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

Potential drug targets of SARS-CoV-2: From genomics to therapeutics

Anas Shamsi a,e,⁎, Taj Mohammad a,⁎, Saleha Anwar a, Samreen Amani b, Mohd Shahnawaz Khan c, Fohad Mabood Husain d, Md. Tabish Rehman e, Asimul Islam a, Md Imtaiyaz Hassan a

a Centre for Interdisciplinary Research in Basic Sciences, Jamia Millia Islamia, Jamia Nagar, New Delhi 110025, India
b Department of Biochemistry, F/O Life Science, Aligarh Muslim University, Aligarh, India
c Department of Biochemistry, College of Sciences, King Saud University, Riyadh 11451, Saudi Arabia
d Department of Food Science and Nutrition, Faculty of Food and Agricultural Sciences, King Saud University, Riyadh 11451, Saudi Arabia
e Department of Pharmacognosy, College of Pharmacy, King Saud University, Riyadh, Saudi Arabia

Abstract

The emergence of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) from China has become a global threat due to the continuous rise in cases of Coronavirus disease 2019 (COVID-19). The problem with COVID-19 therapeutics is due to complexity of the mechanism of the pathogenesis of this virus. In this review, an extensive analysis of genome architecture and mode of pathogenesis of SARS-CoV-2 with an emphasis on therapeutic approaches is performed. SARS-CoV-2 genome consists of a single, ~29.9 kb long RNA having significant sequence similarity to BαT-CoV, SARS-CoV and MERS-CoV genome. Two-third part of SARS-Cov-2 genome comprises of ORF (ORF1ab) resulting in the formation of 2 polyproteins, pp1a and pp1ab, later processed into 16 smaller non-structural proteins (NSPs). The four major structural proteins of SARS-CoV-2 are the spike surface glycoprotein (S), a small envelope (E), membrane (M), and nucleocapsid (N) proteins. S protein helps in receptor binding and membrane fusion and hence plays the most important role in the transmission of CoVs. Priming of S protein is done by serine 2 transmembrane protease and thus plays a key role in virus and host cell fusion. This review highlights the possible mechanism of action of SARS-CoV-2 to search for possible therapeutic options.

© 2021 Elsevier B.V. All rights reserved.

ARTICLE INFO

Article history:
Received 19 December 2020
Received in revised form 8 February 2021
Accepted 8 February 2021
Available online 9 February 2021

Keywords:
Severe acute respiratory syndrome coronavirus-2
Coronavirus disease 19
Drug targets

Contents

1. Introduction .......................................................... 2
2. Genome structure of SARS-CoV-2 .................................. 2
3. Insight into Pathogenesis of SARS-CoV-2 ................... 2
4. Potential therapeutic drug targets .................................. 2
4.1. Spike (S) protein ............................................. 2
4.2. Envelope (E) protein ....................................... 5
4.3. Membrane (M) protein ....................................... 5
4.4. Nucleoprotein (N) ........................................... 5
4.5. Replicase polyprotein ......................................... 5
4.6. Main protease (Mpro or 3CLpro) and papain-like protease (PLpro) ............................................. 5
5. RNA-dependent RNA polymerase ................................. 6
6. Non-structural protein ............................................. 6
7. Host cell proteases .................................................. 6
8. Conclusion and future directions ................................. 6
9.Declaration of competing interest. ................................. 7
Acknowledgements .................................................. 7
CRediT authorship contribution statement. ...................... 7
References ................................................................ 7

⁎ Corresponding author.
E-mail address: mshamsi1@jmi.ac.in (A. Shamsi).
1 These authors contributed equally to this work.

https://doi.org/10.1016/j.ijbiomac.2021.02.071
0141-8130/© 2021 Elsevier B.V. All rights reserved.
1. Introduction

Coronaviruses (CoVs) are exceptionally diverse, positive-sense, enveloped, and single-stranded RNA viruses from Nidovirales (order) of the Coronaviridae family [1]. These cause a wide array of diseases affecting respiratory, hepatic, neurological, and enteric systems with varied intensity amid animals and humans [2]. Based on serotype and genotypic type, CoVs are sorted into four genera viz. α, β, γ, and δ [3]. The first two CoVs i.e. α- and β- affect mammals while γ-CoVs infect avian species and δ-CoVs mammals as well as aves. In the past, six CoVs species had the potential to cause human diseases however, in immune-competent individuals, four of these six can only cause common cold [4,5]. The other two species were severe acute respiratory syndrome CoV (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV). These two were much more morbid than others and can ensure deadly disease. The newly identified CoV referred to as SARS-CoV-2 is far more contagious as compared to other earlier recognized human CoVs SARS-CoV (2002) and MERS-CoV (2013). Additionally, SARS-CoV-2 is also initiated by a zoonotic transmission as were SARS-CoV and MERS-CoV. SARS-CoV-2 is very fatal as it is transferred among humans through sneezing and making contacts [5].

SARS-CoV and MERS-CoV are very similar in sequence and encode crucial enzymes and proteins like spike (S), envelope (E), membrane (M), and nucleocapsid (N) proteins. This similarity provides a clue that both these share the same pathogenesis mechanism which can be further used in common therapeutic targeting.

On 11th February 2020, the World Health Organization (WHO) has given the title of Coronavirus disease (COVID-19) to the infection caused by this species of CoV. According to WHO, this disease is now pandemic worldwide and is a health emergency of international concern as it has taken more than two million lives around the globe. Due to the lack of understanding of the mechanism of pathogenesis, an effective therapeutic option is unavailable. Only timely diagnosis of this disease will help to control its spread by social distancing and quarantine [6,7].

At present, the entire world is dealing with this global pandemic condition and hence, there is an immediate need to develop strategies that can be implicated in therapeutics of COVID-19 [7,8]. Thus, the need of the hour is to perform an extensive analysis of its genomics and its comparison with other pathogenic human CoVs as this detailed insight will provide a platform to figure out the origin of SARS-CoV-2 pathogenesis. This will further help us in designing therapeutics strategies to combat COVID-19.

2. Genome structure of SARS-CoV-2

The genome size of the SARS-CoV-2 is ~29.9 kb [9]. It consists of approximately 13 to 15 ORFs which are flanked by 5’ and 3’ UTRs [10] (Fig. 1). These ORFs are ordered in such a way that they constitute a replicase assembly which has the potential of encoding around 27 distinct structural and non-structural proteins (NSPs) [11]. First ORF (ORF1ab) covers two-third part of the genome and translates to 7096 residues long polyprotein (PP) [9]. PP pp1a and pp1b are encoded by ORF1a and ORF1b respectively, which further produces 16 NSPs after processing through virally encoded proteases. All these 16 proteins, mainly involved in viral replication, assembly, and immune response modulation are conserved in all SARS viruses that belong to the same family [9]. The four major structural proteins of SARS-CoV-2 are spike surface glycoprotein (S) a small envelope protein (E), membrane (M), and nucleocapsid (N) proteins. There is a difference in epidemiological dynamics of SARS-CoV-2 and SARS-CoV and MERS-CoV and hence former is comparatively more infectious and lethal. 

3. Insight into Pathogenesis of SARS-CoV-2

Various organs like lungs, heart, arteries, kidney, intestine, etc. possess Angiotensin-converting enzyme 2 (ACE2) attached to their cell membrane. ACE2 is a monooxidase enzyme that plays an essential role in maintaining the balance of the renin-angiotensin system [12]. This enzyme acts as a cell receptor for SARS-CoV [13]. It is through this enzyme that SARS-CoV virion particles enter into the cell and its genome is then translated into PPVs, via the host ribosome, which further gets processed by proteolysis. PP degradation is mediated by main protease (Mpro) and papain-like protease (PLpro) which chop it into smaller fragments to support replication and helps in forming new virions [14]. The resulting PPIa and PPIab are processed into the individual NSPs that form the viral replication and transcription complex [15]. The PPVs move into endoplasmic reticulum (ER) membranes and transfer through the ER-to-Golgi intermediate compartment, where contact with N-encapsidated, newly formed genomic RNA fallouts in budding into the lumen of secretory vesicular compartments, and then virions are secreted from the infected cell by the process of exocytosis [15]. Similarly, SARS-CoV-2 also uses ACE2 as a way in the ACE2-expressing cells, thus it can be assumed that both targets a similar spectrum of cells. According to literature, macrophages and pneumocytes are the main targets of SARS-CoV [16]. The extra-pulmonary spread of SARS-CoV was also reported as ACE2 expression is not restricted to lungs only. A high transmission rate of SARS-CoV-2 in contrast to SARS-CoV can be attributed to more efficient exploitation of cellular attachment factors that mediates vigorous infections of ACE2 expressing cells in the upper respiratory tract. ACE2 occurs mainly in type II alveolar epithelial cells (pneumocytes). It is with these ACE2 molecules, the S-protein of SARS-CoV-2 interact Expression of ACE2 is slightly elevated in males in comparison to females which can be attributed to higher incidences of COVID-19 in men. CoVs have been known to modulate the affecting cells by their cytocidal activity and immune-mediated mechanisms [17]. CoVs infection results in cytopathic effects, as well as apoptosis and cell degradation.

The binding of ACE to SARS-CoV-2, as was in SARS-CoV, may lead to its increased expression, resulting in alveolar damage. Based on biophysical and structural analysis, in contrast to SARS-CoV, SARS-CoV-2 showed approximately 10–20 times higher affinity for ACE2 [18]. Finally, it can be said that SARS-CoV-2 entry in the host cell is initiated by binding to the ACE2 receptor. Apart from ACE2, there is another host cell factor that plays a key role in virion entry is the serine protease, Transmembrane protease serine 2 (TMPRSS2) [19]. TMPRSS2 plays a key role in pathogenesis by cleaving S protein at S1/S2 and the S2’ site and helps in S protein priming that allows virion entry into the host cell and ultimately viral and cellular membrane fusion. Thus, TMPRSS2 can also serve as a potential drug target. Post binding, ss RNA of SARS-CoV-2 gets attached to the host cell ribosomes which lead to the translation of two co terminals large PPVs. Fig. 2 shows a brief outline of various factors performing different responsibilities in the life cycle of SARS-CoV-2. The PPVs thus formed are further processed into 16 distinct NSPs by the action of two main proteolytic enzymes namely SARS-CoV-2 Mpro and PLpro [14]. These smaller components further play different roles such as viral assembly, immune response modulation, and others. Mpro can be considered as an important drug target among CoVs owing to its function in the processing of PPVs that result from the translation of viral RNA. CoVs Like all pathogens, the mechanism of SARS-CoV-2 pathogenesis follows innate and the adaptive immune system as well [20].

4. Potential therapeutic drug targets

4.1. Spike (S) protein

The envelope spike (S) protein plays a crucial role in CoVs infection and pathogenesis [21,22]. SARS-CoV-2 has a highly glycosylated S protein that belongs to trimeric class I viral fusion glycoprotein. S protein has 1273 amino acid residues forming three subunits namely, S1, S2, and S2’. They undergo structural changes during the process of viral and host membrane fusion (Li, 2016). S1 and S2 domains are
responsible for receptor binding and membrane fusion respectively, while S2', a cleaved subunit of S protein, acts as a fusion peptide [23]. The process of viral infection is initiated when the virus binds to the human ACE2 cell surface receptor with its S1 subunit [19].

The head region of S1 is known as receptor binding domain (RBD) which recognizes ACE2, with Glu394 of RBD and Lys31 of ACE2 playing a crucial role in receptor and S protein interaction [24]. This binding destabilizes the pre-fusion trimer that leads to the release of the S1 subunit and ultimately the transition of the S2 subunit to a more stable post-fusion conformation [25]. For attachment to host cell receptor, RBD of S1 endures hinge-like conformational changes that can conceal or reveal determinants of receptor binding transiently. These two states are known as “down” and “up” conformation with down conformation being referred to as stable and receptor unapproachable while up conformation as less stable receptor approachable state [26]. The function of the S2 subunit is to cause membrane fusion between the virions and the host cell. For this, S2 exists in three different conformations namely; pre-fusion native state, hairpin intermediate state, and post-fusion hairpin state viz. RBD is the most unpredictable feature of SARS-CoV-2 with maximum variation in the receptor-binding motif. This variation can be directly linked to variation in the mechanism of pathogenesis among different CoVs. An understanding of changing conformations of S protein that results in the entry of the virion into the mammalian cell can provide a breakthrough in COVID-19 therapeutics.
Our recent study suggested that out of six residues of RBD of S1 subunit of SARS-CoV-2 (Leu455, Phe486, Gln493, Ser494, Asn501, and Tyr505) that are crucial for binding to ACE2, five differ from SARS-CoV [27]. The similarity between S protein of SARS-CoV-2, SARS-CoV, and MERS-CoV is that in all these it exists in homologous trimeric conformation consisting of three chains named chain A, B, and C. These chains are aligned and visualized in PyMOL which revealed a high degree of structural deviations in the N-terminal domains (NTDs) and receptor-binding domains (RBDs) of the chain A and C compared to that of chain B. The structural coordinates of the S protein of the SARS-CoV-2, SARS-CoV, and MERS-CoV were taken from the RCSB Protein Data Bank with PDB IDs: 6VSB, 6ACD, and 5W9H, respectively.

Recent research reported the preclinical development of two BNT162b vaccine candidates, which contain lipid-nanoparticle (LNP) formulated nucleoside-modified mRNA encoding SARS-CoV-2 spike glycoprotein-derived immunogens [30].

---

**Fig. 3.** Structural comparison of spike (S) proteins trimeric conformation. The S protein of the SARS-CoV-2, SARS-CoV, and MERS-CoV exists in homologous trimeric conformation consisting of three chains named chain A, B, and C. These chains are aligned and visualized in PyMOL which revealed a high degree of structural deviations in the N-terminal domains (NTDs) and receptor-binding domains (RBDs) of the chain A and C compared to that of chain B. The structural coordinates of the S protein of the SARS-CoV-2, SARS-CoV, and MERS-CoV were taken from the RCSB Protein Data Bank with PDB IDs: 6VSB, 6ACD, and 5W9H, respectively.

**Fig. 4.** Structural representation of the SARS-CoV-2 main protease (3CL\textsuperscript{pro} or M\textsuperscript{pro}) (PDB ID: 6YB7).
4.2. Envelope (E) protein

The E protein plays a key part in the virion life cycle, particularly during viral morphogenesis and assembly [31,32]. It is another structural protein that can serve as a potential drug target in COVID-19 therapeutics because it can form a pentamer and function as an ion channel, thus named as E channel or viroporin. Moreover, there are regions of high similarity in E proteins of BAT-CoV, SARS-CoV, and SARS-CoV-2. Interestingly, there is a subtle disparity among SARS-CoV-2 and MERS-CoV E proteins. These proteins can form ion channels, an important function in virus-host interaction [33]. Ion conductivity is considered a beneficial feature for viruses and thus in its pathogenesis [34]. E protein ion channel activity is also required for inflammasome activation [35]. E protein also plays an important role in intracellular protein trafficking as well as regulation and hence this is another aspect of E protein that can be used in COVID-19 therapeutics.

4.3. Membrane (M) protein

Another example of transmembrane glycoprotein is M protein which has three transmembrane domains, a characteristic trait of the membrane proteins. These are the structural blocks made up of 222 amino acids that provide a framework to the virion particle and aids in the structural assembly of the virus. This structural protein functions in harmony with E, N, and S proteins and aids in RNA packaging [36]. Another role played by M protein is intracellular homeostasis. M protein also helps in viral-specific humoral response and can elicit efficient neutralizing antibodies in SARS patients [37]. Like E proteins, there is a similarity among M proteins of BAT-CoV, SARS-CoV, and SARS-CoV-2 and subtle differences in M proteins of SARS-CoV-2 and MERS-CoV.

4.4. Nucleoprotein (N)

Structurally, N proteins comprise three domains arranged as an N-terminal RNA-binding domain (NTD), followed by a Ser/Arg (SR)-rich central linker region and a C-terminal dimerization domain (CTD). These N proteins help in packing viral RNA into viral nucleocapsid. The viral genomes are released into the network of the endoplasmic reticulum–Golgi intermediate compartment (ERGIC) with the help of these N proteins. The central core of N protein has ~140 amino acid residues long RNA binding domain which is arranged in a ‘bead on string manner’ and helps in its binding with viral RNA. Targeting this RNA-binding domain of N protein can be a possible therapeutic option in COVID-19 treatment. In addition to this, N proteins also assist viral RNA transcription and replication. It also plays an important role in cellular processes like cell cycle, cytoskeleton reorganization, and host cell apoptosis. All these important functional aspects of N proteins enable them as potential drug targets that can be implicated in COVID-19 therapeutics. There is a high sequence similarity among N protein of BAT-CoV, SARS-CoV, and SARS-CoV-2 which implicates that Abs recognizing N proteins of BAT-CoV, SARS-CoV, and SARS-CoV-2 also and therefore N proteins can be used as a diagnostic tool.

4.5. Replicase polyprotein

Apart from the above discussed structural proteins, Replicase proteins also play a key part in SARS-CoV-2 pathogenesis. The major part of the replicase genome is covered by the ORF1ab gene which encodes two major PPs i.e., pp1a and pp1ab. The replicase PP is made up of mainly three domains; the macro domain, the papain-like domain, and the main protease. These multifunctional proteins guide the degradation of host RNA and replication of viral RNA [38]. These PPs are subsequently processed into 16 smaller NSPs through proteolytic enzymes, mainly the Mpro and PLpro [14], which cleaves the C and N terminal ends of these PPs respectively [39]. The function of viral RNA transcription and replication is performed by a specific RNA dependent RNA polymerase (RdRp) region present in ORF 1ab. As per the literature, similar to structural proteins, high sequence similarity is observed in 3CLpro of BAT-CoV, SARS-CoV, and SARS-CoV-2, unlike MERS-CoV which doesn’t show such similarity.

4.6. Main protease (Mpro or 3CLpro) and papain-like protease (PLpro)

Mpro is a vital cog of SARS-CoV-2 that can be an attractive drug target in COVID-19 therapeutics [14]. It is a crucial enzyme that works in cooperation with other components and helps in the replication and transcription of viral RNA Mpro is also the most studied target for CoVs drugs as it is responsible for the processing of polyproteins required for the assembly of virus drugs. Fig. 4 depicts the structure of this main protein depicting the secondary structural content. Mpro of SARS-CoV-2 is a 306 amino acid residues long protein, having three domains in its structure viz. domain I (residues 8–101), domain II (residues 102–184), and domain III (residues 201–303). Domain III and II are connected by a 15 residues long loop (residues 185–200) having conserved nidovirus component with endoribonuclease activity and an exonuclease activity, involved in the unique transcription.

| S. No | Protein | Function |
|-------|---------|----------|
| 1     | Nsp1    | The exact function is biologically unique and unknown. Forms a previously unknown complex β-barrel fold with several unique structural features and contributes to the degradation of mRNA [51]. It is also involved in innate immune response antagonism [52]. |
| 2     | Nsp2    | A replicase product, has no special known function but fund to be involved in modulation of host cell survival signaling pathway by interacting with host PHB and PHB2 [53] |
| 3     | Nsp3    | Binds to viral RNA, nucleocapsid protein, as well as other viral proteins, and contributes in polyprotein processing [54]. It has also an important role in innate immune response antagonism. |
| 4     | Nsp4    | Plays a role in membrane rearrangement in association with Nsp3 thereby affecting viral replication. |
| 5     | Nsp5    | 3C-like proteinase and main proteinase involved in viral polyprotein processing during replication [55]. |
| 6     | Nsp6    | Transmembrane domain, plays a role in the initial induction of autophagosomes from the host endoplasmic reticulum. |
| 7     | Nsp7    | An RNA-dependent RNA polymerase works in association with Nsp8 [55]. Also stimulates the polymerase activity of Nsp12. The stoichiometric ratio of nsp7 and nsp8 is found to be less than nsp12 [56]. Nsp7 and nsp8 increase nsp12 binding to the template-primer RNA [56]. |
| 8     | Nsp8    | Replicase capable of de novo initiation and has been proposed to operate as a primase in complex with nsp7. Crystalized together with the 10-kDa nsp2, forming a hexadecameric dsRNA-encoding ring structure (i.e. Nsp 7 + 8), consisting of 8 copies of both Nsp [57]. The nsp12-nsp7-nsp8 complex also showed RNA polymerization activity on a poly-U template upon addition of ATP [56]. |
| 9     | Nsp9    | Single-stranded RNA-binding protein mediate both viral replication and virulence [58]. |
| 10    | Nsp10   | Acts as a stimulatory factor along with Nsp16 to execute its MTase activity, therefore plays an essential role in viral mRNAs cap methylation [59]. It also forms complex with Nsp14 and stimulates MTase activity. |
| 11    | Nsp11   | Unknown |
| 12    | Nsp12   | RNA-dependent RNA polymerase and also has nucleotidyltransferase activity [46] |
| 13    | Nsp13   | The helicase unwinds the double-stranded RNA segment into single strands by hydrolyzing NTPs, involved in replication and transcription. |
| 14    | Nsp14   | Nsp14 has two enzymatic activities, an N7 methyltransferase activity and an exonuclease activity, involved in the unique proofreading system of CoVs [60]. |
| 15    | Nsp15   | Mn(2+)-dependent Endoribonuclease activity [61]. A highly conserved nidovirus component with endoribonuclease activity acts in conjunction with the viral replication complex to limit the exposure of viral dsRNA to host dsRNA sensors [62]. |
| 16    | Nsp16   | 2′-O-ribosyl methyltransferase involved in MTase activity [63]. |
Cys-His catalytic dyad. SARS-CoV and SARS-CoV-2 share ~96% similarity in their M\textsuperscript{pro} sequence. Just like its counterparts in other CoVs, the substrate-binding site of M\textsuperscript{pro} in SARS-CoV-2 is located in the cleft between domain I and II [40,41]. It cleaves PPs to generate NSPs that form a replicate-transcriptase complex (RTC). The ability of M\textsuperscript{pro} to control this vital step of SARS-CoV-2 highlights its importance that can be used in drug discovery. SARS-CoV-2 M\textsuperscript{pro} recognizes and acts remarkably at fewer than 11 cleavage sites of Leu-Gln ↓ (Ser, Ala, Gly) of the polyprotein replicase 1ab [42]. A recent study reported the crystal structure of SARS-CoV-2 providing a platform to develop COVID-19 therapeutics targeting M\textsuperscript{pro} [14]. Design and development of safe and potential for SARS-CoV-2 can be achieved via targeting conserved enzymes residues of M\textsuperscript{pro} [43–45].

5. RNA-dependent RNA polymerase

A multisubunit replication and transcription complex of NSPs arbitrate the replication of SARS-CoV-2. The catalytic subunit (Nsp12) of the RNA-dependent RNA polymerase (RdRp) enzyme is the core component of this complex. Nsp12 has little activity by itself and requires additional factors (Nsp7 and Nsp8) for proper functioning [46]. RdRp is treated as a potential drug spot for developing antivirals. It is a well-defined target for Remdesivir, an antiviral drug. The active site region of SARS-CoV-2 RdRp is consists of many critical residues including Lys545, Arg555, Asp623, Ser682, Thr687, Asn691, Ser759, Asp760, and Asp761. These residues where Remdesivir is bound can be served as a platform for the development of powerful, effective, and selective inhibitors of SARS-CoV-2 RdRp [47].

6. Non-structural protein

These refer to proteins that are not included in viral particles but function inside the infected cell playing several important roles viz.

viral replication, controlling early transcription and modulates immune response [9]. NSPs are also found to be associated with helicase activity. Table 1 lists all the NSPs along with their proposed functions.

7. Host cell proteases

SARS-CoV-2 enters into the host cell with the help of its S protein whose S1 subunit binds to the receptor on the cell surface. In the cell line, other key players involved in this entry were found to be endosomal cysteine proteases cathepsin B and L (CatB/L) [48] and the serine protease TMPRSS2 [49] which are involved in the priming of S protein. The important fact is that only the activity of TMPRSS2 is crucial for viral spread and pathogenesis unlike the dispensable activity of CatB/L activity [50] (Fig. 2). Another study suggested that TMPRSS2 is a critical host cell SARS-CoV-2 also playing a role in its spread and pathogenesis [50]. A TMPRSS2 inhibitor approved for clinical use, camostat mesylate, blocked the entry and can provide a new avenue in COVID-19 therapeutics.

8. Conclusion and future directions

At present, SARS-CoV-2 is offering a major risk across the globe and there is no specific drug available to treat COVID-19. In the past, antiviral drugs have been used to treat other human CoVs, but these have been rendered ineffective due to structural differences in SARS-CoV-2 as compared to other human CoVs. SARS-CoV-2 is far more pathogenic in contrast to other human CoVs. The SARS-CoV-2 vaccine is still under development, and there is no specific drug at present and all other trial drugs have also been unsuccessful. A comparative genomics-based approach with earlier known human CoVs can provide a breakthrough in COVID-19 therapeutics. Fig. 5 shows the SARS-CoV-2 life cycle with the depiction of different target sites that can be retorted in COVID-19 therapeutics. Thus,
the need of the hour is to perform an extensive analysis of its genomics and its comparison with other pathogenic human CoVs as this detailed insight will provide a platform to understand the molecular basis of SARS-CoV-2 pathogenesis. This review article provides an extensive investigation of the genome of SARS-COV-2 and its comparison with other human CoVs that enable us to identify the molecular way of pathogenesis. It also provides a brief insight into the differences in the structure of these important proteins of SARS-CoV-2 that can act as possible drug targets and enlightened the differences in the structure of these proteins in comparison to other human CoVs. Table 2 lists various drugs against different targets of SARS-CoV-2 having clinical effectiveness and can be implicated in COVID 19 therapeutics.

### Table 2

| S. No. | Drug                     | Use          | Target                          | Mechanism of action                                                                 | Clinical trial                          | Ref.   |
|--------|--------------------------|--------------|---------------------------------|--------------------------------------------------------------------------------------|----------------------------------------|--------|
| 1      | Atazanavir               | HIV          | SARS-CoV-2 Mpro                 | Atazanavir could adjust in Mpro active site and can inhibit its activity resulting in a disruption in viral replication. | Phase 2: NCT04459286                   | [64]   |
| 2      | Baricitinib              | Rheumatoid arthritis | Human AP2-associated protein kinase 1 (AAK1); Janus kinase [JAK] 1 and 2 | Can block the entry and infectivity of SARS CoV-2 in pneumocytes by impairing AAK1 that are involved in virus endocytosis; also inhibit the intracellular signaling pathway of cytokines IL-2, 6, 10 and INF-γ, a granulocyte-macrophage colony-stimulating factor that is enhanced in severe SARS CoV-2 infection | Phase 2: NCT04321993; Phases 2 and 3: NCT04320277; Phase 3 NCT04421027 | [65–77]|
| 3      | Meflopirid hydrochloride | Cancer       | N protein; poly-ADP-ribose polymerase 1 (PARP1) | Can target N protein to reduce its RNA binding and thus impede viral replication; inhibit the production of IL-6 by Cpg oligodeoxynucleotide 1826 in peripheral blood mononuclear cells | Phase 1: NCT04292899 | [68]   |
| 4      | Pemirolast, nitrofurantoin| Numerous     | ACE2 receptor                   | Can interact with ACE2 receptor more efficiently and inhibit undesirable S protein to ACE2 interaction. | –                                     | [69]   |
| 5      | Remdesivir               | Ebola        | RdRP                            | Nucleoside (adenosine) analogue RdRP inhibitor which inhibits RNA synthesis and can result in premature termination | Phase 3: NCT04353336                   | [70,71]|
| 6      | Chloroquine/hydroxychloroquine | Malaria, lupus | Affect both early and late stage of viral replication 3CLpro | keep the virus out of host cells by disturbing ACE2 glycosylation and breaking down the production of viral proteins by inhibiting endosomal acidification. | Phase 2 and 3: NCT04535336 | [70,72–74]|
| 7      | Lopinavir/ritonavir combination | HIV | 3CLpro                           | Disrupt the process of viral replication and release from the cell. | Phase 2: NCT0427668 | [77,78]|
| 8      | Nafamostat or camostat   | Pancreatitis | Serine protease TMPRSS2          | Acts as an antagonist to the serine protease TMPRSS2; Prevents membrane fusion by reducing the release of cation B. | Phase 2 and 3: NCT04418128 | [79]   |
| 9      | Famotidine               | Heartburn    | PL$$\alpha$$                    | Possibly bind PL$$\alpha$$ which is known to be essential to the entry of SARS-CoV-2 | Phase 2: NCT04625114                   | [80]   |
| 10     | Umifenovir               | Influenza    | Viral lipid membrane            | Can bind viral lipid membrane and affect cellular trafficking of the virus | Phase 4: NCT04504240                  | [81]   |
| 11     | Nitazoxanide             | Influenza; diarrhea | Not known                       | Can suppress maturation of the viral hemagglutinin and the viral transcription factor immediate-early 2 (IE2) as well as by activating the translation INF20s. | Phase 2: NCT04552483                  | [82]   |
| 12     | Ivermectin               | Influenza; dengue; broad-spectrum antiparasitic | Not known                       | Can inhibit expression of the viral N protein and IL-6; Inhibit viral IMPα/1-mediated nuclear import, causing a reduction in viral replication; Can also work by binding and destabilizing cell-transport proteins used to enter the nucleus. | Phase 1: NCT04343092                  | [83]   |
| 13     | Teicoplanin              | Gram-positive bacterial infection | Not known                       | Inhibit the activity of cathepsin L which potentially plays an important role in blocking viral entry in the cells | –                                     | [84]   |
| 14     | Tocilizumab/sarilumab (mAb) | Rheumatoid arthritis | IL-6 receptor antagonists       | Inhibition of IL-6 may attenuate pulmonary inflammation and fibrosis | Phase 3 and Phase 2: NCT04315298     | [84]   |
| 15     | Anti TNF-α agents        | Rheumatoid arthritis | TNF-α                            | TNF-α blockage leads to down-regulation of pro-inflammatory mediators, including IL-1, IL-6, and granulocyte-macrophage colony-stimulating factor as well as cytokines and acute-phase proteins | –                                     | [84]   |

### Declaration of competing interest

None.

### Acknowledgements

A.S. is thankful to Dr. D.S. Kothari Postdoctoral Fellowship ((BSR)/BL/19-20/0119). M.S.K., F.M.H., and M.T.R. acknowledge the generous support from the Deanship of Scientific Research at King Saud University, Riyadh, Kingdom of Saudi Arabia (Grant No. RGP-215).

### CRediT authorship contribution statement

**Anas Shamsi**: Conceptualization; Data curation; Formal analysis Validation; Visualization; Roles/Review – original draft; Writing – review & editing. **Taj Mohammad**: Resources; Software; Validation; Visualization; Roles/Review – original draft; Writing – review & editing.
A. Shamsi, T. Mohammad, S. Anwar et al. International Journal of Biological Macromolecules 177 (2021) 1–9

**References**

[1] A. Zulma, J.F. Chan, E.J. Azhar, D.S. Hui, K.Y. Yuen, Coronavirus—drug discovery and therapeutic options, Nat. Rev. Drug Discov. 15 (5) (2016) 327.

[2] J.F.-W. Chan, S.K.-P. Lau, P.C.-Y. Woo, The emerging novel Middle East respiratory syndrome coronavirus: the “knowns” and “unknowns”, J. Formos. Med. Assoc. 112 (7) (2013) 372–381.

[3] A.C. Wong, X. Li, S.K. Lau, P.C. Woo, Global epidemiology of bat coronaviruses, Viruses 11 (2) (2019) 174.

[4] S. Su, G. Wang, W. Shi, J. Liu, A.C. Lai, J. Zhou, W. Liu, Y. Bi, G.F. Gao, Epidemiology, genetic recombination, and pathogenesis of coronaviruses, Trends Microbiol. 24 (6) (2016) 490–502.

[5] P. Arsanjani, G.M. Hasan, S.S. Sohal, M.I. Hassan, Molecular basis of pathogenesis of coronaviruses: a comparative genomics approach to planetary health to prevent zoonotic outbreak in the 21st century, OMICS: A Journal of Integrative Biology 17 (2013) 634–644.

[6] P. Arsanjani, M.S. Eapen, C. Chia, G. Haug, H.C. Weber, I. Hassan, S.S. Sohal, Diagnostic Approaches in COVID-19: Clinical Updates, Expert Review of Respiratory Medicine, 2020 (null-null).

[7] P. Kumari, A. Singh, M.R. Ngasainao, I. Shakeel, S. Kumar, S. Lal, A. Singhal, S.S. Sohal, I.K. Singh, M.I. Hassan, Potential diagnostics and therapeutic approaches in COVID-19, Clin. Chim. Acta 510 (2020) 488–497.

[8] U. Fatima, S.S.A. Rizvi, S. Fatima, M.I. Hassan, Impact of hydroxychloroquine/chloroquine in COVID-19 therapy: two sides of the coin, Int. J. Interferon. 10 (2020) 469–471.

[9] 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 (2020) 265–269.

[10] R. Li, X. Zhao, J. Li, P. Liu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu, Genomic characterization and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding, Lancet (2020) 565–574.

[11] Z. Liu, X. Xiao, X. Wei, J. Li, Y. Yang, H. Tan, J. Zhu, Q. Jiang, Z. Wu, L. Liu, Composition and divergence of coronavirus spike proteins and host ACE2 receptors predict potential intermediate hosts of SARS-CoV-2, Med. Virol. 92 (6) (2020) 595–601.

[12] M. Tabish, Alternative splicing of ACE2 possibly generates variants that may limit the entry of SARS-CoV-2: a potential therapeutic approach using SSOs, Sci. Clin. Sci. 134 (10) (2020) 1143–1150.

[13] G. Oudit, Z. Kassiri, C. Jiang, P. Liao, S. Poutanen, J. Penninger, J. Butany, SARS-coronavirus modulation of myocardial ACE2 expression and inflammation in patients with SARS, Eur. J. Clin. Invest. 39 (7) (2009) 618–625.

[14] L. Zhang, D. Lin, X. Sun, U. Curth, C. Drost, L. Sauerhöfer, S. Becker, K. Röss, R. Hilgenfeld, Crystal structure of SARS-CoV-2 main protease provides a basis for designing improved inhibitors, Science 368 (6489) (2020) 402–412.

[15] P. Vökkers, A. Kratzel, S. Steiner, H. Stalder, V. Thiel, Coronavirus biology and replication: implications for SARS-CoV-2, Nat. Rev. Microbiol. (2020) 1–16.

[16] W.-J. Shihe, C.-H. Hsiao, C.D. Paddock, J. Guaner, C.S. Goldsmith, K. Tatts, M. Packard, L. Martin, M.-Z. Wu, P. Rollin, Immunohistochemical, in situ hybridization, and ultrastructural localization of SARS-associated coronavirus in lung of a fatal case of severe acute respiratory syndrome in Taiwan, Hum. Pathol. 36 (3) (2005) 303–309.

[17] T. Suzuki, Y. Otake, S. Uchimoto, A. Hasebe, Y. Goto, Genomic characterization and phylogenetic classification of bovine coronaviruses through whole genome sequence analysis, Viruses 12 (2020) 183.

[18] D. Wrapp, N. Wang, K.S. Corbett, J.A. Goldsmith, C.-L. Hsieh, O. Abiona, B.S. Graham, D.W. Stull, B.A. Williamson, D. McGeoch, C.V. Verheugen, W. Hsu, C. Cheng, Crystal structure of SARS-CoV-2 main protease provides a basis for designing improved drugs for COVID-19: crystal structure of SARS-CoV-2 main protease M pro, Science Translational Medicine 5 (2020) 1–2.

[19] R. Raj, J.G. Park, K.-H. Cho, P. Choi, T. Kim, J. Ham, J. Lee, Assessment of antiviral potencies of a series of 2-benzylaminomethyl-1,4-benzodioxin compounds and in vitro and computational approaches, Int. J. Biol. Macromol. 168 (2021) 474–485.

[20] A. Shamsi, T. Mohammad, S. Anwar, M.F. AlAjmi, A. Hussain, M. Rehman, A. Islam, M. Hassan, Cleavageprir and Maraviroc are high-affinity inhibitors of SARS-CoV-2 main protease: possible implication in COVID-19 therapy, Biocis. Rep. 40 (6) (2020), e1000256.

[21] T. Mohammad, A. Shamsi, S. Anwar, M.U. Miah, M. Rehman, M.F. AlAjmi, A. Islam, M.I. Hassan, Identification of high-affinity inhibitors of SARS-CoV-2 main protease: towards the development of effective COVID-19 therapy, Res. Viral. 8 (2020) 198102.

[22] A.T. Ton, F. Gentile, M. Hsing, F. Ban, A. Cherkasov, Rapid identification of potential inhibitors of SARS-CoV-2 main protease by deep docking of 1.3 billion compounds, Molecular Informatics 39 (2020), e2000208.

[23] R.N. Kirchdoerfer, A.B. Ward, Structure of the SARS-CoV npol2 polymerase bound to npsp7 and npsp8 co-factors, Nat. Struct. Mol. Biol. 10 (1) (2019) 1–9.

[24] W. Yin, C. Mao, X. Lu, D.-D. Shen, Q. Shen, H. Su, X. Wang, F. Zhou, W. Zhao, M. Gao, Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir, Science 369 (2020) 1499–1504.

[25] G. Simmons, D.N. Cossia, A.J. Rennekamp, J.D. Reeves, S.L. Diamond, P. Bates, Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry, Proc. Natl. Acad. Sci. 102 (33) (2005) 11876–11881.

[26] I. Glowacka, S. Bertram, M.A. Müller, P. Allen, E. Souleix, S. Pfefferle, I. Steffen, T.S. Tsegaye, Y. He, K. Gniirs, Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral replication by the humoral immune response, J. Virol. 85 (2021) 4112–4124.

[27] N. Iwata-Yoshikawa, T. Okumura, Y. Shimizu, H. Hasegawa, M. Takeda, N. Nagata, TMPRSS2 contributes to virus spread and immunopathology in the airways of mice models after coronavirus infection, J. Virol. 93 (6) (2019) e01815–e01818.
