



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



## Review

## Corona virus versus existence of human on the earth: A computational and biophysical approach

Zainy Zehra <sup>a,1</sup>, Manav Luthra <sup>b,\*</sup>, Sobia Manaal Siddiqui <sup>a,c,1</sup>, Anas Shamsi <sup>a</sup>, Naseem A. Gaur <sup>d</sup>, Asimul Islam <sup>a,\*\*</sup><sup>a</sup> Centre for Interdisciplinary Research in Basic Sciences, Jamia Millia Islamia, Jamia Nagar, New Delhi 110025, India<sup>b</sup> Department of Orthopedics, Medical College, Jalaun, UP, India<sup>c</sup> Jawaharlal Nehru Medical College, Faculty of Medicine, Aligarh Muslim University, UP, India<sup>d</sup> Yeast Biofuel Group, International Centre for Genetic Engineering and Biotechnology, Aruna Asaf Ali Marg, New Delhi-110067, India

## ARTICLE INFO

## Article history:

Received 1 May 2020

Received in revised form 31 May 2020

Accepted 1 June 2020

Available online 5 June 2020

## Keywords:

COVID-19

SARS-CoV-2

NSPs

Antiviral drugs

Homology

## ABSTRACT

SARS-CoV-2 has a positive sense RNA genome of 29.9 kb in size, showing high sequence similarity to the BAT-CoV, SARS-CoV, MERS-CoV. SARS-CoV-2 is composed of 14 open reading frames (ORFs), which encodes for a total of 27 proteins divided into structural and non-structural proteins (NSPs). The fundamental structural protein-encoding genes are a spike protein (S) gene, envelope protein (E) gene, a membrane protein (M) gene, and a nucleocapsid protein (N) gene. They make about 33% of the entire genome and are vital for the viral life cycle. Rest 67% is distributed among different NSPs (such as M<sup>pro</sup>, helicase, and RNA-dependent RNA polymerase) encoding genes across the ORFs, which are involved in virus-cell receptor interactions during viral entry. Researchers are trying to formulate vaccines, therapeutic antibodies or protein-targeted antiviral drugs to control the spread. This review proceeds stepwise through the COVID-19 outbreak, structural and genomic organization, entry mechanism, pathogenesis, and finally highlighting the essential proteins involved at each step that might be potential targets for drug discovery. Currently, approved treatment modalities consist of only supportive care and oxygen supplementation. This review is established on the current knowledge that has expanded on structural motifs and topology of proteins and their functions.

© 2020 Published by Elsevier B.V.

## Contents

1.	Introduction . . . . .	272
1.1.	General concept . . . . .	272
1.2.	Genomic organization . . . . .	273
1.3.	Pathogenesis . . . . .	274
1.4.	Clinical features . . . . .	274
1.5.	Disease transmission . . . . .	274
2.	Structural organization . . . . .	275
	Structural proteins . . . . .	275
2.1.	N protein . . . . .	275
2.1.1.	Structural details of N protein . . . . .	275
2.1.2.	Functional aspects of N protein . . . . .	275
2.1.3.	Potential of N protein in drug discovery. . . . .	275
2.2.	Spike protein (S). . . . .	275
2.2.1.	Structural details of the spike (S) protein . . . . .	276
2.2.2.	Functional aspects of the spike protein . . . . .	277
2.2.3.	Potential of spike protein in drug discovery . . . . .	277
2.3.	Envelope protein (E protein) . . . . .	277

\* Corresponding author.

\*\* Correspondence to: A. Islam, National Academy of Sciences, India (NASI), Centre for Interdisciplinary Research in Basic Sciences, Jamia Millia Islamia, India.

E-mail addresses: [luthramanav1@gmail.com](mailto:luthramanav1@gmail.com) (M. Luthra), [aislam@jmi.ac.in](mailto:aislam@jmi.ac.in) (A. Islam).<sup>1</sup>Both the authors have contributed equally for the manuscript

2.3.1.	Functional aspects of the envelope protein . . . . .	277
2.3.2.	Potential of envelope protein in drug discovery . . . . .	277
2.4.	Membrane protein (M protein) . . . . .	277
2.4.1.	Functional aspects and potential of membrane protein in drug discovery . . . . .	277
2.5.	Non-structural proteins . . . . .	277
2.6.	Main protease (M <sup>Pro</sup> ) . . . . .	277
2.6.1.	Structure of M <sup>Pro</sup> . . . . .	277
2.7.	Papain-like protease (PL <sup>Pro</sup> ) . . . . .	278
2.7.1.	Potential in drug discovery. . . . .	278
2.8.	Accessory NSPs. . . . .	278
2.8.1.	NSP13/helicase. . . . .	278
2.8.2.	NSP12/RNA-dependent RNA polymerase . . . . .	278
2.8.3.	NSP14/N-terminal exoribonuclease and C-terminal guanine-N7 methyltransferase . . . . .	278
2.8.4.	NSP15/uridylylate-specific endoribonuclease . . . . .	278
2.8.5.	NSP16/2'-O-methyltransferase. . . . .	279
2.8.6.	NSP10 . . . . .	279
2.8.7.	Hemagglutinin esterase (HE) . . . . .	279
2.8.8.	Trans membrane protease serine 2 (TMPRSS 2) . . . . .	279
3.	Treatment and testing . . . . .	279
4.	Emerging treatment options. . . . .	279
5.	Conclusions and future directives . . . . .	279
	Declaration of competing interest . . . . .	279
	Acknowledgements . . . . .	279
	References. . . . .	279

## 1. Introduction

*“The World Health Organization announced the novel coronavirus pneumonia epidemic caused by SARS-CoV-2 as a public health emergency of international concern on 30 January, and a pandemic on 11 March 2020”*

In December 2019, Wuhan city located in China reported 27 cases of pneumonia of unknown cause. Wuhan, which is among the most populated cities in China, has a seafood market that deals in fish, bats, poultry, marmots, snakes, etc. To the surprise of the administrators, all the above 27 cases had a link to the seafood market. These patients most commonly presented with cough, shortness of breath, sore throat, high fever; further the imaging tools showed bilateral lung infiltrates. The Chinese Centre for Disease Prevention and Control (CCDC) collected swabs from throat of all the above patients and confirmed that a new organism called Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) was the causative agent [1]. Subsequently, the disease was named Coronavirus Infectious Disease 2019 or COVID-19 by the World Health Organization (WHO) on 11 February 2020. On 31 December 2019, Chinese government officially reported to the WHO that the first confirmed case was diagnosed on 8 December 2019. The majority of the patients that had a cough, sore throat, common cold, fever, and other system symptoms, recovered spontaneously, but some of them landed into serious complications such as severe pneumonia, bronchitis, septic shock, organ failure, and acute respiratory distress syndrome (ARDS) [1]. This virus did not limit itself in China and soon spread to other countries also, causing global havoc.

According to WHO as of May 31, 2020, this pandemic disease has more than 6,258,870 confirmed cases of infection with more than 373,688 confirmed deaths across as many as 213 countries, since the first patient was hospitalized on 12 December 2019. The United States accounts for almost 30% of all cases with 1,836,872 positive cases, followed by Brazil, Russia, Spain, United Kingdom, Italy, India, France and Germany and Peru. In India, 190,609 cases have been reported till 31 May with more than 5400 deaths. However, the testing in India remains low with just 262 tests per 1 million populations as against 4718 in Germany while in Bahrain and Iceland more than 17,800 per million [2].

The global fatality rate of this new disease was found to be 2.4% with a large range of 0.5–5% for different places. However, precise measurement of the mortality rate is possible only once the pandemic is over, as the formula for total mortality is valid only when all cases have had an outcome. In other words, the case fatality rate calculations are misleading when an epidemic is still on going and the outcome of a large number of cases is unknown. According to data from China and Italy, surely, the highest sufferers are among people who are over 80 years indicating that the aged section of the society is more prone to it. According to the data from reports of secondary transmission as a result of short-term events such as a lunch or holiday visit, the Secondary Attack Rate is calculated to be about 35%. However, more data is needed to accurately assess the within-household and between-household transmission rates of COVID-19 [3]. The median incubation period is believed to be 5–6 days, with a range of one to fourteen days. Among the patients who died in Italy till 19 Mar 2020, 73.8% were hypertensive, 33.9% diabetic, 30.1% had a history of ischemic heart disease, 22.0% atrial fibrillation, 19.5% had cancer diagnosed in the last five years.

### 1.1. General concept

Viruses are the most notorious microscopic pathogens that can grow and multiply only within the host organism. They are highly contagious and cause infections of ranging severity from the common cold to AIDS. Each viral particle (virion) has either RNA or DNA genome, surrounded by a nucleocapsid, which is composed of multiple copies of coat proteins [4]. In many animal viruses, this nucleocapsid is further enveloped by a phospholipid bilayer [4]. Viruses are classified on the basis of genetic composition, morphology, the mechanism used for replication, and other criteria such as host organism. Human infecting viruses are divided into 21 families out of the spectrum of different host ranges [5].

In particular, the coronaviruses belong to the family *Coronaviridae* in turn subfamily *Coronavirinae*. This subfamily is classified on the basis of their evolutionary relationships and gene sequence into 4 genera: *alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus* and *Deltacoronavirus* (according to the International Committee on Taxonomy of Viruses/ICTV). The *alphacoronavirus* and *betacoronavirus* genera primarily infect mammals, whereas the *Gammacoronavirus* and *Deltacoronavirus* predominantly infect birds [6]. This century witnessed the outbreak of three previously unidentified coronaviruses: severe

acute respiratory syndrome coronavirus (SARS-CoV) in 2003, Middle East respiratory syndrome coronavirus (MERS-CoV) in 2012, and 2019 novel coronavirus in 2019–2020. All of them belong to the Coronaviridae, a family of viruses that possess a positive-sense single-stranded RNA genome. Earlier in 2002–2003, a severe acute respiratory syndrome (SARS-CoV) outbreak confirmed its transmission in eighty thousand patients globally and more than seven hundred death during the initial period of the outbreak. This outbreak led to the detection of the bat and civet SARS-CoV and also human coronaviruses such as NL63 and HKU1 [7,8]. The patients showed pneumonia like symptoms which later results into acute respiratory distress syndrome (ARDS). In 2012, another member of *betacoronaviruses* named MERS-CoV caused an endemic in Middle Eastern countries. The infection of MERS-CoV was more epidemic in Saudi Arabia. Similar manifestations were detected as acute lung injury often followed by pulmonary and renal failure. Dromedary camels were involved in the infection, which were thought to be the most common source of transmission from animal to human. Nevertheless, its risk factors in human remained unclear [9,10]. The infection of MERS-CoV further virus spread to France, UK, Spain, Italy and Tunisia. The fatality rate of MERS-CoV infected persons was about 35% [11]. High fatality rates of 9.5% and 34.4% were reported for SARS-CoV and MERS-CoV respectively. Fortunately, the fatality rate of SARS-CoV-2 is only 2.4% reported which is significantly low [12]. All of these three pathogenic viruses belong to the genus *betacoronavirus*. Other previously known coronaviruses that can infect human includes human coronavirus 229E (HCoV-229E), OC43 (HCoV-OC43), NL63 (HCoV-NL63), HKU1 (HCoV-HKU1) [13].

The metagenomics next-generation sequencing technology identified that the genetic material of the new SARS-CoV-2 virus exhibit approximately 88% relatedness with the two Coronaviruses SARS and MERS that have also originated from bats [13]. According to homology modelling, amino acid variation was found in some key residues, the spike proteins of SARS-CoV and MERS-CoV bind to different host receptors via different receptor-binding domains (RBDs). SARS-CoV uses angiotensin-converting enzyme 2 (ACE2) as one of the main receptors [14] with CD209L as an alternative receptor [15], whereas MERS-CoV uses dipeptidyl peptidase 4 (DPP4), which is also known as CD26, as the primary receptor. Initial analysis suggested that 2019-nCoV has a close evolutionary association with the SARS like bat coronaviruses [16]. Further, it was revealed that both the SARS CoV-2 and SARS-CoV had similar receptor-binding domains [17]. The SARS-CoV has 14 binding residues with the human angiotensin-converting enzyme-2 (ACE-2) receptor. 8 residues out of these 14 have been observed in the new SARS-CoV-2. The main protease is highly conserved between the two with an overall identity of 96%. In fact, according to Zhou et al., 2020, the homology between the SARS-CoV-2 and RaTG13 (SARS-like coronavirus in bats) is 96% at the whole-genome level [18], depicting bat as the zoonotic source of newly evolved SARS-CoV-2 [19].

1.2. Genomic organization

SARS-CoV-2 was found to be an enveloped virus with a particle size of 100–160 nm in diameter and as the name suggests “corona” (means crown), it has crowned like projections when seen under electron microscopy [20]. It has a positive sense RNA as a genetic material having size of 29.9 kb (Fig. 1) roughly distributed among adenosines (30%), cytosines (18%), guanines (20%), and rest thymines (32%) [21]. The genome is compacted into a helical nucleocapsid often enclosed by a lipid bilayer. Also, the RNA genome of CoVs is largest among all the RNA viruses [22].

Recent studies showed that the genomic organization of SARS-CoV-2 is similar to other *betacoronaviruses*. It is composed of 14 open reading frames (ORFs) arranged as main replicase assembly flanked by 5' untranslated region and a 3' untranslated region (UTRs) (Fig. 1). A replicase complex encodes a total of 27 proteins divided into structural and non-structural proteins (NSPs) (Fig. 2). The structural proteins encoding genes are present in a fixed order as, a spike protein (S) gene, envelope protein (E) gene, a membrane protein (M) gene, and a nucleocapsid protein (N) gene. They make about 33% of the entire genome and are vital for the viral life cycle. Rest 67% is distributed among 16 different NSPs (such as M<sup>pro</sup>, helicase, and RNA-dependent RNA polymerase) encoding genes across the ORF [23]. They are involved in virus-cell receptor interactions during viral entry. Some unidentified non-structural open reading frames are also present [23].

The *orf1ab* is the largest gene in SARS-CoV-2, encoding a polyprotein (PP1ab). Another gene *orf1a* encodes for a polyprotein (PP1a). Two-thirds of viral RNA situated in the first ORF (ORF1ab) encodes a 7096 residues long polyprotein. Thus, both *orf1a* and *orf1ab* are translated to produce PP1a and PP1ab polyproteins, which are cleaved by the proteases that are encoded by ORF1a to yield 15 non-structural proteins (NSPs). Sequence studies have noted variations between SARS-CoV-2 and SARS-CoV such as the lack of 8a gene or in variation in the number of base pairs in 8b and 3c genes in SARS-CoV-2 [24]. In addition, around 380 substitutions are recognized within the genome of SARS-CoV-2

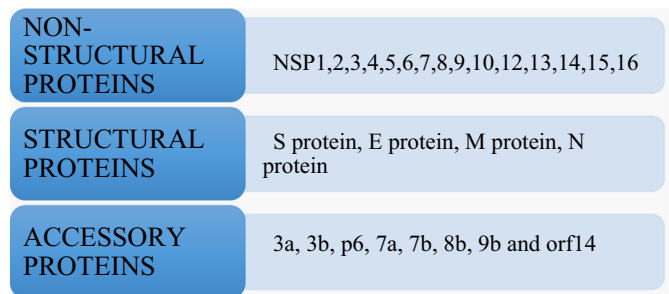


Fig. 2. Classification of 27 proteins of SARS-CoV-2.

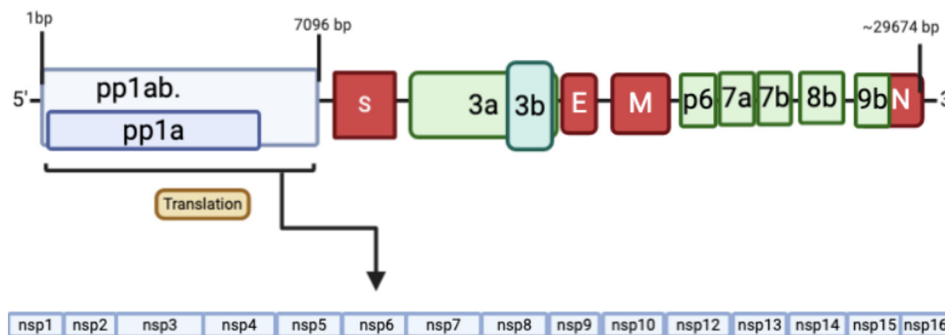


Fig. 1. Depiction of the genomic organization of SARS-CoV-2.

when compared to previously known coronaviruses in a systemic study, which may have affected the functionality and pathogenicity of this novel virus [24].

These said proteins are extensively studied for devising new antiviral agents against COVID-19 because the genome, as well as 3D structures, indicate that main drug-binding pockets are probably conserved across SARS-CoV-2, SARS, and MERS [23].

### 1.3. Pathogenesis

Viral infection at the cellular level initiates when a virus effectively engages with surface receptors expressed by the host and terminates with the release of the new infectious virion. The first step employed by all enveloped viruses during the entry into the host cell is the fusion of their lipid bilayer with host cell membranes through clathrin-mediated endocytosis [7].

For coronaviruses, entry is found to be often biphasic in nature, which means it can enter the cell when near the cell surface or in the late endosomal phase [8]. In addition, a characteristic coronavirus fusion peptide (fragment of the fusion protein) that works in a calcium-dependent manner is employed during membrane fusion. Different receptors are identified for different CoVs for example, ACE2 receptor for SARS-CoV (22), DPP4 (aka CD26) receptor for MERS-CoV [9] and receptor of HCoV-229E is found to be aminopeptidase N (aka CD13). Lung epithelial cells are the primary targets for SARS-CoV-2, and binding takes place between receptor binding domain of class I viral fusion protein called spike (S) glycoprotein and ACE-2 receptor of lung epithelial cells [10] (Fig. 3).

According to the findings of Walls et al., 2020, it is found that 3-D structures of the receptor-binding domain of spike proteins of SARS-CoV-2 and SARS-CoV are strikingly similar [11]. Molecular interactions studies and crystal structure analysis, also confirms that spike protein of SARS-CoV-2 binds to human angiotensin-converting enzyme 2 (ACE2) receptor and readily infects ciliated bronchial cells and type II pneumocytes found in lungs. Unfortunately, SARS-CoV-2 has more affinity for ACE2 receptors than SARS-CoV due to a single N501T mutation in the S protein of SARS-CoV-2 [25], which aids in their transmission from host to host. It is noteworthy that the expression level as well as allele frequency of ACE2 receptors varies among populations. These factors may correlate the disease susceptibility and sequence polymorphism. The binding of the viral S protein to ACE2 induces a negative feedback loop. As the level of ACE2 drops, its counterpart enzyme ACE picks up its role of converting substrate angiotensin I to angiotensin II. Elevated levels of angiotensin II results in the profound binding to its receptor, AGTR1A resulting in the increased pulmonary vascular

permeability [26]. A study of 99 patients infected with SARS-CoV-2 showed that females were less susceptible to infection than males, and older males with comorbidities were more likely to be infected with SARS-CoV-2, and highlighted insights into the role of ACE2 in the SARS-CoV-2 pandemic [27,28]. Advance research in the field of structural and molecular biology shall reveal the complexities of the molecular mechanism that prompts coronavirus S-mediated membrane fusion.

### 1.4. Clinical features

SARS-CoV-2, the seventh member of the coronavirus family has the characteristic manifestation of both lower and upper respiratory tracts with symptoms like dry or productive cough, sore throat, lassitude followed by fever. Serious cases develop pneumonia, pulmonary edema, sepsis, respiratory distress, or septic shock. Elderly people or persons already suffering from comorbidities such as diabetes, cardiovascular disease, hypertension, chronic respiratory illness, etc. are more prone to severe complications. Children are found to be at a lower risk for serious illness, with the symptomatic illness being rare and mild. Infections in pregnant women appear to be like non-pregnant women. There is no concrete evidence to support the risk of vertical transmission. However, Alzamora et al. reported a severe COVID-19 disease in a 41-year-old pregnant woman with risk factors of diabetes mellitus and obesity, with respiratory distress, mechanical invasive ventilation, and preterm delivery. The neonate tested positive on RT-PCR for nasal swab 16 h post caesarean section [29]. The sterility of a C-section coupled with isolation measures immediately after, rule out the possibility of infection during or after birth, strongly raising the suspicion of in utero transmission of the virus. The neonate needed external ventilator support for 12 h, then, he was extubated and put on CPAP. The outcome was favourable and he did not require antibiotic support. Lab tests and imaging were found to be normal throughout [30].

Infected patients have shown high leukocyte counts, abnormal respiratory functioning, or boosted pro-inflammatory cytokines such as TNF- $\alpha$ , MIP1 $\alpha$ , Interleukin-2, 7, and 10. Additionally, C-reactive protein, erythrocyte sedimentation rate, and D-dimer formation also peaks with infection [31].

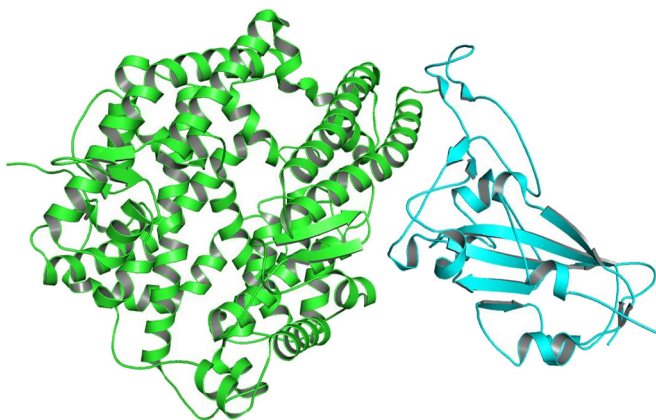
COVID-19 showed some unique clinical features that include the targeting of the lower airway as is evident by symptoms like rhinitis, sneezing, and sore throat (Fig. 4). Also, patients infected with COVID-19 were found to have developed intestinal symptoms like diarrhoea [32], while only a few of MERS or SARS patients had diarrhoea [31].

### 1.5. Disease transmission

Upon entering the host, SARS-CoV-2 has been reported to stay in the respiratory tract for a few days, which is the asymptomatic period. It has been found that the virus persists for up to 14 days in severe cases and 8 days in non-severe cases. The viral load profile is found to peak at the time of onset of symptoms, and so is transmissible easily at an early stage of infection. A high viral load is found in severe cases and can be potentially used as a marker of case severity and prognosis. The viral RNA particle is found in the faeces for 5 days after the onset of the symptom and persists up to 4 to 5 weeks, as well as in total blood, urine, serum, and saliva [32].

The main mode of transmission is through fomites and respiratory droplets. Rates of transmission appear to be similar for asymptomatic and symptomatic patients. According to the data revealed by European Centre for Disease Prevention and Control, the environmental sustainability of active SARS-CoV-2 is up to 3 h in the air after aerosolization, around 24 h on chipboard, and almost 3 days on stainless steel or plastic.

Based on the assessment made by the Chinese government, the WHO indicated that about 80% of affected individuals have symptoms similar to those of mild seasonal influenza. The remaining 20%,



**Fig. 3.** Structure of SARS-CoV-2 spike receptor-binding domain (depicted in cartoon representation in cyan color) bound with ACE2 (depicted in cartoon representation in green color) (PDB ID: 6M0J).



Fig. 4. The systemic and pulmonary disorders caused by COVID-19 infection.

however, develop viral interstitial pneumonia, requiring hospitalization. In approximately 5% of these patients, the pneumonia is critical and requires intensive care. Overall mortality ranges between 0.5% and 5%, with a clear positive correlation with age and other comorbidities [33].

## 2. Structural organization

### Structural proteins

As mentioned earlier, the CoVs typically have 4 conserved structural proteins classified as the matrix (M), the small envelope (E) proteins, the trimeric spike (S) glycoproteins, and the nucleocapsid (N) proteins [23] (Fig. 5). An additional glycoprotein called hemagglutinin esterase (HE), is often present in *betacoronaviruses* including SARS-CoV-2 [34]. Homology modelling enables to reveal of the tentative structure of various structural/non-structural proteins based on previously known proteins of SARS coronavirus (Fig. 5).

### 2.1. N protein

The N protein or nucleocapsid protein of coronavirus is a subnuclear structure. It is generally found in abundance in primary host cells as well as virions carrying plasmids that express N protein [35]. N protein plays a key role in forming a long helical nucleocapsid structure or ribonucleoprotein (RNP) complex composed of viral RNA and N-terminal domain of N Protein (N-NTD) [36]. These proteins release the viral genomes into the membranous network of the endoplasmic reticulum-Golgi intermediate compartment (ERGIC) where they assemble with other structural proteins to form the mature virion. It is discovered in immunological experiments that the RNP complex is involved in maintaining stable RNA conformation required for the replication and transcription of the viral genome [37].

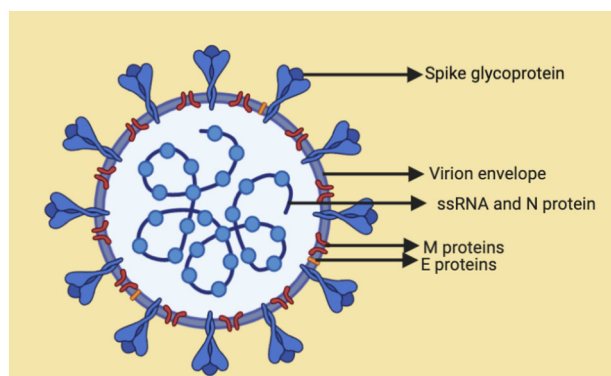


Fig. 5. Structure of SARS-CoV-2 virion.

### 2.1.1. Structural details of N protein

Structurally, N proteins comprise of three domains arranged as a N-terminal RNA-binding domain (NTD), followed by a Ser/Arg (SR)-rich central linker region and a C-terminal dimerization domain (CTD) [37]. The NTD and CTD of the N proteins execute the binding and oligomerization of RNA respectively, whereas the central linker region of the N protein is the RNA-binding sites as it is chief phosphorylation site [37]. The previously published structural analysis showed that the NTD is a fist-shaped structure rich in basic amino acids while the core is composed of loops of  $\beta$ -sheets providing flexibility there [37].

Lin et al., 2014, carried out research to identify ribonucleotide-binding sites using ribonucleoside monophosphate instead of RNA. It showed that Arg122, Tyr124, Tyr126, and Arg164 side chains participate in the formation of ribonucleotide-binding pocket and also reach out to RNA via covalent/non-covalent interactions. These critical residues of N protein are conserved in another human coronavirus also. Few other identified amino acid residues in the N protein involved in RNA binding are Y124, R122, and R164 [35].

### 2.1.2. Functional aspects of N protein

The N protein is employed in regulating viral RNA replication, transcription, and genome packaging [37]. During virion assembly, N protein and RNA compacts to form the helical nucleocapsid [35]. They regulate cellular processes like cell cycle, cytoskeleton reorganization or host cell apoptosis [35]. They are also capable of eliciting protective immune responses from the host due to its highly immunogenic phosphoprotein [37].

### 2.1.3. Potential of N protein in drug discovery

Previous studies proposed that N proteins have a big role to play in the life cycle of coronaviruses thus can positively be seen as antiviral drug targets [37]. Mutation in the RNA-binding sites at the N terminal domain may result in a reduced affinity for RNA binding [37]. Thus, it would be reasonable to target the RNA-binding domain of coronavirus N protein for broad range anti-coronavirus drugs. Competitive inhibitors of this site may be employed to combat CoV infection. Further, by the virtue of the conservation of N protein sequence and its strong immunogenicity, the N protein can be seen as a diagnostic tool. Viral copy numbers can be quantified in the test sample via quantitative real-time RT-PCR (qRT-PCR) followed by determination of N protein expression in the cells using immunofluorescence microscopy after 48h of infection [30].

## 2.2. Spike protein (S)

The spike protein of coronaviruses is a glycoprotein. It contains clove-shaped protrusions that can bind only to their unique receptors present on the host cell surface. The spike protein has three segments: (i) ectodomain (ED) region, (ii) transmembrane (TM) region, and (iii)

intracellular (IC) domain, which comprise the intracellular short tail part [38]. It is a large transmembrane protein consisting of two subunits, S1 trimeric stalk and S2 trimeric stalk on C-terminal together comprise the ED [39], both are crucial for the successful entry of coronavirus into a host cell. S1 initiates the first step of viral entry, as it contains the receptor-binding domain. S2 later mediate the fusion of the cell and viral membranes. The critical step of cleavage of S protein is generally mediated by host protein convertase that allows for the fusion sequences to be exposed as required for the fusion. Cleavage of S protein at S1/S2 site mediates cell-cell fusion and is an essential step for the entry of S protein into human lung cells. This vital step is governed by cellular protease furin that performs this cleavage. SARS-CoV-2, like MERS-CoV depends on this furin-mediated pre-cleavage of S proteins at the S1/S2 site for the activation of S protein by TMPRSS2 in lung cells, which fail to express robust levels of cathepsin L. Inhibitors of furin and TMPRSS2 can serve as possible therapeutic options in COVID-19. A recent study reported that the inhibitor of TMPRSS2 blocks SARS-CoV-2 infection [40]. However, an interesting point that should be noted is that furin is needed for normal development unlike TMPRSS2 and hence, obstructing furin for longer periods can be associated with unwanted toxic effects [40].

Genomic analysis of the novel coronavirus revealed that its spike protein differs from those of close bat CoV relatives, in the sense that it has a specific site, which gets activated by an enzyme called furin (protease) and three adjacent O-linked glycans. Although, few commonalities have also been proved between SARS-CoV-2 and MERS-CoV stressing on their furin-mediated cleavage of S protein at specific site and subsequent activation by TMPRSS2 of host cell. Furin belongs to subtilisin-like proprotein convertase family, which facilitate the conversion of precursor proteins into their biologically active form. Structurally, furin is a type I trans membrane protease expressed by various host tissues including humans. The S1/S2 cleavage site is the target for furin during infection. The reported furin cleavage sequence found to be polybasic, predominantly arginine residues (PRRARS♦V) but the function is still unclear [41,42]. Since, furin is present in various human organs like the lungs, small intestines, or liver, it aids virus to metastasize in the body, says Li Hua, a structural biologist at Huazhong University of Science and Technology. Viruses that reach target sites (like respiratory tract) convert and activate their own surface glycoproteins using furin. This makes their role in viral protein processing significant. There is an implied possibility of cleavage of SARS-CoV-2 S glycoprotein once it exits from lung epithelial cells and consequently can efficiently infect other cells [43]. This aspect makes SARS-CoV-2 more dangerous than its other counterparts in the same genus. Acquisition of the furin cleavage site can be expressed as 'gain of function' that allowed a bat CoV to infect humans [41]. It can also transmit between different hosts through gene recombination or mutations in the receptor-binding domain, leading to a higher mortality rate.

The SARS-CoV-2 virus can enter human bronchial epithelia through contact with the ACE2 receptors on them, making this virus highly host-specific. Recently, the crystal structure of the spike protein's receptor-binding domain of SARS-CoV-2, in complex with human ACE2, was released by Wang and Zhang's group [25]. Interestingly, a study conducted revealed that ACE2 expression is higher in males than females. Supporting this notion 61.8% deaths were reported in males in a sample study carried out in New York. Additional findings suggest that the binding of SARS-CoV-2 to ACE2 is synergistic to the expression of ACE2 [44].

### 2.2.1. Structural details of the spike (S) protein

The main components of the S1 domain are the N-terminal domain (NTD) and the C-terminal domain (CTD). This S1 domain is marked as a major antigen on the surface of the coronavirus. Receptor binding domain (RBD) at the head region of S1 recognizes the host cell receptor (ACE2). It is reported that the glutamine 394 and lysine 31 are the key

residues of RBD and ACE2 respectively that are involved in this interaction [25] initiating the membrane fusion between viral particles and the host's receptor. There are 18 residues of ACE-2 that interact with the RBD (contain 14 amino acids) of spike protein and for this contact, K341 of ACE-2 and R453 residue of RBD play the most important role (38). A small section of the S1 region formed between 318 and 510 amino acid residues are plenty for strong attachment to the peptidase domain of ACE2 [25]. According to another report six RBD residues: L455, F486, Q493, S494, N501 and Y505 are actively involved in ACE2 receptor binding and determining host range of SARS-CoV-2 [45]. S2 contains basic elements like hydrophobic fusion peptide needed for membrane fusion (Fig. 6). X-ray crystallographic analysis demonstrates that HR1 and HR2 domains combinedly form the 6-HB fusion core in the S2 subunit of SARS-CoV-2. The structure is crucial for membrane fusion in SARS-CoV-2 pathogenesis [46]. Fig. 6 below shows the structure of SARS-CoV-2 6-HB as cartoon and surface representation with HR1 colored in red and HR2 in green. Using the Wimley-White interfacial

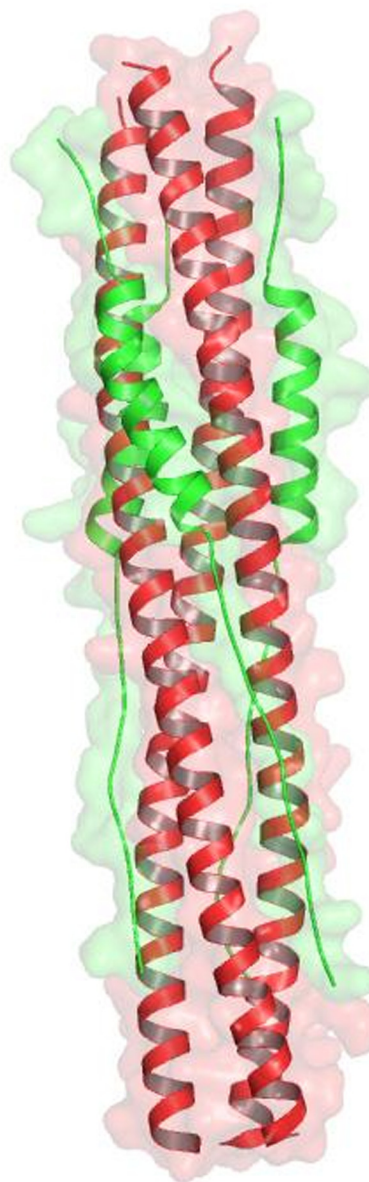


Fig. 6. Structure of SARS-CoV-2 6-HB depicted in cartoon and surface representation with HR1 colored in red and HR2 in green. (PDB ID: 6LXT).

hydrophobic interface scale, it was reported that fusion peptide SARS-CoV is a part of N-terminal domain which is conserved across the *Coronaviridae* [47].

### 2.2.2. Functional aspects of the spike protein

The spike protein of SARS-CoV-2 mediates the binding of the virus particle to the host cell surface initiating the infection process [39]. It has a key role in neutralizing antibody and T-cell responses of the host thus compromise its immunity [47]. These proteins define the range of host specificity of the virus i.e. they determine the host tropism and transmission capacity.

### 2.2.3. Potential of spike protein in drug discovery

The spike (S) glycoprotein is a promising target for the development of entry inhibitor drugs. Currently, the best-characterized antibodies against coronaviruses target the receptor-binding domain (RBD) of the S protein and prevent binding to host cells [30]. The RBD, however, is the most variable part of the spike protein, and antibodies that target this domain are unlikely to be cross-reactive. Point mutations on D454 or R441 site of RBD said to disturb the binding pocket for ACE2 [39] or another option is an anti-ACE-2 antibody that could block viral entry and its replication in host cells [48]. The 6-helices or a 5-helices structure is used as a receptor for protein or peptide docking [39]. Scientists are also looking at molecules that could block furin, which could be investigated as possible therapies [41,49–51].

## 2.3. Envelope protein (E protein)

It is the smallest membrane protein with only 8.4–12 kDa size [52]. Two distinct domains comprise E protein i.e. the hydrophobic domain, and the charged cytoplasmic tail. It forms pentamers and functions as an ion channel, thus named as E channel or viroporin. It is found near the ER and Golgi body regions [52].

### 2.3.1. Functional aspects of the envelope protein

The E protein has a special role in viral morphogenesis, especially during assembly and viral progress [53]. E protein also helps in increasing viral titre and mature progenies [53]. Many other studies infer that the E protein acts in coordination with other intracellular proteins and modulates the activity of those proteins [53]. E protein can also act as a virulence factor [54].

### 2.3.2. Potential of envelope protein in drug discovery

As E protein is reported to be involved in intracellular protein trafficking and regulation, it would be an excellent approach to devise drugs to intervene in its expression. According to an in-vivo study on the mammalian cell line, hexamethylene amiloride drug blocks the ion channel [53].

## 2.4. Membrane protein (M protein)

M protein is the most abundant structural protein, responsible for providing the framework to the virion particle [55]. It spans the envelope thrice with a long C terminal and short N terminal at both ends. M protein co-operates with other essential proteins namely spike proteins and nucleocapsid in special assembly [55].

### 2.4.1. Functional aspects and potential of membrane protein in drug discovery

M protein aids in protein-protein interaction. Viral intracellular homeostasis is also maintained by M protein [55]. It acts as an epitope in

humoral responses [55]. Since M protein is determined as a recognisable antigen, the development of antibody therapy against is a good option.

## 2.5. Non-structural proteins

Among several proteins encoded by the SARS-CoV-2 genome, 16 NSPs are encoded by replicase, scattered between two polyproteins set i.e., PP1a and PP1ab. There are two kinds of cysteine proteases namely chymotrypsin-like cysteine protease or (3CL<sup>pro</sup>) and papain-like protease (PL<sup>pro</sup>) act on these PPs to release the NSPs [56]. They cleave at the C-terminal end and the N-terminal end of these PPs respectively [57]. The first three cleavage sites of the PPs are cut by PL<sup>pro</sup>, releasing four NSPs and the rest 12 NSPs are released after CL<sup>pro</sup>, by cleaving the remaining 11 sites [57].

## 2.6. Main protease (M<sup>pro</sup>)

M<sup>pro</sup> also called chymotrypsin-like protease (3CL<sup>pro</sup>) and the presence of cys-his dyad on the active site is responsible for its protease activity (47). The most commonly found recognition sequence of M<sup>pro</sup> is Leu,Gln♦Ser,Ala,Gly (where ♦ denotes the cleavage site) [58]. M<sup>pro</sup> is majorly involved in the cleavage of PPs to generate (NSPs) which later compile into the replicase-transcriptase complex (RTC) [59].

### 2.6.1. Structure of M<sup>pro</sup>

M<sup>pro</sup> exists as a homodimer having a molecular mass of 33.797 kDa, as determined by mass spectroscopy [59]. The 2.1 Å high-resolution crystal structure reveals the two-fold symmetry of the M<sup>pro</sup> molecule where the two subunits are held together by a salt-bridge interaction between Glu290 of one protomer and Arg4 of the other [59] (Fig. 7). Each monomer has three distinct domains: Domains I and II display antiparallel β-barrel structures whereas domain III is a large antiparallel globular structure made up of 5 α-helices arranged. Domains II and III, are connected using a long loop like segment of domain III. The substrate binds to a pocket created between Domains I and II. It is highly conserved among all the CoVs. On the other hand, the most variable regions were α-helices and surface loops domain III [59]. The two monomers are arranged in a perpendicular orientation such that the domain II of one is in contact with the NH<sub>2</sub>-terminal residues aka 'N-finger' of others. This dimerization is essential for the formation of substrate binding sites, thus also affect catalytic efficiency of the protease, and thereby maintain the shape of the substrate-binding pocket of S1 protein [59]. SARS-CoV-2 Mpro is made up of 306 amino acid residues. Crystal structure analysis suggests that Thr25, Thr26, Leu27, His41, Ser46, Met49, Tyr54, Phe140, Leu141, Asn142, Gly143, Cys145, His163, Met165, Glu166, Leu167, Pro168, Phe185, Asp187, Gln189, Thr190, Ala191, and Gln192 are found in the active site pocket [60].

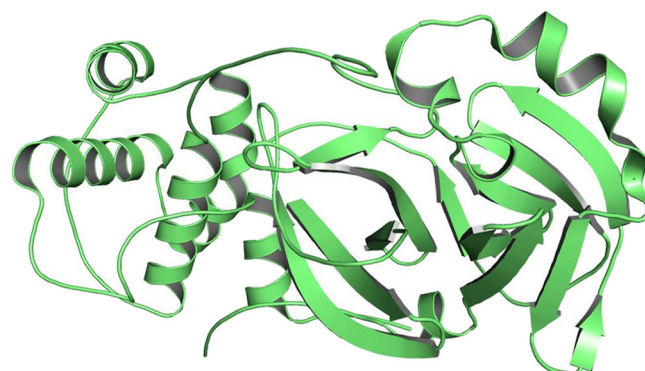


Fig. 7. Structure of SARS-CoV-2 main protease (PDB ID: 6YB7).



## 2.7. Papain-like protease (PL<sup>pro</sup>)

PL<sup>pro</sup> cleaves at the borderlines of NSP1/2, NSP2/3, and NSP3/4. It works in association with M<sup>pro</sup> to cleave the poly protein into NSPs [56]. PL<sup>pro</sup> at its catalytic core domain contains 316 amino acids, which is responsible for cleaving replicase substrates, and a consensus sequence at cleavage sites usually found to be LXGG [59] (Fig. 8).

### 2.7.1. Potential in drug discovery

Major antiviral drugs are designed against M<sup>pro</sup> and PL<sup>pro</sup> proteases. Inhibiting the activity of these enzymes could hamper viral replication. Many studies have recently reported that drugs targeting the main protease can be possible therapeutics in COVID-19. A recent literature used repurposing of FDA approved drugs to identify potential leads that inhibited the SARS-CoV-2 M<sup>pro</sup> and can play a key role in COVID-19 therapeutics [60]. The absence of any similar human protease makes M<sup>pro</sup> an easy target for antiviral drug development [59]. Nonetheless, care should be taken to avoid the chances of host toxicity by ensuring that the cleavage sequence for the host protease should be different [59]. Another approach could be to mutate Ser139 and phe140 positions, which are crucial sites involved in the dimerization of 3CLpro [58].

## 2.8. Accessory NSPs

### 2.8.1. NSP13/helicase

The helicase unwinds the double-stranded RNA segment into single strands by hydrolysing NTPs. This enzyme prefers ATP, dATP, and dCTP as substrates. Due to its sequence conservation in all coronaviruses, this helicase is an easy target to develop antiviral drugs [61]. The ADP binding site and the nucleic acid binding site of NSP 13 are considered in molecular docking studies while designing inhibitors but a serious limitation is the non-specificity of inhibitors that may cause host toxicity.

### 2.8.2. NSP12/RNA-dependent RNA polymerase

RdRP is the structurally conserved RNA polymerase that executes RNA replication and transcription with the help of its cofactors NSP7 and NSP8 (Fig. 9). This complex forms the largest (160 kDa) unit

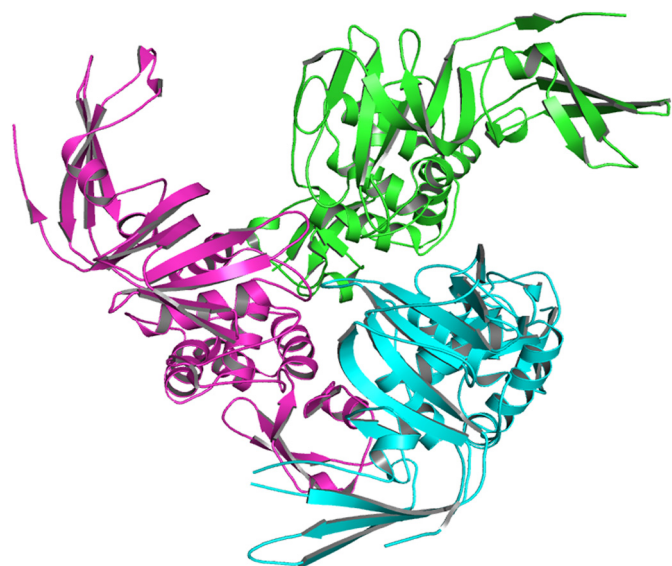


Fig. 8. Structure of papain-like protease of SARS-CoV-2 (PDB ID: 6W9C).

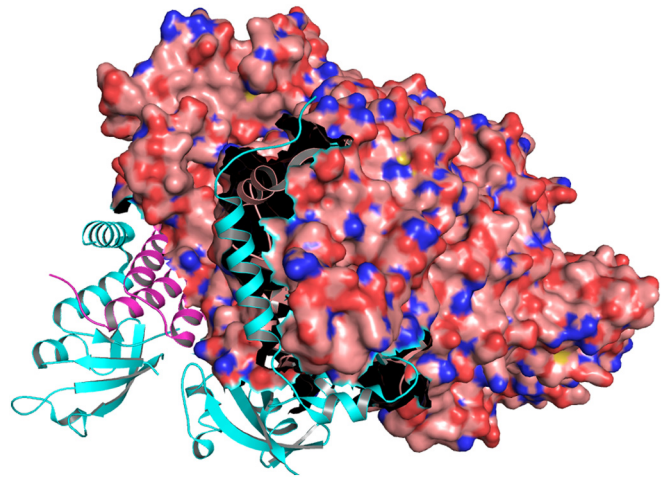


Fig. 9. Structure of RNA dependent RNA polymerase of SARS-CoV-2 (NSP12) in association with NSP7 and NSP8. NSP12 is depicted as surface while NSP7 (purple) and NSP8 (cyan) are depicted in cartoon form. (PDB ID: 7BTF).

involved in RNA synthesis. It also possesses the characteristic *nidovirus* NSP12 N-terminal domain having kinase-like structural fold [62].

### 2.8.3. NSP14/N-terminal exoribonuclease and C-terminal guanine-N7 methyltransferase

It is involved in the various steps of the central dogma of novel coronavirus. It has two main domains: The N-terminal exoribonuclease (ExoN) domain has an important part to play in proofreading and preventing any fatal mutations, and the other is the C-terminal domain act as a guanine-N7 methyltransferase (N7-MTase) required for mRNA capping [63].

### 2.8.4. NSP15/uridylylate-specific endoribonuclease

It is one of the RNA-processing enzymes encoded by the coronavirus genome. It forms a six-unit endoribonuclease that cleaves 3' end of uridines [64].

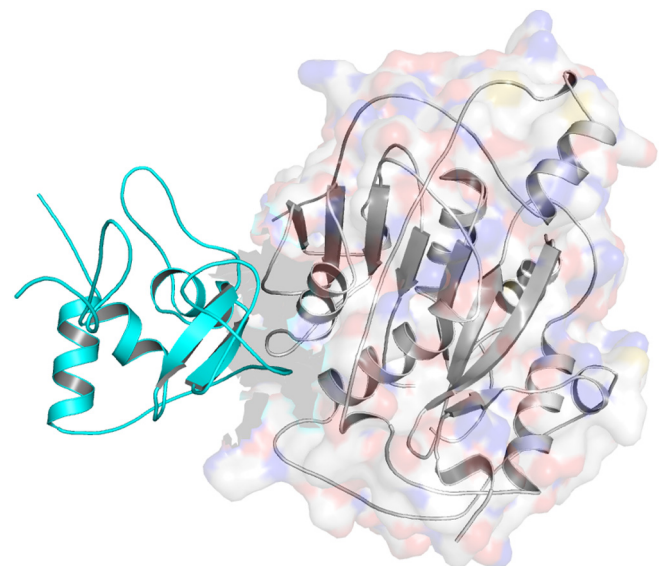


Fig. 10. Structure of NSP16 (depicted in surface and cartoon form) and NSP10 (depicted in cartoon form, cyan color) complex from SARS-CoV-2. (PDB ID: 6W61).

### 2.8.5. NSP16/2'-O-methyltransferase

NSP16 is a S-adenosylmethionine (SAM) dependent nucleoside-2'-O-methyltransferase. Activation of MTase required binding with NSP10 along with the presence of a bivalent cation mostly  $Mg^{+2}$  (Fig. 10). It regulates cellular processes like toxin neutralization, signal transduction, chromatin remodelling, or other post-transcriptional mechanisms.

### 2.8.6. NSP10

It acts as a cofactor and forms an assembly with NSP14 and NSP16. In the heterodimer conformation NSP16 is positioned over the NSP10 monomer. NSP 10 has an overall conserved structure comprising of a N-terminus made up of two  $\alpha$ -helices, a central  $\beta$ -sheet domain, and a C-terminus domain having various loops and helices [65].

### 2.8.7. Hemagglutinin esterase (HE)

This HE enzyme when present is localized in the envelope of CoV. Usually, in *betacoronavirus*, both the spike proteins and HEs moiety bind to O-acetylated-sialic-acids, where the Hemagglutinin esterase domain encourages virus elution by destroying the receptor for attachment. Nonetheless, this action of binding of O-acetylated-sialic-acids is reversible [66]. This prevents the permanent damage of receptors in the respiratory or gastrointestinal tract. Moreover, HE also helps in the freeing of nascent virus particles from the infected host cells [66].

### 2.8.8. Trans membrane protease serine 2 (TMPRSS2)

According to the study of Hoffman et al. trans membrane protease serine 2 (TMPRSS2) mediated priming of spike protein is required for the efficient entry of virus particles into the host cell. It eventually leads to the infection of other cells that possess ACE2 receptors [40,65]. Camostat mesylate (serine protease inhibitor) is well known for its action against TMPRSS2 activity, which makes it a potential drug candidate to test for SARS-CoV-2 [65].

## 3. Treatment and testing

As per Centres for diseases control and prevention CDC, two types of tests are available for COVID-19. Viral test confirms current infection state (at the time of testing). It is primarily based on reverse transcribed quantitative PCR analysis of the sample of patient along with serological analysis for reliable results. Primer sets for diagnosis should be cautiously designed, since rapidly occurring variants would affect the performance of the RT-qPCR. Another, antibody test that detect if a person had any infection history. An antibody test may not detect current infection, as it takes 15–20 days make antibodies after infection. It should be noted that having antibodies against infection, do not necessarily protect a person from getting infection again with the virus. ELISA based immunoassays are also engaged along with viral testing to increase detection sensitivity. Early detection by the means of radiological examinations such as chest X ray (CXR) and chest computed tomography (CT) scan are also employed in the testing of COVID-19 [26].

## 4. Emerging treatment options

Currently approved treatment modalities consist of supportive care teamed with oxygen supplementation performed via non-invasive ventilation, or through mechanical ventilation. Severe cases may also be treated with antibiotics against bacterial infection and vasopressor support. Clinicians from the USA and Italy have reported several other complications such as acute myocardial damage, thromboembolic phenomena such as PE, and sudden death. Around 12 trials are in progress to test possible treatments against COVID-19. Wang et al. from Wuhan Institute of Virology screened some of the FDA approved anti-virus or anti-infection drugs and found that remdesivir and chloroquine

could effectively prevent the viral growth in cell-based assay with EC50 of 0.77 and 1.13  $\mu M$ , respectively [67]. Repurposing these drugs happens to be the most practical approach in response to rapidly increasing positive cases.

Other antiviral drugs, monoclonal antibodies, and antibody-rich plasma from recovered patients are also under trials. On-going therapies can be broadly classified into two on the basis of their targets; first approach targets the virus entity by blocking the key enzyme responsible for viral life processes (replication, translation etc.) or by preventing the entry of virus particle into human cells. The other approach stress on immune system functioning either by boosting innate/humoral immunity or by inhibiting inflammatory response that led to pulmonary failure [26]. On 20 March, WHO announced the launch of SOLIDARITY, a coordinated effort to collect scientific data rapidly. The study design has been kept simple so that even hospitals with an overwhelming number of patients can participate. The physician simply records the date the patient left the hospital or died, the duration of the entire stay, and whether the patient required ventilation or oxygen. The design is not blinded, and the WHO website randomizes the patient to one out of four drug regimes, or local standard care. For this study WHO has chosen an experimental antiviral remdesivir; the antimalarial drug – hydroxychloroquine, a combination of antiretroviral drugs lopinavir and ritonavir; and this combination plus IFN $\beta$ . Kong et al., recently introduced a docking server for docking small molecules, peptides and antibodies against potential targets of COVID-19. This can be very useful for researchers across the world for predicting target-ligand interactions for COVID-19 [68].

## 5. Conclusions and future directives

The evolution of novel coronavirus and the eventual outbreak of COVID-19 has put the burden on the medical fraternity and the scientific community to identify the components of this viral entity including its genome, replication cycle, proteome as well as modes of transmission. This can only be achieved with collaborative efforts backed up by advances in molecular and structural virology. Viral proteins are identified as potential targets since they are involved in each phase of the viral life cycle. In the present review, we discussed the structure, function, and host-virus interactions of various SARS-CoV-2 proteins based on a thorough review of the latest literature available. We hope our review will give a useful groundwork for the devising new drugs against COVID-19.

Although many potential therapies are under clinical trial, no pharmaceutical products have yet been offered to be reliable as well as effective against COVID-19 (WHO, 31st March 2020). We need to focus on bridging the gaps in knowledge of COVID-19 biology and pathogenesis. This can be achieved by application of newer approaches like metagenomics sequencing, 3D Cryo-EM structures, and extensive in-vivo experimentation along with molecular docking for anti-viral drugs, which may provide basis for devising new safe drugs and reliable testing kits. This will greatly contain the spread and survival of SARS-CoV-2. On the other hand, this whole ordeal of COVID-19 indicates the possibility of the evolution of new pathogens that can cross the species barrier and causing serious damage to mankind. In conclusion, we must have an open-eyed approach in our progress against such pandemics.

## Declaration of competing interest

The authors declare no conflict of interest.

## Acknowledgements

This work was supported by the grant from the Indian Council of Medical Research (ICMR) BIC/12(16)/2014. The authors are indebted to FIST Program (SR/FST/LSI-541/2012), Jamia Millia Islamia (a Central

University) and Department of Orthopedics, Medical College, Jalaun, UP, for facilitating the research.

## References

- [1] C. Sohrabi, Z. Alsafi, N. O'Neill, M. Khan, A. Kerwan, A. Al-Jabir, C. Iosifidis, R. Agha, World Health Organization declares global emergency: a review of the 2019 novel coronavirus (COVID-19), *Int. J. Surg.* 76 (2020) 71–76, <https://doi.org/10.1016/j.ijssu.2020.02.034>.
- [2] W.H. Organization, WHO Coronavirus Disease (COVID-19) Dashboard, World Health Organization, Geneva, Switzerland. <https://covid19.who.int/>, Accessed 5 (2020).
- [3] Y. Liu, R.M. Eggo, A.J. Kucharski, Secondary attack rate and superspreading events for SARS-CoV-2, *Lancet* 395 (10227) (2020) e47.
- [4] H. Lodish, A. Berk, S.L. Zipursky, P. Matsudaira, D. Baltimore, J. Darnell, *Viruses: structure, function, and uses*, Molecular Cell Biology, 4th edition WH Freeman, 2000.
- [5] H.R. Gelderblom, Structure and classification of viruses, *Medical Microbiology*, 4th edition University of Texas Medical Branch at Galveston, 1996.
- [6] P.C. Woo, S.K. Lau, C.S. Lam, C.C. Lau, A.K. Tsang, J.H. Lau, R. Bai, J.L. Teng, C.C. Tsang, M. Wang, Discovery of seven novel mammalian and avian coronaviruses in the genus deltacoronavirus supports bat coronaviruses as the gene source of alphacoronavirus and betacoronavirus and avian coronaviruses as the gene source of gammacoronavirus and deltacoronavirus, *J. Virol.* 86 (7) (2012) 3995–4008.
- [7] H. Wang, P. Yang, K. Liu, F. Guo, Y. Zhang, G. Zhang, C. Jiang, SARS coronavirus entry into host cells through a novel clathrin- and caveolae-independent endocytic pathway, *Cell Res.* 18 (2) (2008) 290–301.
- [8] J.K. Millet, G.R. Whittaker, Physiological and molecular triggers for SARS-CoV membrane fusion and entry into host cells, *Virology* 517 (2018) 3–8.
- [9] G. Lu, Y. Hu, Q. Wang, J. Qi, F. Gao, Y. Li, Y. Zhang, W. Zhang, Y. Yuan, J. Bao, Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26, *Nature* 500 (7461) (2013) 227–231.
- [10] S. Belouzard, J.K. Millet, B.N. Licitra, G.R. Whittaker, Mechanisms of coronavirus cell entry mediated by the viral spike protein, *Viruses* 4 (6) (2012) 1011–1033.
- [11] A.C. Walls, Y.-J. Park, M.A. Tortorici, A. Wall, A.T. McGuire, D. Velesler, Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein, *Cell* 2 (2020) 281–292, <https://doi.org/10.1016/j.cell.2020.02.058>.
- [12] N. Petrosillo, G. Viceconte, O. Ergonul, G. Ippolito, E. Petersen, COVID-19, SARS and MERS: are they closely related? *Clin. Microbiol. Infect.* 6 (2020) 729–734, <https://doi.org/10.1016/j.cmi.2020.03.026>.
- [13] F. Wu, S. Zhao, B. Yu, Y.-M. Chen, W. Wang, Z.-G. Song, Y. Hu, Z.-W. Tao, J.-H. Tian, Y.-Y. Pei, A new coronavirus associated with human respiratory disease in China, *Nature* 579 (7798) (2020) 265–269.
- [14] X.-Y. Ge, J.-L. Li, X.-L. Yang, A.A. Chmura, G. Zhu, J.H. Epstein, J.K. Mazet, B. Hu, W. Zhang, C. Peng, Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor, *Nature* 503 (7477) (2013) 535–538.
- [15] S.A. Jeffers, S.M. Tusell, L. Gillim-Ross, E.M. Hemmila, J.E. Achenbach, G.J. Babcock, W.D. Thomas, L.B. Thackray, M.D. Young, R.J. Mason, CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus, *Proc. Natl. Acad. Sci.* 101 (44) (2004) 15748–15753.
- [16] F. Zhou, T. Yu, R. Du, G. Fan, Y. Liu, Z. Liu, J. Xiang, Y. Wang, B. Song, X. Gu, Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study, *Lancet* 395(10229) (2020) 1054–1062, [https://doi.org/10.1016/S0140-6736\(20\)30566-3](https://doi.org/10.1016/S0140-6736(20)30566-3).
- [17] S. Khan, R. Siddique, M.A. Shereen, A. Ali, J. Liu, Q. Bai, N. Bashir, M. Xue, Emergence of a novel coronavirus, severe acute respiratory syndrome coronavirus 2: biology and therapeutic options, *J. Clin. Microbiol.* 58 (5) (2020).
- [18] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang, A pneumonia outbreak associated with a new coronavirus of probable bat origin, *Nature* 579 (7798) (2020) 270–273.
- [19] A. Rezaibakhsh, A. Ala, S.H. Khodaie, Novel coronavirus (COVID-19): a new emerging pandemic threat, *Journal of Research in Clinical Medicine* 8 (1) (2020) 5–5.
- [20] S.G. Siddell, The Coronaviridae, the Coronaviridae, Springer, 1995 1–10.
- [21] R. Sah, A.J. Rodriguez-Morales, R. Jha, D.K. Chu, H. Gu, M. Peiris, A. Bastola, B.K. Lal, H.C. Ojha, A.A. Rabaan, Complete genome sequence of a 2019 novel coronavirus (SARS-CoV-2) strain isolated in Nepal, *Microbiology Resource Announcements* 9 (11) (2020).
- [22] B. Li, H.-R. Si, Y. Zhu, X.-L. Yang, D.E. Anderson, Z.-L. Shi, L.-F. Wang, P. Zhou, Discovery of bat coronaviruses through surveillance and probe capture-based next-generation sequencing, *MSphere* 5 (1) (2020).
- [23] J. Cui, F. Li, Z.-L. Shi, Origin and evolution of pathogenic coronaviruses, *Nat. Rev. Microbiol.* 17 (3) (2019) 181–192.
- [24] A. Wu, Y. Peng, B. Huang, X. Ding, X. Wang, P. Niu, J. Meng, Z. Zhu, Z. Zhang, J. Wang, Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China, *Cell Host Microbe* 27(3) (2020) 325–328, <https://doi.org/10.1016/j.chom.2020.02.001>.
- [25] H. Zhang, J.M. Penninger, Y. Li, N. Zhong, A.S. Slutsky, Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target, *Intensive Care Med.* (2020) 1–5.
- [26] Y.-F. Tu, C.-S. Chien, A.A. Yarmishyn, Y.-Y. Lin, Y.-H. Luo, Y.-T. Lin, W.-Y. Lai, D.-M. Yang, S.-J. Chou, Y.-P. Yang, A review of SARS-CoV-2 and the ongoing clinical trials, *Int. J. Mol. Sci.* 21 (7) (2020) 2657.
- [27] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, *Lancet* 395 (10223) (2020) 497–506.
- [28] A. Remuzzi, G. Remuzzi, COVID-19 and Italy: what next? *Lancet* 395 (2020) 1225–1228, [https://doi.org/10.1016/S0140-6736\(20\)30627-9](https://doi.org/10.1016/S0140-6736(20)30627-9).
- [29] M.C. Alzamora, T. Paredes, D. Caceres, C.M. Webb, L.M. Valdez, M. La Rosa, Severe COVID-19 during pregnancy and possible vertical transmission, *Am. J. Perinatol.* (2020) <https://doi.org/10.1055/s-0040-1710050>.
- [30] F.F. Rousseau, Du L. Zhao G, Lin Y, Chan C, He Y, Jiang S, Wu C, Jin DY, Yuen KY, Zhou Y, Zheng BJ. Priming with rAAV encoding RBD of SARS-CoV S protein and boosting with RBD-specific peptides for T cell epitopes elevated humoral and cellular immune responses against SARS-CoV infection. *Vaccine.* 2008 Mar 20; 26 (13): 1644–51.
- [31] L. Du, G. Zhao, Y. Lin, H. Sui, C. Chan, S. Ma, Y. He, S. Jiang, C. Wu, K.Y. Yuen, D.Y. Jin, Y. Zhou, *genome* 367 (2) (2007) 428–439.
- [32] H.A. Rothan, S.N. Byrareddy, The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak, *J. Autoimmun.* 109 (2020), 102433. <https://doi.org/10.1016/j.jaut.2020.102433>.
- [33] C.I. COVID, Convalescent Plasma-Emergency INDs, 2020.
- [34] C.P.E.R.E. Novel, The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China, *Zhonghua liu xing bing xue za zhi = Zhonghua liuxingbingxue zazhi* 41 (2) (2020) 145.
- [35] Q. Zeng, M.A. Langereis, A.L. van Vliet, E.G. Huizinga, R.J. de Groot, Structure of coronavirus hemagglutinin-esterase offers insight into corona and influenza virus evolution, *Proc. Natl. Acad. Sci.* 105 (26) (2008) 9065–9069.
- [36] S.-Y. Lin, C.-L. Liu, Y.-M. Chang, J. Zhao, S. Perlman, M.-H. Hou, Structural basis for the identification of the N-terminal domain of coronavirus nucleocapsid protein as an antiviral target, *J. Med. Chem.* 57 (6) (2014) 2247–2257.
- [37] C. Zhao, Z. Lou, Y. Guo, M. Ma, Y. Chen, S. Liang, L. Zhang, S. Chen, X. Li, Y. Liu, Nucleoside monophosphate complex structures of the endonuclease domain from the influenza virus polymerase PA subunit reveal the substrate binding site inside the catalytic center, *J. Virol.* 83 (18) (2009) 9024–9030.
- [38] R. McBride, M. Van Zyl, B.C. Fielding, The coronavirus nucleocapsid is a multifunctional protein, *Viruses* 6 (8) (2014) 2991–3018.
- [39] S. Duquerry, A. Vigouroux, P.J. Rottier, F.A. Rey, B.J. Bosch, Central ions and lateral asparagine/glutamine zippers stabilize the post-fusion hairpin conformation of the SARS coronavirus spike glycoprotein, *Virology* 335 (2) (2005) 276–285.
- [40] L. Du, Y. He, Y. Zhou, S. Liu, B.-J. Zheng, S. Jiang, The spike protein of SARS-CoV—a target for vaccine and therapeutic development, *Nat. Rev. Microbiol.* 7 (3) (2009) 226–236.
- [41] M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, T.S. Schiergens, G. Herrler, N.-H. Wu, A. Nitsche, SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor, *Cell* 181 (2020) 271–280, <https://doi.org/10.1016/j.cell.2020.02.052>.
- [42] B. Coutard, C. Valle, X. de Lamballerie, B. Canard, N. Seidah, E. Decroly, The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade, *Antivir. Res.* 176 (2020), 104742.
- [43] M. Hoffmann, H. Kleine-Weber, S. Pöhlmann, A multifasic cleavage site in the spike protein of SARS-CoV-2 is essential for infection of human lung cells, *Mol. Cell* 78 (2020) 779–784, <https://doi.org/10.1016/j.molcel.2020.04.022>.
- [44] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, A novel coronavirus from patients with pneumonia in China, 2019, *N. Engl. J. Med.* 382 (2020) 727–733, <https://doi.org/10.1056/NEJMoa2001017>.
- [45] C. Wu, S. Zheng, Y. Chen, M. Zheng, Single-cell RNA expression profiling of ACE2, the putative receptor of Wuhan 2019-nCoV, in the nasal tissue, medRxiv (2020) <https://doi.org/10.1101/2020.02.11.20022228>.
- [46] Y. Wan, J. Shang, R. Graham, R.S. Baric, F. Li, Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS coronavirus, *J. Virol.* 94 (7) (2020).
- [47] S. Xia, M. Liu, C. Wang, W. Xu, Q. Lan, S. Feng, F. Qi, L. Bao, L. Du, S. Liu, Inhibition of SARS-CoV-2 (previously 2019-nCoV) infection by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion, *Cell Res.* 30 (4) (2020) 343–355.
- [48] M. Prajapat, P. Sarma, N. Shekhar, P. Avti, S. Sinha, H. Kaur, S. Kumar, A. Bhattacharyya, H. Kumar, S. Bansal, Drug targets for corona virus: a systematic review, *Indian Journal of Pharmacology* 52 (1) (2020) 56.
- [49] B.J. Bosch, R. van der Zee, C.A. de Haan, P.J. Rottier, The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex, *J. Virol.* 77 (16) (2003) 8801–8811.
- [50] Z. Ahmad, S.S. Hassan, S. Azim, A therapeutic connection between dietary phytochemicals and ATP synthase, *Curr. Med. Chem.* 24 (35) (2017) 3894–3906.
- [51] A. Amini, M. Liu, Z. Ahmad, Understanding the link between antimicrobial properties of dietary olive phenolics and bacterial ATP synthase, *Int. J. Biol. Macromol.* 101 (2017) 153–164.
- [52] Z. Ahmad, T.F. Laughlin, I.O. Kady, Thymoquinone inhibits Escherichia coli ATP synthase and cell growth, *PLoS One* 10 (5) (2015).
- [53] D. Schoeman, B.C. Fielding, Coronavirus envelope protein: current knowledge, *Virol. J.* 16 (1) (2019) 69.
- [54] P. Venkatogopalan, S.M. Daskalova, L.A. Lopez, K.A. Dolezal, B.G. Hogue, Coronavirus envelope (E) protein remains at the site of assembly, *Virology* 478 (2015) 75–85.
- [55] M.L. DeDiego, J.L. Nieto-Torres, J.M. Jimenez-Guardeño, J.A. Regla-Nava, C. Castaño-Rodríguez, R. Fernandez-Delgado, F. Usera, L. Enjuanes, Coronavirus virulence genes with main focus on SARS-CoV envelope gene, *Virus Res.* 194 (2014) 124–137.
- [56] Y. Wang, L. Liu, The membrane protein of severe acute respiratory syndrome coronavirus functions as a novel cytosolic pathogen-associated molecular pattern to promote beta interferon induction via a toll-like-receptor-related TRAF3-independent mechanism, *MBio* 7 (1) (2016) e01872-15.
- [57] H.A. Lindner, N. Fotouhi-Ardakani, V. Lytyyn, P. Lachance, T. Sulea, R. Ménard, The papain-like protease from the severe acute respiratory syndrome coronavirus is a deubiquitinating enzyme, *J. Virol.* 79 (24) (2005) 15199–15208.

- [57] J. Shi, J. Song, The catalysis of the SARS 3C-like protease is under extensive regulation by its extra domain, *FEBS J.* 273 (5) (2006) 1035–1045.
- [58] M. Kawase, K. Shirato, L. van der Hoek, F. Taguchi, S. Matsuyama, Simultaneous treatment of human bronchial epithelial cells with serine and cysteine protease inhibitors prevents severe acute respiratory syndrome coronavirus entry, *J. Virol.* 86 (12) (2012) 6537–6545.
- [59] Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Structure of Mpro from COVID-19 virus and discovery of its inhibitors, *Nature* (2020)<https://doi.org/10.1038/s41586-020-2223-y>.
- [60] A. Shamsi, T. Mohammad, S. Anwar, M. Alajmi, A. Hussain, M. Rehman, A. Islam, M. Hassan, Glecaprevir and Maraviroc Are High-Affinity Inhibitors of SARS-CoV-2 Main Protease: Possible Implication in COVID-19 Therapy, *Biosci. Rep.* 40 (2020)<https://doi.org/10.1042/BSR20201256> BSR20201256.
- [61] Z. Jia, L. Yan, Z. Ren, L. Wu, J. Wang, J. Guo, L. Zheng, Z. Ming, L. Zhang, Z. Lou, Delicate structural coordination of the Severe Acute Respiratory Syndrome coronavirus Nsp13 upon ATP hydrolysis, *Nucleic Acids Res.* 47 (12) (2019) 6538–6550.
- [62] R.N. Kirchdoerfer, A.B. Ward, Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors, *Nat. Commun.* 10 (1) (2019) 1–9.
- [63] Y. Ma, L. Wu, N. Shaw, Y. Gao, J. Wang, Y. Sun, Z. Lou, L. Yan, R. Zhang, Z. Rao, Structural basis and functional analysis of the SARS coronavirus nsp14–nsp10 complex, *Proc. Natl. Acad. Sci.* 112 (30) (2015) 9436–9441.
- [64] K. Bhardwaj, S. Palaninathan, J.M.O. Alcantara, L.L. Yi, L. Guarino, J.C. Sacchettini, C.C. Kao, Structural and functional analyses of the severe acute respiratory syndrome coronavirus endoribonuclease Nsp15, *J. Biol. Chem.* 283 (6) (2008) 3655–3664.
- [65] I. Glowacka, S. Bertram, M.A. Müller, P. Allen, E. Soilleux, S. Pfefferle, I. Steffen, T.S. Tsegaye, Y. He, K. Gnirss, Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response, *J. Virol.* 85 (9) (2011) 4122–4134.
- [66] M. Desforges, J. Desjardins, C. Zhang, P.J. Talbot, The acetyl-esterase activity of the hemagglutinin-esterase protein of human coronavirus OC43 strongly enhances the production of infectious virus, *J. Virol.* 87 (6) (2013) 3097–3107.
- [67] M. Wang, R. Cao, L. Zhang, X. Yang, J. Liu, M. Xu, Z. Shi, Z. Hu, W. Zhong, G. Xiao, Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro, *Cell Res.* 30 (3) (2020) 269–271.
- [68] R. Kong, G. Yang, R. Xue, M. Liu, F. Wang, J. Hu, X. Guo, S. Chang, COVID-19 Docking Server: An Interactive Server for Docking Small Molecules, Peptides and Antibodies against Potential Targets of COVID-19, arXiv preprint arXiv:2003.00163, 2020.