



Potential protease inhibitors and their combinations to block SARS-CoV-2

Chandran S. Abhinand^a, Achuthsankar S. Nair^a, Anand Krishnamurthy^b, Oommen V. Oommen^a and Perumana R. Sudhakaran^a

^aDepartment of Computational Biology and Bioinformatics, University of Kerala, Thiruvananthapuram, Kerala, India; ^bDassault Systemes India Pvt Ltd, Chennai, Tamil Nadu, India

Communicated by Ramaswamy H. Sarma

ABSTRACT

COVID-19, which has emerged recently as a pandemic viral infection caused by SARS-coronavirus 2 has spread rapidly around the world, creating a public health emergency. The current situation demands an effective therapeutic strategy to control the disease using drugs that are approved, or by inventing new ones. The present study examines the possible repurposing of existing anti-viral protease inhibitor drugs. For this, the structural features of the viral spike protein, the substrate for host cell protease and main protease of the available SARS CoV-2 isolates were established by comparing with related viruses for which antiviral drugs are effective. The results showed 97% sequence similarity among SARS and SARS-CoV-2 main protease and has same cleavage site positions and ACE2 receptor binding region as in the SARS-CoV spike protein. Though both are N-glycosylated, unlike SARS-CoV, human SARS-CoV-2 S-protein was O-glycosylated as well. Molecular docking studies were done to explore the role of FDA approved protease inhibitors to control SARS-CoV-2 replication. The results indicated that, Ritonavir has the highest potency to block SARS-CoV-2 main protease and human TMPRSS2, a host cell factor that aids viral infection. Other drugs such as Indinavir and Atazanavir also showed favourable binding with Cathepsin B/L that helped viral fusion with the host cell membrane. Further molecular dynamics simulation and MM-PBSA binding free energy calculations confirmed the stability of protein-drug complexes. These results suggest that protease inhibitors particularly Ritonavir, either alone or in combination with other drugs such as Atazanavir, have the potential to treat COVID 19.

ARTICLE HISTORY

Received 20 June 2020
Accepted 1 September 2020

KEYWORDS

COVID-19; spike protein; main protease; TMPRSS2; molecular docking

Introduction

COVID 19, is a zoonotic infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). It is a positive sense single stranded RNA virus, which belongs to the genus β -coronavirus. SARS-COV-2 is a member of Coronaviridae family like severe acute respiratory syndrome coronavirus (SARS) and the Middle-East respiratory syndrome coronavirus (MERS HCoV) (Cui et al., 2019). Its genome is about 30kb in size and has been sequenced from several isolates; the whole genome showed very high similarity to that of SARS virus and appeared to be of bat origin (Chen et al., 2020). The outbreak of SARS-CoV-2 was first reported in Wuhan, China in December 2019 and now has assumed pandemic proportions, spread globally in about 188 countries and territories (Li et al., 2020). According to the latest updates on 30th July 2020 from Johns Hopkins Corona virus resource Centre, the total number of positive cases reached 17.05 million with 668,589 deaths all over the world.

SARS-CoV-2 is transmitted from human to human through direct contact with infected persons with an incubation

period, average of 11.2 days (Sanche et al., 2020). This viral infection shows a range of clinical presentation in humans, including fever, headache, sore throat, cough and runny nose. Besides respiratory manifestations, cardiovascular symptoms have also been reported in some COVID 19 patients. However, in some individuals especially young children, the infection does not show any symptom and in others it shows only upper respiratory tract symptoms (Huang et al., 2020; Jin et al., 2020; Zhu et al., 2020). Presently, the infection has been treated by FDA approved general anti-viral drugs against SARS (Chan et al., 2003; Chu et al., 2004; Lai, 2005). The clinical trials reported some potent drugs against COVID 19 and researchers are still working for better alternatives (Zhang & Liu, 2020).

Studies in the past have identified viral target proteins for controlling corona virus infection rate. Among the different targets, proteases, which have a critical role in promoting viral infection, constitute an important target molecule for antiviral drugs (Yang et al., 2006; Ziebuhr, 2005). A critical early event in the life cycle of corona virus is the attachment

CONTACT Perumana R. Sudhakaran ✉ prslab@gmail.com Department of Computational Biology and Bioinformatics, University of Kerala, Thiruvananthapuram, Kerala 6995581, India

Supplemental data for this article can be accessed online at <https://doi.org/10.1080/07391102.2020.1819881>.

This article has been republished with minor changes. These changes do not impact the academic content of the article.

© 2020 Informa UK Limited, trading as Taylor & Francis Group

to host cell mediated by viral surface spike protein (S protein). This involves host cell protease-mediated proteolytic cleavage and priming of the S protein. This is followed by the binding of the virus particle to host cell receptor, fusion with the cell membrane and entry into endosomes (Ou et al., 2020). Another key event, following the translation of the viral RNA genome employing the host translational machinery, is the proteolytic cleavage of the viral polyprotein by viral protease. This generates key proteins required for viral replication, assembly, transport and secretion (Fung & Liu, 2019). Important host cell proteases involved in aiding corona virus infection include membrane serine proteases, furin, trypsin and cathepsins B/L (Millet & Whittaker, 2015). SARS virus S-protein cleavage and priming are mediated through serine protease TMPRSS2 (Iwata-Yoshikawa et al., 2019; Shirato et al., 2018; Zhou et al., 2015). In a recent study, cleavage of S-protein of SARS-CoV-2 has also been shown to be mediated through TMPRSS2, although an additional role for cathepsin B/L was not excluded (Hoffmann et al., 2020). CoV main protease, also known as 3CL- protease, is the key viral protease involved in maturation -cleavage events within the polyprotein precursor that regulates viral replication and gene expression (Bhardwaj et al., 2020; Islam et al., 2020; Khan et al., 2020). Availability of the SARS-CoV-2 viral genome sequence enabled the expression systems for the main protease. The 3D structure of SARS-CoV-2 main protease has been deposited in protein Data Bank (PDB ID: 6LU7) in the month of February 2020 from China, which may help in developing new drugs that block CoV-2 main protease (Havranek & Islam, 2020).

Although it is important to develop new drugs to target SARS-CoV-2, it is also essential to control the present speeding infection rate. There are a number of drugs that prevent proteolytic cleavage of polyprotein by inhibiting viral main proteases, which block several viral infections (Chen et al., 2020). It has been suggested that inhibitors of other viral proteases such as HIV protease, HCV proteases, either individually or in combination can be used to treat COVID-19. However, evidence for such re-purposing of existing drugs with statutory approvals for treatment of COVID-19 is lacking. Therefore, the structural characteristics of proteolytic components of SARS-COV-2, particularly its main protease and the S-protein were analysed. This helps examine whether the presently available antiviral drugs that target proteases of other viruses particularly, corona virus- mediated infection can be re-purposed against SARS-COV-2, using molecular docking studies. It predicts the potential of such drugs against COVID-19. The main objective was to identify molecule or molecules that can target the proteolytic process during both, the initial entry phase and later intracellular events. Our results show that, of the different antiviral drugs that target proteases, Ritonavir, has better docking characteristics to both COV-2 protease and the host protease, particularly TMPRSS2.

Materials and methods

Data collection and preparation

The protein sequence and structure of different strains of coronaviruses main protease were downloaded from RCSB

Protein Data Bank. Raw proteins were then prepared using Discovery studio 2020 by deleting hetero atoms, ligands and water molecules. FDA approved HIV and HCV protease inhibitors were collected from PubChem compound database and were further prepared for the molecular docking studies using Discovery studio 2020.

PDB ID of proteases and Pubchem ID of protease inhibitors are listed as follows: NL63 (PDB ID-3TLO), HKU1 (PDB ID-3D23), SARS-CoV (PDB ID-2AMQ), MERS-CoV (PDB ID-4WMD), SARS-CoV-2 (PDB ID-6LU7), CATHEPSIN L (PDB ID- 2XU1), CATHEPSIN B (PDB ID- 1CSB), Ritonavir (CID_392622, IUPAC Name: 1,3-thiazol-5-ylmethyl N-[(2S,3S,5S)-3-hydroxy-5-[[[(2S)-3-methyl-2-[[methyl-[(2-propan-2-yl-1,3-thiazol-4-yl)methyl]carbamoyl]amino]butanoyl]amino]-1,6-diphenylhexan-2-yl]carbamate), Indinavir (CID_5362440, IUPAC Name: (2S)-1-[(2S,4R)-4-benzyl-2-hydroxy-5-[[[(1S,2R)-2-hydroxy-2,3-dihydro-1H-inden-1-yl]amino]-5-oxopentyl]-N-tert-butyl-4-(pyridin-3-ylmethyl)piperazine-2-carboxamide), Atazanavir (CID_148192, IUPAC Name: methyl N-[(2S)-1-[2-[(2S,3S)-2-hydroxy-3-[[[(2S)-2-(methoxycarbonylamino)-3,3-dimethylbutanoyl]amino]-4-phenylbutyl]-2-[(4-pyridin-2-ylphenyl)methyl]hydrazinyl]-3,3-dimethyl-1-oxobutan-2-yl]carbamate), Lopinavir (CID_92727, IUPAC Name: (2S)-N-[(2S,4S,5S)-5-[[2-(2,6-dimethylphenoxy)acetyl]amino]-4-hydroxy-1,6-diphenylhexan-2-yl]-3-methyl-2-(2-oxo-1,3-diazinan-1-yl)butanamide), Saquinavir (CID_441243, IUPAC Name: (2S)-N-[(2S,3R)-4-[(3S,4aS,8aS)-3-(tert-butylcarbamoyl)-3,4,4a,5,6,7,8,8a-octahydro-1H-isoquinolin-2-yl]-3-hydroxy-1-phenylbutan-2-yl]-2-(quinoline-2-carboxylamino)butanediamide), Tipranavir (CID_54682461, IUPAC Name: N-[3-[(1R)-1-[(2R)-4-hydroxy-6-oxo-2-(2-phenylethyl)-2-propyl-3H-pyran-5-yl]propyl]phenyl]-5-(trifluoromethyl)pyridine-2-sulfonamide), Nelfinavir (CID_IUPAC Name: (3S,4aS,8aS)-N-tert-butyl-2-[(2R,3R)-2-hydroxy-3-[(3-hydroxy-2-methylbenzoyl)amino]-4-phenylsulfanylbutyl]-3,4,4a,5,6,7,8,8a-octahydro-1H-isoquinoline-3-carboxamide), Fosamprenavir (CID_131536, IUPAC Name: [(3S)-oxolan-3-yl] N-[(2S,3R)-4-[(4-aminophenyl)sulfonyl-(2-methylpropyl)amino]-1-phenyl-3-phosphonoxybutan-2-yl]carbamate), Darunavir (CID_213039, IUPAC Name: [(3aS,4R,6aR)-2,3,3a,4,5,6a-hexahydrofuro[2,3-b]furan-4-yl] N-[(2S,3R)-4-[(4-aminophenyl)sulfonyl-(2-methylpropyl)amino]-3-hydroxy-1-phenylbutan-2-yl]carbamate), Amprenavir (CID_65016, IUPAC Name: [(3S)-oxolan-3-yl] N-[(2S,3R)-4-[(4-aminophenyl)sulfonyl-(2-methylpropyl)amino]-3-hydroxy-1-phenylbutan-2-yl]carbamate), Grazoprevir (CID_44603531, IUPAC Name: (1R,18R,20R,24S,27S)-24-tert-butyl-N-[(1R,2S)-1-(cyclopropylsulfonylcarbamoyl)-2-ethenylcyclopropyl]-7-methoxy-22,25-dioxo-2,21-dioxo-4,11,23,26-tetrazapentacyclo[24.2.1.03,12.05,10.018,20]nonacos-3,5(10),6,8,11-pentaene-27-carboxamide), Boceprevir (CID_10324367, IUPAC Name: (1R,2S,5S)-N-(4-amino-1-cyclobutyl-3,4-dioxobutan-2-yl)-3-[(2S)-2-(tert-butylcarbamoylamino)-3,3-dimethylbutanoyl]-6,6-dimethyl-3-azabicyclo[3.1.0]hexane-2-carboxamide), Simeprevir (CID_24873435, IUPAC Name: (1R,4R,6S,7Z,15R,17R)-N-cyclopropylsulfonyl-17-[7-methoxy-8-methyl-2-(4-propan-2-yl-1,3-thiazol-2-yl)quinolin-4-yl]oxy-13-methyl-2,14-dioxo-3,13-diazatricyclo[13.3.0.04,6]octadec-7-ene-4-carboxamide), Paritaprevir (CID_45110509, IUPAC Name: (1S,4R,6S,7Z,14S,18R)-N-cyclopropylsulfonyl-14-[(5-methylpyrazine-2-carbonyl)amino]-2,15-dioxo-18-phenanthridin-6-yloxy-3,16-diazatricyclo[14.3.0.04,6]nonadec-7-ene-4-carboxamide), Telaprevir (CID_3010818, IUPAC Name: (3S,3aS,6aR)-2-[(2S)-2-

[[[(2S)-2-cyclohexyl-2-(pyrazine-2-carboxylamino)acetyl]amino]-3,3-dimethylbutanoyl]-N-[(3S)-1-(cyclopropylamino)-1,2-dioxohexan-3-yl]-3,3a,4,5,6,6a-hexahydro-1H-cyclopenta[c]pyrrole-3-carboxamide), Asunaprevir (CID_16076883, IUPAC Name: tert-butyl N-[(2S)-1-[(2S,4R)-4-(7-chloro-4-methoxyisoquinolin-1-yl)oxy-2-[[[(1R,2S)-1-(cyclopropylsulfonylcarbamoyl)-2-ethenylcyclopropyl]carbamoyl]pyrrolidin-1-yl]-3,3-dimethyl-1-oxobutan-2-yl]carbamate).

Protein sequence alignment and analysis

Sequences were downloaded from UniProtKB and NCBI databases and alignment was performed using PRALINE, multiple sequence alignment tool (<http://www.ibi.vu.nl/programs/pralinewww/>) (Simossis & Heringa, 2005). Alignment helps identify the conserved amino acid region of closely related viral proteases. Highly conserved regions were indicated as red and less in blue color. Most homology sharing sequences were retrieved to do pairwise alignment for identifying the percentage sequence similarities. The pairwise sequence alignment was done using EMBOSS Needle tool in EMBL-EBI (Rice et al., 2000).

Characterization and comparison of SARS-CoV-2 spike protein cleavage site with SARS-CoV spike protein was also done using pairwise sequence alignment. In order to identify the similarities of main protease and spike protein among the reported isolates of SARS-CoV-2, multiple sequence alignment was done using Clustal Omega (Sievers et al., 2011) tool.

Phylogenetic analysis

Protein sequences were subjected to phylogenetic analysis to identify closely related main protease among SARS-coronavirus species. Phylogenetic tree construction was performed using phelogeny.fr (<http://www.phylogeny.fr/>) using maximum likelihood (Dereeper et al., 2008). Briefly, a predefined pipeline using (1) MUSCLE (2) Gblocks (3) PhyML and (4) TreeDyn was run to build a tree. In that, generating multiple sequence alignment and curation were done by MUSCLE and Gblocks. PhyML was used to carry out phylogeny and TreeDyn to draw trees.

Prediction of N-linked and O-linked glycosylation sites

NetNGlyc 1.0 (Gupta & Brunak, 2002) and NetOGlyc 3.1 (Julenius et al., 2005) server were used to predict the N-glycosylation and O-glycosylation of the viral protein. Both predict the glycosylation sites using artificial neural networks and fix potential >0.5 as cut-off value. Instead of a single cut off score, two scores, G-score and I-score were considered in NetOglyc server for O-linked glycosylation site. Here, an additional score is used for threonines: a case in which G-score <0.5 but I-score >0.5 and since there are no predicted neighbouring sites threonine residue is also predicted as glycosylated.

Protein structure alignment and analysis

The structures of proteins which share maximum sequence similarity were then superimposed using Superimpose function in Discovery studio 2020. Even though this gives an idea about similarities among proteins, it was also useful for identifying the dissimilarities and divergence among the structures. Moreover, the superimposed structure analysis would help identify the commonness between active sites of proteins.

Homology modeling

The 3D structure of TMPRSS2 protein in human is not available till now. Therefore, homology modeling was performed to develop the 3D structure of TMPRSS2. The amino acid sequence of the protein was downloaded from UniProtKB [Uniprot Accession: O15393] in FASTA format and protein homology modeling was done in Discovery studio 2020. Amino acid sequence similarity was compared against data stored in PDB depository using BLASTp program and identified the template protein structures based on the percentage of sequence similarity. 3D structure of TMPRSS2 was built using MODELLER program in Discovery studio 2020 to verify the model. The stereo-chemical quality of the modeled structure was assessed by PROCHECK (Laskowski et al., 1993) program through analyzing its Ramachandran plot.

Molecular docking

Molecular docking was performed against SARS-CoV-2 protease and three human target proteins (TMPRSS2, cathepsin L and cathepsin B) using Discovery studio 2020. Target proteins were docked against 16 protease inhibitors listed above using LIBDOCK protocol in Discovery studio 2020 (Rao et al., 2007). LIBDOCK allowed flexible docking and generate all possible poses for each ligand. The inhibitor interacting amino acid residues was selected and defined as active site before running the docking protocol. After docking, the docked poses were sorted based on the 'LibDock' docking score and corresponding intermolecular H-bond interactions were analyzed.

Molecular dynamics simulation

To analyze the stability of top ranked ligands with its receptor protein, molecular dynamics (MD) studies were done using Standard Dynamics Cascade protocol in Discovery studio 2020. All the selected protein-ligand complexes were subjected to Charmm36 force-field before doing MD simulation. The MD simulations were performed for 130 ns with the following conditions; system target temperature was kept as 300 K, equilibration steps were set to 500 and the results were saved at intervals of 0.05 ns. Remaining parameters were set to their default values. Analyze trajectory protocol in the Discovery studio 2020 was used to determine the structural properties of ligand-protein complexes by generating the RMSD, RMSF of each conformation.

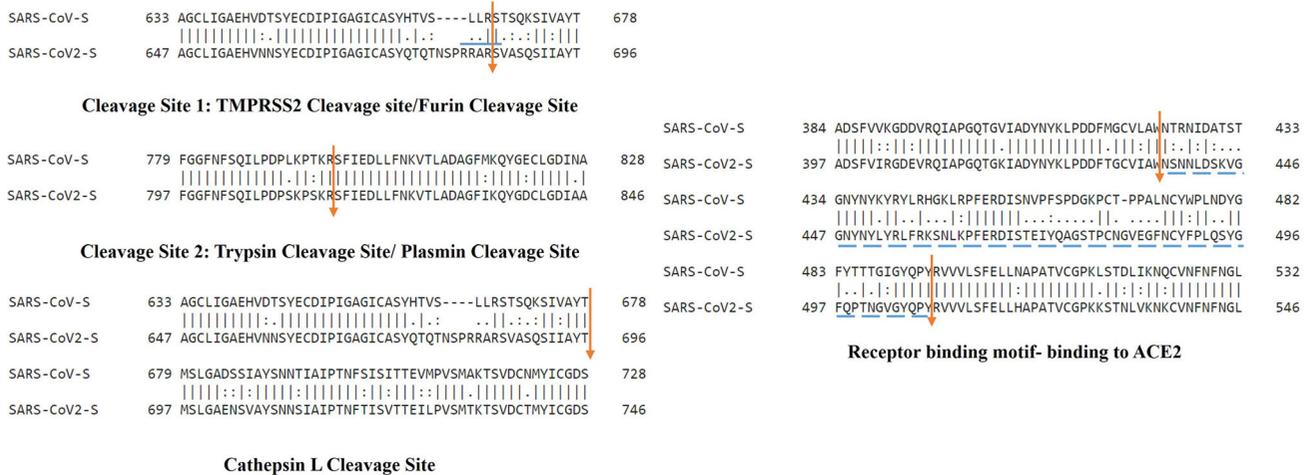


Figure 1. Schematic representation of SARS-CoV-2 spike protein with cleavage sites and receptor binding sites. The S protein sequence of SARS-CoV and SARS-CoV-2 cleavage sites were aligned using EMBOSS Needle tool. Cleavage site 1(TMPRSS2/Furin), Cleavage site 2(Trypsin/Plasmin) and ACE2 binding sites corresponding to SARS-CoV and SARS-CoV-2 spike proteins are indicated by coloured lines. Polybasic furin recognition site 'RRAR' in S1 cleavage site is highlighted with blue colour lines.

Binding free energy of protease-ligand complex

MM-PBSA (Mechanic/Poisson-Boltzmann Surface Area) method was used to calculate the binding free energy of the protein-ligand complexes. This was done by utilizing the 'Binding energy calculation' protocol of Discovery studio 2020 and was performed after molecular dynamics simulation. It provides an overview about the stability of the interactions between protein and ligand. The binding free energy (ΔG_{bind}) was calculated by using the following equation.

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{ligand}} - G_{\text{protein}}$$

where, G_{complex} , G_{protein} and G_{ligand} are the total free energy of the complex, free receptor and ligand respectively.

Statistical analysis

The results are expressed as mean \pm SEM. Statistical significance of difference was analyzed by one way ANOVA using Graph Pad Prism 5.01. A p value of less than 0.05 was considered statistically significant.

Results

Comparison of SARS-CoV and SARS-CoV-2 spike proteins

Spike protein sequence alignment and analysis

Proteolytic cleavage of the spike protein being an important early event in viral attachment and entry into host cells, analyses of the structural features of the S protein of SARS-CoV-2 that makes it a substrate for host cell protease for its activation and priming, were done. The protein sequences corresponding to CDS (protein formed by gene 'S') were downloaded and performed multiple sequence alignment using Clustal Omega, from the genome sequence of hundred isolates of SARS-CoV-2 available in NCBI data base. The sequences of S- protein among the hundred isolates were relatively conserved and found only nine amino acid

substitutions (Supplementary Table S1). Further, pairwise sequence alignment using the sequences downloaded from UniProtKB and NCBI protein data base showed about 87% similarities in amino acid sequence between SARS-CoV and SARS-CoV-2 spike proteins.

Host cell protease cleaves SARS-CoV spike protein to two functional subunits S1 and S2. In that, S1 is responsible for host cell receptor binding and S2 for fusion of virus to host cell membrane. The amino acid cleavage sites for host proteases such as TMPRSS2, Furin, Trypsin and Plasmin are also noted in between the S1/S2 subunits. Pairwise sequence alignment was performed to compare the cleavage site position between SARS-CoV and SARS-CoV-2 spike proteins. The results showed that all the four proteases cleavage sites present in SARS-CoV are also present in SARS-CoV-2 S-protein (Figure 1). However, their locations in SARS-CoV-2 were mapped to a few residues towards C-terminal of cleavage sites in SARS-CoV spike protein. While the site 1 where TMPRSS2/furin acts, is 667(RS) in SARS-CoV, it was 685(RS) in SARS-CoV-2; the site 2 where trypsin/plasmin cleaves is 797 (RS) in SARS-CoV, but it was 815(RS) in SARS-CoV-2. However, in the light of the data on cleavage of both R667 and R797 sites by TMPRSS2 (Reinke et al., 2017) and the recent report on the role of TMPRSS2 in CoV-2S protein cleavage (Hoffmann et al., 2020), both the R667 and R 797 sites in CoV-2S protein may also be the cleavage site for TMPRSS2. Confirming the earlier reports, we found that the S1 cleavage site of SARS-CoV-2, was a polybasic furin recognition site 'RRAR' unlike that in SARS-Cov S. A cathepsin L cleavage site 678 (TM), present in the extracellular domain that proteolytically activates the fusion of SARS-CoV S-protein, was also found in SARS-CoV-2 at position 696 (TM). These protease cleavage sites were fully conserved in all the hundred isolates and the sequences were (680)SPRRAR↓SVA(688) for site1 (TMPRSS2/furin), (813)SKR↓SFI(818) for site 2 (Trypsin/plasmin) and (694)AYT↓MSL(699) for cleavage by Cathepsin L. SARS-CoV spike proteins show a strong affinity with the host cell receptor ACE2 by binding through a receptor binding motif of around 71 amino acids in length. Sequence

A N-linked glycosylation sites				B O-linked glycosylation sites			
Position	Amino acid Sequence	Potential	N-Glyc result	Position	Amino acid	G-score	I-score
17	NLTT	0.6606	+	323	T	0.193	0.707
61	NVTW	0.7820	+++	385	T	0.151	0.547
74	NGTK	0.7192	++	618	T	0.192	0.659
122	NATN	0.6781	+	732	T	0.247	0.613
149	NKSW	0.6318	+				
165	NCTF	0.6220	+				
234	NITR	0.7613	+++				
282	NGTI	0.7378	++				
331	NITN	0.5970	+				
343	NATR	0.5671	+				
603	NTSN	0.5783	+				
616	NCTE	0.7163	++				
717	NFTI	0.6426	++				
801	NFSQ	0.6146	+				
1098	NGTH	0.5496	+				
1134	NNTV	0.5800	+				
1194	NESL	0.6791	++				

Figure 2. Glycosylation sites of SARS-CoV-2 spike protein (A) N-glycosylation sites of SARS-CoV-2 spike protein (NCBI Reference Sequence: YP_009724390.1) was predicted using NetNGlyc 1.0 Server. 17 N-linked glycosylation sites were predicted with threshold value > 0.5 . (B) O-linked glycosylation sites of SARS-CoV-2 spike protein (NCBI Reference Sequence: YP_009724390.1) was predicted using NetOGlyc 3.1 server and the results showed 4 threonine glycosylation sites with G-score < 0.5 , I-score > 0.5 and no predicted neighbouring residues with in a distance ten.

alignment results revealed that in SARS-CoV-2, it starts from position 436 rather than 423 in SARS-CoV. The sequence of the ACE2 receptor binding motif ((437) NSNNLDSKVGNGNY NYLRLFRKSNLKPFRDISTEIYQAGSTPCNGVEGFNCYFPLQSYGF-QPTNGVGYQPY (508)) is also conserved in the S-protein of all the hundred isolates sequenced.

Identification of N and O-linked glycosylation sites

The S-protein is post-translationally modified by glycosylation of select asparagine residues to form N-linked glycans, critical for its function. SARS-CoV is N-glycosylated at multiple asparagine residues. The extent of N glycosylation of SARS-CoV-2 was predicted and compared with that of SARS-CoV. A total of 17 N-linked glycosylation sites were predicted in SARS-CoV-2 S- protein (threshold value > 0.5) (Figure 2). All these N-glycosylation sites are fully conserved in all the hundred isolates. Specifically, 12 glycans are in S1 subunit and 5 in S2 subunit. The validity of the prediction was tested by analyzing N-glycosylation of SARS-CoV S-protein. Our prediction results on SARS-CoV S-protein showed N glycosylation of asparagines at 29, 65, 109, 119, 158, 227, 269, 318, 330, 357, 589, 602, 699, 783, 1080, 1116 and 1176 and showed good similarities in the experimentally reported N-linked glycosylation site of SARS-CoV spike protein. It appeared that the N-linked glycosylation sites in SARS-CoV and SARS-CoV-2 are conserved.

Another post translational modification is formation of O-linked glycans which provides mucin-like characteristics. By considering the importance of O-linked glycosylation, we also predict the O glycosylation sites in SARS-CoV-2 spike protein. Interestingly, we observed four threonine O-linked glycosylation sites in SARS-CoV-2 spike protein at positions 323, 385, 618 and 732 (Figure 2). These O-glycosylation sites are conserved in all the hundred isolates. Of the four sites, three are in S1 subunit and one in S2 subunit. Similar

prediction analysis showed no O-glycosylation sites in SARS-CoV S-protein.

Phylogenetic and sequence-based analysis of SARS-COV-2 main protease

The sequence similarity of SARS-CoV-2 main protease which plays a key role in polyprotein processing critical to viral replication, assembly and progression of infection was studied. For this, the selected hundred isolates of SARS-CoV-2 coding regions for main protease from NCBI and the protein sequence corresponding to CDS (protein formed by gene 'orf1ab') were then downloaded and saved in a single file for further analysis. Multiple sequence alignment was then performed using Clustal Omega and the results showed that sequences of main protease among the hundred isolates are relatively conserved and found only two amino acid substitutions among them (Supplementary Table S2).

Amino acid sequence alignment for the comparison of main proteases in different strains was done by multiple sequence alignment. Confirming the earlier results, sequence alignment showed more divergence among the protein sequences of coronavirus strains than similarity (Figure 3A). The amino acid residues that form active sites including Thr26, His41, Gly143, Ser144, Cys145, His172, Arg188, Ala191 and Glu192 are conserved in all the strains. The alignment also revealed that amino acids sequences at 33-35, 43-47 in domain I and 143-148 at domain II and 217-220 and 290-293 in domain III are highly conserved. Moreover, protease sequence of SARS-CoV and SARS-CoV-2 share some more similarities than others. To confirm the sequence similarities between protease sequence of SARS-CoV and SARS-CoV-2, pairwise sequence alignment was performed. The main protease sequences of these two viruses share 97% similarity, and at their active sites comprising of Cys(145) His (41) dyad,

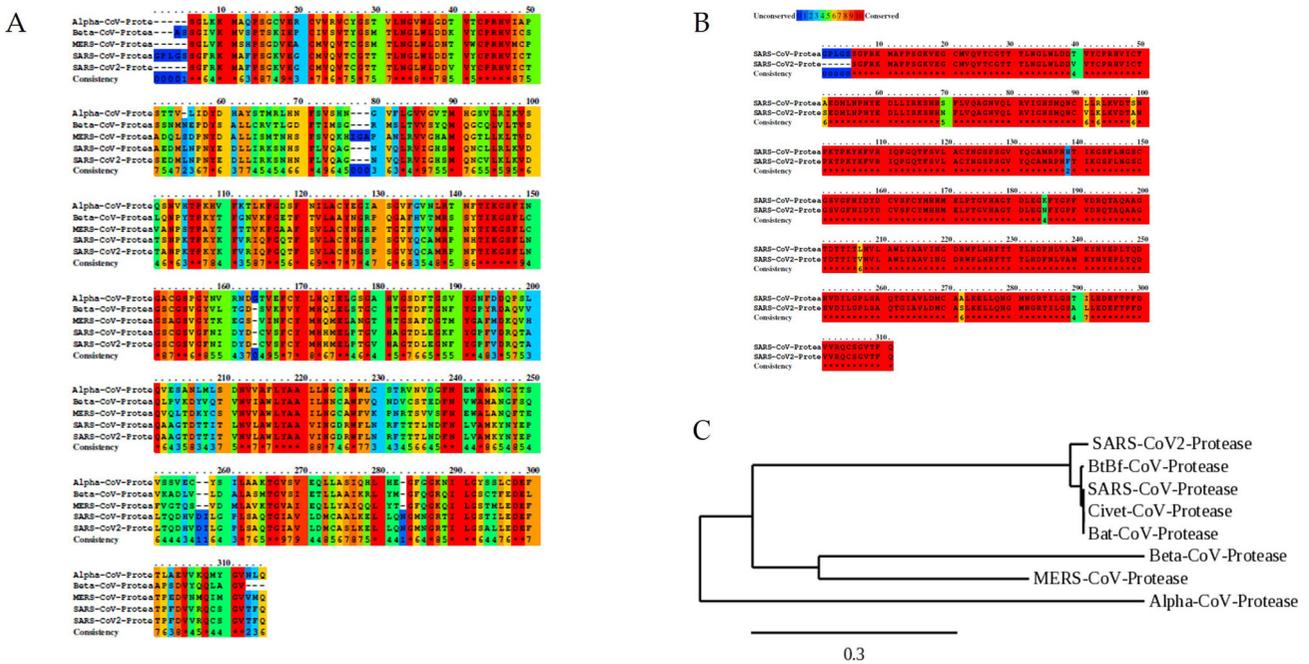


Figure 3. Phylogenetic and sequence alignment of SARS-CoV2protease. (A) Multiple sequence alignment showing conserved and non-conserved regions in main protease of different corona virus strains [(Alpha coronaviruses -NL63: PDB ID-3TLO), (Beta coronavirus- HKU1: PDB ID-3D23), (SARS-CoV: PDB ID-2AMQ), (MERS-CoV: PDB ID-4WMD), (SARS-CoV-2:PDB ID-6LU7)]. Sequence alignment was done using PRALINE tool and observed some active site residues similarities among all the strains. Blue to red color indicates un-conserved residues to conserved residues. (B) Sequence alignment between SARS and SARS-CoV-2 proteases. Pairwise sequence alignment revealed more than 97% sequence similarities among the sequences. (C) Phylogenetic tree of SARS related main protease. Phylogenetic tree was constructed using phelogeny.fr. and identified that bat- derived SARS main protease is conspecific with human SARS-CoV-2 main protease.

amino acids residues share 100% sequence homology. Consistent with the initial genomic studies, sequence analysis studies exposed wide range of similarities between SARS-CoV and SARS-CoV-2 proteases (Figure 3B).

Phylogeny tree was constructed using phelogeny.fr, to further examine the evolutionary relationship between SARS-CoV-2 main protease in comparison with other corona virus main protease. From the phylogenetic tree of main protease, it appeared that the route of its evolution is from Alpha corona virus. In agreement with earlier report (Ul Qamar et al., 2020), our results also suggest that Beta-CoV and MERS-CoV protease share more similarity than others (Figure 3C). The study also revealed that bat-derived SARS main protease is conspecific with human SARS-CoV-2 main protease and in sister relationships to Civet, BtBf and SARS-CoV-main proteases.

Superimposition of SARS and SARS-CoV-2 proteases structures

Pairwise sequence alignment determines the sequence similarities between proteases of SARS and SARS-CoV-2. The structural similarities between these proteases were analysed by superimposition of their 3-D structures. SARS-CoV main protease is reported to contain two chymotrypsin-like β -domain (residues 8-101, and 102-184) and an α -helical domain (residues 201-303) (Yang et al., 2003). Superimposed proteins exhibit high structural similarities among these domains of the two proteases with the active sites located between the domains I and II (Figure 4A). The inhibitor binding pockets of SARS-CoV and SARS-CoV-2 main protease were further closely examined to validate the binding region

among the structures. Study indicated that amino acids that forms the binding pocket of protease inhibitor His41, Met49, Pro140-Cys145, His163-Leu167, His169, Val186, Gln189, Thr190 and Gln 192 are conserved and form similar type of binding pockets in SARS-CoV and SARS-CoV-2 proteases (Figure 4B and C). The typical catalytic dyad Cys (145) His (41) present in the active site of SARS-CoV protease (Yang et al., 2003) superimposed with the corresponding dyad present in SARS-CoV-2 protease.

Protease inhibitors binding to SARS-CoV-2 and host cell proteases

Docking of anti-viral protease inhibitors to SARS-CoV-2 main protease

Molecular docking studies were performed using a set of antiviral drugs belonging to class of protease inhibitors, on 3-CL protease of SARS-CoV-2 and the results are given in Table 1. 2123 poses were generated, when 16 protease inhibitors were docked against SARS-CoV-2 protease, of which 199 poses were found when docking was done with Ritonavir to SARS-CoV-2 protease. Ritonavir also showed highest LibDock score to the target protein followed by Indinavir, Atazanavir and Lopinavir in that order (Table 1). The intermolecular H-bonds formed between each protease inhibitor and protease, were further analyzed and captured. The H-Bonds formed between the protease inhibitors and the amino acids in target proteins are having similarities with the inhibitor N3 binding residues. Among the different protease inhibitors, Ritonavir showed H-bond interaction with maximum number of active site residues unlike other

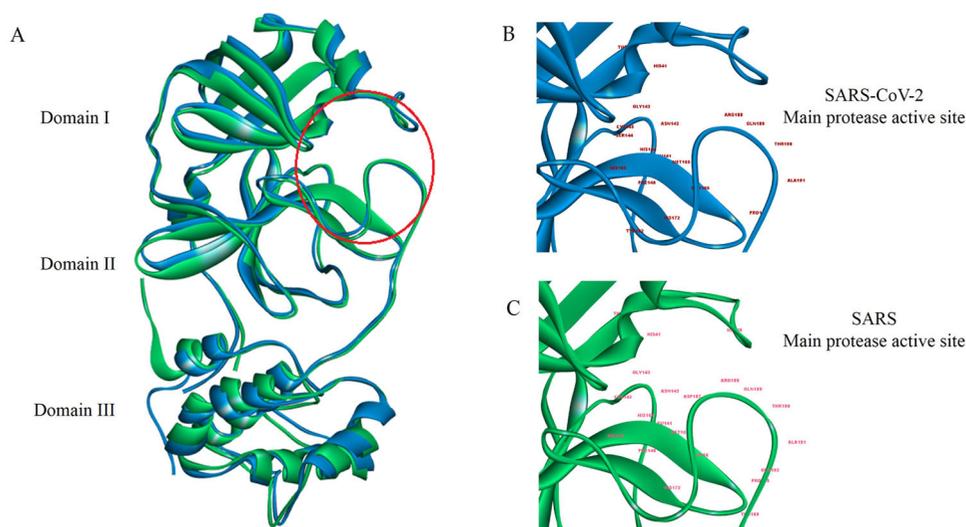


Figure 4. Comparison of crystal structures of main protease from SARS-CoV (PDB ID-2AMQ) and SARS-CoV-2 (PDB ID-6LU7). (A) Superimposition of SARS (green colour) and SARS-CoV-2 (blue colour) main proteases using 2020. Domains are labeled by numbers and the residues 8-101, 102-184, 201-303 represent domain I, II and III respectively. (B) Active site residues labelled in SARS-CoV-2 main protease. (C) Active site residues labelled in SARS main protease.

Table 1. Docking score and binding site residues of SARS-CoV-2 main protease against 16 protease inhibitors.

Target	Ligand	LibDock Score	Active site residues interacted through H-bond
SARS-Cov-2 protease	Ritonavir (CID_392622)	196.615	HIS41, GLY143, CYS145, GLU166, HIS164, GLN189, THR190
	Indinavir (CID_5362440)	182.293	LEU141, GLU166
	Atazanavir (CID_148192)	182.122	LEU141, GLU166, GLN189
	Lopinavir (CID_92727)	177.816	HIS41, LEU141, GLU166
	Saquinavir(CID_441243)	172.61	HIS164, GLU166, GLN189, THR190, GLN192
	Boceprevir * (CID_10324367)	171.794	HIS41, GLY143, CYS145, HIS164, GLU166, GLN189
	Fosamprenavir(CID_131536)	167.101	HIS41, PHE140, ASN142, SER144, HIS164, HIS163, GLU166
	Darunavir(CID_213039)	166.872	ASN142, HIS163, GLN189, THR190, HIS164, GLY143
	Tipranavir(CID_54682461)	166.864	HIS41, SER144, HIS163
	Nelfinavir(CID_64143)	165.774	LEU141, SER144, GLU166, HIS163, GLN189
	Amprenavir(CID_65016)	164.012	HIS41, ASN142, GLY143, SER144, CYS145, HIS164, GLN189, GLU166
	Grazoprevir* (CID_44603531)	163.342	LEU141, ASN142, GLY143, HIS164
	Simeprevir* (CID_24873435)	162.375	HIS41, ASN142, HIS164
	Paritaprevir* (CID_45110509)	153.766	ASN142, GLY143, CYS145, GLU166
	Telaprevir* (CID_3010818)	151.137	GLU166, ARG188
	Asunaprevir* (CID_16076883)	147.93	ASN142, SER144, HIS163, GLU166, GLN189, PHE140

Molecular docking studies were done using LibDock protocol in Discovery studio 2020 and the pose having highest docking score corresponding to each ligand is documented for further analysis. Study revealed that Ritonavir showed highest docking score and favourable interactions with amino acid residues in defined active site. * indicate HCV inhibitors.

protease inhibitors. It showed interaction with key active site residues including the Cys (145) His (41) dyad and other residues such as His (164), Gln (189), Glu (166) and Thr (190). The 2D interaction map of the H-bond formed between the Ritonavir and SARS-CoV-2 main protease docking result is illustrated in Figure 5.

In order to compare the molecular interactions of 16 protease inhibitors on SARS-CoV-2 protease with SARS-CoV main protease, molecular docking studies were performed against SARS-CoV protease. Docking generated 2155 poses of which 200 poses were found when docking of Ritonavir to SARS-CoV protease was done. The docking results also revealed that Ritonavir has the highest LibDock score (197.137) and favourable interacting bonds compared with other protease inhibitors.

Docking of protease inhibitors with host cell proteases

Further, we considered human proteases including TMPRSS2, Cathepsins B and L which play an important role in activating viral spike protein required for viral particle attachment

and fusion to the host cell, as targets to examine whether compounds that bind to viral proteases also dock to these host proteases.

TMPRSS2 has specific role in cutting spike protein of SARS-CoV and SARS-CoV-2 to cause viral infection. Studies have shown that blocking TMPRSS2 can prevent the entry of SARS-CoV to the host cell. By considering its role in SARS-CoV-2 anti-viral drug discovery, a homology model of human TMPRSS2 was made using Discovery studio 2020 (Figure 6A). The verified score of the modeled three dimensional structure of TMPRSS2 (159.001) was higher than the expected low score (71.5501). The model was then validated using PROCHECK and the results revealed that 87.4% residues are within the most favoured regions, 10.9% residues are within the additional allowed regions, 1.3% residues are within the generously allowed regions and 0.3% residues are within the disallowed regions. The Ramachandran Plot indicated that the modeled 3D structure of TMPRSS2 is an acceptable model and so it has been further used for molecular docking studies.

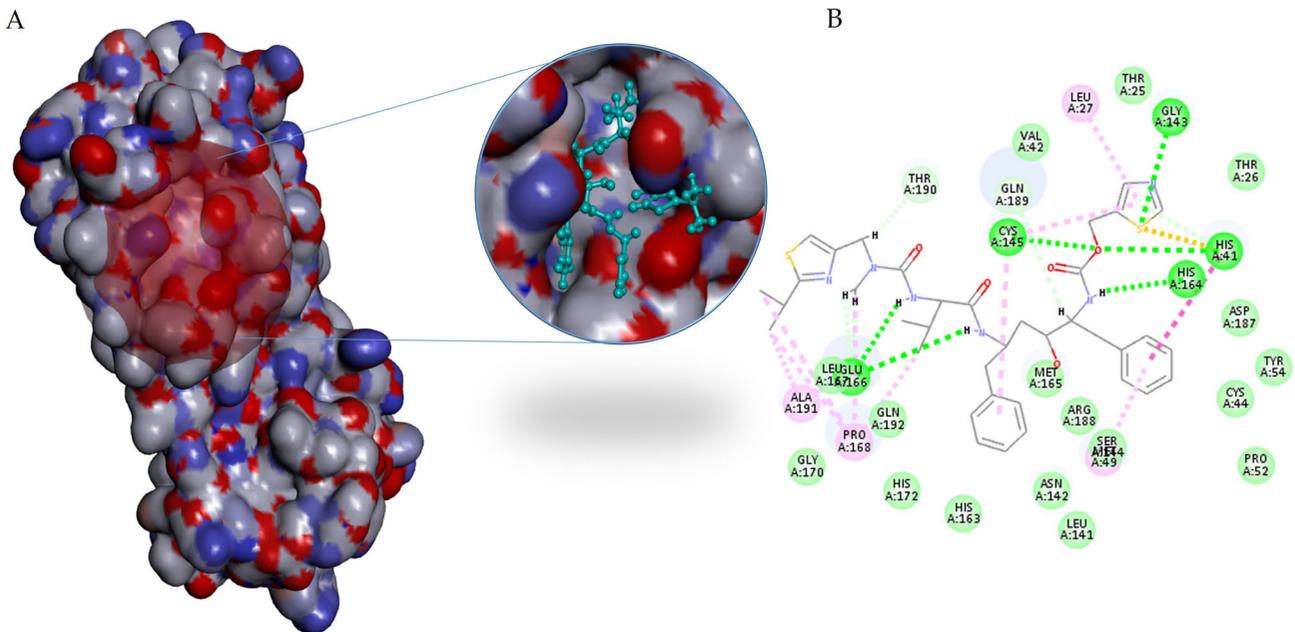


Figure 5. Docking of SARS-CoV-2 main protease against Ritonavir. Molecular docking of SARS-CoV-2 main protease (PDB ID:6LU7) against Ritonavir (CID_392622) showed highest docking score and favourable intermolecular interactions with in the 16 ligands. The image is a representative of docked pose with highest docking score (A) Ritonavir binds in the active site of SARS-CoV-2 main protease. (B) 2D interaction map of H-bonds formed between Ritonavir and amino acids in SARS-CoV-2 protease.

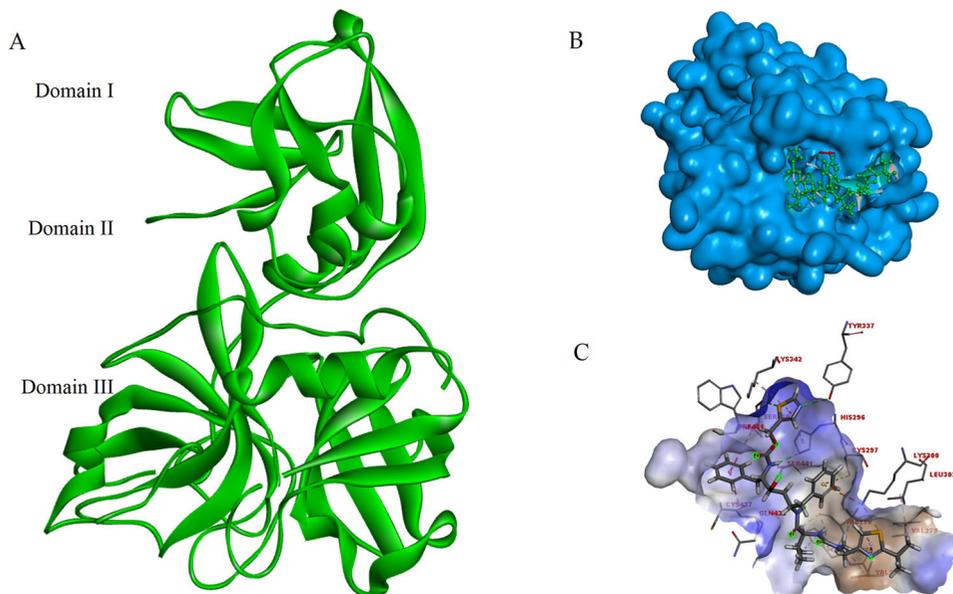


Figure 6. Modelled structure of TMPRSS2 and docking of TMPRSS2 protease against Ritonavir. (A) Homology modelled 3D structure of TMPRSS2 using Discovery studio 2020. Domains are labeled by numbers and the residues 1-84, 85-105, 106-492 represent domain I, II and III respectively. Amino acid residues His296, Asp345, Ser441 and Asp435 which constitute the active site were located in domain III (B) Ritonavir (CID_392622) binds in the active site of TMPRSS2 protease. The image is a representative of docked pose with highest docking score (C) Intermolecular interactions formed between Ritonavir and amino acids in TMPRSS2 protease.

Amino acid residues His296, Asp345, Ser441 and Asp435 which constitute the active sites were selected as defined active site and performed molecular docking. The docking results revealed that protease inhibitor Ritonavir, Atazanavir and Indinavir docked to TMPRSS2 with the LibDock score of 195.831, 185.41 and 178.747, respectively (Supplementary Table S3). The intermolecular interaction formed between Ritonavir and TMPRSS2 protease in docking is illustrated in

Figure 6B and C. Molecular docking was also performed with camostat mesylate, a clinically proven serine protease inhibitor active against TMPRSS2 (Kawase et al., 2012). The results showed that camostat mesylate has favourable amino acid interaction with LibDock score 116.771, which is less than that of the binding scores of protease inhibitors, particularly Ritonavir, indicating better binding of the selected protease inhibitors to the key host cell protease as well.

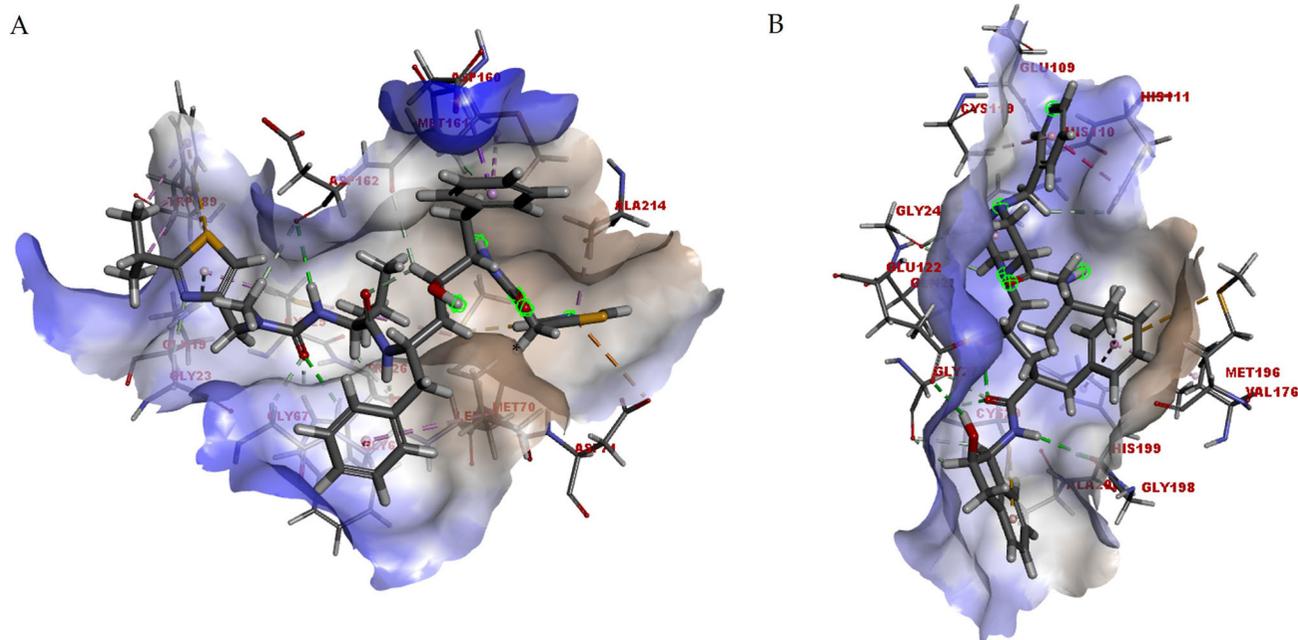


Figure 7. Docking view of Atazanavir in binding site of Cathepsin L (A) and Indinavir in binding site of Cathepsin B (B). Intermolecular interactions between the receptor in surface view and ligand corresponding to each interactions were drawn in Discovery studio 2020. Molecular docking of Cathepsin L (PDB ID- 2XU1) against Atazanavir (CID_148192) and Cathepsin B (PDB ID- 1CSB) against Indinavir (CID_5362440) showed highest docking score and favourable intermolecular interactions with in the 16 ligands. Each image is a representative of docked pose of the respective ligand with highest docking score.

Recent studies showed that cathepsin L and cathepsin B have important role in SARS-CoV-2 entry (Hoffmann et al., 2020). Molecular docking studies were carried out to investigate whether the listed protease inhibitors can block the enzymatic activity of cathepsin L and cathepsin B. Inhibitor binding regions in the PDB 3D structures corresponding to cathepsin L and cathepsin B were set as active sites in the molecular docking process. The results showed that Atazanavir, Ritonavir and Indinavir are the potential protease inhibitors of cathepsin L with the LibDock score of 166.186, 158.542 and 157.285, respectively (Supplementary Table S4). Potential protease inhibitors of cathepsin B are Indinavir, Darunavir and Amprenavir with the LibDock score of 169.347, 157.177 and 152.165, respectively (Supplementary Table S5). Analysis of the intermolecular interactions showed that the highly potential protease inhibitors indicated for each target protein has most favourable bonding with the same amino acids indicated in the PDB 3D structure. The intermolecular interaction formed between the Atazanavir with Cathepsin L and Indinavir with cathepsin B are illustrated in Figure 7.

Molecular dynamics

The molecular dynamics simulation using standard dynamics cascade in Discovery studio 2020, was performed to examine the structural stabilities of target proteins with its top ranked ligand. The trajectory analysis of simulated files provides root mean square deviations (RMSD) and root mean square fluctuation (RMSF) of each conformation. In order to identify the stability of selected three target-ligand complexes RMSD graph was plotted separately (Figure 8). SARS-CoV-2 main protease- Ritonavir complex showed equilibrium in about

40 ns and was stable around 1.50 Å (Figure 8A). TMPRSS2-Ritonavir also showed equilibrium in about 40 ns and was stable around 1.40 Å (Figure 8B). For Cathepsin L- Atazanavir complex, it attained stable state around 1.85 Å (Figure 8C) and showed equilibrium in about 40 ns. From the three RMSD graphs it showed interactions of ligand with their specific receptors and were stable during the simulation.

RMSF graph of each complex was plotted to identify the fluctuation at amino acid level. RMSF plots (Figure 9) of complexes showed the residue fluctuations were within an RMSF of 2 Å. Residues in SARS-CoV-2 main protease- Ritonavir (Figure 9A), TMPRSS2- Ritonavir (Figure 9B) complex and Cathepsin L- Atazanavir (Figure 9C) complex showed stable throughout the dynamics simulation.

MM-PBSA binding free energy

MM-PBSA analysis was performed with three selected protein-ligand complexes to elucidate the affinities between ligand and to the target protein. These calculations were done for the opted MD trajectories after its RMSD attained equilibrium. Here the snapshots were collected with an increment of 10 from the total conformations generated between 40 to 130 ns and the binding energy calculations done using 10 randomly selected snapshots of each. Changes in binding free energy of protein-ligand complexes at various time intervals are shown in Figure 10.

The SARS-CoV-2 main protease- Ritonavir complex showed the least net binding energy (-19.237 ± 0.529 kcal/mol) among the three complexes. Cathepsin L- Atazanavir complex showed net binding free energy (-18.934 ± 0.560 kcal/mol) and the TMPRSS2- Ritonavir complex showed the net binding energy of (-7.177 ± 0.691 kcal/mol). Comparison of

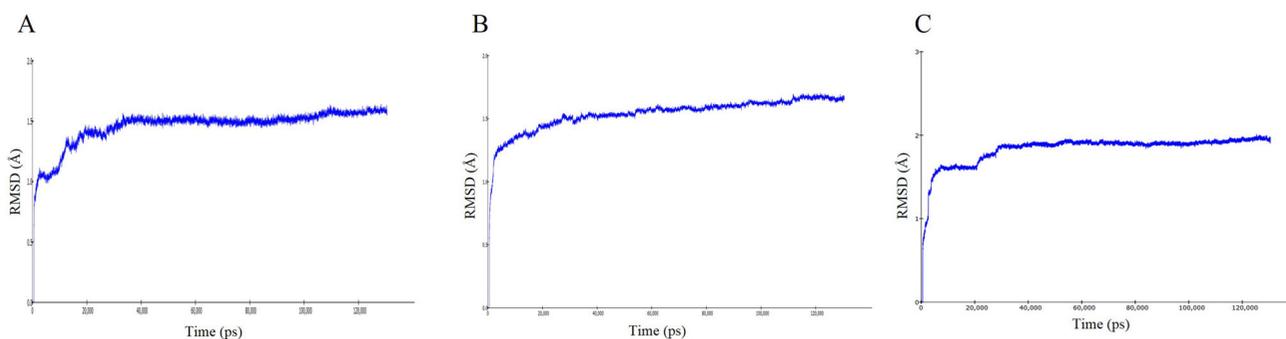


Figure 8. RMSD plot of top ranked protease inhibitors complexed with three different targets. (A) RMSD of SARS-CoV-2 main protease- Ritonavir complex showed that the structure was stable around 1.50 Å for about 130 ns. (B) RMSD of TMPRSS2- Ritonavir complex showed the complex was stable around 1.40 Å for about 130 ns. (C) Cathepsin L- Atazanavir complex showed the complex was stable around 1.85 Å for about 130 ns.

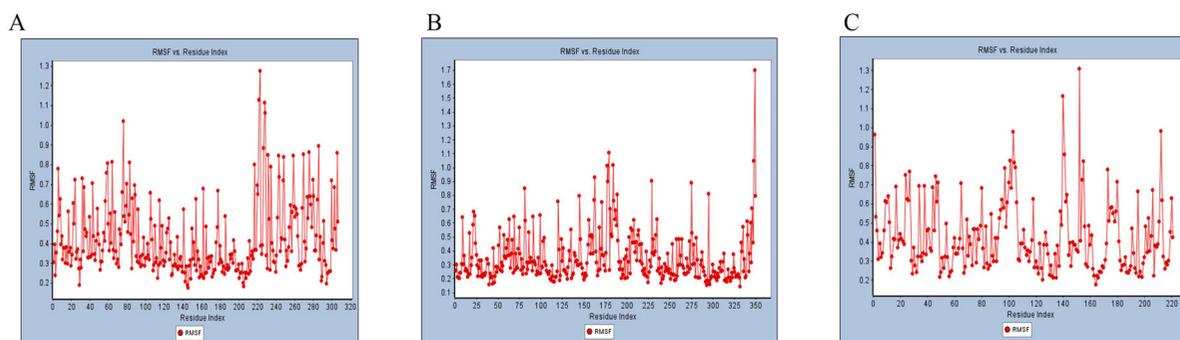


Figure 9. RMSF plot of three different targets with potential protease inhibitors. (A) RMSF plot of SARS-CoV-2 main protease- Ritonavir complex showed the amino acid residue fluctuation is within 1.3 Å. (B) RMSF plot of TMPRSS2- Ritonavir complex showed the residue fluctuation is within 1.7 Å. (C) RMSF plot of Cathepsin L- Atazanavir complex showed the fluctuation is within 1.3 Å.

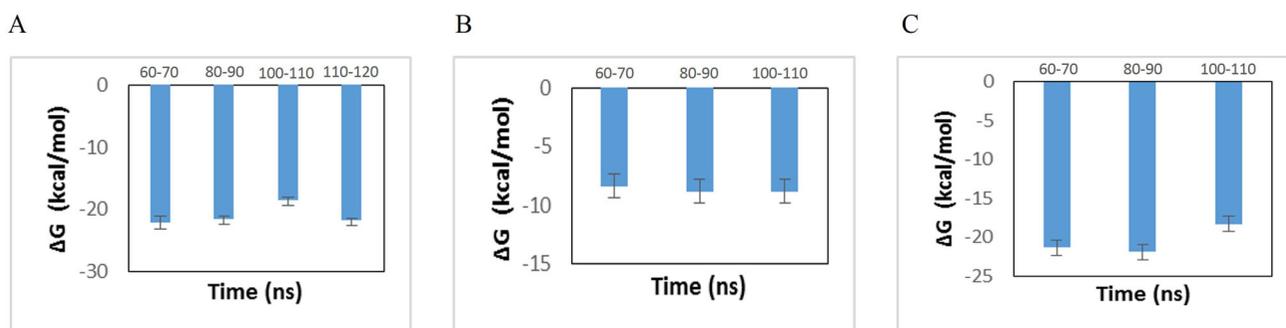


Figure 10. MM-PBSA of protein -ligand complexes at different time intervals. MM_PBSA calculations were done by binding energy calculation protocol in Discovery Studio 2020. A-Main protease –Ritonavir complex B-TMPRSS2-Ritonavir complex, C-Cathepsin-L-Atazanavir complex. Results given are the mean of multiple binding free energy values (in units of kcal/mol) \pm SEM at 10 ns intervals. Comparison of the mean values of each complex at different time intervals was made by one way ANOVA analysis. $p > 0.05$ not significant.

the binding free energies of each of the three complexes at various time intervals showed that they are not significantly different.

Hydrogen bond analysis

Hydrogen bonds contribute significantly to determining the strength of interaction between the target protein and ligand. Therefore, Hydrogen bond analyses of a range of conformations of the protein-ligand complex formed between 40 ns to 130 ns simulation were performed. This investigated the stability of hydrogen bond between the ligand and the residues in the active site of proteins (Figure 11). In SARS-CoV-2 main protease- Ritonavir complex (Figure 11A), out of the eleven hydrogen bonds formed between the target and

the ligand, four hydrogen bonds showed consistency in the whole simulation. For TMPRSS2- Ritonavir (Figure 11B) complex, two hydrogen bonds (out of seven) retained its consistency, while in the case of Cathepsin L- Atazanavir complex (Figure 11C) six hydrogen bonds (out of nine) showed consistency during the whole simulation. The hydrogen bond analysis also revealed that the key active site residues that formed the hydrogen bond in the protein-ligand complex showed no variation during the simulation.

Discussion

The lack of adequate information about the pathophysiology of COVID-19, absence of a potential prophylactic vaccine and

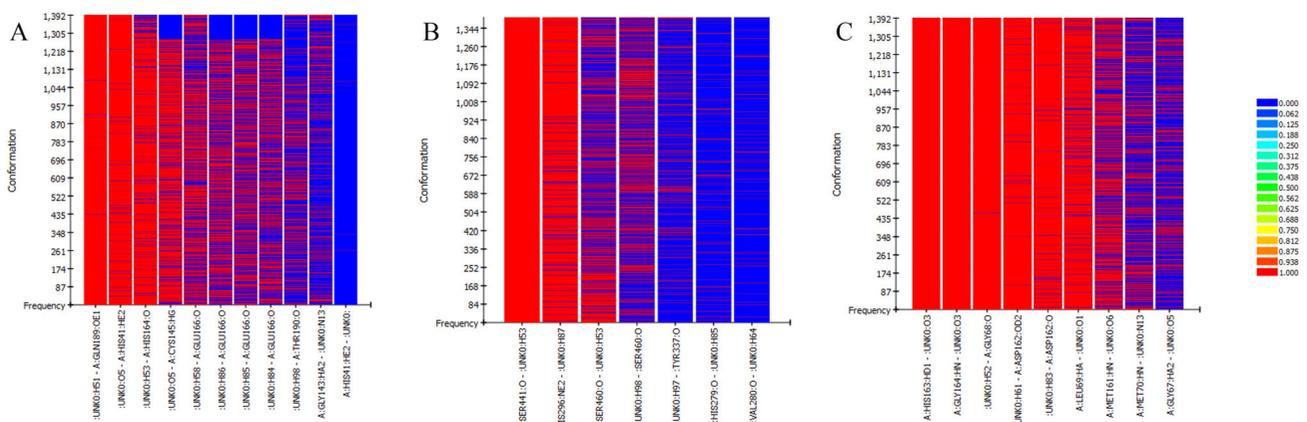


Figure 11. Heat map representing the consistency among the hydrogen bonds formed in three complexes. (A) SARS-CoV-2 main protease- Ritonavir complex showed four consistent hydrogen bonds during the simulation. (B) TMPRSS2- Ritonavir complex showed two consistent hydrogen bonds during the simulation. (C) Cathepsin L- Atazanavir complex showed six consistent hydrogen bonds during the simulation.

a well-defined treatment protocol or a potential drug to control SARS-CoV-2 infection, have led to the recently emerged life-threatening viral disease to spread at an alarming rate, to assume pandemic proportions. Even as no potential therapeutic agents were identified, the emergency of the outbreak lead to the use of FDA approved anti-viral drugs as a trial, without enough supporting evidence for their efficacy against SARS-COV-2 infection. Drugs such as Chloroquine, hydroxychloroquine, Japan flu drug, EIDD-2801, HIV drugs, arthritis drug, remdesivir, a failed Ebola drug, are now used to control the disease. Though in a recent multicentric clinical trial, remdesivir treated patients show speedy recovery, there is no change in the rate of mortality (Wang et al., 2020). Several drugs belonging to the category of protease inhibitors that target the key viral protease, which cleaves the viral polyprotein that generate functional proteins required for viral maturation, have been approved for medical use, particularly against HIV/HCV viral infection. Given the critical role of the host proteases such as TMPRSS2/cathepsins and the viral main protease, 3-C-like protease in SARS-COV-2 infection, we have examined the possibility of re-purposing the known antiviral protease inhibitors against SARS-COV-2. The results presented above show that protease inhibitors such as Ritonavir, Indinavir, Atazanavir can target both host protease and viral main protease. The potential of these molecules for therapeutic use against SARS-COV-2 is evidenced by the following. (a) The sequence of CoV-2 S-protein, whose proteolytic cleavage by host protease is critical in virus attachment and entry into the cell, is relatively conserved across different isolates. This showed the presence of all the four protease cleavage sites that are present in SARS-CoV S -protein with which it shares 87% sequence similarity. (b) All the four proteolytic cleavage sites are conserved in different isolates tested. The TMPRSS2 cleavage site in SARS-CoV-2S protein is more basic than that in SARS-COV S-protein. (c) The sequence of SARS-CoV-2 main protease, 3-C-like protease is relatively conserved in different isolates and showed about 97% sequence similarity with that of SARS-CoV main protease. It's domain structure superimposed with that of SARS-CoV main protease, with conserved active site residues that formed similar type of binding pockets (d)

Molecular docking studies showed that anti retroviral compounds, particularly approved for medical use against HIV/HCV infection, docked at the active site of SARS-CoV-2 3-CL-protease with binding affinity in the order Ritonavir > Indinavir > Atazanavir, indicating their potential to inhibit the protease and (e) Molecular Docking studies with TMPRSS2 also showed binding of the antiviral protease inhibitors at its active site with docking scores in the order Ritonavir > Atazanavir > Indinavir > Lopinavir; all of them showed greater docking scores than camostat mesylate, a clinically tested inhibitor of TMPRSS2. The results thus suggest that the protease inhibitor, particularly Ritonavir approved for treatment of HIV infection and has been shown to have antiviral potential against SARS-CoV, can target both viral 3-CL-protease and host TMPRSS2. Based on the molecular docking studies, we prioritized the drug and their combinations using docking scores and favourable interactions formed between drug and the target protein. Ritonavir showed the highest binding affinity with viral main protease and host protease TMPRSS2. Indinavir and Atazanavir are the other protease inhibitors showing best docking score with other host target enzymes such as Cathepsin B/L which are also involved in S-protein priming and fusion (Hoffmann et al., 2020).

Proteolytic cleavage of the SARS-CoV-2 S- protein is a critical event in its activation and binding with the ACE2 receptor and fusion with the host cell membrane. It has been reported that TMPRSS2, belonging to a family of cell surface trans-membrane serine protease, is a key enzyme involved in the activation of SARS-CoV and SARS-CoV-2 S-protein (Hoffmann et al., 2020; Wilson et al., 2005). The TMPRSS2 cleavage site in SARS-COV-2S protein, unlike in SARS-CoV S-protein, is a poly basic site RRARS flanked at the N terminal by a proline residue confirming the earlier reports (Coutard et al., 2020; Wrapp et al., 2020). This is conserved in all the isolates analysed. Such polybasic cleavage sites are found commonly in highly virulent virus spike proteins such as human influenza virus spike protein (Chen et al., 1998). Apart from the proteolytic cleavage site, we also identified the receptor ACE2 binding motif similarities in SARS-CoV and SARS-CoV-2S protein and identified the homology between

them. Multiple sequence alignment done among the S protein of hundred isolates also confirms the sequence conservation of motif in the S- protein of all the isolates. Sequences of a large number of isolates from different locations are deposited in public data base almost on a daily basis and it requires extension of these analyses for further validation.

Besides, proteolytic cleavage sites and receptor binding sites, S proteins are glycosylated at specific asparagine residues and these N-linked glycans are involved in folding of spike protein, antibody recognition and priming by host proteases. The present study revealed that SARS-CoV-2 S protein comprises of 17 N-linked glycosylation sites, which are highly conserved in SARS-CoV and SARS-CoV-2. Experimental data also confirm the position of 17 predicted glycosylation sites in SARS-CoV (Walls et al., 2020; Watanabe et al., 2019; Bagdonaite et al., 2015). Apart from N-linked glycosylation sites, we also identified four possible threonine O-linked glycosylation sites in SARS-CoV-2 S- protein, unlike in SARS-CoV. Studies revealed that O-linked glycosylation also has crucial role for viral particle formation in host cells and infectivity rather than N-linked glycosylation.

The anti-viral compounds that target virus proteases have been designed and tested in HIV infected patients in the last decade. The key role of viral protease was established experimentally by demonstrating loss of infectivity on mutation of critical residues in viral proteases. *In vitro* assays also provided evidence for HIV proteases as a potent target for anti-viral drug (Patick & Potts, 1998). Clinical trials also confirmed the efficacy and safety of protease inhibitors used for HIV infection and the currently approved HIV protease inhibitors are Amprenavir, Atazanavir, Darunavir, Fosamprenavir, Indinavir, Lopinavir, Nelfinavir, Ritonavir, Saquinavir and Tipranavir. Apart from HIV protease inhibitors, some new protease inhibitors have also been approved to treat chronic hepatitis C infection safely. Clinically proven and approved HCV protease inhibitors include Asunaprevir, Boceprevir, Grazoprevir, Paritaprevir, Simeprevir and Telaprevir. Even the protease inhibitors that are used for treating HIV and HCV, are also recommended to treat other viral infections including SARS- CoV and recently against SARS-CoV-2 as well (Tu et al., 2020; Zhou et al., 2020). Previous experimental study (Yamamoto et al., 2004), showed that Nelfinavir inhibits replication of SARS-associated coronavirus *in vitro*. Introduction of a combination antiretroviral therapy has improved significantly the management of viral infection, particularly of HIV patients. Combination of drugs that target different molecules such as viral main protease and RNA replicase or drugs that target viral protease along with a booster of protease inhibitor action appears effective in reducing viral load. Certain control studies have shown the effectiveness of the combination of Lopinavir/Ritonavir against HIV and SARS-CoV infections (Chan et al., 2003). Favourable clinical response of SARS patients to a combination of Lopinavir/Ritonavir and Ribavirin was also reported (Chu et al., 2004). Atazanavir/Ritonavir combination was also reported to be as effective as Lopinavir/Ritonavir (Molina et al., 2008). Ritonavir, apart from inhibiting viral main protease, inhibits cytochrome P4503A4

(CYP3A4) isoenzyme whereby the rate of degradation of partner drugs such as lopinavir is reduced, resulting in increase in the serum levels of lopinavir. Accordingly, Lopinavir/Ritonavir combination therapy for COVID 19 patients has been suggested. Although in one of the recently reported trials, no additional benefit for such a combination therapy was observed (Cao et al., 2020), in the multi country 'Solidarity' trial initiated by WHO, such a combination therapy has also been included in one of the study arms. Hundreds of clinical trials, several of which include such combination drug regimen against COVID-19 are underway for discovering effective treatment (Lythgoe & Middleton, 2020).

The results presented above provide support for re-purposing the protease inhibitor drugs such as Ritonavir and Atazanavir, which are approved for medical use against other viral disease, also against SARS-COV-2 infection. Ritonavir has the highest docking score and favourable interactions to SARS-CoV-2 main protease. It showed binding to key active site residues of the main protease. Our results showed that Ritonavir also targets host (Human) proteases such as TMPRSS2, that has a crucial role in S-protein cleavage and priming for entry of virion into the cell apart from viral main protease. Ritonavir showed better docking score and favourable interaction with the active site residues than the clinically tested inhibitor of TMPRSS2. Atazanavir is an azapeptide that binds to the active site and inhibits the action of the viral protease. It is pharmacologically related but structurally different from other protease inhibitors and shows docking, with parameters comparable to Ritonavir, with both SARS-CoV-2 main protease and TMPRSS2. Moreover, Atazanavir also showed the best docking score and favourable interaction with the active site residues of cathepsin L, which is involved in virus fusion and transport in endosomes.

To obtain more insights into the stability of protein-ligand complexes, we then performed molecular dynamics simulation. The results of dynamics are represented according to the RMSD (a standard measure of the structural distance between coordinates) and RMSF (a measure of residue fluctuation) values obtained during simulation. The SARS-CoV-2 main protease- Ritonavir complex showed fluctuating RMSD between 5 to 40 ns in the range of 1.00 Å to 1.45 Å, and attained stability at 1.50 Å after 40 ns. TMPRSS2- Ritonavir complex also showed fluctuating RMSD up to 40 ns at RMSD range between 1.25 Å to 1.40 Å and large fluctuation observed between 30 ns to 35 ns (RMSD > 1.40 Å), and attained stability at 1.45 Å after 40 ns. The Cathepsin L- Atazanavir complex showed higher stable RMSD 1.85 Å between 40 ns to 130 ns after fluctuating RMSD in the range of 1.50 Å to 1.75 Å. between 15 to 40 ns. These results were also confirmed by calculating residue fluctuations by RMSF plot. SARS-CoV-2 main protease- Ritonavir complex and Cathepsin L- Atazanavir complex showed RMSF less than 1.3 Å at 130 ns simulation while TMPRSS2- Ritonavir complex showed RMSF less than 1.7 Å at 130 ns simulation. RMSF confirms that the residue fluctuations of three protein ligand complexes are within the range of 2 Å, suggesting no significant conformational changes during binding.

Ligand binds with high affinity to the active site on the target protein; binding energy calculations are employed to identify the accurate disposition of ligand in to the active site. Of late, MM-PBSA approach is commonly employed to calculate the absolute ΔG_{bind} when long molecular dynamics simulations are done in association with molecular docking studies, particularly when sufficient information about the bioactive conformation of the protein-ligand complex are not available. Moreover, this approach can be used to screen out active compounds and its structure-activity relationship, virtual screening and improve the results of molecular docking (Poli et al., 2020). In addition to the docking parameters, thermodynamic parameters of molecular dynamics conformations have also been analysed in the present study. Binding free energies are calculated from different conformations generated from the whole molecular dynamics simulation over 130 ns after it reaches equilibrium state. The absolute free energy is obtained by taking the average of energies determined for different conformations. More negative MM-PBSA ΔG_{bind} value indicates stronger binding of the target protein with ligand. The SARS-CoV-2 main protease-Ritonavir complex showed greater absolute binding free energy than Cathepsin L-Atazanavir complex and the TMPRSS2-Ritonavir complex. The MM-PBSA result revealed that all three complexes are stable but among that SARS-CoV-2 main protease-Ritonavir complex is more stable and energetically favourable. The stability of interaction between the ligand and protein was further examined by hydrogen bond analysis. The consistency of intermolecular hydrogen bond during the MD simulation among the three complexes also suggested the binding stability of the complexes. In SARS-CoV-2 main protease-Ritonavir complex, active site residues HIS41, GLY143, CYS145, GLU166, HIS164, GLN189 and THR190 of the protein formed eleven hydrogen bonds. Of these, four hydrogen bonds formed by the key residues GLN189, HIS41, HIS164 and CYS145 showed consistency among the conformations during the whole simulation. For TMPRSS2-Ritonavir complex, two hydrogen bonds formed by the residues HIS296 and SER441 in TMPRSS2 retain its consistency during the whole simulation. Six hydrogen bonds formed by the key active site residues (ASP162, GLY68, HIS163 and GLY67) in Cathepsin L with Atazanavir showed consistency among the nine hydrogen bonds during the whole simulation. The protein-ligand stability determined by molecular dynamics simulation, hydrogen bond analysis, absolute binding energy and the relative docking scores suggest that a combination of Atazanavir/Ritonavir would target not only SARS-CoV-2 main protease, but also the host proteases such as TMPRSS2.

Prioritization of the results of the docking study on various protease targets critically involved in the life cycle SARS-CoV-2 therefore, suggests that a combination therapy with Atazanavir/Ritonavir would be useful against COVID-19. Though the inhibitory effect of Ritonavir and Atazanavir on SARS main protease is proven, experimental validation of the effect of these molecules on TMPRSS2 was lacking hitherto. Further pharmacokinetic analysis would be required to fix the relative dose to obtain an optimal effect on these targets so that an effective reduction in viral load is achieved.

Acknowledgements

The authors acknowledge the AICADD and DBT-BIF support at the Department of Computational Biology and Bioinformatics, University of Kerala, India for providing the necessary facilities to carry out this study.

Disclosure statement

The authors declare no conflict of interest.

Funding

This study was supported by the SIUCEB at the Department of Computational Biology and Bioinformatics, University of Kerala, India. P.R.S. was supported by ISCA, Kolkata by Asutosh Mookerjee Fellowship.

Author contributions

CSA and PRS conceived and designed the study. CSA and AK performed the experiments. CSA, ASN, AK, OVO and PRS analyzed the data. Manuscript written by CSA and PRS, edited by OVO and ASN. All authors read and approved the final manuscript.

References

- Bagdonaite, I., Nordén, R., Joshi, H. J., Dabelsteen, S., Nyström, K., Vakhrushev, S. Y., Olofsson, S., & Wandall, H. H. (2015). A strategy for O-glycoproteomics of enveloped viruses—the O-glycoproteome of herpes simplex virus type 1. *PLoS Pathogens*, 11(4), e1004784. <https://doi.org/10.1371/journal.ppat.1004784>
- Bhardwaj, V. K., Singh, R., Sharma, J., Rajendran, V., Purohit, R., & Kumar, S. (2020). Identification of bioactive molecules from tea plant as SARS-CoV-2 main protease inhibitors. *Journal of Biomolecular Structure and Dynamics*, 1–10. <https://doi.org/10.1080/07391102.2020.1766572>
- 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., ... Wang, C. (2020). A trial of lopinavir-ritonavir in adults hospitalized with severe covid-19. *The New England Journal of Medicine*, 382(19), 1787–1799. <https://doi.org/10.1056/NEJMoa2001282>
- Chan, K. S., Lai, S. T., Chu, C. M., Tsui, E., Tam, C. Y., & Wong, M. M. (2003). Treatment of severe acute respiratory syndrome with lopinavir/ritonavir: A multicentre retrospective matched cohort study. *Hong Kong Medical Journal*, 9(6), 399–406.
- Chen, J., Lee, K. H., Steinhauer, D. A., Stevens, D. J., Skehel, J. J., & Wiley, D. C. (1998). Structure of the hemagglutinin precursor cleavage site, a determinant of influenza pathogenicity and the origin of the labile conformation. *Cell*, 95(3), 409–417. [https://doi.org/10.1016/S0092-8674\(00\)81771-7](https://doi.org/10.1016/S0092-8674(00)81771-7)
- Chen, Y., Liu, Q., & Guo, D. (2020). Emerging coronaviruses: Genome structure, replication, and pathogenesis. *Journal of Medical Virology*, 92(4), 418–423. <https://doi.org/10.1002/jmv.25681>
- Chen, Y. W., Yiu, C. B., & Wong, K. Y. (2020). Prediction of the SARS-CoV-2 (2019-nCoV) 3C-like protease (3CL pro) structure: virtual screening reveals velpatasvir, ledipasvir, and other drug repurposing candidates. *F1000Research*, 9, 129. doi: <https://doi.org/10.12688/f1000research.22457.1>
- Chu, C. M., Cheng, V. C., Hung, I. F., Wong, M. M., Chan, K. H., & Chan, K. S. (2004). Role of lopinavir/ritonavir in the treatment of SARS: Initial virological and clinical findings. *Thorax*, 59(3), 252–256. <https://doi.org/10.1136/thorax.2003.012658>
- Coutard, B., Valle, C., de Lamballerie, X., Canard, B., Seidah, N. G., & Decroly, E. (2020). The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. *Antiviral Research*, 176, 104742. <https://doi.org/10.1016/j.antiviral.2020.104742>

- Cui, J., Li, F., & Shi, Z. L. (2019). Origin and evolution of pathogenic coronaviruses. *Nature Reviews Microbiology*, 17(3), 181–192. <https://doi.org/10.1038/s41579-018-0118-9>
- Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.-F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.-M., & Gascuel, O. (2008). Phylogeny.fr: Robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research*, 36(Web Server issue), W465–W469. <https://doi.org/10.1093/nar/gkn180>
- Fung, T. S., & Liu, D. X. (2019). Human coronavirus: Host-pathogen interaction. *Annual Review of Microbiology*, 73(1), 529–557. <https://doi.org/10.1146/annurev-micro-020518-115759>
- Gupta, R., & Brunak, S. (2002). Prediction of glycosylation across the human proteome and the correlation to protein function. In *Pacific Symposium on Biocomputing*, 7, 310–322.
- Havranek, B., & Islam, S. M. (2020). An in silico approach for identification of novel inhibitors as potential therapeutics targeting COVID-19 main protease. *Journal of Biomolecular Structure and Dynamics*, 38, 1–12. <https://doi.org/10.1080/07391102.2020.1776158>
- Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N.-H., Nitsche, A., Müller, M. A., Drosten, C., & Pöhlmann, S. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*, 181(2), 271–280. <https://doi.org/10.1016/j.cell.2020.02.052>
- Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., ... Cao, B. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan. *Lancet*, 395(10223), 497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5)
- Islam, R., Parves, M. R., Paul, A. S., Uddin, N., Rahman, M. S., & Mamun, A. A. (2020). A molecular modeling approach to identify effective antiviral phytochemicals against the main protease of SARS-CoV-2. *Journal of Biomolecular Structure and Dynamics*, 1–12. <https://doi.org/10.1080/07391102.2020.1761883>
- Iwata-Yoshikawa, N., Okamura, T., Shimizu, Y., Hasegawa, H., Takeda, M., & Nagata, N. (2019). TMPRSS2 contributes to virus spread and immunopathology in the airways of murine models after coronavirus infection. *Journal of Virology*, 93(6), 18. <https://doi.org/10.1128/JVI.01815-18>
- Jin, Y. H., Cai, L., Cheng, Z. S., Cheng, H., Deng, T., & Fan, Y. P. (2020). A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version). *Military Medical Research*, 7, 1–23. <https://doi.org/10.1186/s40779-020-0233-6>
- Julenius, K., Mølgaard, A., Gupta, R., & Brunak, S. (2005). Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology*, 15(2), 153–164. <https://doi.org/10.1093/glycob/cwh151>
- Kawase, M., Shirato, K., van der Hoek, L., Taguchi, F., & Matsuyama, S. (2012). Simultaneous treatment of human bronchial epithelial cells with serine and cysteine protease inhibitors prevents severe acute respiratory syndrome coronavirus entry. *Journal of Virology*, 86(12), 6537–6545. <https://doi.org/10.1128/JVI.00094-12>
- Khan, S. A., Zia, K., Ashraf, S., Uddin, R., & Ul-Haq, Z. (2020). Identification of chymotrypsin-like protease inhibitors of SARS-CoV-2 via integrated computational approach. *Journal of Biomolecular Structure and Dynamics*, 1–10. <https://doi.org/10.1080/07391102.2020.1751298>
- Lai, S. T. (2005). Treatment of severe acute respiratory syndrome. *European Journal of Clinical Microbiology & Infectious Diseases*, 24(9), 583–591. <https://doi.org/10.1007/s10096-005-0004-z>
- Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26(2), 283–291. <https://doi.org/10.1107/S0021889892009944>
- 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., ... Feng, Z. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *The New England Journal of Medicine*, 382(13), 1199–1207. <https://doi.org/10.1056/NEJMoa2001316>
- Lythgoe, M. P., & Middleton, P. (2020). Ongoing Clinical Trials for the Management of the COVID-19 Pandemic. *Trends in Pharmacological Sciences*, 41(6), 363–382. <https://doi.org/10.1016/j.tips.2020.03.006>
- Millet, J. K., & Whittaker, G. R. (2015). Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. *Virus Research*, 202, 120–134. <https://doi.org/10.1016/j.virusres.2014.11.021>
- Molina, J.-M., Andrade-Villanueva, J., Echevarria, J., Chetochisakd, P., Corral, J., David, N., Moyle, G., Mancini, M., Percival, L., Yang, R., Thiry, A., & McGrath, D. (2008). Once-daily atazanavir/ritonavir versus twice-daily lopinavir/ritonavir, each in combination with tenofovir and emtricitabine, for management of antiretroviral-naïve HIV-1-infected patients: 48 week efficacy and safety results of the CASTLE study. *The Lancet*, 372(9639), 646–655. [https://doi.org/10.1016/S0140-6736\(08\)61081-8](https://doi.org/10.1016/S0140-6736(08)61081-8)
- Ou, X., Liu, Y., Lei, X., Li, P., Mi, D., Ren, L., Guo, L., Guo, R., Chen, T., Hu, J., Xiang, Z., Mu, Z., Chen, X., Chen, J., Hu, K., Jin, Q., Wang, J., & Qian, Z. (2020). Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nature Communications*, 11(1), 1620. <https://doi.org/10.1038/s41467-020-15562-9>
- Patick, A. K., & Potts, K. E. (1998). Protease inhibitors as antiviral agents. *Clinical Microbiology Reviews*, 11(4), 614–627. <https://doi.org/10.1128/CMR.11.4.614>
- Poli, G., Granchi, C., Rizzolio, F., & Tuccinardi, T. (2020). Application of MM-PBSA methods in virtual screening. *Molecules*, 25(8), 1971. <https://doi.org/10.3390/molecules25081971>
- Rao, S. N., Head, M. S., Kulkarni, A., & LaLonde, J. M. (2007). Validation studies of the site-directed docking program LibDock. *Journal of Chemical Information and Modeling*, 47(6), 2159–2171. <https://doi.org/10.1021/ci6004299>
- Reinke, L. M., Spiegel, M., Plegge, T., Hartleib, A., Nehlmeier, I., Gierer, S., Hoffmann, M., Hofmann-Winkler, H., Winkler, M., & Pöhlmann, S. (2017). Different residues in the SARS-CoV spike protein determine cleavage and activation by the host cell protease TMPRSS2. *PLoS One*, 12(6), e0179177. <https://doi.org/10.1371/journal.pone.0179177>
- Rice, P., Longden, I., & Bleasby, A. (2000). EMBOS: The European Molecular Biology Open Software Suite. *Trends in Genetics: TIG*, 16(6), 276–277. [https://doi.org/10.1016/S0168-9525\(00\)02024-2](https://doi.org/10.1016/S0168-9525(00)02024-2)
- Sanche, S., Lin, Y. T., Xu, C., Romero-Severson, E., Hengartner, N., & Ke, R. (2020). High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2. *Emerging Infectious Diseases*, 26(7), 1470–1477. <https://doi.org/10.3201/eid2607.200282>
- Shirato, K., Kawase, M., & Matsuyama, S. (2018). Wild-type human coronaviruses prefer cell-surface TMPRSS2 to endosomal cathepsins for cell entry. *Virology*, 517, 9–15. <https://doi.org/10.1016/j.virol.2017.11.012>
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7, 539. <https://doi.org/10.1038/msb.201175>
- Simossis, V. A., & Heringa, J. (2005). PRALINE: A multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. *Nucleic Acids Research*, 33(Web Server issue), W289–W294. <https://doi.org/10.1093/nar/gki390>
- Tu, Y.-F., Chien, C.-S., Yarmishyn, A. A., Lin, Y.-Y., Luo, Y.-H., Lin, Y.-T., Lai, W.-Y., Yang, D.-M., Chou, S.-J., Yang, Y.-P., Wang, M.-L., & Chiou, S.-H. (2020). A review of SARS-CoV-2 and the ongoing clinical trials. *International Journal of Molecular Sciences*, 21(7), 2657. <https://doi.org/10.3390/ijms21072657>
- Ul Qamar, M. T., Alqahtani, S. M., Alamri, M. A., & Chen, L. L. (2020). Structural basis of SARS-CoV-2 3CL(pro) and anti-COVID-19 drug discovery from medicinal plants. *Journal of Pharmaceutical Analysis*, 10(4), 313–319. <https://doi.org/10.1016/j.jpha.2020.03.009>
- Walls, A. C., Park, Y. J., Tortorici, M. A., Wall, A., McGuire, A. T., & Veersler, D. (2020). Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell*, 181(2), 281–292. <https://doi.org/10.1016/j.cell.2020.02.058>
- Wang, Y., Zhang, D., Du, G., Du, R., Zhao, J., Jin, Y., Fu, S., Gao, L., Cheng, Z., Lu, Q., & Hu, Y. (2020). Remdesivir in adults with severe COVID-19: A randomised, double-blind, placebo-controlled, multicentre trial. *Lancet*, 395, 1569–1578. [https://doi.org/10.1016/S0140-6736\(20\)31022-9](https://doi.org/10.1016/S0140-6736(20)31022-9)

- Watanabe, Y., Bowden, T. A., Wilson, I. A., & Crispin, M. (2019). Exploitation of glycosylation in enveloped virus pathobiology. *Biochimica et Biophysica Acta. General Subjects*, 1863(10), 1480–1497. <https://doi.org/10.1016/j.bbagen.2019.05.012>
- Wilson, S., Greer, B., Hooper, J., Zijlstra, A., Walker, B., Quigley, J., & Hawthorne, S. (2005). The membrane-anchored serine protease, TMPRSS2, activates PAR-2 in prostate cancer cells. *The Biochemical Journal*, 388(Pt 3), 967–972. <https://doi.org/10.1042/BJ20041066>
- Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C.-L., Abiona, O., Graham, B. S., & McLellan, J. S. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, 367(6483), 1260–1263. <https://doi.org/10.1126/science.abb2507>
- Yamamoto, N., Yang, R., Yoshinaka, Y., Amari, S., Nakano, T., Cinatl, J., Rabenau, H., Doerr, H. W., Hunsmann, G., Otaka, A., Tamamura, H., Fujii, N., & Yamamoto, N. (2004). HIV protease inhibitor nelfinavir inhibits replication of SARS-associated coronavirus. *Biochemical and Biophysical Research Communications*, 318(3), 719–725. <https://doi.org/10.1016/j.bbrc.2004.04.083>
- Yang, H., Bartlam, M., & Rao, Z. (2006). Drug design targeting the main protease, the Achilles' heel of coronaviruses. *Current Pharmaceutical Design*, 12(35), 4573–4590. <https://doi.org/10.2174/138161206779010369>
- Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., Sun, L., Mo, L., Ye, S., Pang, H., Gao, G. F., Anand, K., Bartlam, M., Hilgenfeld, R., & Rao, Z. (2003). 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*, 100(23), 13190–13195. <https://doi.org/10.1073/pnas.1835675100>
- Zhang, L., & Liu, Y. (2020). Potential interventions for novel coronavirus in China: A systematic review. *Journal of Medical Virology*, 92(5), 479–490. <https://doi.org/10.1002/jmv.25707>
- Zhou, Y., Hou, Y., Shen, J., Huang, Y., Martin, W., & Cheng, F. (2020). Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discovery*, 6, 14. <https://doi.org/10.1038/s41421-020-0153-3>
- Zhou, Y., Vedantham, P., Lu, K., Agudelo, J., Carrion, R., Nunneley, J. W., Barnard, D., Pöhlmann, S., McKerrow, J. H., Renslo, A. R., & Simmons, G. (2015). Protease inhibitors targeting coronavirus and filovirus entry. *Antiviral Research*, 116, 76–84. <https://doi.org/10.1016/j.antiviral.2015.01.011>
- Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G. F., Tan, W., & China Novel Coronavirus Investigating and Research Team. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *The New England Journal of Medicine*, 382(8), 727–733. <https://doi.org/10.1056/NEJMoa2001017>
- Ziebuhr, J. (2005). The coronavirus replicase. *Current Topics in Microbiology and Immunology*, 287, 57–94. https://doi.org/10.1007/3-540-26765-4_3