

Targeting Novel Coronavirus 2019: A Systematic Drug Repurposing Approach to Identify Promising Inhibitors Against 3C-like Proteinase and 2'-O-Ribose Methyltransferase

Rameez Jabeer Khan^{1,§}, Rajat Kumar Jha^{1,§}, Gizachew Muluneh Amera¹, Monika Jain¹, Ekampreet Singh¹, Amita Pathak², Rashmi Prabha Singh³, Jayaraman Muthukumaran^{1,*}, Amit Kumar Singh^{1,*}

¹Department of Biotechnology, School of Engineering and Technology, Sharda University, P.C. 201310, Greater Noida, U.P., India

²Department of Chemistry, Indian Institute of Technology, Hauz Khas, New Delhi, 110016, India

³Department of Biotechnology, IILM College of Engineering & Technology, P.C. 201306, Greater Noida, U.P., India

[§]Authors Contributed equally

*Correspondence author E-mail address: amitk.singh@sharda.ac.in, j.muthukumaran@sharda.ac.in

Running Title: Identification of anti-viral drugs against 3CLpro and 2'-O-MTase from 2019- Novel Coronavirus

Abstract

The recent pandemic associated with 2019-nCoV, a virus of the Coronaviridae family, has resulted in an unprecedented number of infected people. The highly contagious nature of this virus makes it imperative for us to identify potential inhibitors from pre-existing antiviral drugs. Two druggable targets, namely 3C-like proteinase (3CLpro) and 2'-O-ribose methyltransferase (2'-O-MTase) were selected in this study due to their indispensable nature in the viral life cycle. 3CLpro is a cysteine protease responsible for the proteolysis of replicase polyproteins resulting in the formation of various functional proteins, whereas 2'-O-MTase methylates the ribose 2'-O position of the first and second nucleotide of viral mRNA, which sequesters it from the host immune system. The selected drug target proteins were screened against an in-house library of 123 antiviral drugs. Two promising drug molecules were identified for each protein based on their estimated free energy of binding (ΔG), the orientation of drug molecules in the active site and the interacting residues. The selected protein-drug complexes were then subjected to MD simulation, which consists of various structural parameters to equivalently reflect their physiological state. From the virtual screening results, two drug molecules were selected for each drug target protein [Paritaprevir ($\Delta G = -9.8$ kcal/mol) & Raltegravir ($\Delta G = -7.8$ kcal/mol) for 3CLpro and Dolutegravir ($\Delta G = -9.4$ kcal/mol) and Bictegravir ($\Delta G = -8.4$ kcal/mol) for 2'-OMTase]. After the extensive computational analysis, we proposed that Raltegravir, Paritaprevir, Bictegravir and Dolutegravir are excellent lead candidates for these crucial proteins and they could become potential therapeutic drugs against 2019-nCoV.

Keywords:

Drug repurposing; 2019-nCoV; 3C-like proteinase; 2'-O-ribose methyltransferase ; Docking; MD simulation

Abbreviations:

3CLpro, 3C-like proteinase; 2'-O-MTase, 2'-O-ribose methyltransferase; 2019-nCoV, 2019 Novel Coronavirus

1. Introduction

Members of the family Coronaviridae are enveloped and maintain a single-strand, positive-sense RNA genome ranging from 26 to 32 kb in length [1]. They can be classified into four genera: alpha, beta, delta, and gamma, out of which alpha and beta Coronavirus (CoVs) are known to infect humans [2]. They are circulated among humans, other mammals, and birds and can cause respiratory, enteric, hepatic, and neurologic diseases [3, 4]. Even though, the majority of human coronavirus infections are mild, the epidemics of two Betacoronaviruses (β CoV), namely Severe Acute Respiratory Syndrome coronavirus (SARS-CoV) and Middle East Respiratory Syndrome coronavirus (MERS-CoV), have caused more than 10,000 cumulative cases in the past two decades, with mortality rates of 10% for SARS-CoV and 37% for MERS CoV [5, 6]. In late December 2019, several local health facilities reported groups of patients with pneumonia of unknown cause, and they were found to be epidemiologically connected with the Huanan wholesale seafood market in Wuhan, Hubei province, China[7]. Next-generation sequencing analysis from bronchoalveolar lavage fluid samples and cultured isolates from the patients indicated a novel human-infecting coronavirus[7], provisionally named 2019 novel coronavirus (2019-nCoV).

As of Feb 16, 2020, China has reported more than 72,000 confirmed cases of 2019-nCoV infection across 30 Chinese provinces, with 1750 fatalities. Besides, 2019-nCoV has also been reported in Thailand, Japan, Singapore, South Korea, Malaysia, Hong Kong, Australia, Germany, Taiwan, Macau and various other countries [8]. On Jan 30, 2020, the World Health Organisation (WHO) declared 2019-nCoV pandemic as a Public Health Emergency of International Concern [9]. In order to control the menace caused by 2019-nCoV, we suggest two of its indispensable druggable targets, namely 3C-like proteinase and 2'-O-ribose methyltransferase [10, 11]. 3C-like proteinase (3CLpro) is a cysteine protease present in the Coronavirus replicase polyprotein. This protease plays a central role in viral replication and transcription functions through extensive proteolysis of two replicase polyproteins, pp1a and pp1ab [10]. It cleaves at least 11 inter-domain sites on the pp1a and pp1ab polyproteins to generate individual functional proteins, including an RNA-directed RNA polymerase, a helicase, an exoribonuclease, an endoribonuclease, and a 2'-O-ribose methyltransferase. Also, it is excised from polyproteins by its proteolytic activity and forms a homodimer with one active site per subunit [10]. All of these features of 3CLpro makes it a principal drug target for 2019-nCoV.

Another crucial protein responsible for viral replication and expression in host cells is non-structural protein 16 (nsp16) or 2'-OMTase. Most viral mRNAs possess a 5'-terminal cap structure (m⁷GpppN) which is essential for efficient splicing, nuclear export, translation and stability [12]. This structure undergoes methylation catalysed by 2'-O-ribose methyltransferase at the ribose 2'-O position of the first and second nucleotide of the mRNA[12, 13]. Nsp 16 encodes the 2'-OMTase which provides the viral mRNA with the ability to camouflage and obscure itself from the host cell, thus preventing recognition and activation of the host immune response which is essential for successful viral infection [14]. This protein can, therefore, act as another potential drug target for the 2019-nCoV.

The present study utilises the systematic drug repurposing approach to identify antiviral drugs which can act as promising inhibitors against 3CLpro and 2'-OMTase and also understand their inhibitory mechanism through extensive *in silico* approach. We have employed molecular docking and molecular dynamics (MD) simulation

studies to calculate various structural parameters including the estimated binding free energy (ΔG) of the drugs, their estimated inhibition constant (K_i), Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), Radius of Gyration (R_g), Solvent Accessible Surface Area (SASA), Principal Component Analysis (PCA) and the intermolecular hydrogen bonds (H-bonds) for free and drug bounded protein molecules.

2. Material and Methods

2.1 Homology modelling and model evaluation

The amino acid sequence of 2019-nCoV orflab polyprotein was retrieved from NCBI (Tax id: 2697049, Accession No: QHD43415). It was then subjected to pairwise sequence alignment against 3CLpro and 2'-OMTase from its close relative SARS-CoV (tax id: 694009) to figure out their amino acid sequence in 2019-nCoV. Homology modeling of these proteins was performed using SWISS-MODEL [15] with the user template option. Template structures were selected from Protein BLAST [16] against Protein Data Bank (PDB) [17] based on query coverage, percentage of sequence identity and crystal structure resolution. Energy minimisation of modeled proteins was achieved using YASARA [18] and GalaxyWEB [19]. Model evaluation was performed before and after minimisation using Structure Analysis and Verification Server (SAVES) v5.0 meta server [20], which includes ERRAT [21], Verify-3D [22] and PROCHECK [23] programs.

2.2 Screening and selection of Antiviral Drugs against 3CLpro and 2'-OMTase

PyRx virtual screening software v0.8 (<http://pyrx.sourceforge.net/downloads>) was used for the screening of antiviral drugs against 3CLpro and 2'-OMTase proteins. One hundred twenty-three entries belonging to the category of antiviral drugs were downloaded from DrugBank database [24] and were subsequently converted into a single structure data file (SDF) library (see supplementary data, **Table 3**). AutoDockVina [25] docking wizard inbuilt in PyRx was utilised for all of the docking calculations. In PyRx, blind docking with the exhaustiveness 32 option was used to carry out the virtual screening and docking of antiviral drugs against 3CLpro (Grid Box: center_x = 75.98, center_y = -15.53, center_z = 18.01, size_x = 70.46, size_y = 90.53, size_z = 61.34) and 2'-OMTase (Grid Box: center_x = 64.12, center_y = 70.35, center_z = 71.46, size_x = 67.11, size_y = 66.13, size_z = 68.36). Once we got the results obtained from virtual screening, we selected two drug molecules for each protein based on the estimated binding free energy (ΔG), estimated inhibition constant (K_i), their orientation at the catalytic site and the interacting residues. The selected protein-drug complexes were further subjected to MD simulations for understanding the structural stability of protein-drug complexes at the long-interval.

2.3 Molecular dynamics (MD) simulations

Molecular docking studies is based on semi-flexible approach which considers ligand as a flexible entity and protein as a rigid entity, which is one of the main limitations of this study. As a result, docking calculations might not equivalently reflect the physiological state of protein or protein-ligand complexes. Thus, to make the docking calculations statistically significant, the best protein-drug complexes were subjected to MD simulation in an implicit solvent model. MD simulation was performed using GROMACS 2019 [26] with GROMOS96 43a1 force

field parameters [27]. The topology of the drug molecules was created using PRODRG webserver [28]. The simulation of all of the proteins and protein-drug complexes were run for a period of 100 ns. To make the system electrostatically neutral, counter ions were added to the protein-drug complexes. The complexes were solvated within 10 Å SPC/E water cube [29]. The protein-drug complexes were minimised in multiple steps using steepest descent method, where minimisation of the whole system, water cube and non-heavy atoms of the complexes were accomplished. The entire systems were then progressively heated up to 300 K on a time scale of 100 ps. The equilibration steps were performed in two different phases, one with constant pressure and temperature (NPT) and the other with steady volume and temperature (NVT). Various structural parameters, like Root Mean Square Deviation (RMSD), Root Mean Square Fluctuations (RMSF), Radius of Gyration (Rg), Principal Component Analysis based on essential dynamics approach, Inter-molecular Hydrogen Bonding (H-bonding) and Solvent-Accessible Surface Area (SASA) were calculated as a function of time to explore the structural behaviour of the proteins and protein-drug complexes.

3. Results and discussion

3.1 Homology Modeling and Model Evaluation

The 3D models of 3CLpro and 2'-OMTase were predicted using SWISS-MODEL [15]. For the best results, we used the user template option present in SWISS-MODEL. Protein BLAST [16] results were taken into account for the selection of suitable template structures. The X-ray crystal structure of 3CLpro from Human SARS coronavirus (SARS-CoV) solved at 1.9 Å (PDB ID: 1UJ1) was selected as the template structure for 2019-nCoV 3CLpro. For 2'-OMTase, we used a 2 Å X-ray crystal structure of nsp10/nsp16 complex (PDB ID: 3R24) of SARS-CoV. The template 1UJ1 was selected because it has 96.08% sequence identity and 100% query coverage towards 2019-nCoV 3CLpro. Also, the template 3R24 has a sequence identity of 93.29% and query coverage of 100% against 2019-nCoV. In SWISS-MODEL, Global Model Quality Estimation (GMQE) [15] score of 0.99 out of 1 and QMEAN [11] score of 0.09 was achieved for 3CLpro, whereas for 2'-OMTase, GMQE score was 0.97 out of 1, and QMEAN score was -2.57. These quality estimations depict the reliability of the modeled protein structure, target-template alignment and the template search method. It is also worth noting that 3CLpro makes a homodimeric structure, but since only one protomer is active in each dimer, we only modeled the single protomer, and further used it for *in silico* interaction and MD simulation studies[30].

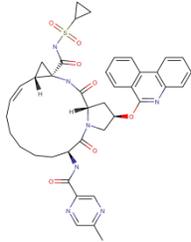
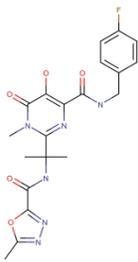
Once the models were obtained, they were evaluated using SAVES v5.0 metaserver based on ERRAT, Verify 3D and PROCHECK. The models were further refined and minimised using YASARA Energy Minimization Server and GalaxyWEB server. The final model of 3CLpro had a Verify 3D score of 100% and ERRAT quality factor of 98.19. In the case of PROCHECK, 93.9% residues were present in the most favoured regions of Ramachandran Plot [31], and more importantly, only one residue (Tyr154) was present in the disallowed regions. In the case of 2'-OMTase, the final model had a Verify 3D score of 87.33%, and ERRAT quality factor of 94.50. According to PROCHECK, 90% of the residues were in the most favoured regions of the Ramachandran plot, and notably, no residues were present in the disallowed regions. We used PDBsum [32] to calculate the overall average

G-Factor score, which is 0.12 for both 3CLpro and 2'-OMTase, which also reflected the accuracy of the refined modeled structure (see supplementary data, **Table 1-2**).

3.2 Screening of anti-viral Drugs against 3CLpro and 2'-OMTase

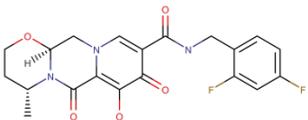
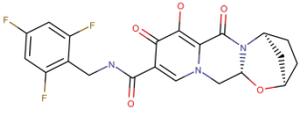
PyRx was used for the virtual screening of antiviral drugs against 3CLpro and 2'-OMTase proteins. One hundred twenty-three compounds belonging to the category of antiviral drugs were downloaded from DrugBank [24] and were subsequently converted into a single SDF library. This in-house SDF library was used to screen both proteins. Two drugs were selected against each protein based on their estimated free energy of binding (ΔG), binding site orientation and the protein residues with which they were found to be interacting with drug molecules. In the case of 3CLpro (see **Table 1**), we chose two molecules, namely Paritaprevir (DB09297) and Raltegravir (DB06817). Similarly, for 2'-OMTase (see **Table 2**), the best two drugs were found to be Dolutegravir (DB08930) and Bictegravir (DB11799).

Table 1: Selected drug molecules obtained from virtual screening of 3CLpro against the in-house library of antiviral drugs

S/No	Drug Bank ID	EBE (kcal/mol)	EIC (M)	Molecular formula	Chemical scheme
1	Paritaprevir (DB09297)	-9.8	65.52 nM	C ₄₀ H ₄₃ N ₇ O ₇ S	
2	Raltegravir (DB06817)	-7.8	1.916 μ M	C ₂₀ H ₂₁ FN ₆ O ₅	

EBE: Estimated Binding Free Energy; EIC: Estimated Inhibition Constant

Table 2: Selected drug molecules obtained from virtual screening of 2'-OMTase against the in-house library of antiviral drugs

S/No	Drug Bank ID	EBE (kcal/mol)	EIC(M)	Molecular formula	Chemical scheme
1	Dolutegravir (DB08930)	-9.4	128.7 nM	C ₂₀ H ₁₉ F ₂ N ₃ O ₅	
2	Bictegravir (DB11799)	-8.4	0.696 μM	C ₂₁ H ₁₈ F ₃ N ₃ O ₅	

EBE: Estimated Binding Free Energy; EIC: Estimated Inhibition Constant.

3.3 Binding Pattern of Drug Molecules against 3CLpro and 2'-OMTase

3.3.1 nCoV 3CLpro

As SARS-CoV 3CLpro (PDB ID: 1UJ1), the predicted model also forms a homo-dimer with both protomers (signified as "A" and "B") oriented nearly at right angles to each other (see supplementary data, **Fig. 1**). Each monomer of the modeled structure is composed of three distinct globular domains (see **Fig. 1**). Domain I comprises residues 8-101, domain II includes residues 102-184, and domain III consists of residues 201-303 and is attached with domain II through an extended loop region (residues 185-200). Domain I and Domain II form a chymotrypsin fold, whereas domain III is required for the homodimer assembly. The substrate-binding site is positioned in the cleft of domain I and domain II and is composed of a Cys145-His41 catalytic dyad. The substrate-binding pocket consists of His41, Phe140, Asn142, Gly143, Ser144, Cys145, Tyr161, His163, Glu166, and His172 [10].

The estimated binding free energy (ΔG) of Paritaprevir against 3CLpro was observed to be -9.8 kcal/mol, and the estimated inhibitory constant was 65.52 nM. On the other hand, the binding affinity (ΔG) of Raltegravir was measured to be -7.8 kcal/mol, and the inhibitory constant was 1.916 μM. Both Paritaprevir and Raltegravir were found to be interacting with the active site residues (including the Cys145-His41 catalytic dyad) (see **Table 3**) and interacting with the substrate-binding pocket (see **Fig. 1**).

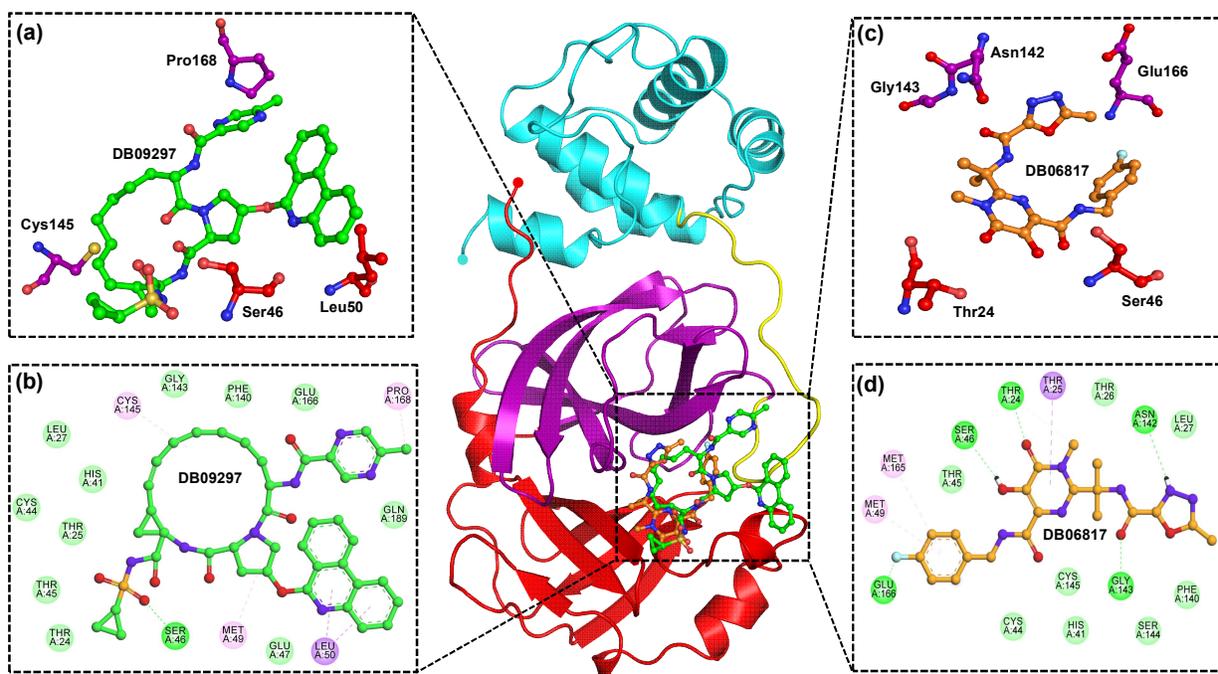


Fig. 1 Interaction of drugs with 3CLpro (Domain I-Red, Domain II-Purple, Domain III-Cyan, Extended loop-Yellow) (a) 3D representation of 3CLpro active site residues interacting with Paritaprevir (DB09297) (Green) (b) 2D representation of 3CLpro active site residues interacting with Paritaprevir (Green) via Van der Waals interactions (slightly green colour), hydrogen bonds (dark green colour), and pi-interactions (light pink colour) (c) 3D representation of 3CLpro active site residues interacting with Raltegravir (DB06817) (Orange) (d) 2D representation of 3CLpro active site residues interacting with Raltegravir (Orange).

Table 3. Interaction details of Paritaprevir and Raltegravir with 3CLpro

S/No	DrugBank ID	HB	D (Å)	Pi-SR	D (Å)	vdWISR
1.	Paritaprevir (DB09297)	Ser46	2.33	Met49 Leu50 Cys145 Pro168	5.27 3.85, 4.61 4.60 3.64	Thr24, Thr25, Leu27, His41, Cys44, Thr45, Glu47, Phe140, Gly143, Glu166, Gln189
2.	Raltegravir (DB06817)	Thr24 Ser46 Asn142 Gly143 Glu166	2.53 2.47 2.39 2.11 2.59	Thr25 Met49 Met165	3.93 4.87 5.14	Thr26, Leu27, His41, Cys44, Thr45, Phe140, Ser144, Cys145

Hydrogen Bond residues (HB), Distance (D), Pi-Interaction Sharing Residues (Pi-SR), and van der Waals Interaction Sharing Residues (vdWISR)

3.3.2 nCoV 2'-OMTase

The predicted model of 2'-O methyltransferase (2'-OMTase) (see **Fig. 2**) comprises of a core MTase domain, S-Adenosyl methionine (SAM) binding region and a cap-binding groove. The core MTase domain is found to comprise residues 30-209 of nsp16 as found in SARSCoV [12]. 2'-OMTase catalyses the transfer of SAM methyl group by relying on a conserved K-D-K-E tetrad within the substrate-binding pocket for activity [33]. The surface provided by K-D-K-E (Lys-46, Asp-130, Lys-170 and Glu-203) motif present at the bottom of the central groove might bind the first adenine nucleotide conserved as the acceptor of methyl group during methylation [12].

In the case of 2'-OMTase, Dolutegravir showed an estimated binding free energy (ΔG) of -9.4 kcal/mol, and the inhibition constant was 128.7 nM. Besides, Bictegravir conferred an estimated binding free energy (ΔG) of -8.4 kcal/mol against 2'-OMTase, and the estimated inhibition constant was 0.696 μ M. Moreover, both Dolutegravir and Bictegravir were found to be actively interacting with the catalytic site residues (see **Table 4**) and interacting with the substrate-binding pocket (see **Fig.2**).

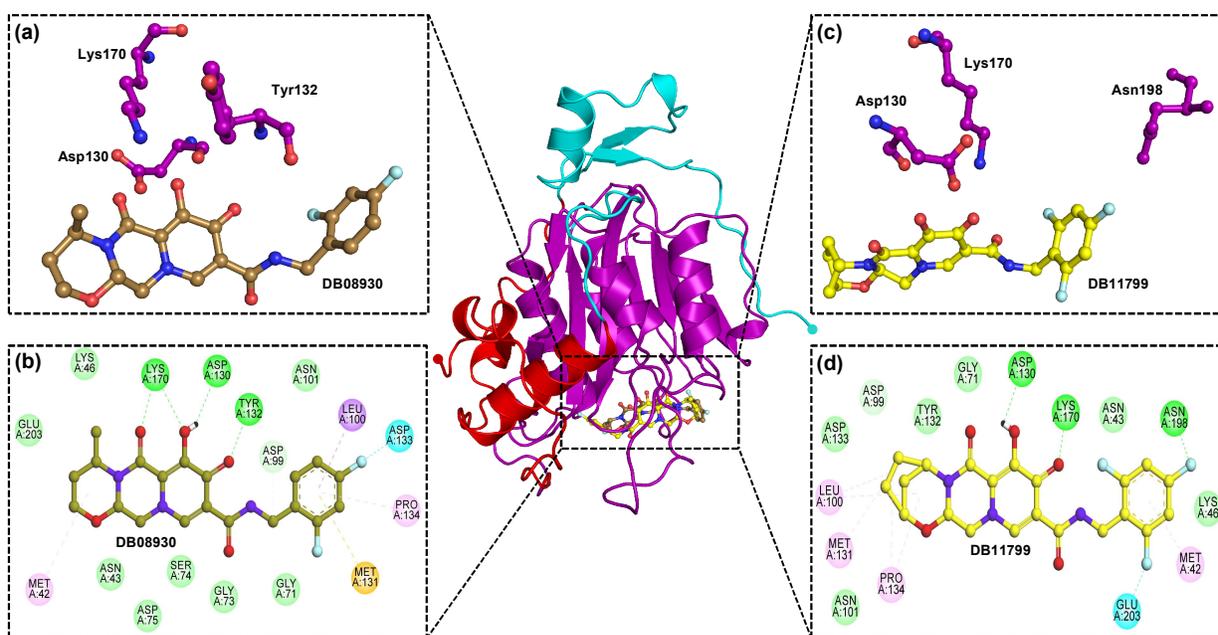


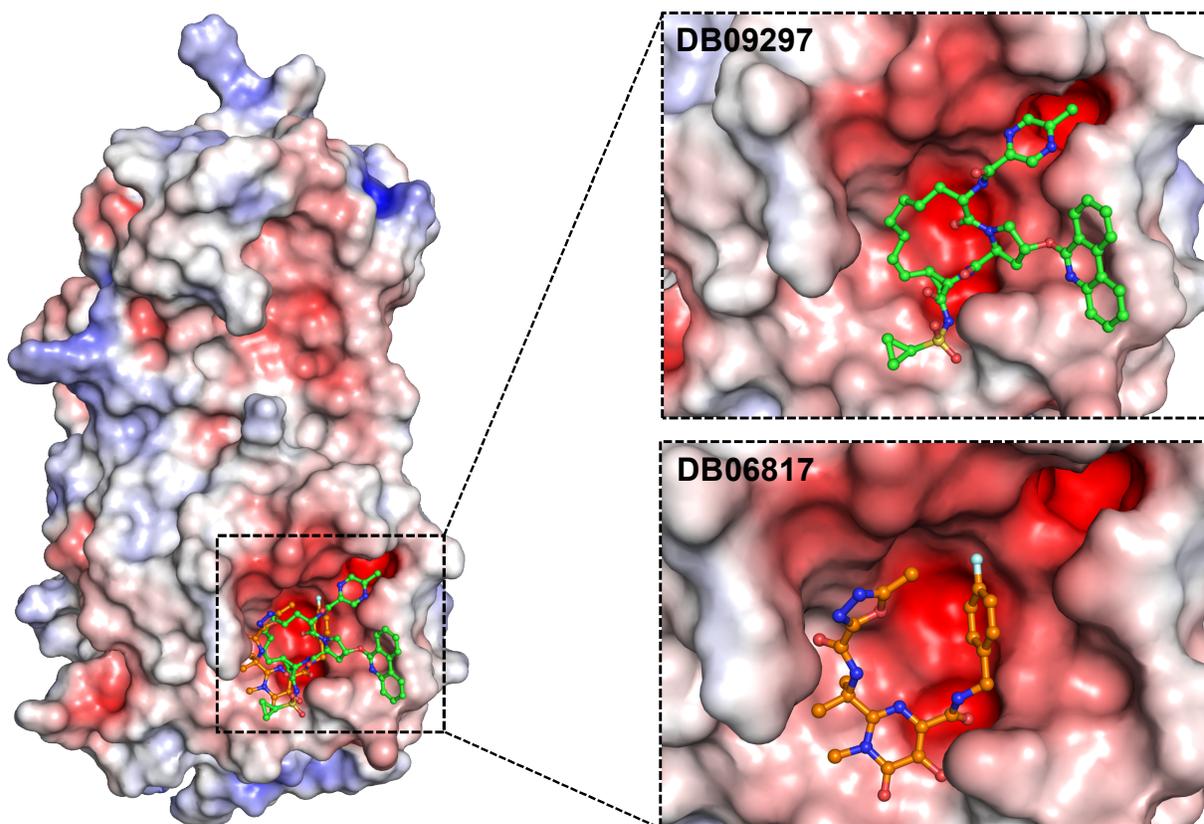
Fig. 2 Interaction of drugs with 2'-OMTase (MTase Domain-Purple) (a) 3D representation of 2'-OMTase active site residues interacting with Dolutegravir(DB08930) (Sand colour) (b) 2D representation of 2'-OMTase active site residues interacting with Dolutegravir (Sand colour) via Van der Waals interactions (slightly green colour), hydrogen bonds (dark green colour), and pi-interactions (light pink colour, orange colour) (c) 3D representation of 2'-OMTase active site residues interacting with Bictegravir (DB11799) (Yellow) (d) 2D representation of 2'-OMTase active site residues interacting with Bictegravir (Yellow).

Table 4. Interaction details of Dolutegravir and Raltegravir with 2'-OMTase

S/No	DrugBank ID	HB	D (Å)	Pi-SR	D (Å)	vdWISR
1	Dolutegravir (DB08930)	Asp130	2.45	Met42	5.45	Asn43, Lys46, Gly71,
		Tyr132	2.94	Leu100	3.49	Gly73, Ser74, Asp75,
		Lys170	1.95,	Met131	5.83	Asp99, Asn101, Glu203
			2.55	Pro134	4.73	
2	Bictegravir (DB11799)	Asp130	2.43	Met42	4.34	Asn43, Lys46, Gly71,
		Lys170	2.28	Leu100	4.46,	Asp99, Asn101, Tyr132,
		Asn198	2.15		5.02	Asp133
				Met131	4.73	
				Pro134	5.12,	
					5.17	

Hydrogen Bond residues (HB), Distance (D), Pi-Interaction Sharing Residues (Pi-SR), and van der Waals Interaction Sharing Residues (vdWISR)

According to the results obtained from electrostatic surface potential, all of the drug molecules were found interacting mostly with the negatively charged residues of both 3CLpro (see Fig. 3) and 2'-OMTase (see Fig. 4), which revealed that these molecules were oriented in the active site of both the proteins.

**Fig. 3** The electrostatic surface potential interaction of Paritaprevir (DB09297) (Green) and Raltegravir (DB06817) (Orange) bound to 3CLpro. The zoomed view is representing the active site cleft.

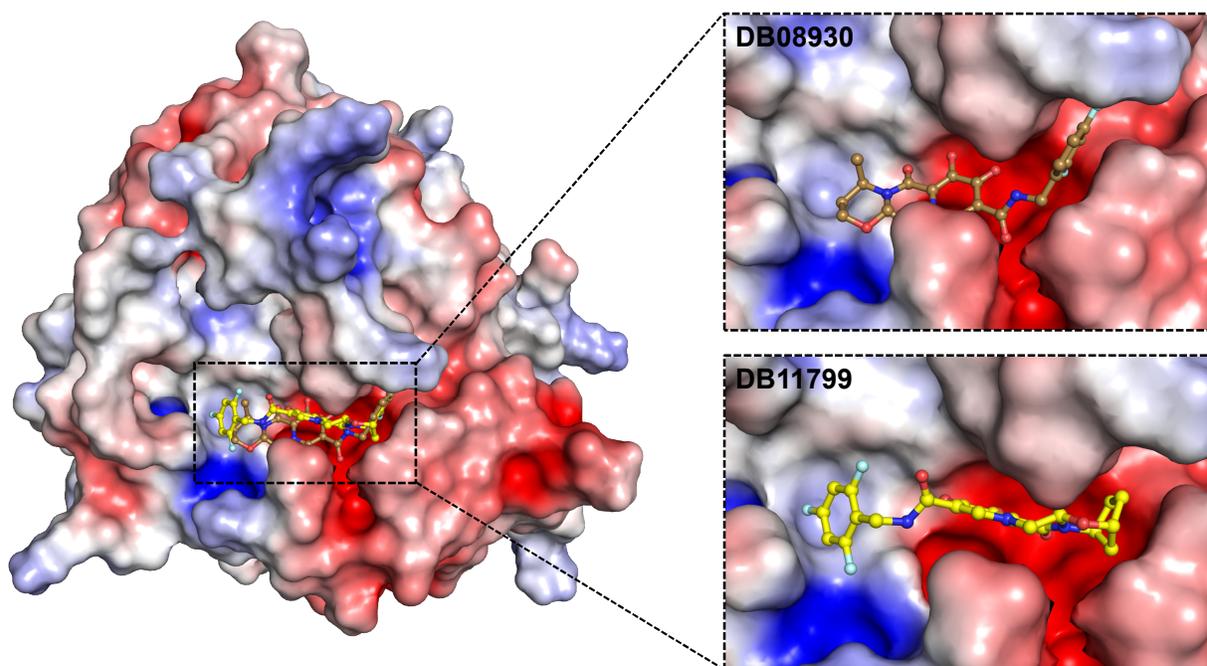


Fig. 4 The electrostatic surface potential interaction of Dolutegravir (DB08930) (Sand) and Bictegravir (DB11799) (Yellow) bound to 2'-OMTase. The zoomed view is representing the active site cleft.

3.4 Molecular dynamics (MD) Simulation

Molecular dynamics simulation is used to study the physical transitions of atoms as well as entire macromolecules and are effectively adopted to comprehend the structure-to-function relevance of a given macromolecule. Moreover, it can also be employed to illustrate the strength, pattern, and properties of protein-drug interactions and the dynamic conformational changes a macromolecule experiences under various physiological conditions. Different structural parameters, including RMSD, RMSF, Rg, PCA (based on ED approach), Intermolecular H-bonding and SASA, were evaluated as a function of time. To elucidate the conformational stability, dynamics, structural stability, folding properties and compactness of protein-drug complexes, MD simulations were performed for a period of 100 ns.

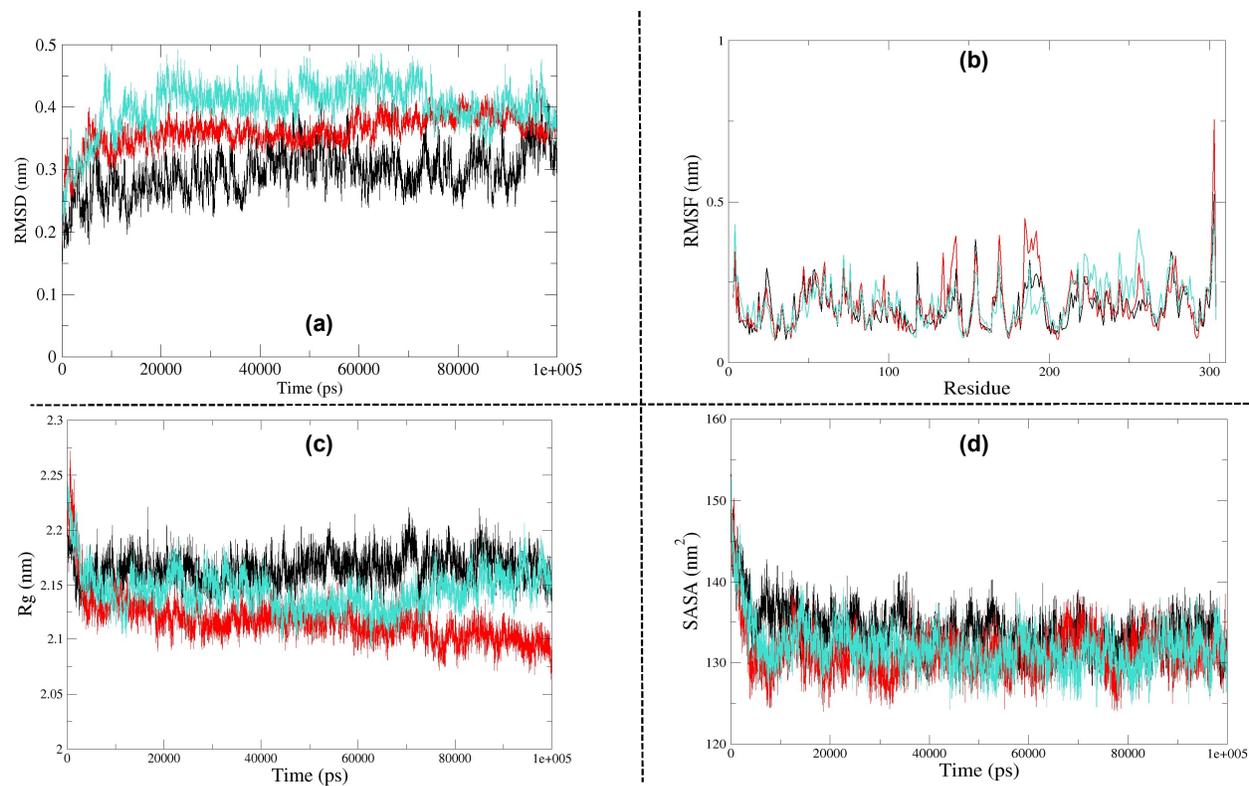


Fig. 5 Analysis of Molecular Dynamics Simulation results of free 3CLpro (Black), 3CLpro-Raltegravir complex (Red) and 3CLpro-Paritaprevir complex (turquoise) (a) Root Mean Square Deviation (RMSD) (b) Root Mean Square Fluctuation (RMSF) (c) Radius of Gyration (Rg) (d) Solvent Accessible Surface Area (SASA).

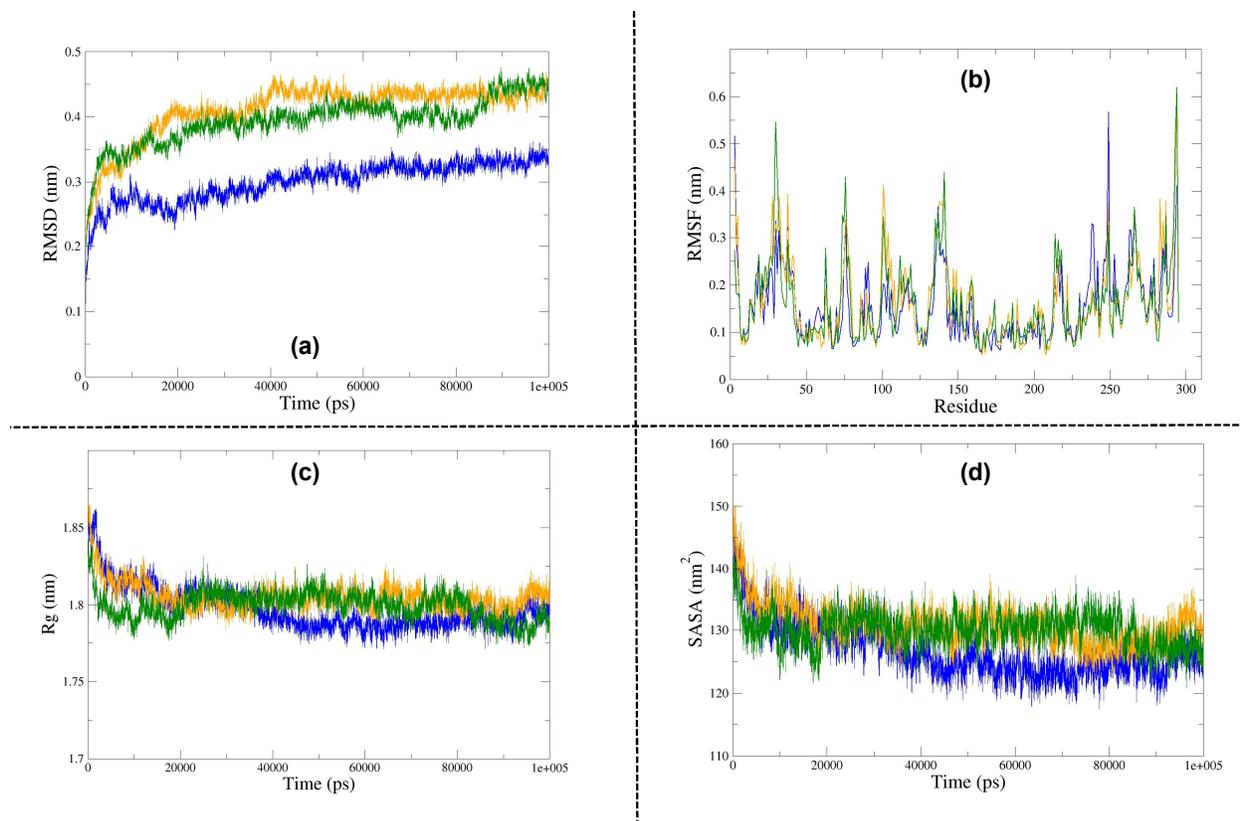


Fig. 6 Analysis of Molecular Dynamics Simulation results of free 2'-OMTase (Blue), 2'-OMTase-Bictegravir complex (Orange) and 2'-OMTase-Dolutegravir complex (Dark Green) (a) Root Mean Square Deviation (RMSD) (b) Root Mean Square Fluctuation (RMSF) (c) Radius of Gyration (Rg) (d) Solvent Accessible Surface Area (SASA).

3.4.1 RMSD Analysis

RMSD is essential to quantify the structural stability of a protein or protein-drug complexes within a regular time frame. RMSD analysis depicted that unbound 3CLpro started stabilising after 15 ns, and it maintained stability until 90 ns. The RMSD pattern varied slightly during the final 10 ns. Besides, the 3CLpro-Raltegravir complex also attained stability at around 10 ns, and it maintained it until 90 ns. The RMSD again decreased during the final 10 ns (90-100 ns). In contrast, 3CLpro-Paritaprevir complex gained stability at around 20 ns, retained that stability until 70 ns except for one minor dip at about 50 ns. After 70 ns, the average RMSD declined until the end of the simulation cycle (see **Fig. 5a**). An average RMSD value of 0.295 nm, 0.361 nm and 0.401 nm was obtained for unbound 3CLpro, 3CLpro-Raltegravir complex and 3CLpro-Paritaprevir complex, respectively. Overall results explained that these two drug molecules did not significantly influence the structural stability of 3CLpro, although they nearly maintained the structural integrity, particularly Raltegravir. Based on the average RMSD values obtained from this analysis, the 3CLpro-Raltegravir complex was found to be closer to the drug-free form of 3CLpro.

RMSD analysis of the free form of 2'-OMTase depicted that the protein molecule achieved stability at around 40 ns, and it almost maintained its stability until 100 ns. The maximum RMSD fluctuations were recorded in

the first 20 ns only. There were minor variations at around 60 ns, but apart from that, the protein was quite stable until the end of the simulation cycle. The 2'-OMTase-Dolutegravir complex attained stability at around 20 ns and maintained that stability until 90 ns, except a minor drop at approximately 70 ns. The average fluctuations spiked a little bit in the final 10 ns (90-100 ns) of this simulation. The 2'-OMTase-Bictegravir complex also took 40 ns to attain stability, but it also maintained it until the end of the simulation cycle (40-100 ns). It showed a minor dip at 60 ns, but it was able to gain its stability back at around 65 ns. As the free protein, the maximum fluctuations of protein-drug complexes were recorded in the first 20 ns only, after that, all of them took a steady path (see **Fig. 6a**) and maintained convergence state until the end of the simulation period. However, in the case 2'-OMTase-Dolutegravir complex, slight changes were observed at 90-100 ns. An average RMSD value of 0.299 nm, 0.393 nm and 0.410 nm was obtained for unbound 2'-OMTase, 2'-OMTase-Dolutegravir complex and 2'-OMTase-Bictegravir complex, respectively. As evident from the average RMSD values, both of the drug complexes were very comparable to each other, but the 2'-OMTase-Dolutegravir complex was found closer to the free protein molecule.

3.4.2 RMSF Analysis

RMSF is an essential structural parameter to identify the flexible and rigid regions of the protein structure. It is a measure of deviations of a particle from its original position. Moreover, it can also be used to identify the flexible residues in the protein, and thus it enables us to explore the conformational flexibility of the protein structure. The RMSF analysis was estimated for C-alpha atom of each residue representing the average displacement of each atom.

In the case of 3CLpro, we observed the highest fluctuations in loop heavy domain II and domain III. The loop regions are very flexible elements of protein molecules, and their flexibility is essential to accommodate the drug at the binding site appropriately. Besides, the 3CLpro-Raltegravir complex was found to be showing more significant fluctuations in comparison to 3CLpro-Paritaprevir complex (see **Fig. 5b**). The average RMSF values of 3CLpro, 3CLpro-Raltegravir complex and 3CLpro-Paritaprevir complex were found to be 0.170 nm, 0.182 nm and 0.184 nm, respectively. Additionally, the majority of the protein residues were stable with RMSF values smaller than 0.3 nm. Also, the catalytic dyad (His41 and Cys145) participating in interactions with the drug molecules remained highly stable throughout the MD simulation.

In the case of 2'-OMTase, the fluctuations were observed throughout the entire protein. 2'-OMTase also has an extensively looped structure, and as previously explained, we did expect to see significant fluctuations across the molecule. Moreover, both protein-drug complexes were also highly flexible (see **Fig. 6b**). The average RMSF values of 2'-OMTase, 2'-OMTase-Dolutegravir complex and 2'-OMTase-Bictegravir complex were found to be 0.155 nm, 0.161 nm and 0.165 nm, respectively. Even though the average fluctuations of both protein-drug complexes were very similar, the 2'-OMTase-Dolutegravir complex was consistently found to be showing the most significant fluctuations. Besides, the conserved K-D-K-E (Lys-46, Asp-130, Lys-170 and Glu-203) remained remarkably stable throughout the MD simulation.

3.4.3 Radius of Gyration (Rg) Analysis

The radius of gyration is a valuable tool to understand the folding properties and compactness of protein and protein-drug complexes. It can also be used to elucidate the influence a drug molecule exerts leading to conformational changes in protein structure. A relatively high Rg value indicates a protein molecule with loose packing, whereas a smaller Rg value indicates a protein structure with tight packing. This analysis is also used to understand whether the drug molecules maintained the folding behaviour of protein or not.

In the case of 3CLpro, both of the protein-drug complexes were found to be more compact than the unbound protein molecule (see **Fig. 5c**). The average Rg value of 3CLpro, 3CLpro-Raltegravir complex and 3CLpro-Paritaprevir complex was found to be 2.165 nm, 2.116 nm and 2.144 nm, respectively. We also observed a few minor variations in Rg, which could be attributed to the conformational shifts that resulted from the changes in the secondary structures of protein during the MD simulation.

In the case of 2'-OMTase, both the protein-drug complexes were slightly less compact in comparison to the free 2'-OMTase. The free protein, as well as the protein-drug complexes, attained stability at 30 ns and remained stable until the end (see **Fig. 6c**). The average Rg values of 2'-OMTase, 2'-OMTase-Dolutegravir complex and 2'-OMTase-Bictegravir complex were found to be 1.796 nm, 1.799 nm and 1.805 nm, respectively. Most of the fluctuations in the Rg values were recorded before 20 ns which again supported that these drug molecules stayed strongly bound to the active site and they helped to maintain the stability and compactness of the protein structure.

3.4.4 Intermolecular H-bonding

Hydrogen bonding is among the most crucial parameters to understand the binding affinity of drug molecules towards a protein molecule. Formation or deformation of H-bonds is a necessary consideration during MD simulations. A large number of H-bonds present in between protein and drug molecule signifies a strong binding affinity.

In the case of 3CLpro, we observed the maximum number of hydrogen bonds in the 3CLpro-Paritaprevir complex (six intermolecular H-bonds) followed by 3CLpro-Raltegravir complex (five intermolecular H-bonds). The average value of intermolecular H-bonds were 2 for both 3CLpro-Paritaprevir complex and 3CLpro-Raltegravir complex, respectively (see **Fig. 7**). This results explained that both the drug molecules interacted effectively towards the active site of 3CLpro with a significant number of hydrogen bonds.

Hydrogen Bonds

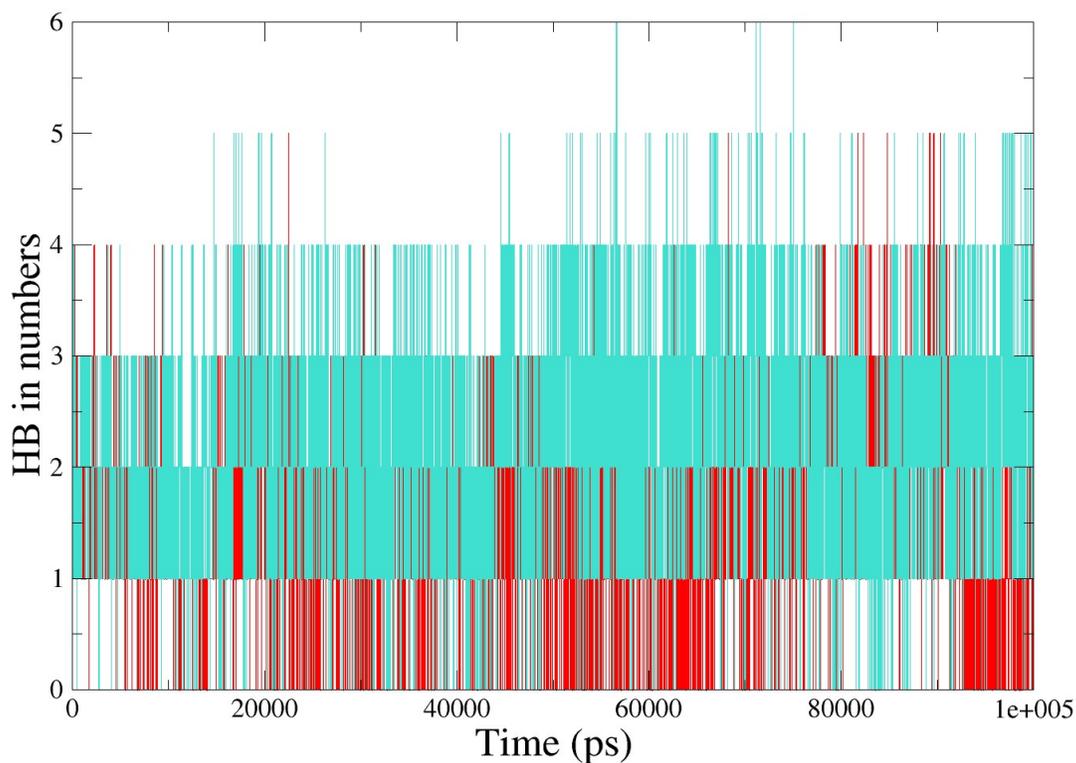


Fig. 7 Intermolecular hydrogen bonds between the drugs and 3CLpro; 3CLpro-Paritaprevir complex (turquoise), 3CLpro-Raltegravir complex (Red).

In the case of 2'-OMTase, we observed a maximum number of hydrogen bonds in the 2'-OMTase-Dolutegravir complex (six intermolecular H-bonds) followed by 2'-OMTase-Bictegravir complex (four intermolecular H-bonds). The average value of intermolecular H-bonds were 2 and 1 for 2'-OMTase-Dolutegravir complex and 2'-OMTase-Bictegravir complex, respectively (see **Fig. 8**). From the results of intermolecular hydrogen bonding, Dolutegravir was found to be interacting with a higher number of hydrogen bonds towards 2'-OMTase than the other drug molecule.

Hydrogen Bonds

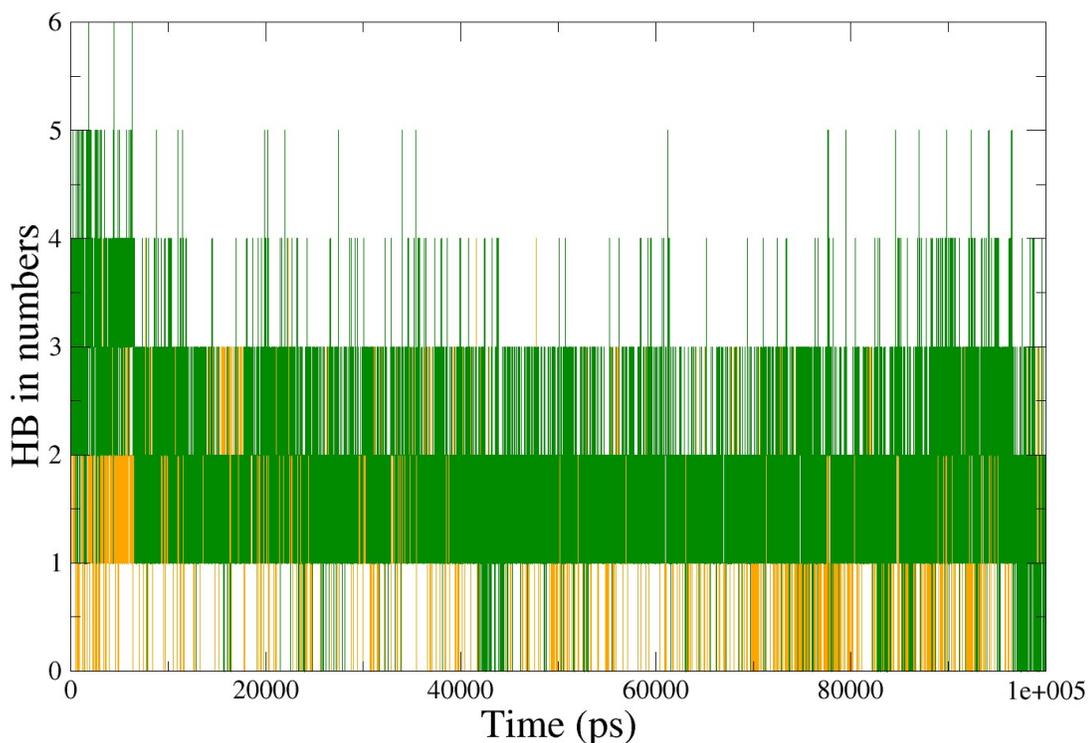


Fig. 8 Intermolecular hydrogen bonds between the drugs and 2'-OMTase; 2'-OMTase-Bictegravir complex (Orange), 2'-OMTase-Dolutegravir complex (Dark Green).

3.4.5 Solvent Accessible Surface Area (SASA)

Additionally, we also performed a SASA analysis of all of the proteins and protein-drug complexes. This analysis is useful in understanding the solvent behaviour (Hydrophilic or Hydrophobic) of a protein molecule as well as protein-drug complexes.

No major differences were observed in the SASA profiles of 3CLpro and its protein-drug complexes (see **Fig. 5d**). The average values of SASA for the free 3CLpro, 3CLpro-Paritaprevir complex and 3CLpro-Raltegravir complex were 133.89 nm², 131.79 nm² and 131.26 nm², respectively.

The average SASA of 2'-OMTase, 2'-OMTase-Dolutegravir complex and 2'-OMTase-Bictegravir complex were found to be 126.77 nm², 130.116 nm² and 131.03 nm², respectively. The results suggested that both of the proteins-drug complexes were impressively stable after the binding of drug molecules to their active sites (see **Fig. 6d**).

3.4.6 Principal Component Analysis (PCA)

PCA of MD simulations is a standard technique widely used to identify the essential motion of protein molecules. The PCA was employed to reveal the biologically significant and relevant motions from the global trajectories of unbound proteins as well as the protein-drug complexes using Essential Dynamics (ED) approach. ED

analyses the combined fluctuations of the most unsteady regions of protein molecules into two variables, namely Principal Component 1 (PC1) and Principal Component 2 (PC2), which represents the majority of fluctuations observed in MD simulation.

PCA graph of unbound 3CLpro, 3CLpro-Raltegravir complex and 3CLpro-Paritaprevir complex revealed that the protein molecule was slightly more flexible in the presence of both drug molecules (see **Fig. 9**). The trace of the covariance matrix for 3CLpro, 3CLpro-Raltegravir complex and 3CLpro-Paritaprevir complex was found to be 115.482 nm², 136.366 nm² and 132.059 nm², respectively. The 3CLpro-Paritaprevir complex was found to be occupying lesser conformational space than the 3CLpro-Raltegravir complex. This result, along with previous analyses such as RMSD, RMSF, Rg and SASA additionally supported that Raltegravir and also Paritaprevir could be the most promising candidates which might be inhibiting the function of 3CLpro.

2D projection of trajectory

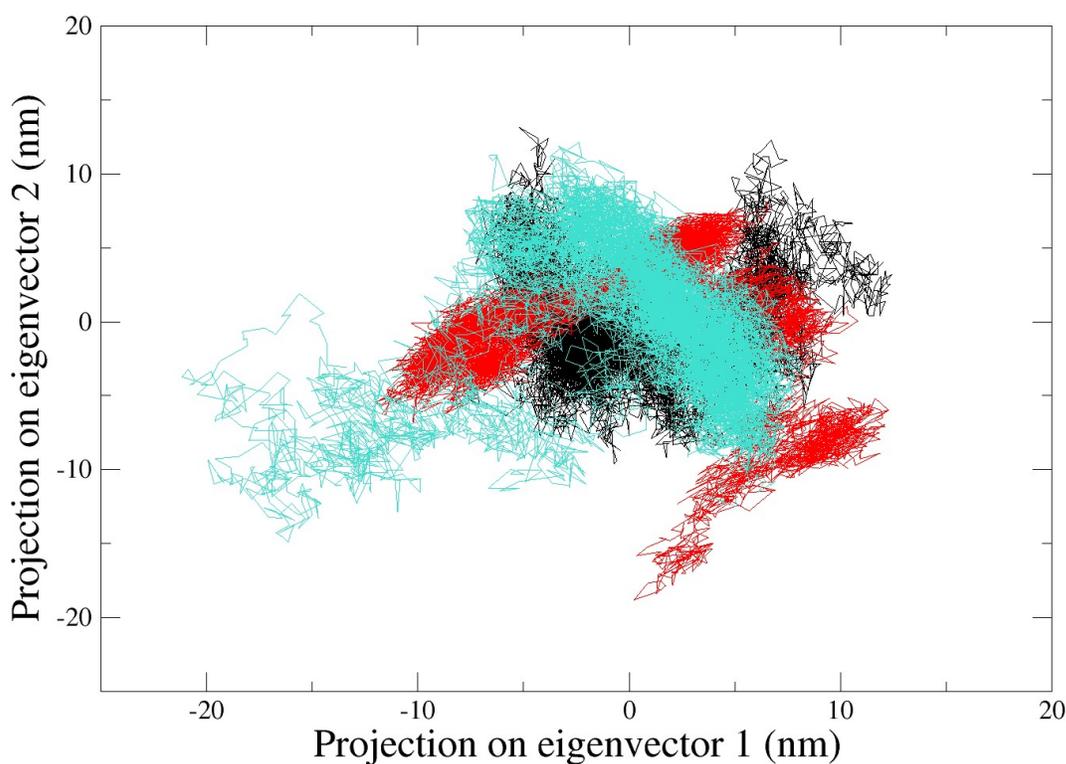


Fig. 9 Principal Component Analysis (PCA) of free 3CLpro (Black), 3CLpro-Paritaprevir complex (turquoise) and 3CLpro-Raltegravir complex (Red).

PCA of free 2'-OMTase, 2'-OMTase-Dolutegravir complex and 2'-OMTase-Bictegravir complex depicted that the protein, as well as the protein-drug complexes, were occupying almost equal conformational space (see **Fig. 10**). The trace of the covariance matrix for 2'-OMTase, 2'-OMTase-Dolutegravir complex and 2'-OMTase-Bictegravir complex was found to be 106.463 nm², 123.878 nm² and 117.187 nm², respectively. Together with all of

the MD simulation analysis, PCA supported that Dolutegravir and Bictegravir could be the most promising drugs capable of inhibiting the function of 2'-OMTase.

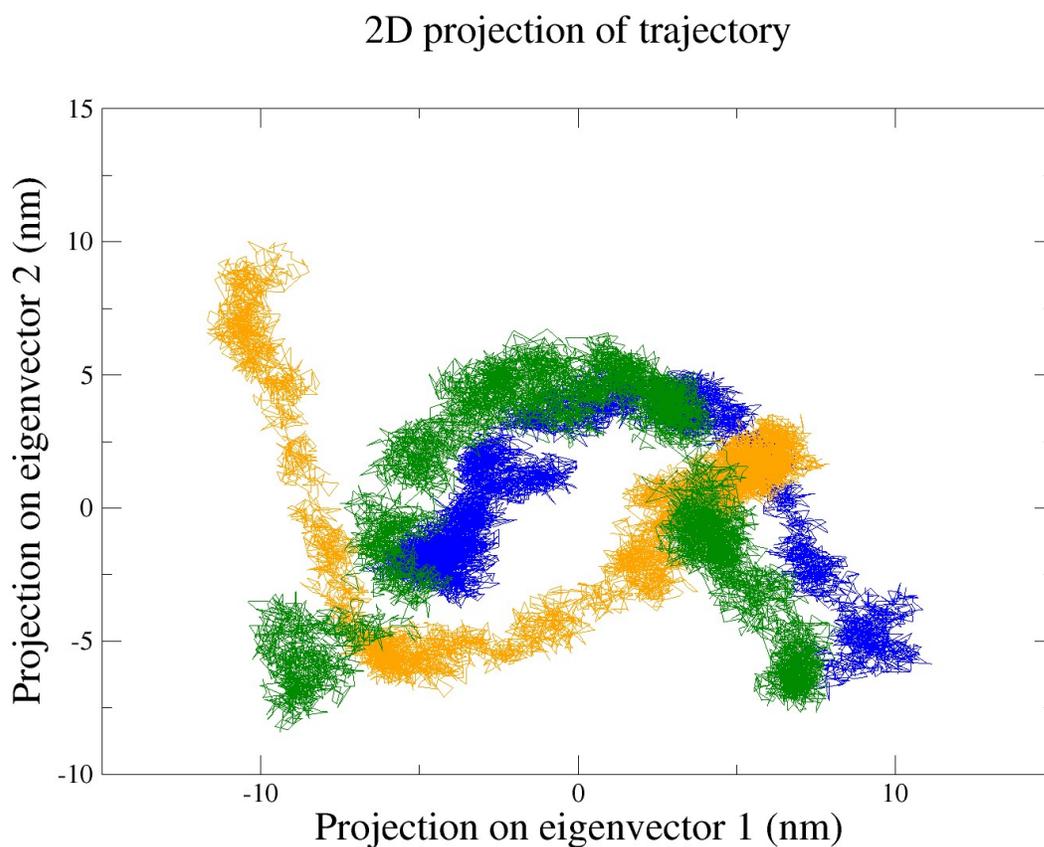


Fig. 10 Principal Component Analysis (PCA) of free 2'-OMTase (Blue), 2'-OMTase-Dolutegravir complex (Dark Green) and 2'-OMTase-Bictegravir complex (Orange).

The average values of the structural parameters including RMSD, RMSF, Rg, SASA, PCA and Intermolecular H-bonding has been described in **Table 5.** (3CLpro) and **Table 6.** (2'-OMTase).

Table 5. Time averaged structural properties obtained from MD simulation of 3CLpro

Analysis	3CLpro	3CLpro-Raltegravir	3CLpro-Paritaprevir
Average RMSD (nm)	0.295	0.361	0.401
Average RMSF (nm)	0.170	0.182	0.184
Average Rg (nm)	2.165	2.116	2.144
PCA (Trace of Covariance Matrix nm ²)	115.482	136.366	132.059
Average Intermolecular Hydrogen Bonds (Protein-Ligand)	-	2	2

Average SASA (nm ²)	133.89	131.26	131.79
---------------------------------	--------	--------	--------

Table 6. Time averaged structural properties obtained from MD simulation of 2'-OMTase

Analysis	2'-OMTase	2'-OMTase-Dolutegravir	2'-OMTase-Bictegravir
Average RMSD (nm)	0.299	0.393	0.410
Average RMSF (nm)	0.155	0.161	0.165
Average Rg (nm)	1.796	1.799	1.805
PCA (Trace of Covariance Matrix nm ²)	106.463	123.878	117.187
Average Intermolecular Hydrogen Bonds (Protein-Ligand)	-	2	1
Average SASA (nm ²)	126.77	130.116	131.03

3.4.7 Comparative analysis of pre and post MD simulation

Structural superposition before and after the MD simulation of drug-free 3CLpro structures showed an RMSD of 2.41 Å. It indicates that the MD simulated 3CLpro structure was consistent and geometrically well optimized, which did not deviate, in comparison to the initial protein structure. Moreover, we also performed the structural superimposition analysis of drug bounded 3CLpro structures which explained that the RMSD before and after MD simulation of 3CLpro-Raltegravir complex and 3CLpro-Paritaprevir complex was reliable and stable with the values of 2.15 Å and 2.47 Å, respectively (see Fig. 11).

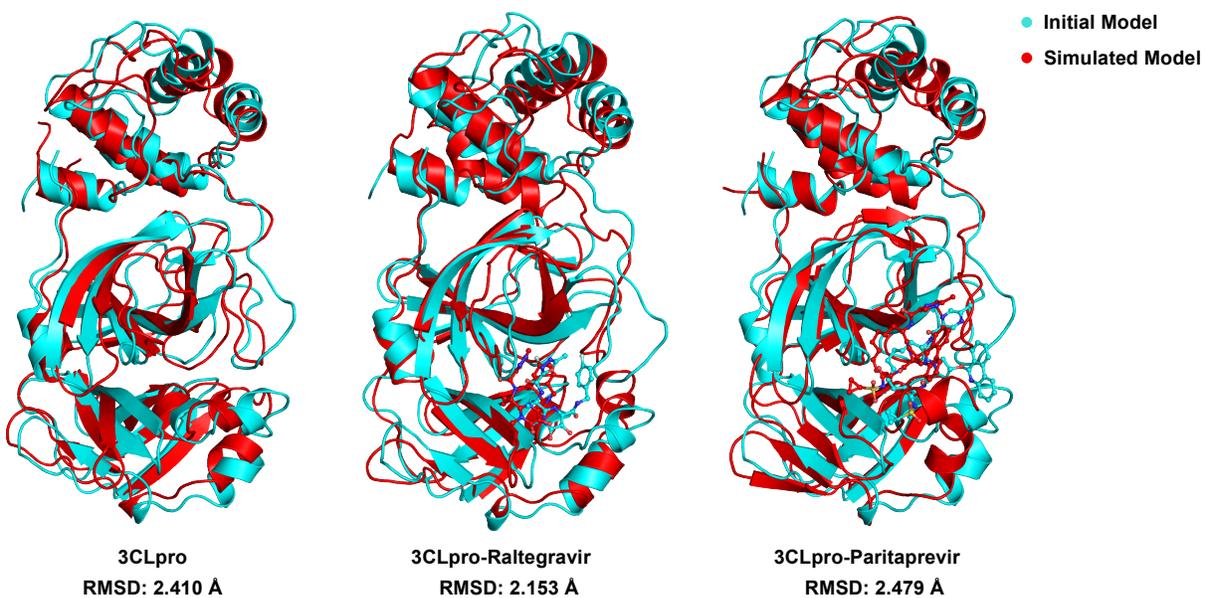


Fig. 11. Structural superimposition of the initial and MD simulated free 3CLpro, 3CLpro-Raltegravir complex and 3CLpro-Paritaprevir complex.

Structural superposition of pre and post MD simulation of drug-free 2'-OMTase structures showed an RMSD of 2.115 Å. Besides, we also carried out the same analysis for both drug bounded structures of 2'-OMTase. Also, we performed the same structural superimposition analysis of drug bounded structures 2'-OMTase, which explained that the RMSD before and after MD simulation of 2'-OMTase-Dolutegravir complex and 2'-OMTase-Bictegravir complex was reliable and stable with the values of 1.767 Å and 2.568 Å, respectively (see **Fig. 12**).



Fig. 12. Structural superimposition of the initial and MD simulated free 2'-OMTase, 2'-OMTase-Bictegravir complex and 2'-OMTase-Dolutegravir complex.

This is additionally supported that these two drugs were oriented in the active site of both proteins throughout the 100 ns MD simulations.

4. Conclusion

The purpose of this study was to screen pre-existing drugs against 2019-nCoV. The advantage of repurposing drugs is that the clinical trials, including efficacy studies, toxicity studies, and adverse effect studies, can be entirely avoided. Developing novel drugs against 2019-nCoV will demand a lot of time, which we cannot afford at present. 2019-nCoV is spreading at a rapid rate; therefore, it is rational to look for pre-existing drugs and formulations, which can control this highly contagious virus. In the present study, we have performed molecular modeling of two druggable targets from 2019-nCoV, their docking against a library of antiviral drugs and finally, the simulation of selected drugs with the target proteins. We have identified two drugs against 3CLpro (Raltegravir and Paritaprevir) and two drugs against 2'-OMTase (Bictegravir and Dolutegravir), which can act as potential inhibitors against 2019-nCoV. Raltegravir [34], Bictegravir [35] and Dolutegravir [36, 37] belong to the family of integrase strand transfer inhibitors (INSTI). They are already being used in antiretroviral therapy (ART) against the Human Immunodeficiency Virus (HIV) infections. The final drug, Paritaprevir [38, 39], is being used to treat chronic Hepatitis C Virus (HCV) infections. It suppresses viral replication by inhibiting the NS3 protease of HCV. To our knowledge, we are first to employ the extensive computer aided drug-repurposing approach against this deadly virus and the proposed drug molecules are very significant and they open up a gateway for 2019-nCoV therapeutics.

Conflict of interest

None

Ethical standards

Ethical standards are compulsory for studies relating to human and animal subjects.

Acknowledgement

The authors thank Prof. B. Jayaram (Coordinator of the Supercomputing Facility for Bioinformatics & Computational Biology, IIT Delhi) for supporting and providing computational facilities to carry out the MD simulation of proteins and protein-drug complexes. Dr. Amit Kumar Singh thanks the Department of Science and Technology (DST) and Indian National Science Academy (INSA), New Delhi, India. Gizachew Muluneh Amera thanks the College of Natural Science, Wollo University, Dessie, Ethiopia for the sponsorship.

References:

1. Su, S., et al., *Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses*. Trends Microbiol, 2016. **24**(6): p. 490-502.
2. de Wilde, A.H., et al., *Host Factors in Coronavirus Replication*. Curr Top Microbiol Immunol, 2018. **419**: p. 1-42.
3. Weiss, S.R. and J.L. Leibowitz, *Coronavirus pathogenesis*. Adv Virus Res, 2011. **81**: p. 85-164.
4. Zhu, N., et al., *A Novel Coronavirus from Patients with Pneumonia in China, 2019*. 2020.
5. World Health, O., *Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003*. http://www.who.int/csr/sars/country/table2004_04_21/en/index.html.
6. World Health, O. *Middle East respiratory syndrome coronavirus (MERS-CoV)*. November, 2019; Available from: <https://www.who.int/emergencies/mers-cov/en/>.
7. Huang, C., et al., *Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China*. The Lancet, 2020. **395**(10223): p. 497-506.
8. CSSE, J.H. *Coronavirus COVID-19 Global Cases by Johns Hopkins CSSE*. 2020; Available from: <https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6>.
9. World Health, O. *Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV)*. 2020; Available from: [https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-\(2019-ncov\)](https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov)).
10. Yang, H., et al., *The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor*. Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(23): p. 13190-13195.
11. Benkert, P., M. Biasini, and T. Schwede, *Toward the estimation of the absolute quality of individual protein structure models*. Bioinformatics, 2011. **27**(3): p. 343-50.
12. Chen, Y., et al., *Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2'-O-methylation by nsp16/nsp10 protein complex*. PLoS pathogens, 2011. **7**(10): p. e1002294-e1002294.
13. Lugari, A., et al., *Molecular Mapping of the RNA Cap 2'-O-Methyltransferase Activation Interface between Severe Acute Respiratory Syndrome Coronavirus nsp10 and nsp16*. The Journal of biological chemistry, 2010. **285**: p. 33230-41.
14. Menachery, V.D., et al., *Attenuation and restoration of severe acute respiratory syndrome coronavirus mutant lacking 2'-o-methyltransferase activity*. J Virol, 2014. **88**(8): p. 4251-64.
15. Waterhouse, A., et al., *SWISS-MODEL: homology modelling of protein structures and complexes*. Nucleic Acids Research, 2018. **46**(W1): p. W296-W303.
16. Camacho, C., et al., *BLAST+: architecture and applications*. BMC Bioinformatics, 2009. **10**: p. 421.
17. Berman, H.M., et al., *The Protein Data Bank*. Nucleic acids research, 2000. **28**(1): p. 235-242.

18. Krieger, E., et al., *Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8*. Proteins, 2009. **77 Suppl 9**: p. 114-22.
19. Ko, J., et al., *GalaxyWEB server for protein structure prediction and refinement*. Nucleic acids research, 2012. **40**(Web Server issue): p. W294-W297.
20. MBI, U. *The Structure Analysis and Verification Server*. Available from: <https://servicesn.mbi.ucla.edu/SAVES/>.
21. Colovos, C. and T.O. Yeates, *Verification of protein structures: patterns of nonbonded atomic interactions*. Protein Sci, 1993. **2**(9): p. 1511-9.
22. Luthy, R., J.U. Bowie, and D. Eisenberg, *Assessment of protein models with three-dimensional profiles*. Nature, 1992. **356**(6364): p. 83-5.
23. Laskowski, R.A., et al., *PROCHECK: a program to check the stereochemical quality of protein structures*. Journal of Applied Crystallography, 1993. **26**(2): p. 283-291.
24. Wishart, D.S., et al., *DrugBank 5.0: a major update to the DrugBank database for 2018*. Nucleic Acids Res, 2018. **46**(D1): p. D1074-d1082.
25. Trott, O. and A.J. Olson, *AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading*. Journal of computational chemistry, 2010. **31**(2): p. 455-461.
26. M.J. Abraham, D.v.d.S., E. Lindahl, B. Hess, and the GROMACS development team, *GROMACS User Manual version 2019*.
27. Chiu, S.-W., et al., *An Improved United Atom Force Field for Simulation of Mixed Lipid Bilayers*. The Journal of Physical Chemistry B, 2009. **113**(9): p. 2748-2763.
28. Schuttelkopf, A.W. and D.M. van Aalten, *PRODRG: a tool for high-throughput crystallography of protein-ligand complexes*. Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 8): p. 1355-63.
29. Berendsen, H.J.C., J.R. Grigera, and T.P. Straatsma, *The missing term in effective pair potentials*. The Journal of Physical Chemistry, 1987. **91**(24): p. 6269-6271.
30. Chen, H., et al., *Only one protomer is active in the dimer of SARS 3C-like proteinase*. J Biol Chem, 2006. **281**(20): p. 13894-8.
31. Ramachandran, G.N., C. Ramakrishnan, and V. Sasisekharan, *Stereochemistry of polypeptide chain configurations*. Journal of Molecular Biology, 1963. **7**(1): p. 95-99.
32. Laskowski, R.A., et al., *PDBsum: Structural summaries of PDB entries*. Protein science : a publication of the Protein Society, 2018. **27**(1): p. 129-134.
33. Menachery, V.D., K. Debbink, and R.S. Baric, *Coronavirus non-structural protein 16: Evasion, attenuation, and possible treatments*. Virus Research, 2014. **194**: p. 191-199.
34. Markowitz, M., et al., *Rapid and Durable Antiretroviral Effect of the HIV-1 Integrase Inhibitor Raltegravir as Part of Combination Therapy in Treatment-Naive Patients With HIV-1 Infection: Results of a 48-Week Controlled Study*. 2007. **46**(2): p. 125-133.

35. Sax, P.E., et al., *Coformulated bictegravir, emtricitabine, and tenofovir alafenamide versus dolutegravir with emtricitabine and tenofovir alafenamide, for initial treatment of HIV-1 infection (GS-US-380-1490): a randomised, double-blind, multicentre, phase 3, non-inferiority trial*. *The Lancet*, 2017. **390**(10107): p. 2073-2082.
36. Walmsley, S.L., et al., *Dolutegravir plus Abacavir-Lamivudine for the Treatment of HIV-1 Infection*. 2013. **369**(19): p. 1807-1818.
37. Cahn, P., et al., *Dolutegravir versus raltegravir in antiretroviral-experienced, integrase-inhibitor-naive adults with HIV: week 48 results from the randomised, double-blind, non-inferiority SAILING study*. *The Lancet*, 2013. **382**(9893): p. 700-708.
38. Feld, J.J., et al., *Sustained virologic response of 100% in HCV genotype 1b patients with cirrhosis receiving ombitasvir/paritaprevir/r and dasabuvir for 12weeks*. *Journal of Hepatology*, 2016. **64**(2): p. 301-307.
39. Sulkowski, M.S., et al., *Ombitasvir, Paritaprevir Co-dosed With Ritonavir, Dasabuvir, and Ribavirin for Hepatitis C in Patients Co-infected With HIV-1: A Randomized Trial*. *JAMA*, 2015. **313**(12): p. 1223-1231.