Non-canonical RNA dependent RNA polymerase (RdRp) activity of ORF8 protein in SARS-CoV-2: Equivocal biological significance

Ramya Venkatesan[1], Prakash Shyam Karuppiah[2], Rajkumar Prabhakaran[2], Vignesh Kalimuthu[1], Sathya Priya Krishnamoorthy[1], and Kadalmani Balamuthu[1]

[1]Bharathidasan University
[2]VVD and Sons Private Limited

April 1, 2022

Abstract

No current treatment options were successful in containing the ongoing pandemic COVID-19 caused by SARS-CoV-2. It is essential to understand the molecular players of SARS-CoV-2 to find a suitable treatment method and to develop an effective antiviral drug as early as possible. Global researchers have undertaken accelerated structural studies of key proteins involved in host-virus interaction, replication, and transcription. In silico studies support structural biologist with preliminary information to efficiently drive further studies and characterization. From the genome sequence, most SARS-CoV-2 annotated ORF has a conserved sequence similar to SARS-CoV-1, except for ORF8 and ORF10. The function of ORF8 protein in SARS-CoV-2 is uncertain. Herein, we had modelled the ORF8 protein and studied its putative function using various substrates as a probe to determine its biological significance. The modelled SARS-CoV-2 (mORF8^{7A}) protein shows IgG characteristic folds and thus may belong to IgG superfamily. Further, we studied the binding efficacy of various antiviral drugs against the modelled ORF8 of SARS-CoV-2 (mORF8) to repurpose the drug and to use them as a probe to study its function by studying the binding/active sites interaction. Remdesivir had the highest binding affinity to ORF8 protein of SARS-CoV-2. The high affinity of the adenosine analogue yields critical information about the non-canonical RNA dependent RNA polymerase (RdRp) function of ORF8 protein. We hypothesize that the ORF8 protein may be a non-canonical RdRp in SARS-CoV-2 with ability to bind to canonical nsp12 complex.

Keywords: **SARS-CoV-2; ORF8 protein; RNA dependent RNA polymerase (RdRp)**

INTRODUCTION

An outbreak of viral pneumonia (later known as COVID-19) emerged in the city of Wuhan, in the Chinese province of Hubei in December 2019 and became pandemic[1]. On January 9, 2020, The Chinese health authorities and the World Health Organization (WHO) declared the discovery of a novel coronavirus known as SARS-CoV-2 as the agent responsible for the pneumonia cases[2,3]. Coronaviruses (CoV) are a large family of viruses found in both animals and humans[4]. CoV can cause simple cold to severe diseases like Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS)[5,6]. SARS epidemic (2002-2003), also began in China and killed 774 people worldwide[7]. As of March 31, 2020, COVID-19 caused by SARS-CoV-2 had infected more than 750,000 people and caused death in more than 36,000 cases. The mortality rate is approximately 4.79%. The Maximum-Likelihood (ML) value of the reproductive number (R_0) for SARS-CoV-2 is 2.28 to 2.35 (95% CI, 1·15–4·77)[8,9]. Symptoms of SARS-CoV-2 include running nose, sore throat, cough, and fever. However, it can be more severe in some persons leading to pneumonia

or breathing difficulties and fatal[10]. Understanding the structure and molecular features of the SARS-CoV-2 proteins are critical to understand the pathogenicity of the virus. Studies using computational tools laid foundation and guidelines for further experimental works to understand the genome and pathogenicity of novel SARS-CoV-2. Beal and colleagues had identified the tentative number of amino acid sequences that distinguish SARS-CoV-2 from other known viruses in *Coronaviridae* family using FAST-NA – a software tool for screening DNA synthesis order for pathogens. The authors had identified three main regions of SARS-CoV-2[11]. Of which two unique sequence regions belong to 1ab polyprotein (QHO60603.1), and one unique sequence region in surface glycoprotein region (QHO60594.1). The authors showed the entry of SARS-CoV-2 into the host cells might be similar to SARS-CoV-1 using glycosylated spike (S) protein. Xu *et al* analyzed the evolution of the CoV and computationally modelled the SARS-CoV-2- 'S' protein interacting with human ACE2[12]. Following this, Yan and colleagues from Prof. Jason McLellan's lab at the University of Texas at Austin (USA) expressed the ectodomain residues of spike 'S' protein (1-1208 based on the genome sequence of Wuhan-Hu-1 whole genome sequence) of SARS-CoV-2 along with two stabilizing proline mutations in the C-terminal[13]. The Cryo-EM structure of the SARS-CoV-2 'S' protein confirmed the computational results and reveals the higher binding affinity to angiotensin-converting enzyme 2 (ACE2) when compared to SARS-CoV-1[13]. The binding affinity of SARS-CoV-2 'S' protein to ACE2 is ˜15 nM, which is approximately ˜10 to 20 fold higher than the binding affinity of SARS-CoV-1 'S' protein to ACE2. Further, the SARS-CoV-2 receptor-binding domain (RBD) is highly homologous to SARS-CoV-1 RBD. Notably, RBD-directed monoclonal antibodies (mAbs) of SARS-CoV-1 *viz* S230, m396, and 80R showed no binding to the SARS-CoV-2 RBD even at 1mM concentration[13]. Based on the studies, 'S' glycoprotein of SARS-CoV-2 is one of the viable target in developing vaccines, therapeutic antibodies, and diagnostics. SARS-CoV-2 entry in is humans driven by interaction between the 'S' protein and human ACE2 and subsequent protease cleavage drives the fusion. Further, Zhang and colleagues from the Institute of Biochemistry, University of Lübeck, Germany synthesized an improved α-ketoamide inhibitor to inhibit the SARS-CoV-2 main protease enzyme (M^{pro} otherwise known as $3CL^{pro}$). The crystal structure of SARS-CoV-2 M^{pro} along with the improved α-ketoamide inhibitor at 1.75 Å resolution structure revelled some key features of the enzyme when compared to SARS-CoV-1, such as; in SARS-CoV-2 alanine (Ala285) replaced the threonine (Thr285) of SARS-CoV-1 and leucine (Leu286) replaced the isoleucine (Ile286). The T285A and I286L are very significant changes in enhancing the catalytic activity. Of note, these residues are in the catalytic site of M^{pro}. Earlier reports indicate replacing Ser284, Thr285, and Ile286 by alanine residues in SARS-CoV-1 M^{pro} enhances the catalytic activity of the protease by 3.6-fold[14].

In SARS-CoV-2, the canonical nsp12 RNA-directed RNA polymerase (RdRp) coded within ORF1ab region. The cryo-EM structure of the full-length SARS-CoV-2 RNA-dependent RNA polymerase (RdRp, also named nsp12) in complex with cofactors nsp7 and nsp8 at 2.9-Å resolution shows the 'right hand' RdRp domain and a unique N-terminal extension domain (residues D60-R249) that adopts a nidovirus RdRp-associated nucleotidyltransferase (NiRAN) architectures. ORF1ab sweep the major share in the genome and codes for many non-structural proteins. ORF1ab coded proteins are primarily responsible for the replication and transcription of the viral RNA genome[16]. The nsp12 is made up of 932 amino acids is also coded by ORF1ab polyprotein locus from nucleotides 4393 to 5324. Other than ORF1ab that codes for poly-protein(s) there are other ORF's such as ORF3, 6, 7a, 8 and 10. Of which, ORF3a of SARS-CoV-2 may code a protein similar to a homotetrameric potassium sensitive ion channels (viroporin) of HRV that modulates virus release[17]. In SARS-CoV-1, the ion channel activity of the 3a protein is essential for IL-1β secretion[18]. Furthermore, the 3a protein activates the nod-like receptor family pyrin domain-containing 3 inflammasomes (NLRP3) in lipopolysaccharide-primed macrophages[19]. ORF3a protein in SARS-CoV-2 might have a similar function to ORF3a of SARS-CoV-1 due to sequence similarity. ORF6 protein could be a determinant of virus virulence similar to SARS-CoV-1 ORF 6 protein[20] and ORF7a may suppress the host tethering activity by glycosylation interference[21]. Of note, SARS-CoV-1 and SARS-CoV-2 have conserved sequence for both ORF6 and ORF7a. Among the non-structural proteins, except for ORF8 and 10 proteins, we detected the conserved sequences between SARS-CoV-1 and SARS-CoV-2 in most of the other proteins. The sequence of SARS-CoV-2 ORF8 shared very low similarity with sequences in SARS-CoV-1, and ORF10 had a premature stop codon in SARS-CoV-1. In SARS-CoV-2, a one-base deletion caused a frame-shift mutation in ORF10

protein. We used *in silico* tools to further understand the significance of ORF8 protein and computed the putative structure using homology-modelling approach.

Homology modelling using computational hierarchical clustering approaches had yielded critical information and putative functions of several non-structural proteins (nsp's) present in SARS-CoV-2 whose structures were not yet experimentally resolved. Most of these proteins play a vital role in molecular regulation such as replication and transcription. It is imperative to emphasise that SARS-CoV-2- belongs to the *Coronaviridae* family that are single-stranded (+) RNA viruses[22]. Host ribosome can access the genes of the positive-sense RNA viruses directly with complex transcription mechanism to form proteins immediately. During this process sub-genomic mRNAs, ribosomal frameshifting and proteolytic processing of polyproteins are involved[23]. Most ORFs annotated from SARS-CoV-2 have conserved sequence similar to SARS-CoV-1, except for ORF8 and ORF10. Albeit, nine ORFs (orf1ab, E, M, N, S, ORF3a, ORF6, ORF7a and ORF7b) in SARS-CoV-1 and SARS-CoV-2 are similar, the sequence of SARS-CoV-2 ORF8 shared very low similarity with sequences in SARS-CoV-1 and ORF10 had a premature stop codon in SARS-CoV-1. A one-base deletion caused a frame-shift mutation in ORF10. The nature and function of ORF8 and 10 proteins is uncertain. It is currently unclear whether the ORF10 region translates into a functional protein. Researchers have suggested that SARS-CoV-1 ORF8 protein may be acquired from SARS-related CoV from greater horseshoe bats through recombination[24]. Furthermore, the truncation of ORF8 (29 nt deletion) decreased replication up to 23-fold in SARS-CoV-1[25]. In the present manuscript, to understand more about the mysterious ORF8 protein, we modelled the ORF8 protein coded by ORF8 gene of SARS-CoV-2 using homology modelling approach. Based on the modelled ORF8 structure we anticipate that the ORF8 protein of SARS-CoV-2 may belong to IgG superfamily and might possess non-canonical RNA dependent RNA polymerase (RdRp) activity.

RESULTS AND DISCUSSION

Homology model of SARS-CoV-2 ORF8 protein has IgG super-family structural features.

The open reading frame (ORF)8 of the SARS-CoV-2 encodes a protein of unknown function. The modelled protein mORF8^{7A} of the *ORF8* gene revelled 21.57% sequence identity (GMEQ score = 0.22) with the open reading frame ORF7a of the SARS-associated coronavirus (SARS-CoV-1), a unique type I transmembrane protein whose function is also unknown and 13.51% sequence identity (GMEQ score = 0.10) with Semaphorin-7A protein. Semaphorin 7A protein plays an important role in integrin-mediated signalling and functions both in regulating cell migration and immune responses in humans. Semaphorin 7A initiates a T-cell mediated inflammatory response through α1β1 integrin[26]. We used both the ORF7a of SARS-CoV-1 and Semaphorin 7a as a template to develop the homology models of ORF8 SARS-CoV-2 *viz* mORF^{7A} and mORF^{S}respectively (**Fig. 1a and b**). Of note, in homology modelling, BLAST is relatively faster while less sensitive when compared with profile-profile alignment methods. BLAST can detect the homology with significant sequence identity > 40%. On the other hand, PSI-BLAST (profile-sequence alignment method) is more sensitive than sequence-sequence alignment that can recognize distant homology with lower sequence identity > 20%[27]. PyMod has implemented both the tools BLAST and PSI-BLAST[28]. We used PyMod to build the homology model hence 21.57% sequence identity is acceptable and within the tolerable confidence limit. Hence, based upon these rationales, we chose SARS-CoV-1 ORF7a protein as the template to model the ORF8 protein of SARS-CoV-2.

Estimating the quality of the model is important to gain confidentiality on the model. QMEAN is a composite scoring function that is able to derive absolute quality estimates based on one single model for the entire structure (global) and per residue (local)[29]. The local quality estimate and QMEAN scores were acceptable for the modelled protein, -6.72 for mORF8^{7A} (**Fig. 1a**) and -2.70 for mORF8^{S} (**Fig. 1b**). Other parameters such as Cβ, torsion, and solvation scores are appreciable in both the models. The local quality estimate for mORF8^{7A} (**Fig. 1a,** panel 2 bottom) and mORF8^{S} (**Fig. 1b,** panel 2 bottom) indicate the goodness of the model. We have had tabulated the MolProbity results for both the models (**Table 1**). MolProbity is a structure-validation web service that provides broad-spectrum solidly based evaluation of

model quality at both the global and local levels[30]. In the modelled protein (SARS-CoV-2 mORF8^{7A}), 84.72% of the residues were present in the favoured region of Ramachandran plot (**Fig. 1c**). 6.94% of the amino acid residues were present in the outliers indicating the overall goodness of the model quality. In the modelled protein (SARS-CoV-2 mORF8^{S}), 89.74% of the residues were present in the favoured region of Ramachandran plot (**Fig. 1d**). 5.13% of the amino acid residues were present in the outliers. Further, 2.70% residues were in rotamer outliers. We had tabulated the results of the structure assessment by MolProbity for both the models (**Table 1**). The data indicate the overall goodness of the model quality. Further, we had aligned the mORF8^{7A} and mORF8^{S} to find the similarity between the two models. The two models were partially super-imposable and had similar fold on certain regions (**Fig. 1d**). After analyzing both the models, the model obtained using the ORF7a (PDB: 1XAK) of the SARS-CoV-1 template structure was chosen for further studies. The rationalé behind choosing ORF7a of the SARS-CoV-1 as template are (a.) higher sequence similarity (21.57%) among the selected models (b.) the similarity index is within the acceptable range of homology modelling using PSI-BLAST and (c.) both the target model protein and template protein belong to *Coronaviridae* family viral protein. The secondary structure analysis of mORF8^{7A} using SOPMA tool[31] indicates that 19.83% amino acids are present in the α-helix and 4.96% of amino acids were participating in β-turn. 39.67% amino acids were randomly coiled and 35.54% of amino acids are in extended strands. Although it is difficult to draw conclusions about the function of ORF8 protein from the structure alone, some features are noteworthy. The mORF8^{7A} protein shows a characteristic Ig-fold, a sandwich-like structure formed by two sheets of anti-parallel beta-strands. Further, we calculated the electrostatic potential with Adaptive Poisson-Boltzmann Solver (APBS)[32], the mORF8^{7A} protein is more basic along the membrane-proximal (C-terminal). Examining the hydrophobic surface of the mORF8^{7A} indicates that the protein might be a transmembrane protein similar to ORF7a of SARS-CoV-1.

Apart from the ORF7a of the SARS-CoV-1 and Semaphorin-7A[33], ORF8 of SARS-CoV-2 show structural similarity with the hypothetical protein X4[34], lymphocyte antigen 86 (LY86)[35]. Upon studying these proteins, the ORF7a and LY86 share a common interest. Previous reports indicate the cooperation of LY 86 with CD180 and TLR4 to mediate the innate immune response to bacterial lypopolyysaccharide (LPS) and cytokine production. The LY86 is important for efficient CD180 cell surface expression. In addition, the tissue specificity estimation based upon the information derived from experiments at the mRNA level indicates that LY 86 may be highly expressed in B-cells, monocytes and tonsil. Another similar protein to ORF8 of SARS-CoV-2 is ORF7a of SARS-CoV-1. Visualizing the crystal structure of SARS-CoV-1 ORF7a protein indicate the N-terminal ectodomain consists of a compact seven stranded beta-sandwich that is similar in fold and topology to members of the Ig superfamily[36]. In the mORF8 ^{7A} protein, we were able to visualize domains similar to the Ig fold domain that consists of two β-sheets and form a β-sandwich framework. Therefore, we can interpret that ORF7a protein and ORF8 protein may bind together in elucidating a cellular response and play an essential role in host-virus interaction. To test the hypothesis, we performed protein-protein interaction studies with mORF8^{7A} and C1q. We observed the interaction ability of mORF8^{7A} with C1q (**Fig. 1f**). SARS-CoV-1 infected cells express and retain the ORF7a in the intracellular region[36]. The ORF7a of SARS-CoV-1 may also share similar structural homology to IgG superfamily. Of note, the IgG superfamily proteins are associated with the cell adhesion, binding and recognition[37]. In viruses, these proteins act as the receptors of cellular mediators such as interferon(s) and interleukins or may act as surface proteins, and they may enhance virus virulence and dissemination. It is noteworthy to mention that ORF7a elucidates the viral virulence.

Binding studies using repurposed antiviral drugs indicate the putative RNA dependent RNA polymerase (RdRp) activity of ORF8

To establish structure-activity relationship, we utilized *in silico* PatchDock[38] molecular docking tool and LigPlot[39] to study the binding possibilities of selected antiviral drugs against mORF8^{7A}. The purpose of this binding study is to find the possibility of repurposing the existing antiviral drugs and to gather more information on the function of mORF8^{7A} protein using ligands as a probe. (a) Choroquine (**Fig. 2a**) **-** Zinc ionophore that increases the intracellular zinc concentration and inhibits the viral replication by enhancing

the activity of zinc finger proteins[40,41]. (b) Flavipiravir (**Fig. 2b**) **-** inhibits the RdRp of RNA viruses[42]. (c) Nafamostat (**Fig. 2c**) - a synthetic serine protease inhibitor[43]. (d) Nitazoxaninde (**Fig. 2d**) **-** broad-spectrum antiviral agent suggested to blocks maturation of the viral hemagglutinin at the post-translational stage[44]. (e) Penciclovir (**Fig. 2e**) **-** synthetic acyclic guanine derivative with antiviral activity against HSV, it acts as a delayed chain terminator and blocks the polymerase at an internal position within the viral nucleic acid [40]. (f) Ribavirin (**Fig. 2f**) **-** PB1 transcriptase inhibitor[45]. (g) Remdesivir (**Fig. 2g**) **-** an adenosine analogue[46]. We had tabulated the results of the binding studies of mORF8^{7A}against selected antiviral drugs (**Table 2**). Among the selected drugs, Remdesivir effectively bonded to m ORF8^{7A} ^{7A} (ACE score: **-** 395.95) (**Fig. 2h**). The active amino acid participants between Remdesivir and m ORF8^{7A} are Arg48 and Lys53 (directly binding to the ligand), other participating amino acids include Ala52, Ile58, Glu59 and Leu60. Remdesivir (GS-5734), a broad-spectrum antiviral nucleotide prodrug is being developed as a treatment option against a diverse panel of RNA viruses such as Ebola virus (EBOV), Marburg, MERS-CoV, SARS-CoV-1, respiratory syncytial virus (RSV), Nipah virus (NiV), and Hendra virus[47]. Remdesivir is an adenosine analogue, which incorporates into nascent viral RNA chains and results in premature termination of viral RNA transcription. Recently, Wang *et al*[46] studied the effectiveness of Remdesivir against SARS-CoV-2 and found that the EC_{90} value of Remdesivir in Vero E6 cells as 1.76 μM. Remdesivir also inhibited SARS-CoV-2 infection efficiently in human liver cancer Huh7 (SARS-CoV-2 sensitive) cells[46]. Our results corroborate with Wang *et al* experimental findings. Together, these data demonstrate the effectiveness of Remdesivir with the combination of chloroquine (CQ) or hydroxychloroquine (HCQ) along with zinc supplement in SARS-CoV-2 treatment.

Based on the binding studies, we hypothesize the$m$ ORF8^{7A} protein as non-canonical RdRp in SARS-CoV-2. We were able to observe the unique C-terminus that folds back into the active site cleft in order to regulate the RNA synthesis (**Fig. 3a**). RdRp of *Flaviviridae* family have reported similar structural features. CoV replication and transcription are driven by the 15 to 16 nonstructural proteins that are not usually present in other +RNA viruses[48]. The active amino acids in the Remdesivir binding pockets were Glu23, Ser24, Phe86, Ile88, and Thr87 (**Fig. 3b**). Furthermore,m ORF8^{7A} protein has two conserved motifs similar MERS RdRp and the structure shows favourable conformation to accommodate incoming viral ssRNA template during the replication process (**Fig. 3c**). Amino acids shows the conserved sequence present in the RdRp's of SARS-CoV-2 ORF8 protein, MERS RdRp, norovirus RdRp and mitovirus RdRp (**Fig. 3d**). In addition, it is known that the mechanism of action of Remdesivir is by pre-mature termination of transcriptional activity[46]. Therefore, it is highly possible that the mORF8 will be an RdRp protein. It is obvious that flavipiravir also inhibits the RNA-dependent RNA polymerase (RdRp) of RNA viruses, yet flavipiravir has comparatively week binding (ACE score: -129.05) to mORF^{7A}. Notably, flavipiravir *per se* is not an RdRp inhibitor, rather flavipiravir is phosphoribosylated in the cells to be an active form, favipiravir-RTP is recognized as a purine nucleotide by RdRp, and inhibits the RdRp enzyme activity[42]. Evolution had conserved RdRp's. We performed sequence alignment for ORF8 with three other known RdRp's *viz* MERS RdRp (AKJ75494.1), norovirus RdRp (ASR80891.1), and mitovirus RdRp (CAJ32468.1) (**Fig. 3e**). Of which, one of the amino acid motif 'CX' is conserved across all the 4 RdRp's that were studied. Interestingly, the amino acids are present in the region adjacent to the pocket that accommodates the template RNA strand. The nucleotide-binding site amino acids are diagonally opposite to the conserved amino acid regions indicate that the conserved regions play an essential role in template stabilization and have the least role in the catalysis.

To confirm the binding of RNA strand to ORF8 protein we performed docking studies of ORF8 protein with an RNA strand with UGGAA-UGGAA pentad 'GUAC*UGGAA* CAUGUU UCAUG*UGGAA* GUAC', as expected the RNA stand was docked in the template entry (TE) domain of ORF8 (**Fig. 4a-c**). This indicates the TE domain is able to handle a wide range of nucleotides that are present in the vRNA. In addition, we docked the adenine nucleotide with ORF8 RdRp, the adenine nucleotide bound to the catalytic domain present in the tunnel of TE and NE domains with ACE value -139.60 indicating very good binding affinity (**Fig. 4d-f**). This further confirmed the ORF8 catalytic region accepting the incoming nucleotides of the nascent mRNA strands. Every structural evidence point the ORF8 as RdRp. Notably, SARS-CoV-2 has a dedicated RdRp that is responsible for the replication and transcription of viral RNA genome. The RdRp

is consists of 932 amino acids coded by YP_009725307 locus nucleotides 4393 to 5324 of ORF1ab polyprotein (NCBI Reference Sequence: YP_009725307.1). However, our modelling results indicate the possibility of SARS-CoV-2 having two RdRps, the canonical nsp12 (common to other CoV) and non-canonical RdRp ORF8. Despite the availability of the major RdRp (nsp12) in the CoV genome, the necessity of a minor RdRp protein (with 121 a.a.) is evolutionarily superfluous. However, the transcription in CoV is primer dependent, if so, CoV's has to synthesise the primer using a specialized RdRp that can synthesize nucleotides ranging from 6mer to 20mer oligos. The primer synthesizing RdRp's are supposedly non-canonical. In SARS-CoV-1, a non-canonical RdRp (second RdRp) resides in nsp8 whose primary role is primer synthesis[49]. The nsp12 RdRp utilizes the primers synthesized by nsp8. Whether ORF8 protein does have a similar role in SARS-CoV-2 remains an open question. Nevertheless, the necessity of an ORF8 protein with the RdRp catalytic domain in SARS-CoV-2 remains a major puzzle that appeal to solve the crystal structure experimentally and identify their biochemical significance.

ORF8 protein has a binding affinity to canonical RdRp complex (nsp12-nsp7-nsp8-1/2)

Further, we questioned, whether is it possible that ORF8 protein may complex with the RdRp (nsp12) as a sub-unit? If so, what is the significance of this sub-unit? To answer these questions, we analysed the binding affinity of ORF8 with canonical RdRp. Recently, Gao *et al* deposited the cryo-EM structure of the SARS-CoV-2 nsp12 in complex with cofactors nsp7 and nsp8 at 2.9-Å resolution (**Fig.5a**) in Protein Data Bank (PDB) with the accession numbers PDB 6M71[15]. The polymerase domain composed of three subdomains namely finger, palm and thumb (**Fig. 5b**) similar to other virus RdRp. In SARS-CoV-2, residues from L366-A581 and K621-G679 form the fingers subdomain. Residues T582-P620 and T680-Q815 forms the palm subdomain and residues H816-E920 form the thumb (**Fig. 5c**). To study the interaction ability of mORF8^{7A} with the RdRp complex (nsp12 complexes with nsp-7 and nsp8), we performed protein-protein docking using GRAMM-X simulation. The results of the simulation indicate that the ORF8 has very good binding affinity with canonical RdRp of SARS-CoV-2 (**Fig. 5d**). The accessible surface area calculation for RdRp in the protein complex resulted in the polar area to energy with probe radius 1.4Å as 15264.31. The ratio between apolar area to energy in RdRp is 24360.18 and total area to energy is 39624.49. The total number of surface atoms was 3922 and buried atoms were 3583. Similarly, the accessible surface area calculation for$m$ ORF8^{7A} in the protein complex resulted in the the polar area to energy with probe radius 1.4Å as 2027.62. The ratio between apolar area/energy is 4296.18 and total area to energy is 6323.80. The number of surface atoms in m ORF8^{7A} was 498 and buried atoms were 474. Binding of ORF8 to RdRp induces a conformational change in the N terminal region (**Fig. 5d** Left Panel). The promiscuous binding of ORF8 (indicated in pink) to SARS-CoV-2 RdRp induces a conformational change in the NiRAN domain of RdRp (**Fig. 5d** Right Panel). The ORF8 protein binds in the catalytic core of the canonical RdRp. This might be a good explanation for the 23-fold reduction in the replication after truncation of ORF8 (29 nt deletion) SARS-CoV-1[25]. The results suggest the possibility of ORF8 might complex with the canonical RdRp in SARS-CoV-2.

CONCLUSION

To conclude based on the *in silico* modelling studies we posit the ORF8 protein of the SARS-CoV-2 may structurally belong to IgG super-family and might have RdRp activity inside the cell. The possible RdRp activity of ORF8 in SARS-CoV-2 had opened up interesting avenues in the replication and transcription of SARS-CoV-1 and SARS-CoV-2. Interestingly, mORF8^{7A} protein has a very good binding affinity to nsp12 (canonical RdRp) *in silico*. We appeal future research in SARS-CoV-2 ORF8 protein should consider the possibility of being non-canonical RdRp and its ability to bind to nsp12. Further structural and biochemical studies are required to ascertain and understand the significance of ORF8 as non-canonical RdRp in SARS-CoV-2.

MATERIALS AND METHODS

Retrieval of the target sequence and ligand structures

The whole-genome sequences of the Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1 was retrieved from the National Centre for Biotechnology Information (NCBI) database using accession number (GenBank: MN908947.3). The two-dimensional structure of the inhibitors (a) Choroquine (b) Flavipiravir (c) Nafamstat (d) Nitazoxaninde (e) Penciclovir (f) Ribavirin (g) Remdesivir were obtained from the PubChem database in SDF file format[50]. The cryo-EM structure of canonical RdRp in SARS-CoV-2 (nsp12) bound to nsp7 and nsp8 was retrieved Protein Data Bank (PDB) with the accession numbers PDB 6M71 [15]. We modelled the ORF8 protein using the homology modelling approach as outlined below.

Template Search

Template search with PSI-BLAST[51] and HHBlits[52] were performed against the SWISS-MODEL template library (SMTL)[53]. We used BLAST algorithm to search the target sequence against the primary amino acid sequence contained in the SMTL. Using the procedure outlined in Remmert*et al* [52] an initial HHblits profile was built. We initiated the template search against all profiles of SMTL using the resulting template. At the end of search result, we found 6 templates closely aligned with the target amino acid sequence (ORF8 protein) *viz* SARS ORF7a accessory protein[36], Hypothetical protein X4[34], Lymphocyte antigen 86-B, 86-C, LY 86[54], and Semaphorin-7A[33].

Three-dimensional structure modelling, energy minimization and validation of the model

Based on the SMTL data, we gathered information about the possible annotation of quaternary structure, ligands, and co-factors including their oligomeric structure[53]. Using ProMod3[55], we built the models based on the ORF7a of SARS-CoV-1 protein-template alignment. To model ORF8, we copied the conserved coordinates between the ORF8 of SARS-CoV-2 and the ORF7a of SARS-CoV-1 and rebuilt the side chains. Finally, the geometry of the resulting model is regularized by using a force field[56]. The model quality estimation was performed by assessing the global and per-residue model quality using the Qualitative Model Energy (QMEAN) scoring function[29]. We developed the secondary and 3D structure models of ORF8 protein by SWISS-PROT data bank. We got ten models. Ramachandran plots and QMEAN analysis validated the generated models. After assessing the Ramachandran plot, we refined the atomic structures of the selected model using ModRefner[57]. RAMPAGE[58] was used for validating the modelled structure. Then, we calculated the QMEAN 'Z' scores[29] for the complete assessment and refinement of the predicted model. We designated the best 3D structure of SARS-CoV-2 ORF8 protein as *m* ORF8. The model developed using ORF7a of SARS-CoV-1 protein ad template was denoted as *m* ORF8^{7A} and the model developed using Semaphorin-7A protein as the template is denoted as *m* ORF8^{S} The 3D structure was characterized, visualized and analyzed using Discovery Studio software[59].

Ligand Modelling and Oligomeric State Conservation

We transferred the ligands present in the template structure by homology to the model based on the following criteria. First, we annotated the ligands as biologically relevant in the template library, second, we ensured that the ligand is in contact with the model, third, the ligand should not clash with the protein, and finally, *m*ORF8 and template should have conserved the ligand interacting residues. We excluded the ligands are excluded if one of the criteria is violated. To model the*m* ORF8 in its oligomeric form by a supervised machine learning algorithm, the quaternary structure annotation of the template is used. Briefly, Support Vector Machines (SVM), that combines the interface conservation, structural clustering, and other template features to provide a quaternary structure quality estimate (QSQE)[60]. We modelled the oligomeric form based on the QSQE score.

Molecular docking simulation

We performed the protein docking to identify the selected antiviral inhibitory activity against *m* ORF8 protein. We converted the SDF file of the selected antiviral drugs (ligands) to PDB file format using the Cactus online molecular converter tool[61]. The docking analysis of the 3D structure of the *m*ORF8 protein with inhibitor was performed using Patch-Dock molecular modelling simulation software which is effective for protein-ligand docking[62]. The structures of individual proteins and docked complexes were visualized using UCSF Chimera[63] and PyMol molecular visualization tools[64].

ACKNOWLEDGEMENT

One of the authors, Ms. Ramya Venkatesan greatly acknowledges the Department of Science and Technology – Science and Engineering Research Board (DST–SERB) India, Confederation of Indian Industry (CII) and VVD and Sons Private Limited, Thoothukudi, Tamil Nadu for the award of Prime Minister Research Fellowship for Doctoral Students through DST (SERB)-CII Prime Minister Research Fellowship Scheme.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

REFERENCES

1. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet.* 2020;395(10225):689-697.

2. Lai C-C, Shih T-P, Ko W-C, Tang H-J, Hsueh P-R. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and corona virus disease-2019 (COVID-19): the epidemic and the challenges. *International journal of antimicrobial agents.*2020:105924.

3. Chinazzi M, Davis JT, Ajelli M, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science.* 2020:eaba9757.

4. Song HD, Tu CC, Zhang GW, et al. Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. *Proceedings of the National Academy of Sciences of the United States of America.* 2005;102(7):2430-2435.

5. Wang M, Yan M, Xu H, et al. SARS-CoV infection in a restaurant from palm civet. *Emerging infectious diseases.* 2005;11(12):1860-1865.

6. Gong SR, Bao LL. The battle against SARS and MERS coronaviruses: Reservoirs and Animal Models. *Animal models and experimental medicine.* 2018;1(2):125-133.

7. Sohrabi C, Alsafi Z, O'Neill N, et al. World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). *International Journal of Surgery.* 2020;76:71-76.

8. Kucharski AJ, Russell TW, Diamond C, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The Lancet Infectious Diseases.*2020.

9. Zhang S, Diao M, Yu W, Pei L, Lin Z, Chen D. Estimation of the reproductive number of Novel Coronavirus (COVID-19) and the probable outbreak size on the Diamond Princess cruise ship: A data-driven analysis. *International Journal of Infectious Diseases.* 2020.

10. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet.* 2020;395(10223):497-506.

11. Beal J, Mitchell T, Wyschogrod D, Manthey J, Clore A. Highly Distinguished Amino Acid Sequences of 2019-nCoV (Wuhan Coronavirus). *BioRxiv.* 2020.

12. Xu X, Chen P, Wang J, et al. Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission.*SCIENCE CHINA Life Sciences.* 2020;63(3):4.

13. Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of the SARS-CoV-2 by full-length human ACE2. *Science.* 2020.

14. Zhang L, Lin D, Sun X, et al. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science.* 2020.

15. Gao Y, Yan L, Huang Y, et al. Structure of the RNA-dependent RNA polymerase from COVID-19 virus.*Science.* 2020:eabb7498.

16. Sola I, Mateos-Gomez PA, Almazan F, Zuniga S, Enjuanes L. RNA-RNA and RNA-protein interactions in coronavirus replication and transcription. *RNA biology.*2011;8(2):237-248.

17. Chien TH, Chiang YL, Chen CP, et al. Assembling an ion channel: ORF 3a from SARS-CoV. *Biopolymers: Original Research on Biomolecules.* 2013;99(9):628-635.

18. Nieto-Torres JL, DeDiego ML, Verdia-Baguena C, et al. Severe acute respiratory syndrome coronavirus envelope protein ion channel activity promotes virus fitness and pathogenesis. *PLoS pathogens.* 2014;10(5).

19. Chen I-Y, Moriyama M, Chang M-F, Ichinohe T. Severe Acute Respiratory Syndrome Coronavirus Viroporin 3a Activates the NLRP3 Inflammasome. *Frontiers in Microbiology.*2019;10(50).

20. Kopecky-Bromberg SA, Martínez-Sobrido L, Frieman M, Baric RA, Palese P. Severe acute respiratory syndrome coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins function as interferon antagonists. *Journal of virology.* 2007;81(2):548-557.

21. Taylor JK, Coleman CM, Postel S, et al. SARS Coronavirus ORF7a inhibits BST-2 virion tethering through a novel mechanism of glycosylation interference. *Journal of Virology.* 2015.

22. Egloff M-P, Ferron F, Campanacci V, et al. The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. *Proceedings of the National Academy of Sciences.*2004;101(11):3792-3796.

23. Fehr AR, Perlman S. Coronaviruses: an overview of their replication and pathogenesis. In: *Coronaviruses.* Springer; 2015:1-23.

24. Lau SK, Feng Y, Chen H, et al. Severe acute respiratory syndrome (SARS) coronavirus ORF8 protein is acquired from SARS-related coronavirus from greater horseshoe bats through recombination. *Journal of virology.*2015;89(20):10532-10547.

25. Muth D, Corman VM, Roth H, et al. Attenuation of replication by a 29 nucleotide deletion in SARS-coronavirus acquired during the early stages of human-to-human transmission. *Scientific reports.* 2018;8(1):1-11.

26. Suzuki K, Okuno T, Yamamoto M, et al. Semaphorin 7A initiates T-cell-mediated inflammatory responses through alpha1beta1 integrin. *Nature.* 2007;446(7136):680-684.

27. Altschul SF, Madden TL, Schäffer AA, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research.*1997;25(17):3389-3402.

28. Bramucci E, Paiardini A, Bossa F, Pascarella S. PyMod: sequence similarity searches, multiple sequence-structure alignments, and homology modeling within PyMOL.*BMC bioinformatics.* 2012;13 Suppl 4(Suppl 4):S2-S2.

29. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics.* 2011;27(3):343-350.

30. Chen VB, Arendall WB, 3rd, Headd JJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr.* 2010;66(Pt 1):12-21.

31. Geourjon C, Deleage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Bioinformatics.*1995;11(6):681-684.

32. Baker NA. Poisson–Boltzmann methods for biomolecular electrostatics. In: *Methods in enzymology.* Vol 383. Elsevier; 2004:94-118.

33. Suzuki K, Okuno T, Yamamoto M, et al. Semaphorin 7A initiates T-cell-mediated inflammatory responses through α1β1 integrin. *Nature.* 2007;446(7136):680-684.

34. Kanzawa N, Nishigaki K, Hayashi T, et al. Augmentation of chemokine production by severe acute respiratory syndrome coronavirus 3a/X1 and 7a/X4 proteins through NF-κB activation. *FEBS letters.* 2006;580(30):6807-6812.

35. Begum NA, Tsuji S, Nomura M, et al. Human MD-1 homologue is a BCG-regulated gene product in monocytes: its identification by differential display. *Biochemical and biophysical research communications.* 1999;256(2):325-329.

36. Nelson CA, Pekosz A, Lee CA, Diamond MS, Fremont DH. Structure and intracellular targeting of the SARS-coronavirus Orf7a accessory protein. *Structure (London, England : 1993).* 2005;13(1):75-85.

37. Williams AF, Barclay AN. The immunoglobulin superfamily—domains for cell surface recognition. *Annual review of immunology.* 1988;6(1):381-405.

38. Prabhu S, Rajeswari VD. In silico docking analysis of bioactive compounds from Chinese medicine Jinqi Jiangtang tablet (JQJTT) using patch dock. *J Chem Pharm Res.*2016;5(8):15-21.

39. Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein engineering, design and selection.* 1995;8(2):127-134.

40. Deval J. Antimicrobial strategies: inhibition of viral polymerases by 3'-hydroxyl nucleosides. *Drugs.* 2009;69(2):151-166.

41. Liu J, Cao R, Xu M, et al. Hydroxychloroquine, a less toxic derivative of chloroquine, is effective in inhibiting SARS-CoV-2 infection in vitro. *Cell Discovery.*2020;6(1):16.

42. Furuta Y, Komeno T, Nakamura T. Favipiravir (T-705), a broad spectrum inhibitor of viral RNA polymerase. *Proc Jpn Acad Ser B Phys Biol Sci.* 2017;93(7):449-463.

43. Chen X, Xu Z, Zeng S, et al. The molecular aspect of antitumor effects of protease inhibitor nafamostat mesylate and its role in potential clinical applications. *Frontiers in oncology.* 2019;9:852.

44. Rossignol J-F. Nitazoxanide: A first-in-class broad-spectrum antiviral agent. *Antiviral Research.* 2014;110:94-103.

45. Hayden FG, Shindo N. Influenza virus polymerase inhibitors in clinical development. *Curr Opin Infect Dis.* 2019;32(2):176-186.

46. Wang M, Cao R, Zhang L, et al. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Research.* 2020.

47. Lo MK, Jordan R, Arvey A, et al. GS-5734 and its parent nucleoside analog inhibit Filo-, Pneumo-, and Paramyxoviruses. *Scientific reports.* 2017;7:43395.

48. Snijder EJ, Bredenbeek PJ, Dobbe JC, et al. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *Journal of molecular biology.* 2003;331(5):991-1004.

49. Imbert I, Guillemot J-C, Bourhis J-M, et al. A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J.* 2006;25(20):4933-4942.

50. Bolton EE, Wang Y, Thiessen PA, Bryant SH. PubChem: integrated platform of small molecules and biological activities. In: *Annual reports in computational chemistry.* Vol 4. Elsevier; 2008:217-241.

51. Camacho C, Coulouris G, Avagyan V, et al. BLAST+: architecture and applications. *BMC bioinformatics.* 2009;10(1):421.

52. Remmert M, Biegert A, Hauser A, Söding J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nature methods.* 2012;9(2):173.

53. Biasini M, Bienert S, Waterhouse A, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic acids research.*2014;42(Web Server issue):W252-258.

54. Ohto U, Miyake K, Shimizu T. Crystal structures of mouse and human RP105/MD-1 complexes reveal unique dimer organization of the toll-like receptor family. *Journal of molecular biology.* 2011;413(4):815-825.

55. Wako H, Kato M, Endo S. Pro Mode: a database of normal mode analyses on protein molecules with a full-atom model. *Bioinformatics.* 2004;20(13):2035-2043.

56. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis.*2009;30(S1):S162-S173.

57. Xu D, Zhang Y. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophysical journal.*2011;101(10):2525-2534.

58. Prisant M, Richardson J, Richardson D. Structure validation by Calpha geometry: Phi, psi and Cbeta deviation. *Proteins.* 2003;50:437-450.

59. Wang Q, He J, Wu D, Wang J, Yan J, Li H. Interaction of α-cyperone with human serum albumin: Determination of the binding site by using Discovery Studio and via spectroscopic methods. *Journal of Luminescence.* 2015;164:81-85.

60. Bertoni M, Kiefer F, Biasini M, Bordoli L, Schwede T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology.*Scientific reports.* 2017;7(1):1-15.

61. Wishart DS. Introduction to cheminformatics. *Current protocols in bioinformatics.*2007;18(1):14.11. 11-14.11. 19.

62. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic acids research.*2005;33(suppl_2):W363-W367.

63. Goddard TD, Huang CC, Ferrin TE. Visualizing density maps with UCSF Chimera. *Journal of structural biology.* 2007;157(1):281-287.

64. DeLano WL. Pymol: An open-source molecular graphics tool. *CCP4 Newsletter on protein crystallography.* 2002;40(1):82-92.

LEGENDS

Figure 1. Homology Modelling results, surface features, and C1q interaction ability of *m* ORF8 protein. (a) Cartoon structure of the SARS-CoV-2 ORF8 protein using SARS-CoV-1 ORF7a as template protein (*m* $ORF8^{7A}$), results of global quality estimate (1a right top panel), and local quality estimate (1a right bottom panel) of the modelled *m* $ORF8^{7A}$. (b) Cartoon structure of the SARS-CoV-2 ORF8 protein using Semaphorin 7a as template protein (*m* $ORF8^{S}$), results of global quality estimate (1b right top panel), and local quality estimate (1b right bottom panel) of the modelled *m* $ORF8^{S}$. (b) Structure validation by Ramachandran plot for *m* $ORF8^{7A}$. (c) Structure validation by Ramachandran plot for*m* $ORF8^{S}$. (d) Overlapping image of both the models *m* $ORF8^{7A}$ (blue) and*m* $ORF8^{S}$ (yellow). (h) Cartoon structure

showing the interaction of *m* ORF8^{7A} with C1q protein Chain A (PDB ID code 4LOR). (h) Surface map showing the region of*m* ORF8^{7A} interaction with C1q protein.

Figure 2 Binding studies of *m* ORF8^{7A} to selected antiviral drugs: Ligplot showing selected antiviral drug interacting with *m* ORF8^{7A}. (a) Choroquine. (b) Flavipiravir. (c) Nafamostat. (d) Nitazoxaninde. (e) Penciclovir. (f) Ribavirin. (g) Remdesivir. (h) Remdesivir docked inside the*m* ORF8^{7A} protein adjacent to the arginine horn depicted in the hydrophobic surface map - active amino acid participants between Remdesivir and *m* ORF8 are Arg48 and Lys53 (directly binding to the ligand), other participating amino acids include Ala52, Ile58, Glu59 and Leu60.

Figure 3 Structural studies reveal the function of ORF8 as RNA dependent RNA polymerase (RdRp). (a) Visualization of ORF8 RdRp catalytic core, template entry (TE), and nucleotide entry (NE) domains. (b) The TE tunnel marked with the green surface, the selected amino acids marked in megenta mesh indicates the binding partners of Remdesivir (adenosine analogue). (c) Structure of ORF8 RdRp bonded to Remdesivir (shown in the ball model). (d) Purple marked amino acids shows the conserved sequence present in the RdRp's of SARS-CoV-2 ORF8protein, MERS RdRp, norovirus RdRp and mitovirus RdRp. (e) Multiple sequence alignment of RdRps of SARS-CoV-2 ORF8 protein (QHD43422.1) against MERS RdRp (AKJ75494.1), norovirus RdRp (ASR80891.1), and mitovirus RdRp (CAJ32468.1) showing conserved sequences.

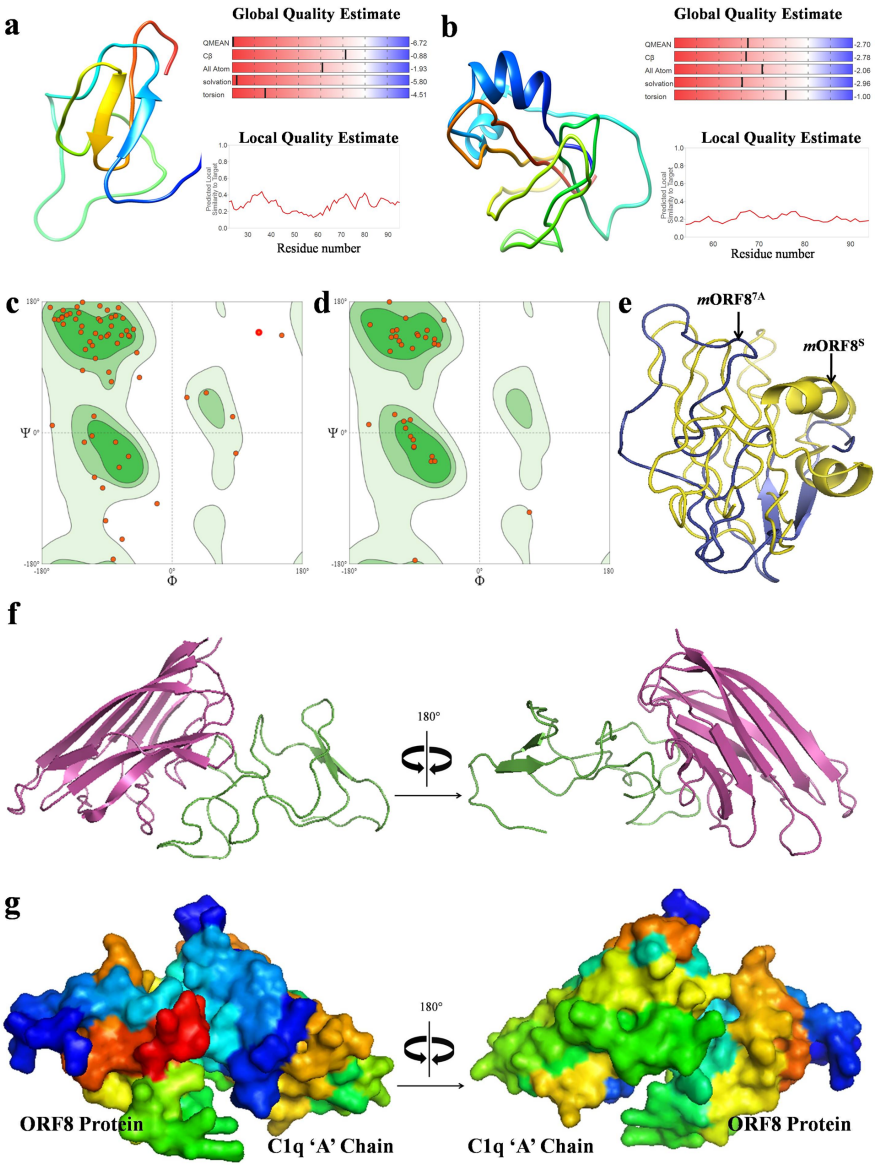
Figure 4 Elucidation of ORF8 protein function. (a) Cartoon structure of E.coli tRNA (6IZP). (b) 6IZP RNA-*m* ORF8^{7A} docked complex generated by PatchDock analysis (c) Nitrogenous base adenine binding with*m* ORF8^{7A}. (d) RaptorX analysis: Lali: 31 (length of alignment.), RMSD: 4.12 (root-mean-square deviation), μGDT(GDT): 17(23) - μGDT is the unnormalized GDT (Global Distance Test) score, TMscore: 0.212 (TMScore is the average TMscore of all the pair-wise alignments in the MSA). (e) Docking score of ORF8 with 6IZP RNA (f). Sequence alignment of ORF8 with 1CSJ (RNA-Dependent RNA Polymerase of Hepatitis C Virus).

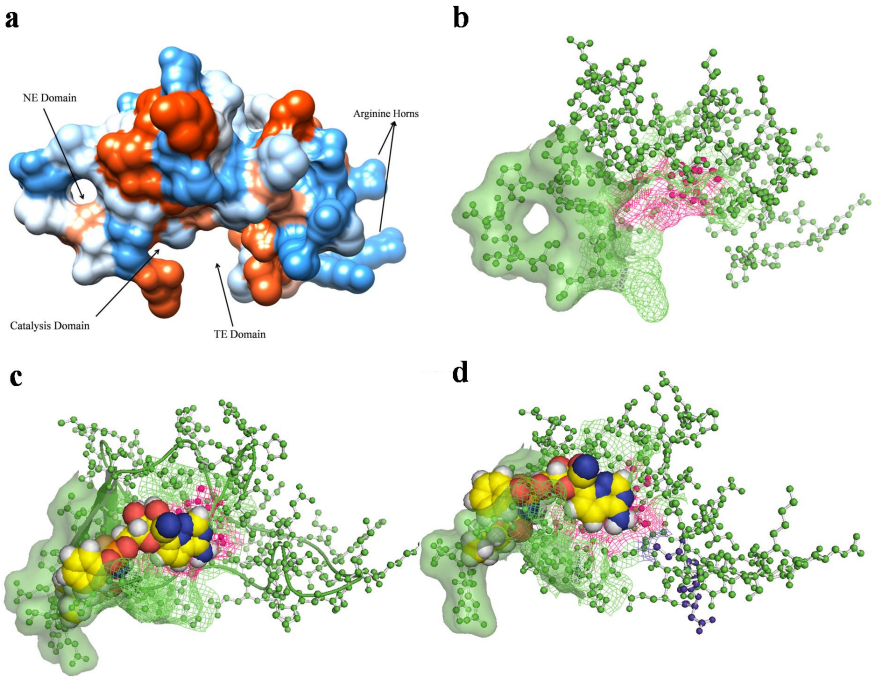
Figure 5. Protein-Protein interaction studies of canonical RdRp complex (nsp12-nsp7-nsp8-1/2) with mORF7A (a) Structure of SARS-CoV-2 virus nsp12-nsp7-nsp8 complex - ribbon diagram of SARS-CoV-2 nsp12 polypeptide chain. (b) Perpendicular view with representative domains indicated: the individual nsp8 (nsp8-1) bound to nsp12 and that in the nsp7-nsp8 pair (nsp8-2) are in grey; the nsp7 is in violet. (c) Domain organization of nsp12 complexes with nsp7 and nsp8-1/2. The inter-domain borders are labeled with residue numbers. (d) Cartoon structure showing binding of ORF8 (yellow ribbon) at the N-terminal region. (e) Structure of SARS-CoV-2 RdRp showing the finger, palm, and thumb domain. (f) Overlapping structural visualisation of SARS-CoV-2 RdRp with RdRp-ORF8 complex displaying the proximity of ORF8 binding to the catalytic core in the N-terminal of RdRp (RdRp represented as cartoon structure, RdRp-*m* ORF8^{7A} complex as tube model coloured magenta).

TABLE LEGENDS

Table 1. MolProbity Results for the modelled proteins

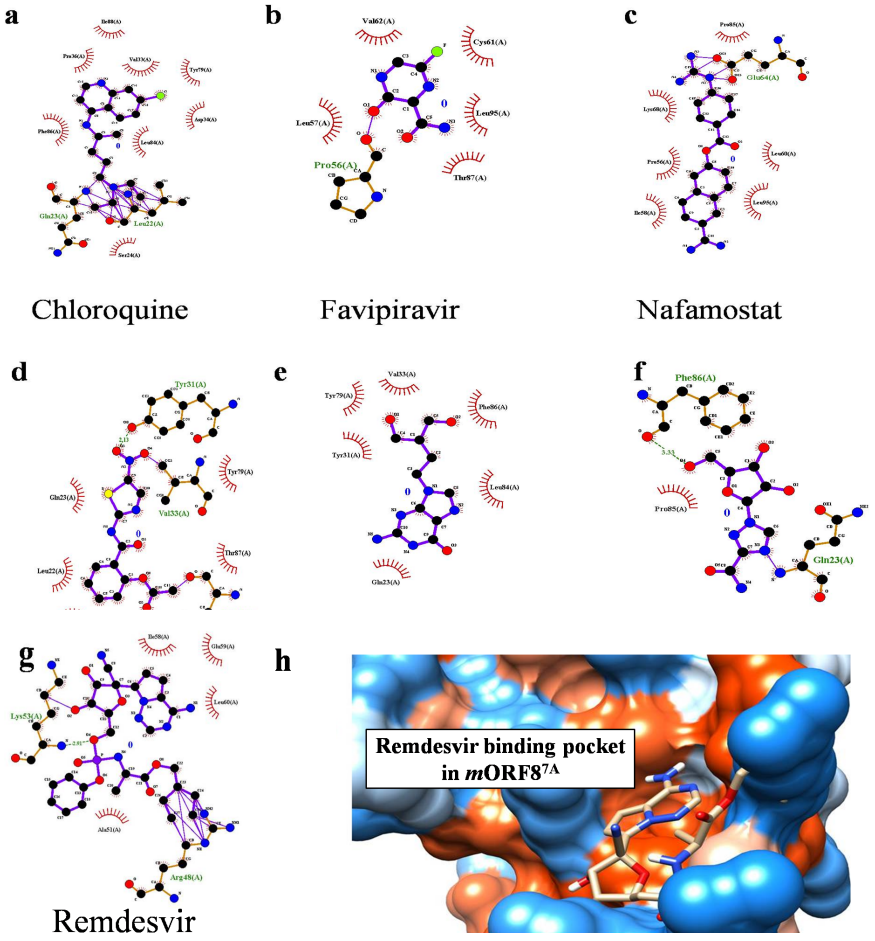
Table 2. Binding studies of the selected antiviral drugs with*m* ORF8^{7A} protein.

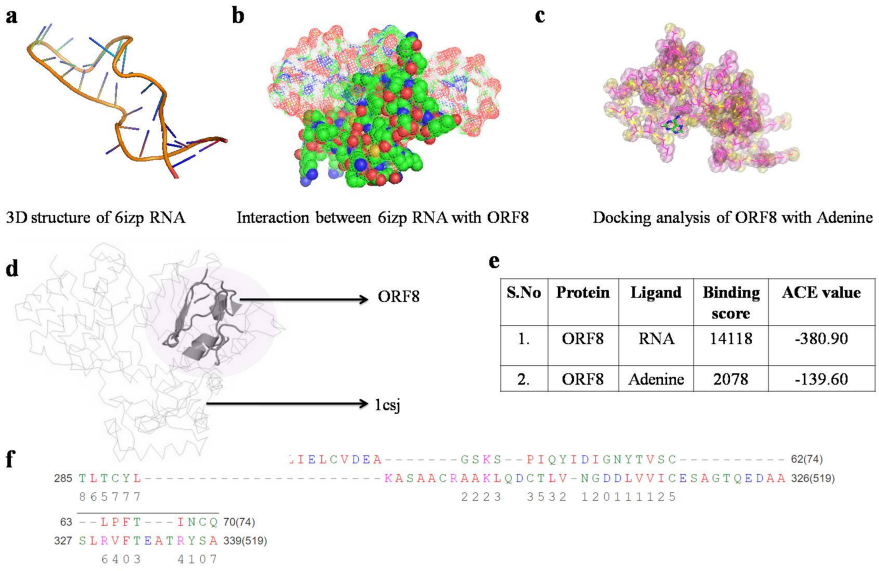




e

```
QHD43422.1      MKFLVFLGIITTVA----------------------------------------------
AKJ75494.1      ---ISAKNRARTVA----------------------------------------------
ASR80891.1      QRAVLATALEIMVK----------------------------------------------
CAJ32468.1      MKNLFTIQRKAVIKNSTTRLPNHLWLGYVKVLVWLYDIPSAHRTYFFELLNRIQSLRRNS
                    :             :

QHD43422.1      ----------------------AFHQEC--SLQSCTQHQPYVVDDPCPIHFYSKWYIRV--
AKJ75494.1      ------------------------------------GVSILSTMTNRQYHQKMLKSM-
ASR80891.1      ---------------------FSPEPDLA-----------QVVAADLLSPSVMDVGDFKI--
CAJ32468.1      GDTWTVGYLKESHRLSSHWLSGNPDVCSGSIRVSVRGLPLIIPGRLRNAIYSRNNINITR
                                                          .           : .

QHD43422.1      ----------------------------GARKSAPLIEL---------------------
AKJ75494.1      ----------------------------AATRGAT-----------------------
ASR80891.1      ----------------------------TINEGLPSGVP--------------------
CAJ32468.1      ALLTLLSTFRVMSAKPKLKLSTITDPFNGVNETLPSLEVSRALKNILQGNKFVFKKEDIN
                                                 .

QHD43422.1      ------------------------------------------------------------
AKJ75494.1      ------------------------------------------------------------
ASR80891.1      ------------------------------------------------------------
CAJ32468.1      DPYSVNRLMDITKAGPNGGVSLRNVIIDAYALMKHPKVYASLLELSNVMSPYLGLLLRNQ

QHD43422.1      ------------------------------------------CVDEAGSK------------
AKJ75494.1      ------------------------------------------CV--------------------
ASR80891.1      ------------------------------------------CTSQWNS-------------
CAJ32468.1      TTIIRAITDFDPIKWERIKPFMAPVYQSIKDQTLNLIPDLSCIGKLSTKEEAAGKVRVFA
```





S.No	Protein	Ligand	Binding score	ACE value
1.	ORF8	RNA	14118	-380.90
2.	ORF8	Adenine	2078	-139.60