan, X.; Wu, P.; Wang, X.; Zhou, L.; Tong, Y.; Ren, R.; Leung, K.S.M.; Lau, E.H.Y.; Wong, J.Y.; Xing, X.; Xiang, N.; Wu, Y.; Li, C.; Chen, Q.; Li, D.; Liu, T.; Zhao, J.; Liu, M.; Tu, W.; Chen, C.; Jin, L.; Yang, R.; Wang, Q.; Zhou, S.; Wang, R.; Liu, H.; Luo, Y.; Liu, Y.; Shao, G.; Li, H.; Tao, Z.; Yang, Y.; Deng, Z.; Liu, B.; Ma, Z.; Zhang, Y.; Shi, G.; Lam, T.T.Y.; Wu, J.T.; Gao, G.F.; Cowling, B.J.; Yang, B.; Leung, G.M.; Feng, Z., Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia, *N. Engl. J. Med.* 2020, 382, 1199-1207.
7. Cao, B.; Wang, Y.; Wen, D.; Liu, W.; Wang, J.; Fan, G.; Ruan, L.; Song, B.; Cai, Y.; Wei, M.; Li, X.; Xia, J.; Chen, N.; Xiang, J.; Yu, T.; Bai, T.; Xie, X.; Zhang, L.; Li, C.; Yuan, Y.; Chen, H.; Li, H.; Huang, H.; Tu, S.; Gong, F.; Liu, Y.; Wei, Y.; Dong, C.; Zhou, F.; Gu, X.; Xu, J.; Liu, Z.; Zhang, Y.; Li, H.; Shang, L.; Wang, K.; Li, K.; Zhou, X.; Dong, X.; Qu, Z.; Lu, S.; Hu, X.; Ruan, S.; Luo, S.; Wu, J.; Peng, L.; Cheng, F.; Pan, L.; Zou, J.; Jia, C.; Wang, J.; Liu, X.; Wang, S.; Wu, X.; Ge, Q.; He, J.; Zhan, H.; Qiu, F.; Guo, L.; Huang, C.; Jaki, T.; Hayden, F.G.; Horby, P.W.; Zhang, D.; Wang, C., A Trial of Lopinavir–Ritonavir in Adults Hospitalized with Severe Covid-19, *N. Engl. J. Med.* 2020, 382, 1787-1799.
8. Yao, T.T.; Qian, J.D.; Zhu, W.Y.; Wang, Y.; Wang, G.Q., A Systematic Review of Lopinavir Therapy for SARS Coronavirus and MERS Coronavirus—A Possible Reference for Coronavirus Disease-19 Treatment Option, *J. Med. Virol.* 2020, 92, 556-563.
9. Chu, C.M.; Cheng, V.C.; Hung, I.F.; Wong, M.M.; Chan, K.H.; Chan, K.S.; Kao, R.Y.; Poon, L.L.; Wong, C.L.; Guan, Y.; Peiris, J.S.; Yuen, K.Y., Role of Lopinavir/Ritonavir in the Treatment of SARS: Initial Virological and Clinical Findings, *Thorax*, 2004, 59, 252-256.
10. Brown, A.J.; Won, J.J.; Graham, R.L.; Dinnon, K.H.; Sims, A.C.; Feng, J.Y.; Cihlar, T.; Denison, M.R.; Baric, R.S.; Sheahan, T.P., Broad Spectrum Antiviral Remdesivir Inhibits Human Endemic and Zoonotic Delta-Coronaviruses with a Highly Divergent RNA Dependent RNA Polymerase, *Antiviral Res.* 2019, 169, 104541.
11. Chen, H.; Zhang, Z.; Wang, L.; Huang, Z.; Gong, F.; Li, X.; Chen, Y.; Wu, J.J., First Clinical Study Using HCV Protease Inhibitor Danoprevir to Treat Naive and Experienced COVID-19 Patients, *medRxiv preprint*, 2020.
12. Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R., Crystal Structure of SARS-CoV-2 Main Protease Provides a Basis for Design of Improved Alpha-Ketoamide Inhibitors, *Science*, 2020, 368, 409-412.
13. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L.W.; Xu, W.; Xiao, G.; Qin, C.; Shi,

- Z.; Jiang, H.; Rao, Z.; Yang, H., Structure of M<sup>pro</sup> from COVID-19 Virus and Discovery of its Inhibitors, *Nature*, 2020, 582, 289-293.
14. Li, Y.; Zhang, J.; Wang, N.; Li, H.; Shi, Y.; Guo, G.; Liu, K.; Zeng, H.; Zou, Q., Therapeutic Drugs Targeting 2019-nCoV Main Protease by High-Throughput Screening, *BioRxiv*, 2020.
  15. Sanders, J.M.; Monogue, M.L.; Jodlowski, T.Z.; Cutrell, J.B., Pharmacologic Treatments for Coronavirus Disease 2019 (COVID-19): A Review, *JAMA*, 2020.
  16. Ton, A.T.; Gentile, F.; Hsing, M.; Ban, F.; Cherkasov, A., Rapid Identification of Potential Inhibitors of SARS-CoV-2 Main Protease by Deep Docking of 1.3 Billion Compounds, *Mol. Inform.* 2020, 39, e2000028.
  17. Xu, Z.; Peng, C.; Shi, Y.; Zhu, Z.; Mu, K.; Wang, X.; Zhu, W., Nelfinavir was predicted to be a potential inhibitor of 2019-nCoV main protease by an integrative approach combining homology modelling, molecular docking and binding free energy calculation, *bioRxiv*, 2020.
  18. Teague, S.J., Implications of Protein Flexibility for Drug Discovery, *Nat. Rev. Drug Discov.* 2003, 2, 527–541.
  19. Kneller, D.W.; Phillips, G.; O'Neill, H.M.; Jedrzejczak, R.; Stols, L.; Langan, P.; Joachimiak, A.; Coates, L.; Kovalevsky, A., Structural Plasticity of SARS-CoV-2 3CL M<sup>pro</sup> Active Site Cavity Revealed by Room Temperature X-ray Crystallography, *Nat. Commun.* 2020, 11, 3202.
  20. Estrada, E., Topological Analysis of SARS CoV-2 Main Protease, *Chaos*, 2020, 30, 061102.
  21. Bzowka, M.; Mitusińska, K.; Raczyńska, A.; Samol, A.; Tuszyński, J.A.; Gora, A., Structural and Evolutionary Analysis Indicate that the SARS-CoV-2 M<sup>pro</sup> is an Challenging Target for Small-Molecule Inhibitor Design, *Int. J. Mol. Sci.* 2020, 21, 3099.
  22. Trott, O.; Olson, A. J., AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading, *J. Comput. Chem.* 2010, 31, 455-461.
  23. Huey, R.; Morris, G. M., Using AutoDock 4 with AutoDocktools: a Tutorial, *The Scripps Research Institute*, 2008, 54-56.
  24. Wishart, D. S.; Knox, C.; Guo, A. C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J., DrugBank: a Comprehensive Resource for in Silico Drug Discovery and Exploration, *Nucleic Acids Res.* 2006, 34, D668-D672.
  25. Pence, H. E.; Williams, A., ChemSpider: an Online Chemical Information Resource, *J. Chem. Educ.* 2010, 87, 1123-1124.

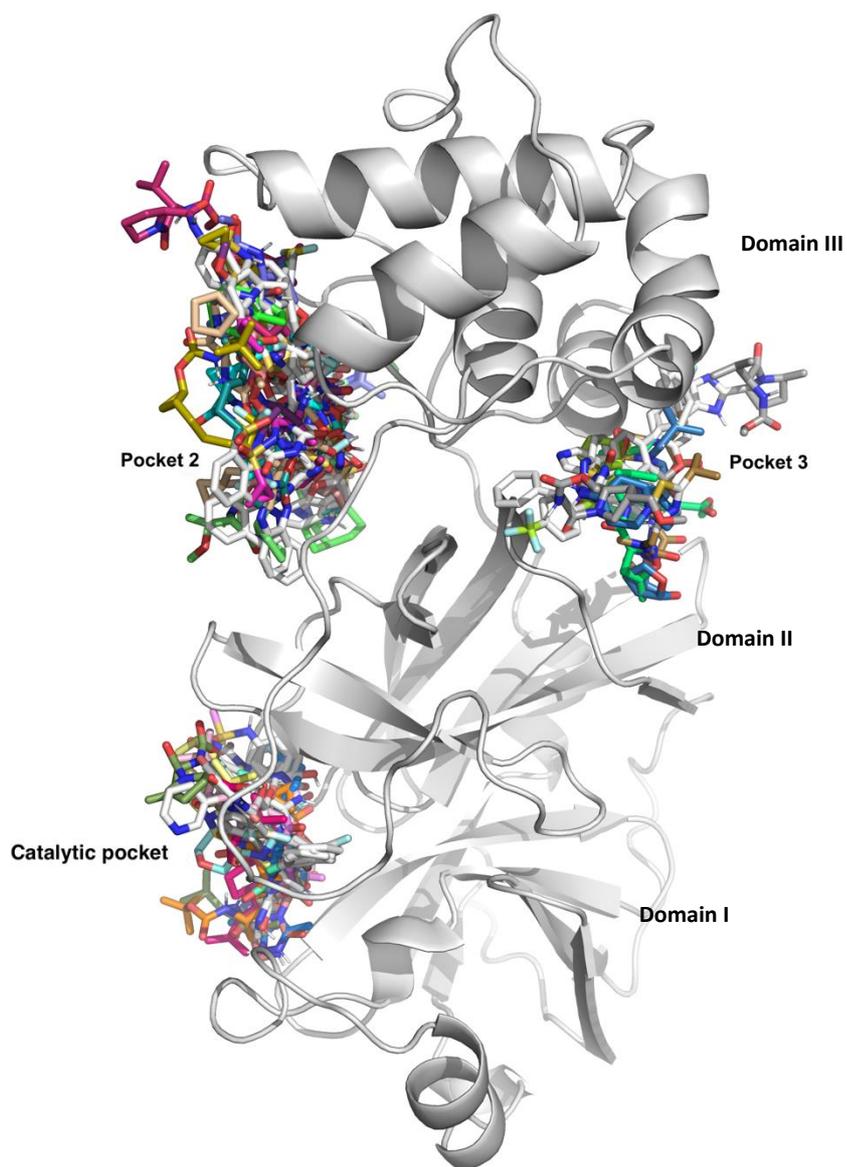
26. O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R., Open Babel: An Open Chemical Toolbox, *J. Cheminform.* 2011, 3, 33.
27. Delano, W. L., PyMOL, DeLano Scientific, 2002.
28. Case, D.A.; Ben-Shalom, I.Y.; Brozell, S.R.; Cerutti, D.S.; Cheatham, T.E.; Cruzeiro, V.W.; Darden, T.A.; Duke, R.E.; Ghoreishi, D.; Gilson, M.K.; Gohlke, H.; Goetz, A.W.; Greene, D.; Harris, R.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T.S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D.J.; Merz, K.M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D.R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C.L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R.C.; Wang, J.; Wei, H.; Wolf, R.M.; Wu, X.; Xiao, L.; York, D.M.; Kollman, P.A., AMBER, 2018, University of California, San Francisco.
29. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A., Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations, *J. Mol. Graph Model.* 2006, 25, 247-260.
30. Jakalian, A.; Jack, D. B.; Bayly, C. I., Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC model: II. Parameterization and Validation, *J. Comput. Chem.* 2002, 23, 1623-1641.
31. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB, *J. Chem. Theory Comput.* 2015, 11, 3696-3713.
32. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and Testing of a General Amber Force Field, *J. Comput. Chem.* 2004, 25, 1157-1174.
33. Darden, T.; York, D.; Pedersen, L., Particle Mesh Ewald: An  $N \cdot \log(N)$  Method for Ewald Sums in Large Systems, *J. Chem. Phys.* 1993, 98, 10089-10092.
34. Van-Gunsteren, W. F.; Berendsen, H. J.C., Algorithms for Macromolecular Dynamics and Constraint Dynamics, *Mol. Phys.* 1977, 34, 1311-1327.
35. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual Molecular Dynamics, *J. Mol. Graph.* 1996, 14, 33-38.
36. Gaillard, T., Evaluation of AutoDock and AutoDock Vina on the CASF-2013 Benchmark, *J. Chem. Inf. Model.* 2018, 58, 1697-1706.
37. He, Y.; Staschke, K. A.; Tan, S. L., HCV NS5A: A Multifunctional Regulator of Cellular Pathways and Virus Replication, *Hepatitis C Viruses: Genomes and Molecular Biology*. Norfolk (UK): Horizon Bioscience, 2006, Chapter 9.
38. Binkowski, T. A.; Naghibzadeh, S.; Liang, J., CASTp: Computed Atlas of Surface Topography of Proteins, *Nucleic Acids Res.* 2003, 31, 3352-3355.

39. Chen, H.; Wei, P.; Huang, C.; Tan, L.; Liu, Y.; Lai, L., Only One Protomer Is Active in the Dimer of SARS 3C-like Proteinase, *J. Biol. Chem.* 2006, 281, 13894–13898.
40. Balasubramaniam, M.; Reis, R. S., Computational Target-Based Drug Repurposing of Elbasvir, an Antiviral Drug Predicted to Bind Multiple SARS-CoV-2 Proteins, *ChemRxiv*, 2020.
41. Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke H.; Roitberg, A. E., MMPBSA.py: An Efficient Program for End-State Free Energy Calculations, *J. Chem. Theory Comput.* 2012, 8, 3314-3321.
42. Massova, I.; Kollman, P.A., Combined Molecular Mechanical and Continuum Solvent Approach (MM-PBSA/GBSA) to Predict Ligand Binding, *Perspect. Drug Discov. Des.* 2000, 18, 113–135.
43. Kumarasamy, N.; Aga, E.; Ribaud, H.J.; Wallis, L.C.; Katzenstein, D.A.; Stevens, S.W.; Norton, M.R.; Klingman, K.L.; Hosseinipour, M.C.; Crump, J.A.; Supparatpinyo, K.; Badal-Faesen, S.; Bartlett, J.; Lopinavir/Ritonavir Monotherapy as Second-line Antiretroviral Treatment in Resource-Limited Settings: Week 104 Analysis of AIDS Clinical Trials Group (ACTG) A5230, *Clinical Infectious Diseases*, 2015, 60, 1552-1558

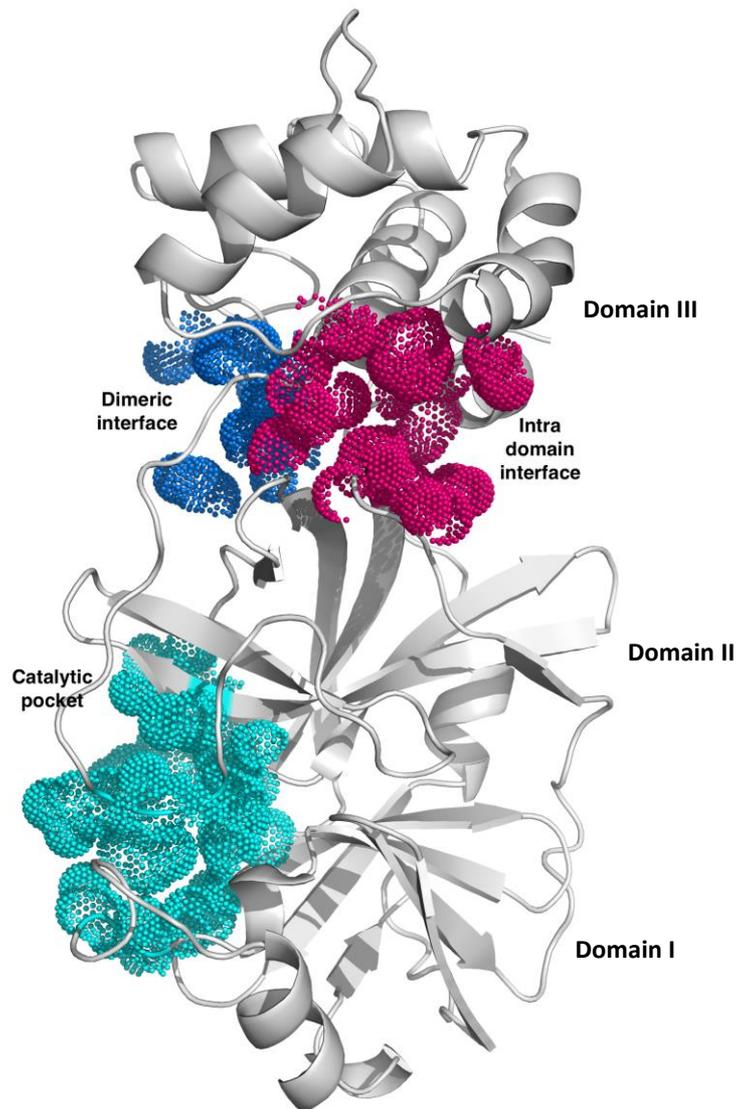
## All Figures



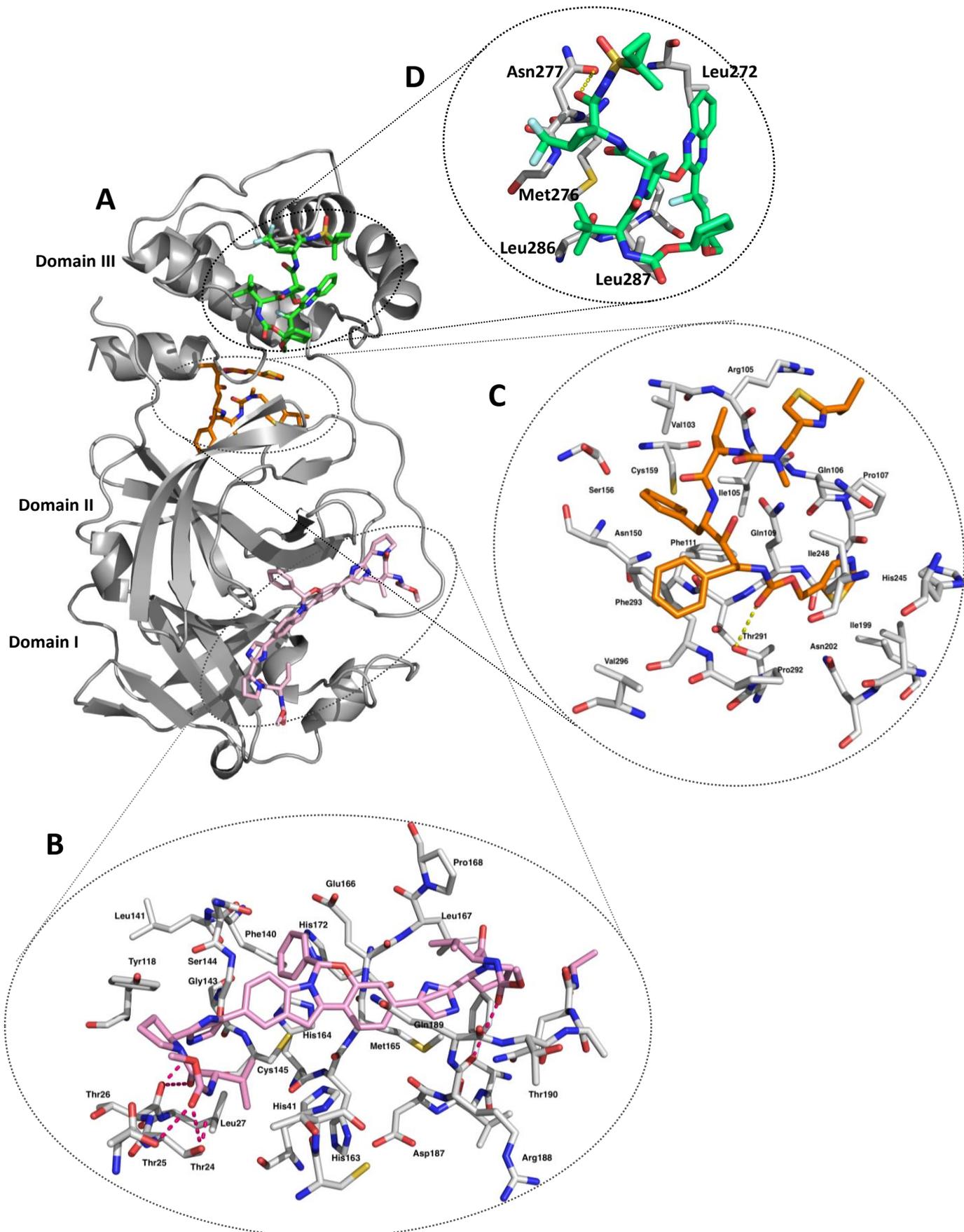
**Figure 1.** (A). The crystal structure of M<sup>Pro</sup> protease monomer consisting of N-terminal anti-parallel  $\beta$ -barrel forming catalytic domain and C-terminal all helical domain. The catalytic dyad His41 and Cys145 are shown in sticks. The catalytic loop (residues Cys44 to Pro52) and the linker Phe185 to Thr201 are shown in red and orange respectively. Highlighted in green are the structural regions of SARS-CoV-2 M<sup>Pr</sup>, which are different from the SARS-CoV M<sup>Pro</sup> protease. (B). Shows the sequence alignment of SARS-CoV-M<sup>Pro</sup> and CoV-2 protease along with their secondary structural regions generated using ESPrict webserver.



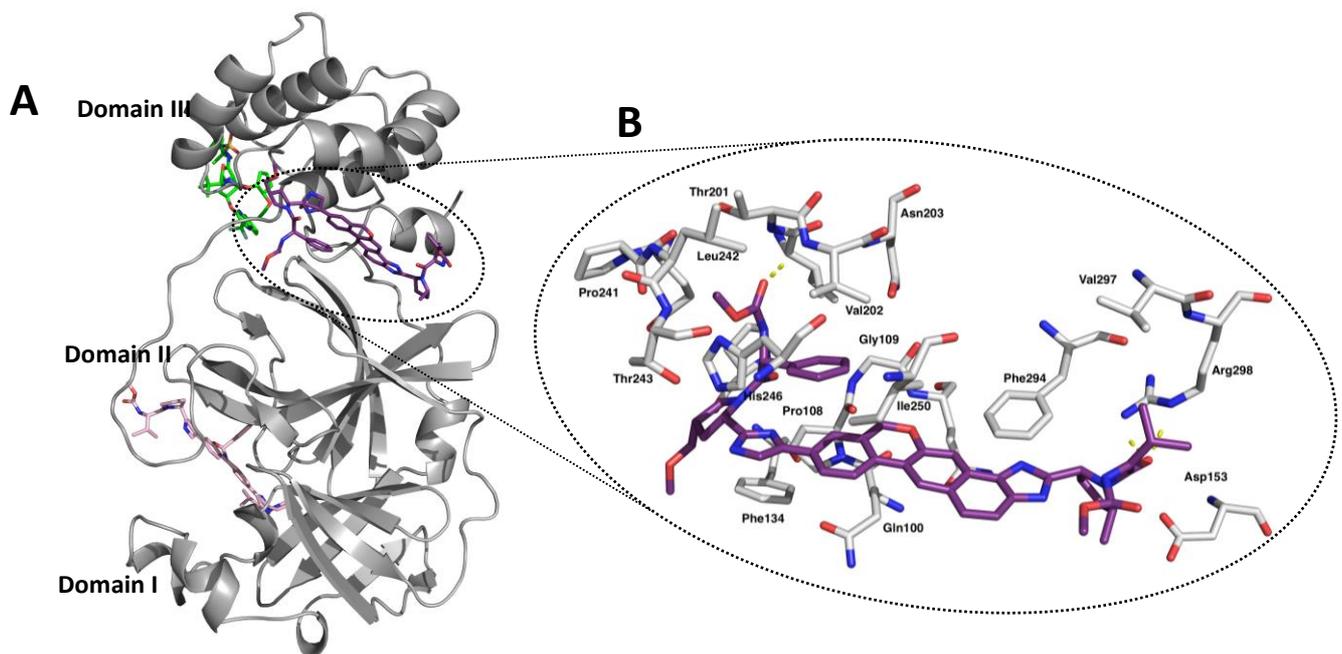
**Figure 2:** The docked poses of all top30 drugs of blind-docking, showing the other potential druggable sites.  $M^{PT0}$  is shown as cartoon and the docked drugs are shown in sticks.



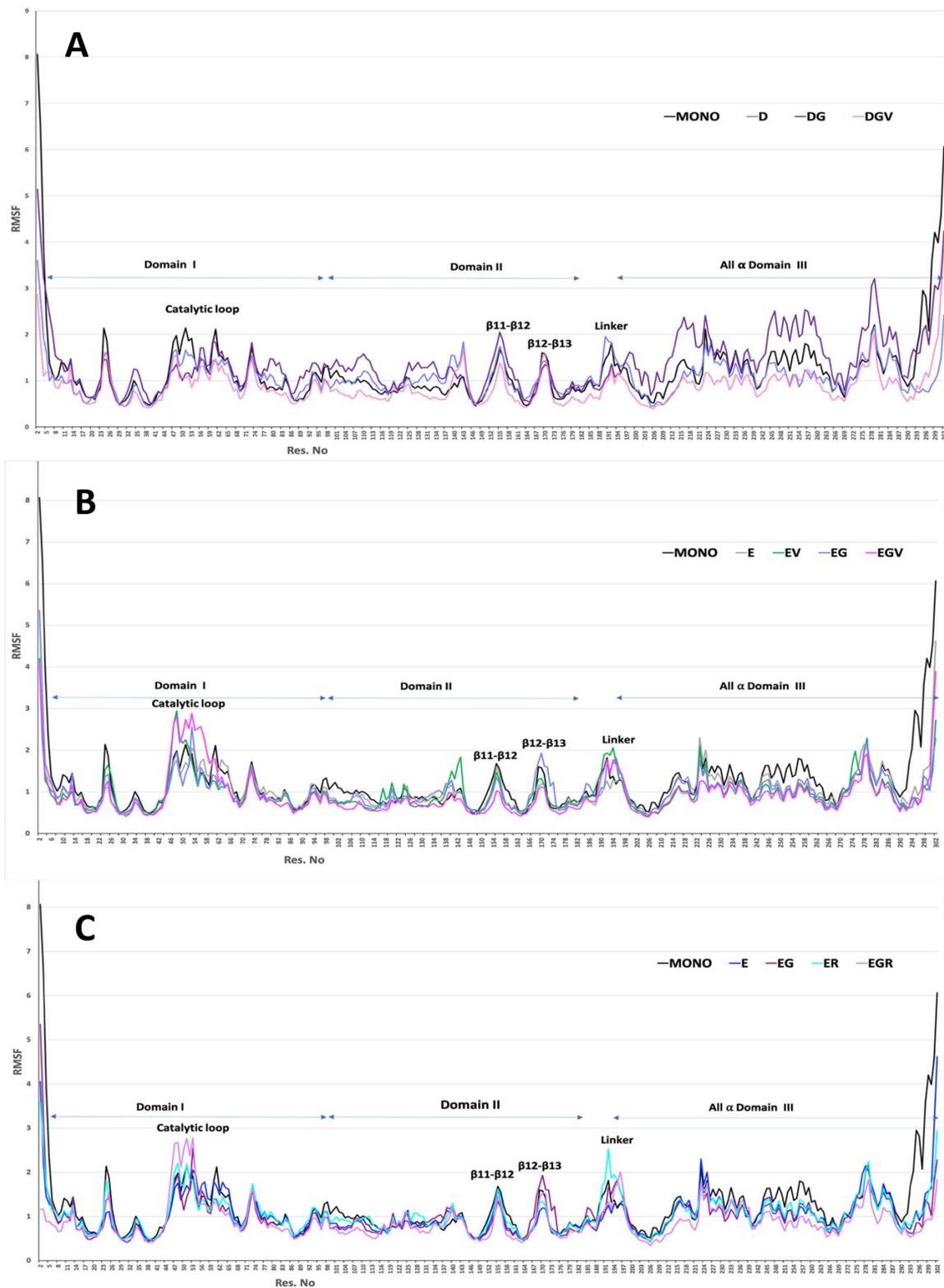
**Figure 3:** Shows the top three predicted pockets from CASTp.  $M^{\text{pro}}$  is represented as cartoon and the catalytic pocket is shown as green, the pocket2 near the dimeric interface in blue, and pink dots represent the pocket3 near the intra-domain interface.



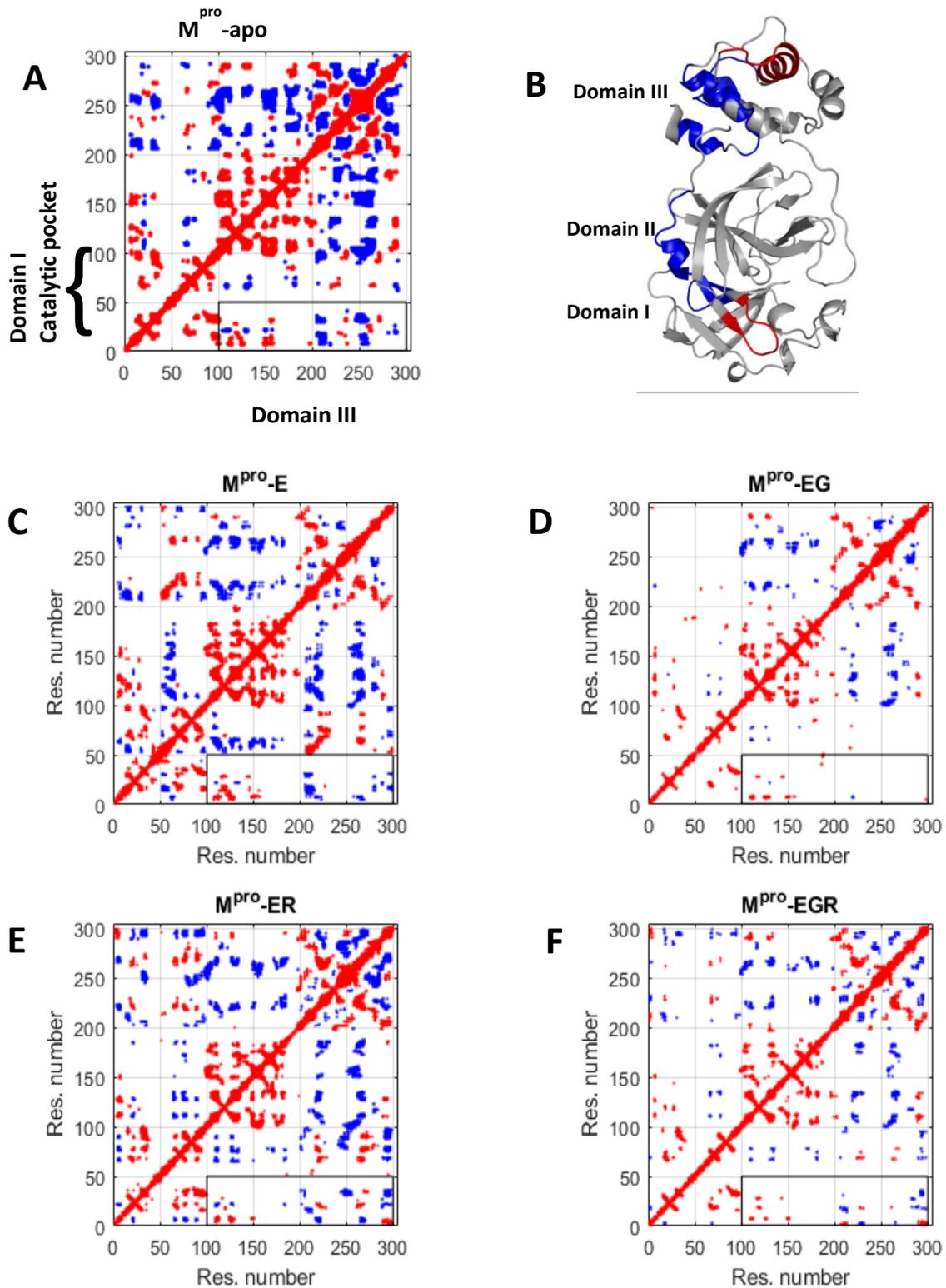
**Figure 4:** (A). The structure of  $M^{PRO}$  with three different drugs Elbasvir, Glecaprevir and Ritonavir at the end of 250ns MD simulations. (B). Shows the residues which contribute for the stability of Elbasvir at the catalytic pocket. (C). Shows the residues which contribute for the stability of Ritonavir at the pocket3. (D). Shows the residues which contribute for the stability of Glecaprevir at the pocket2.



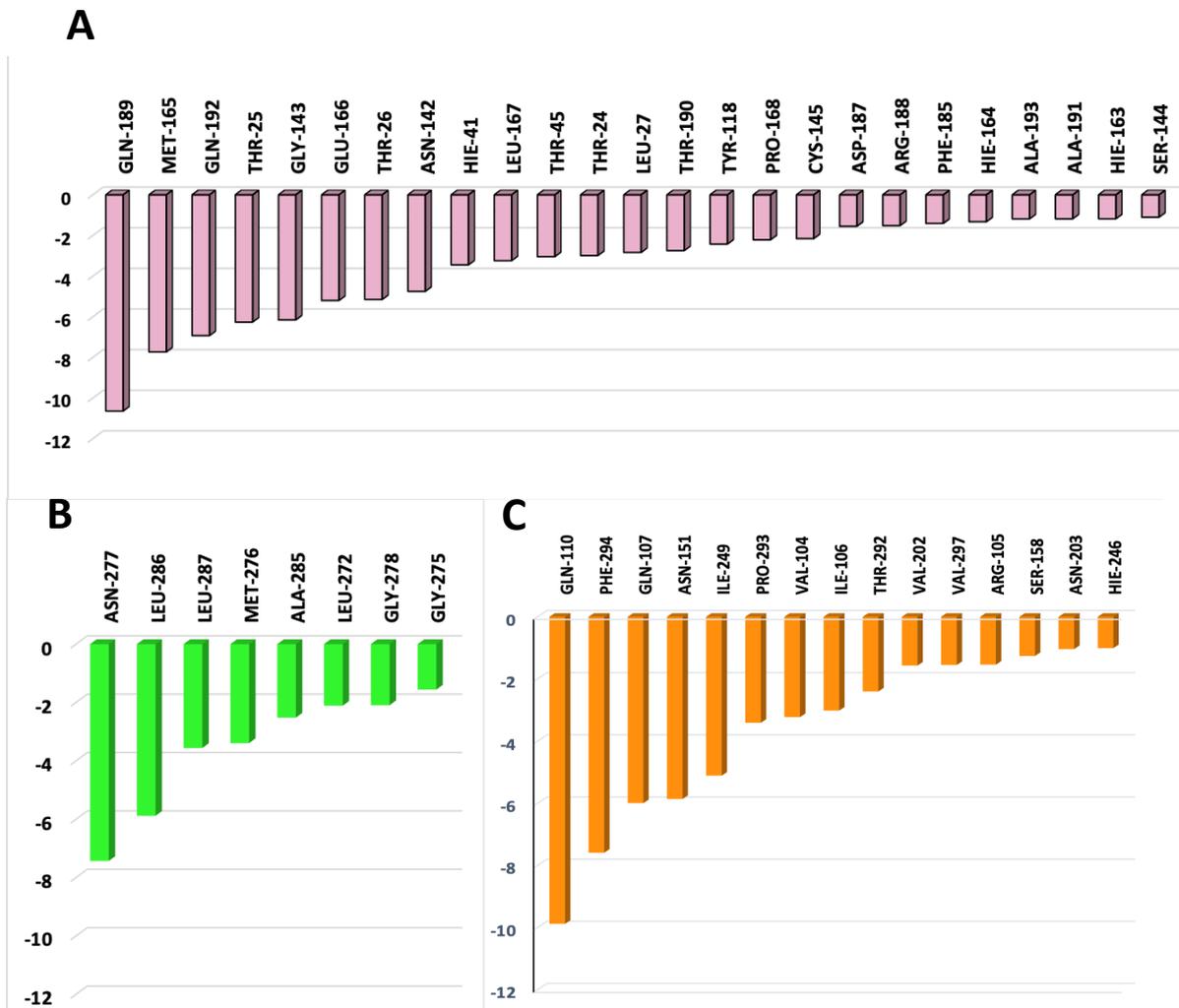
**Figure 5:** (A). The structure of  $M^{\text{PRO}}$  with three different drugs Elbasvir, Glecaprevir and Velpatasvir at their respective pockets at the end of 250ns MD simulations. (B). Shows the interaction of drug Velpatasvir with the residues of pocket3. Velpatasvir is shown in purple colour sticks and  $M^{\text{PRO}}$  residues are shown in grey sticks.



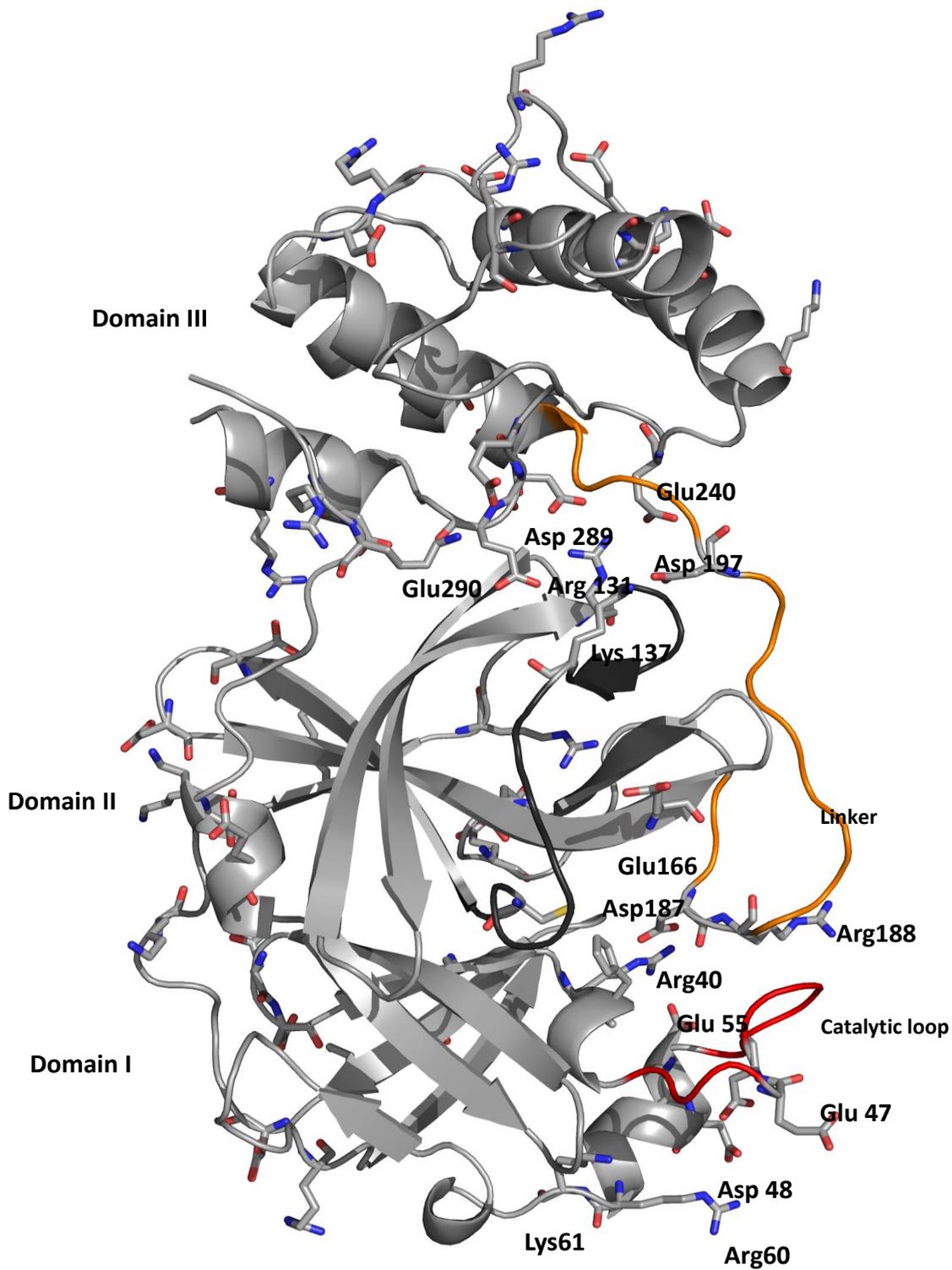
**Figure 6:** Root Mean Square fluctuations [RMSF] of C $\alpha$  atoms of M<sup>pro</sup>-drug complexes. (A). Shows the RMSF of M<sup>pro</sup>- single, two and three drug combinations of D, DG, DGV along with mono(apo)-M<sup>pro</sup> RMSF. (B). Shows the RMSF of M<sup>pro</sup>-single, two and three drug combinations of E, EG, EV, EGV along with mono (apo)-M<sup>pro</sup> RMSF. (C). Shows the RMSF of M<sup>pro</sup>-single, two and three drug combinations of E, EG, ER, EGR along with mono (apo) - M<sup>pro</sup> RMSF.



**Figure 7:** (A). The Dynamic Cross Correlation (DCCM) of  $C^\alpha$  atoms calculated for the snapshots taken for  $M^{\text{pro}}$ -apo simulations. The positive correlation values  $> 0.5$  are coloured in red, Negative correlated values  $< -0.5$  are coloured in blue for the structures taken from the simulation of  $M^{\text{pro}}$ -apo complex. (B). The positively (shown in red) and negatively correlated regions (shown in blue) of catalytic and the alpha helical domain are highlighted in the monomer of  $M^{\text{pro}}$  protease. (C-F) represents the DCCM of  $M^{\text{pro}}$  - E, EG, EV, EGR complexes respectively. Highlighted in box shows the reduced correlated motions; effect of drug in various pockets.



**Figure 8:** Per-residue contribution of  $M^{\text{pro}}$ -EGR complex. (A). Residues contributed for the binding of Elbasvir (ELB). (B). Residues contributed for the binding of Glecaprevir (GLE). (C). Residues contributed for the binding of Ritonavir (RIT).



**Figure 9.** The crystal structure of M<sup>Pro</sup> showing the presence of charged residues Lys, Arg, Glu and Asp. The catalytic loop is coloured in red and linker in orange and the loop forming the catalytic pocket ( $\beta$ 10-  $\beta$ 11) is coloured in black. The charged residues of catalytic loop, linker and catalytic pockets are shown in grey sticks.

**Table 1. Top 10 hits from the screening (targeted docking) of antiviral drugs against the SARS-CoV-2 M<sup>pro</sup> protease.**

S. No.	Drug Name	Known receptor	Binding Energy (Kcal/mol)	Toxicity
1.	Glecaprevir	HCV NS3/4A protease	-9.9	No genotoxicity shown in in vitro or in vivo studies.
2.	Paritaprevir	HCV NS3/4A protease	-9.0	Neither genotoxic nor carcinogenic.
3.	Danoprevir	HCV NS3/4A protease	-9.0	NA
4.	Elbasvir	HCV NS5A	-9.0	Fatigue, headache and nausea.
5.	BMS-955176	HIV maturation protein	-8.7	NA
6.	Grazoprevir	HCV NS3/4A protease	-8.6	Fatigue, headache and nausea.
7.	Simeprevir	HCV NS3/4A protease	-8.4	Fatigue, headache, pruritus, rash and elevated serum bilirubin.
8.	Saquinavir	HIV protease	-8.4	Probable pain in throat.
9.	Vaniprevir	HCV NS3/4A protease	-8.4	NA
10.	Odalasvir	HCV NS5A	-8.3	NA

**Table2A. Top hits for catalytic pocket from blind docking of Antiviral drugs against M<sup>pro</sup> Protease**

S.No	Drug Name	Known Receptor	Binding Site	Binding Energy (Kcal/mol)	Toxicity
1.	Danoprevir	HCV NS3/4A protease inhibitor	Catalytic site	-9.0	NA
2.	Elbasvir	HCV NS5A inhibitor	Catalytic site	-8.9	Fatigue, headache and nausea
3.	BMS-955176	HIV maturation protein inhibitor (Gag protein)	Catalytic site	-8.7	NA
4.	Maraviroc	HIV entry blocker by inhibiting HIV-CCR5 interaction	Catalytic site	-8.5	Liver toxicity
5.	Lopinavir	HIV protease inhibitor	Catalytic site	-8.0	Possible overdose mainly in pediatric patients can cause acute renal failure, cardiomyopathy and lactic acidosis.
6.	Indinavir	HIV protease inhibitor	Catalytic site	-7.9	Overdose symptoms include myocardial infarction and angina pectoris.
7.	Baloxavir marboxil	Influenza A and B, cap - endonuclease inhibitor	Catalytic site	-7.9	No mutagenic, carcinogenic, genotoxic side effects reported in <i>in vivo</i> studies.
8.	Nelfinavir	HIV protease inhibitor	Catalytic site	-7.8	Thirst, hunger, weight loss, frequent urination, fatigue, dry and itchy skin.
9.	Delaviridine	HIV, reverse transcriptase inhibitor	Catalytic site	-7.7	Rashes.
10.	Adafosbuvir	HCV	Catalytic site	-7.7	NA
11.	Raltegravir	HIV integrase inhibitor	Catalytic site	-7.7	NA

**Table2B. Top hits for pocket near dimeric interface from blind docking of antiviral drugs against M<sup>pro</sup> Protease**

S.No	Drug Name	Known Receptor	Binding Site	Binding Energy (Kcal/mol)	Toxicity
1.	Glecaprevir	HCV NS3/4A Protease	Near dimer Interface	-10.3	No genotoxicity shown in <i>in vitro</i> or <i>in vivo</i> studies.
2.	Odalasvir	HCV NS5A inhibitor	Near dimer Interface	-9.6	NA
3.	Paritaprevir	HCV NS3/4A protease inhibitor	Near dimer Interface	-9.3	Neither genotoxic nor carcinogenic.
4.	Simeprevir	HCV NS3/4A protease inhibitor	Near dimer interface	-9.2	Fatigue, headache, nausea, rash, pruritus and elevated serum bilirubin
5.	Ruzasvir	HCV NS5A inhibitor	Near dimer interface	-9.1	NA
6.	Pibrentasvir	HCV NS5A inhibitor	Near Dimer Interface	-8.9	Not shown to be genotoxic in <i>in vitro</i> or <i>in vivo</i> studies.
7.	Vaniprevir	HCV NS3/4A protease inhibitor	Near Dimer Interface	-8.8	NA
8.	L-756423	HIV protease inhibitor	Near dimer interface	-8.7	NA
9.	Beclabuvir	HCV NS5B inhibitor	Near dimer Interface	-8.6	NA
10.	Voxilaprevir	HCV NS3/4A protease inhibitor	Near dimer interface	-8.2	No effects on CNS, respiratory and cardiovascular parameters.
11.	Dasabuvir	HCV NS5B inhibitor	Near dimer interface	-8.1	Pruritus, nausea, insomnia and asthenia
12.	Bictegravir	HIV INST inhibitor	Near dimer interface	-8.1	Lactic acidosis and hepatotoxicity , diarrhea, nausea and headache.
13.	Ledipasvir	HCV NS5A inhibitor	Near dimer interface	-8.0	Headache and fatigue.
14.	Ombitasvir	HCV NS5A inhibitor	Near dimer interface	-7.8	Asthenia, fatigue, nausea, insomnia.

**Table2c. Top hits for intradomain interface pocket from Blind docking of Antiviral drugs against M<sup>pro</sup> Protease.**

S.No	Drug Name	Known Receptor	Binding Site	Binding Energy (Kcal/mol)	Toxicity
1.	Velpatasvir	HCV NS5A inhibitor	Intradomain interface	-8.5	No indication of carcinogenicity
2.	TMC-310911	HIV protease inhibitor	Intradomain interface	-8.2	NA
3.	Rupintrivir	Rhinovirus, 3C protease inhibitor	Intradomain interface	-7.8	NA
4.	Tecovirimat	Smallpox, viral p37 protein inhibitor	Intradomain interface	-7.8	Headache, abdominal pain, nausea and vomiting.
5.	Ritonavir	HIV protease inhibitor	Intradomain interface	-7.6	Overdose side effects include, renal failure, hepatotoxicity, pancreatitis and allergic reactions.
6.	Saquinavir	HIV protease inhibitor	Intradomain interface	-7.3	Probable pain in throat.

**Table 3. Summary of simulations of M<sup>pro</sup> -monomer-drug complexes**

S.No.	Simulated complex <sup>§</sup>	Binding pocket (residues <sup>#</sup> )			Summary
		Catalytic pocket	Dimeric interface Pocket2 (interacting res.)	Intra-domain Pocket 3 (interacting res.)	
1	M <sup>pro</sup> -apo	-	-	-	Very flexible
2	M <sup>pro</sup> -G	G	-	-	Unstable
3	M <sup>pro</sup> -D <sup>§</sup>	D	-	-	Unstable
4	M <sup>pro</sup> -E	E	-	-	Stable for a long time
5	M <sup>pro</sup> -DG	D	G	-	G moved out of the pocket
6	M <sup>pro</sup> -EG	E	G (199, 239, 271-276, 285-289)	-	E unstable
7	M <sup>pro</sup> -EV	E	V (108-111, 153, 154, 201-204, 293-295)	-	E unstable
8	M <sup>pro</sup> -ER	E	R (102-111, 151, 158, 200, 203, 292-298)	-	E mostly stable R unstable
9	M <sup>pro</sup> -DGV	D	G (237, 238, 273, 274)	V (108-111, 200-203, 240-242, 294)	D and G unstable
10	M <sup>pro</sup> -EGV	E	G (237, 239, 272-277, 285-287)	V (107, 111, 243-245, 249, 292-294)	Stable complex With G near helical domain
11	M <sup>pro</sup> -EGR	E	G (237, 272, 275-279, 285-287)	R (105-111, 151, 153, 158, 200, 202, 203, 292-294)	Stable complex with G near helical domain

<sup>#</sup>Residues within 4Å from the ligand in the final MD snapshot

<sup>§</sup>Simulation time was 100 ns for all but M<sup>pro</sup>-G and M<sup>pro</sup>-DG

\*Drugs used for simulation were: Elbasvir (E), Glecaprevir (G), Ritonavir (R) and Velpatasvir (V)

**Table 4. Binding energetics of M<sup>Pro</sup>-drug complexes, calculated for structures taken from respective MD simulations.**

	<b>D (unstable)</b>	<b>G</b>	<b>E</b>	<b>DG</b>	<b>EG</b>	<b>EV</b>	<b>ER</b>	<b>DGV</b>	<b>EGR</b>	<b>EGV</b>
<b>vdW</b>	-	-55.1	-89.7	-94.4	-124.8	-133.4	-138.9	-143.8	-185.5	-182.8
<b>E<sub>EL</sub></b>	-	-6.6	-30.3	-12.7	-47.8	-51.2	-26.8	-68.5	-64.1	-48.5
<b>E<sub>GB</sub></b>	-	32.6	50.9	66.9	85.3	100.6	77.4	127.3	107.0	127.2
<b>E<sub>SURF</sub></b>	-	-6.9	-10.0	-11.2	-14.1	-15.6	-16.2	-16.5	-21.8	-20.6
<b>G<sub>gas</sub></b>	-	-61.6	-120.0	-107.1	-172.6	-184.5	-165.7	-212.3	-249.6	-231.3
<b>G<sub>solv</sub></b>	-	25.7	40.9	55.6	71.2	85.0	61.2	110.8	85.3	106.6
<b>TOTAL</b>	-	<b>-35.9</b>	<b>-79.1</b>	<b>-51.4</b>	<b>-101.4</b>	<b>-99.5</b>	<b>-104.5</b>	<b>-101.4</b>	<b>-164.3</b>	<b>-124.7</b>
<b>Translational</b>	-	-13.7	-13.7	-14.2	-14.3	-14.3	-14.3	-14.6	-14.6	-14.7
<b>Rotational</b>	-	-12.0	-12.4	-13.7	-14.5	-14.5	-14.5	-15.2	-15.2	-15.2
<b>Vibrational</b>	-	-53.6	-56.1	-71.9	-74.7	-76.6	-76.6	-84.6	-83.9	-86.5
<b>TOTAL</b>	-	<b>-79.3</b>	<b>-82.3</b>	<b>-99.9</b>	<b>-103.5</b>	<b>-105.4</b>	<b>-105.4</b>	<b>-114.4</b>	<b>-113.8</b>	<b>-116.4</b>
<b>ΔG</b>	-	<b>43.4</b>	<b>3.2</b>	<b>48.4</b>	<b>2.0</b>	<b>5.9</b>	<b>0.8</b>	<b>13.0</b>	<b>-50.5</b>	<b>-8.3</b>

## The Table of Contents graphic

