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.
Molecular mechanisms of the novel coronavirus SARS-CoV-2 and potential anti-COVID19 pharmacological targets since the outbreak of the pandemic

Dimitrios Vlachakis a,b,c,d, Eleni Papakonstantinou a, Thanasis Mitsis a, Katerina Pierouli a, Io Diakou a, George Chrousos b,c, Flora Bacopoulou b,*

a Laboratory of Genetics, Department of Biotechnology, School of Applied Biology and Biotechnology, Agricultural University of Athens, 75 Iera Odos, Athens, 11855, Greece
b University Research Institute of Maternal and Child Health & Precision Medicine, and UNESCO Chair on Adolescent Health Care, National and Kapodistrian University of Athens, Aghia Sophia Children’s Hospital, 8 Levadias Street, Athens, 11527, Greece
c Lab of Molecular Endocrinology, Center of Clinical, Experimental Surgery and Translational Research, Biomedical Research Foundation of the Academy of Athens, 4 Soranus Ephesius Street, Athens, 11527, Greece
d Department of Informatics, Faculty of Natural and Mathematical Sciences, King’s College London, Strand, London WC2R 2LS, UK

ARTICLE INFO
Keywords:
Coronavirus
SARS-CoV-2
COVID-19
Pandemic
Spike
Protein

ABSTRACT
The novel coronavirus SARS-CoV-2 has emerged as a severe threat against public health and global economies. COVID-19, the disease caused by this virus, is highly contagious and has led to an ongoing pandemic. SARS-CoV-2 affects mainly the respiratory system, with most severe cases primarily showcasing acute respiratory distress syndrome. Currently, no targeted therapy exists, and since the number of infections and death toll keeps rising, it has become a necessity to study possible therapeutic targets. Antiviral drugs can target various stages of the viral infection, and in the case of SARS-CoV-2, both structural and non-structural proteins have been proposed as potential drug targets. This review focuses on the most researched SARS-CoV-2 proteins, their structure, function, and possible therapeutic approaches.

1. Introduction
Coronaviruses (CoVs) are enveloped, positive-sense, single-stranded ribonucleic acid (RNA) viruses that belong to the family Coronaviridae. These viruses harbor the largest genome among RNA viruses, specifically 26–32 kilobases (Ye et al., 2020). Based on phylogenetic analyses, CoVs can be classified into four genera, which include alpha coronaviruses, beta coronaviruses, gamma coronaviruses, and delta coronaviruses (Jaiswal and Saxena, 2020). Out of those genera, beta-CoVs contain the majority of viruses that infect humans, with the aforementioned genera being subdivided into four lineages (A, B, C, D) (Ye et al., 2020). Generally, CoVs can infect the respiratory, hepatic, gastrointestinal, and central nervous system of mammals, birds, and fish (Y Chen et al., 2020). There are seven CoVs known to cause human disease. Four of those CoVs (HCoV229E, NL63, OC43, and HKU1), known as non-severe acute respiratory syndrome (SARS)-like CoVs, induce mild diseases and are globally endemic, while three highly pathogenic CoVs (SARS-CoV-1, MERS, and SARS-CoV-2) can lead to lethal disease (Raoult et al., 2020). Out of the latter three CoVs, SARS-CoV-2, responsible for the Coronavirus Disease 2019 (COVID-19), has proved a serious threat against public health and global economies in 2020 (Zheng, 2020).

SARS-CoV-2 is a beta-coronavirus of the B lineage with a diameter of approximately 60–140 nm (Cascella et al., 2020). Its genome appears to encode as many as 14 open reading frames (Orfs), with the 5′Orf1a/Orf1ab encoding polyproteins, which are auto-protolytically processed into 16 non-structural proteins (NSPs) (Gordon et al., 2020). Those NSPs (Nsp1-16) form the replicase/transcriptase complex (RTC). This complex is composed of various enzymes, including a papain-like protease (Nsp3), the main protease (Nsp5), an Nsp7-Nsp8 primase complex, a primary RNA-dependent RNA polymerase (RdRp/Nsp12), a helicase/triphosphatase (Nsp13), a 3′-5′ exoribonuclease (Nsp14), an uridine-specific endonuclease (Nsp15), and N7- and O-methyltransferases (Nsp10/Nsp16). The remaining Orfs encode four structural proteins and nine putative accessory factors (Gordon et al., 2020). These structural proteins include the spike (S) glycoprotein, the small envelope (E) glycoprotein, the membrane (M) glycoprotein, and a nucleocapsid (N) protein (Astu et al., 2020)

Coronaviruses entry into host target cells is dependent on the binding of S glycoprotein to a specific cell receptor and the subsequent priming
of the aforementioned protein by host cell proteases. SARS-CoV-2 uses the angiotensin-converting enzyme 2 (ACE2) host receptor for internalization and, mainly, the type II transmembrane serine protease (TMPRSS2) for Spike glycoprotein priming (Kumar et al., 2020). Specifically, the S glycoprotein binding domain, which is present at 331 to 524 residues, attaches strongly to the ACE2 receptor (Astuti and Ysrafil, 2020). After attachment, TMPRSS2 and a number of lysosomal proteases cleave and activate the S glycoprotein leading to the entrance of SARS-CoV-2 into the cell through endocytosis or direct fusion of the viral membrane with the host membrane (Romano et al., 2020; Shang et al., 2020). Once inside the host cytoplasm, the viral RNA undergoes translation, where the Orfs that encode Nsps are firstly translated and processed by the N-terminal domain protease (3CLpro) (Astuti and Ysrafil, 2020). These proteases are the papain-like proteases (PLpro) and the zinc-binding helicase (Nsp13), and enzymes, such as a bifunctional NTPase/Helicase (Nsp14), which are related to viral RNA modification, such as messenger RNA (mRNA) capping and RNA proofreading (Viswanathan et al., 2020). The accessory and structural proteins are translated by a specific set of subgenomic RNAs (Romano et al., 2020). Some of those, including M, S, E, are insulated in the endoplasmatic reticulum and then transported to the Endoplasmatic Reticulum-Golgi Intermediate Compartment (ERGIC) (Astuti and Ysrafil, 2020). The replicated genome has the ability to join the N protein to the nucleocapsid form and move into the ERGIC (Astuti and Ysrafil, 2020). The virions are later assembled at ERGIC and are subsequently released through exocytosis out of the cell (Kumar et al., 2020).

Patients infected with SARS-CoV-2 were first reported in December 2019, and, since then, cases have risen to the point that as of 30 August 2020, this virus has led to 24,854,140 cases and 838,924 deaths worldwide (WHO, 2020a,b; Zheng, 2020). The dominant clinical features of COVID-19 include cough, shortness of breath or difficulty breathing, fever or chills, muscle or body aches, vomiting or diarrhea, and new loss of taste or smell (CDC, 2020). Person to person spread of SARS-CoV-2 via respiratory droplets, when a patient coughs, sneezes, or talks is the main way of COVID-19 transmission, while the disease can also occur through indirect means such as touching objects contaminated with SARS-CoV-2 (Lofti et al., 2020). This virus primarily affects the respiratory system, with respiratory symptoms of COVID-19 being quite heterogeneous, varying from minimal symptoms to significant hypoxia with acute respiratory distress syndrome (ARDS) (Yuki et al., 2020). Moreover, severe COVID-19 patients display an acute inflammatory response. Specifically, transcriptomic RNA-seq analysis of COVID-19 patients showed that the viral infection induced several immune pathways and pro-inflammatory cytokines, implying sustained inflammation and cytokine storm (Wu et al., 2020b). COVID19 can be fatal in the presence of risk factors, including the co-existence of chronic diseases, older age, male sex, obesity, and smoking (Michelozzi et al., 2020; Petrakis et al., 2020). Obesity, particularly, is considered a chronic inflammatory state with obese individuals showcasing high concentrations of pro-inflammatory cytokines like interleukin 6 (IL-6). With regard to COVID-19, IL-6 has been shown as an independent predictor of mortality (Yadav et al., 2020). The currently available antiviral option for hospitalized patients is remdesivir, which may inhibit the replication process by targeting the RdRp. Previously proposed treatments for hospitalized patients included hydroxychloroquine, which thought to disrupt virus endocytosis, and lopinavir/ritonavir, which thought to inhibit SARS-CoV-2 main protease (Astuti and Ysrafil, 2020; Magro, 2020). To date no targeted therapy exists, therefore, the development of such therapeutic molecules is a priority for the scientific community.

2. Structural proteins as potential drug targets

2.1. Spike (S) protein

Spike is a transmembrane glycoprotein in CoVs that forms homotrimers that protrude from the viral surface (Walls et al., 2020). Spike is the structural protein that gives CoV viral particles their characteristic crown-like shape, from which their original name was coined (Coutard et al., 2020). Spike is essential in viral entry, a process which is achieved through binding with the ACE2 host receptor (Liu et al., 2020). The
above has led to considering S a promising drug target, therefore the study of this glycoprotein is important in the development of novel therapeutics against COVID-19.

Structure-wise, each monomer of the S glycoprotein contains two subunits termed S1 and S2, which mediate attachment and membrane fusion, respectively (Ou et al., 2020). The S1 is divided into the N-terminal domain (NTD), followed by the receptor-binding domain (RBD) and two structurally conserved subdomains termed SD1 and SD2. The RBD constantly switches between a standing-up position, associated with receptor binding, and a lying down position associated with immune evasion (Shang et al., 2020). The SD1 and SD2 subdomains cap the S2 subunit and, thus, protect the conserved fusion machinery (Henderson et al., 2020). The S2 subunit contains the fusion peptide (FP), the heptad repeat 1 (HR1), the heptad repeat 2 (HR2), the transmembrane domain (TM), and the cytoplasmic domain (CP) (Xia et al., 2020). Viral entry requires both binding of S to the cellular receptor and S glycoprotein priming by cellular proteases (Hoffmann et al., 2020). This priming includes a cleavage at the boundary between the S1 and S2, for the majority of coronaviruses, plus a cleavage at the S2’ site located immediately upstream of the FP, a process present in all coronaviruses (Hoffmann et al., 2020; McCallum et al., 2020; Walls et al., 2020). The latter cleavage has been implicated in activating the glycoprotein for membrane fusion through extensive irreversible conformation changes (McCallum et al., 2020). Specifically, after cleavage, the exposed fusion peptide inserts itself into the plasma membrane while the HR1 and HR2 can interact to form the six helical bundle (6HB) fusion core, which eventually leads to the fusion of the viral membrane with the host plasma membrane (Joshi et al., 2020). The above priming is accomplished through proteases such as TMPRSS2 and lysosomal cathepsins (Jaimes et al., 2020). SARS-CoV-2 S differentiates from CoVs of the same clade since it features a novel furin-like cleavage upstream of the cleavage site 1 which seems to be cleaved during biosynthesis. This characteristic potentially allows for more efficient spreading of SARS-CoV-2 in the human population compared to other lineage b beta coronaviruses (Walls et al., 2020; Coutard et al., 2020).

While using the spike glycoprotein as a potential drug target, a researcher can focus on viral binding to host receptor or membrane fusion. As prime example, monoclonal antibodies can be used to tightly bind the spike glycoprotein’s RBD in order to neutralize SARS-CoV-2. Such a case is the monoclonal antibody CR3022, which binds to S RBD while not competing with the binding of ACE2, allowing the monoclonal antibody to interfere with ACE2 binding (Huo et al., 2020). Another approach, this time focusing on membrane fusion, is EK1C4. EK1C4 is a lipopeptide which disrupts membrane fusion by targeting HR1 (Xia et al., 2020). On the other hand, interrupting S function could also be achieved by targeting host cell proteins with the use of drugs such as ACE2 and TMPRSS2 inhibitors (McKee et al., 2020). Unfortunately, ACE inhibitors, which are used for the treatment of hypertension and chronic heart failure, do not inhibit ACE2. Moreover, scientists are concerned that such drugs could increase the likelihood of severe COVID-19 illness, though current data tends to provide no link between the use of ACE inhibitors and increased severity of COVID-19 (Hippisley-Cox et al., 2020; WHO, 2020a). All the above can be considered the first steps towards COVID-19 therapeutics and indicate that the Spike glycoprotein may indeed be the most promising drug target against SARS-CoV-2 infection.

2.2. Nucleocapsid (N) protein

Coronaviruses’ nucleocapsid protein, or N protein, is the most abundant viral protein that can be detected in the human host from the early stages of infection (Che et al., 2004). Its main function is to bind to viral RNA to form a ribonucleoprotein nucleus, which contributes to its entry into the host cell as well as to the interaction with cellular processes after fusion, such as cell cycle regulation, and antigen processing and presentation of the virus (Huang et al., 2004; Wang et al., 2010).

The sequence of N protein of SARS-CoV-2 is approximately 90% similar to the same protein of SARS-CoV (Gralinski and Menachery, 2020). The N protein genome comprises a serine-rich binding region (SR) between an N terminal domain (NTD) and a C terminal domain (CTD). N structural protein forms the replication transcription complex (RTC), which play an important role in the synthesis of the genome virus (Xia et al., 2020).

In SARS-CoV, the N protein contributes to the activation of cyclooxygenase-2 (COX-2), thus promoting inflammation in the lungs (Yan et al., 2006). In addition, it participates in processes that affect the cell cycle, such as inhibiting the phosphorylation of phosphoprotein B23, a protein that is essential for cell cycle evolution during centrosome replication (Zeng et al., 2008). Moreover, studies have shown the inhibitory effect of N protein on the cyclin-cyclin-dependent kinase complex (C-CDK), resulting in a reduction of S-phase progression (Surjit et al., 2006). A different function of the N protein is its interaction with the protease subunit p42, which degrades viral proteins (Wang et al., 2010). Specifically, this interaction may impair proteolysis of viral proteins and their presentation to Cytotoxic T-Lymphocytes, thus promoting viral evasion from immune effectors (Wang et al., 2018). Concerning the immune system, N protein causes limitations in the immune responses generated by the body due to viral infections as it inhibits type I interferon (IFN) (Lu et al., 2011). According to another study, due to the aggregation of the Human elongation factor 1 a (EF1a) induced by the N protein, the proliferation of the cells of the cell line under study, specifically 293 T cells, was reduced (Zhou et al., 2008). In contrast, N protein interacts with heterogeneous nuclear ribonucleoprotein (hnRNPA1), causing an increase in viral RNA synthesis (Luo et al., 2005).

The structural study of the NTD of SARS-CoV-2 N protein showed that it resembles a wrist consisting of acidic moieties with a palm of basic components, and the core of the beta-sheet extends like fingers (Kang et al., 2020). The NTD of the N protein binds to the RNA genome in the N45-181 region that exists as a monomer (Chang et al., 2006), while the presence of the amino acids arginine at position 94 and tyrosine at position 122 appears to be necessary for viral RNA binding (McBride et al., 2014). According to further studies, the enhanced binding capacity of viral RNA requires the combination of the binding region and NTD and CTD of N protein (Chang et al., 2009). N protein function regulation is mediated through the central linker region, which includes serine and arginine residues (SR region), having essential phosphorylation sites. In this region, the major phosphorylation site is formed, which enhances protein interactions in proteins and the localization of binding proteins within the cell (Chang et al., 2014). The CTD is hydrophobic and rich in helical region in which dimerization occurs as it contains residues that self-bond to form homodimers (McBride et al., 2014). Structurally, each asymmetric CTD group consists of four individual homodimers from the compound of which an octamer is formed. The arrangement of these sections is schematically similar to the letter ‘X’ which forms symmetrical folding structures, perpendicular to the middle point of the structure. The N terminus appears to be basic due to the positive charges in this region, which could be implied to be a site for nucleic acid binding (Chen et al., 2007).

N protein, is a very promising area in the development of effective therapeutics to prevent the proliferation of viral offspring as it participates in activities necessary for the function and proliferation of the virus and is, therefore, another key ingredient after Spike protein (Satarker and Nampoothiri, 2020). Currently, the only potential drug targeting the N protein is pj34 which inhibits the viral protein’s function and has shown some promise both in vitro and in animal studies (Saxena, 2020).

2.3. Envelope (E) protein

The coronaviruses’ envelope protein, or E protein, is a vital protein, which is tiny and consists of the N-terminal domain, a hydrophobic
domain, and the C terminal domain (Ruch and Machamer, 2011). The N terminal extends to the first nine amino acids, the hydrophobic region extends between 10 and 37 amino acids and the C terminal between 38 and 76 amino acids, where the first 11 amino acids are located in virion while the hydrophobic tail is in the cytoplasm, which through its oligomerization creates an ionic pore across the membrane (Shen et al., 2003; Verdia-Baguena et al., 2013). Through structural studies, SARS-CoV-2 form is presented, which includes 35 α-helical regions and 40 looped regions (Gupta et al., 2020).

A unique characteristic that occurs in SARS-CoV-2 E protein and not in other homologous Sars-CoV E protein is the replacement of the 69th amino acid of the protein sequence, from arginine to alanine, glutamine, and aspartate. Furthermore, a deletion specific to Sars-CoV-2 envelope protein flanks this position. These changes may have an effect on viral conformation, protein-protein interaction and, possibly, affect the oligomerization process necessary to form a transmembrane ion channel. In addition, at positions 55 and 56 of the amino acid sequence, amino acid valine and threonine have been identified (Bianchi et al., 2020).

In coronaviruses, the E protein forms viroporins which are small hydrophobic proteins that are necessary for viral assembly and release, while also participating in pathogenic processes that cause cytotoxicity (Ye and Hogue, 2007). The co-expression of M and E proteins facilitates the production of spherical infectious particles, while the heterotypic interactions of nsp2 and nsp3 cause the desired curvature in the Endoplasmic Reticulum (ER) membrane (Schoeman and Fielding, 2019).

The hydrophobic tail of the protein in the cytoplasm through the proline residues incorporated in it targets the region of the cis-Golgi complex, as well as the N-terminal of protein E targets the Golgi complex through additional elements (Cohen et al., 2011). The diffusion of the ionic gradient from the E protein into the ERGIC and Golgi may lead to the exit of the virion (Liu et al., 2007). The last four amino acids of the sequence that comprise the C terminal domain include a pattern called posttranscriptional density protein/Disc Large/Zonula occludans-1 (PDZ) binding pattern (PDM). Therefore, E protein through binding of Protein Associated with Caenorhabditis elegans Lin-7 protein 1 (PALS1) to the PDM, aids in the disruption of the lumen epithelium, thus being a potential pharmacological target for improved treatment against SARS-CoV-2 (Teoh et al., 2010; Satarker and Nampoothiri, 2020). Regarding therapeutics, in silico studies have proposed a number of drugs that target the E protein, which include Belachinal, Macaflavanone E, and Vibsanol (Gupta et al., 2020).

2.4. Membrane (M) protein

Membrane protein, or M protein, is present in greater amounts than all other proteins in coronaviruses (EA and Jones, 2019). In coronaviruses, this protein has a short length N terminal domain, as its total length is 220–260 amino acids, belongs to the N-linked glycosylated proteins with a conserved region of 12 amino acids and is attached to tripartite transmembrane regions which are further linked to the C terminal domain (Arndt et al., 2010). According to structural studies, it exists in two forms, one compact and one long. Initially, these two forms are homodimers of the N-terminal ectodomain and the C-terminal endodomain, where the endodomain can be elongated or compressed, resulting in a long and compact form. Tyrosine residues at position 211 are necessary for the stability of the long form of the M protein, as well as the spike protein is observed mainly in the long form of the M protein, suggesting that this form promotes the establishment of the spike proteins. The long form of the M protein bends the membrane, thus creating a spherical structure surrounding the ribonucleoprotein (Neuman et al., 2011).

M protein is organized in a two-dimensional lattice and provides a scaffold in viral assembly, while its translation takes place in the poly-somes connected to the membranes following its fusion to the endoplasmic reticulum and transport to the Golgi complex, where it interacts with the E proteins to create virions (Neuman et al., 2006; Tseng et al., 2010).

M protein appears to affect the immune response to pathogens, as it inhibits Nuclear Factor Kappa B (NFκB) through interactions with IκKB (I Kappa B Kinase) leading to a decrease in COX-2 levels and resulting in an increase in the proliferation of the pathogenic virus (Fang et al., 2007). In addition, the C terminus of the M protein inhibits the interaction of 3-phosphoinositide-dependent protein kinase 1 (PDK1) and protein kinase B (PKB), thus leading to the release of the caspases 8 and 9, which eventually cause cell death or apoptosis (Tsoi et al., 2014).

Based on the above information regarding the M protein of coronaviruses, SARS-CoV-2 protein can be a potential pharmacological target for limiting and inhibiting the formation of virions and preventing inflammatory reactions in host cells, as the presence of the M protein is crucial in the viral life cycle (Satarker and Nampoothiri, 2020). Currently, broad spectrum antiviral drugs that target the Membrane protein of coronaviruses, such as JL103, have been proposed for possible use against SARS-CoV-2 (Khan et al., 2020).

3. Non-structural proteins as potential drug targets

3.1. Proteases

Viruses in the Coronaviridae family code for two types of cysteine proteases, including the 3C-like protease (3CLpro) and up to two papain-like proteases. With these enzymes’ use, the polyprotein generated through the translation of the viral genome of the virus is cleaved into sixteen nonstructural proteins. This processing step is crucial for the generation of the replicase complex that is required for RNA replication. SARS-CoV-2, in particular, encodes a single papain-like protease (henceforth referred to as PLpro) (Freitas et al., 2020).

The PLpro domain is part of the multi-domain nsp3 protein. It processes the nsp1/2, nsp2/3, and nsp3/4 cleavage sites. The PLpro domain, which is membrane-associated, recognizes a cleavage sequence in the threshold areas between nsp1/12, nsp2/3, and nsp3/4. After the cleavage, the nsp1 nsp2, nsp3 are separated from the viral polyprotein and the next steps of the formation of the replicase complex may be executed, allowing the production of new virions and the establishment of the viral infection. Studies on the enzymatic function of PLpro have shown its capacity for recognition and hydrolysis of the UBL protein ISG15 (interferon-induced gene 15) and the protein ubiquitin (Ub), both of which are cellular proteins. Thus, PLpro constitutes a deubiquitinating (DUB) and delISGylating enzyme (Báez-Santos et al., 2015). This type of activity has important ties to the course of the innate immune response in the setting of SARS-CoV infection. Numerous instances of ubiquitination and ISGylation events are observed in various stages of the innate immune response. The interference of PLpro in different parts of these pathways can be achieved through the recognition and the interaction with delISGylating and/or deubiquitinating proteins that exist therein. Specifically, PLpro appears to act as an antagonist, blocking the generation of cytokines of importance in the host innate immune response. These include Type I interferon (IFN-I), as well as CCL5 and CXCL10 (Freitas et al., 2020).

Given its crucial role in the virus replication cycle and the suppression of the host innate defenses, PLpro has had a significant existing history as a drug target against a number of coronaviruses, such as MERS-CoV and SARS-CoV. Similarly, amid the COVID-19 pandemic, PLpro constitutes a promising pharmacological target. Within the scientific community, there is a race for the development of inhibitors against this particular viral enzyme, as well as for the possible employment of existing protease inhibitors. Numerous studies are dedicated to the exploration of pre-existing, FDA-approved drugs with the goal of repurposing them for the inhibition of SARS-CoV-2. Through a number of methods, among which is molecular dynamics and virtual screening, FDA-approved drugs, which offer the advantage of already completed preclinical and clinical phases, are combed through in order to select the best candidates for the inhibition of the SARS-CoV-2 papain-like protease.
protease. A prime example of a drug candidate provided through such an approach is phenformin (Kandeel et al., 2020).

As mentioned above, the main protease, also known as 3C-like protease (3CLpro), is among popular targets against SARS-CoV-2. 3CLpro executes proteolytic cleavage of the overlapping pp1a and pp1ab translated viral polyproteins, allowing the production of functional proteins during the coronavirus replication process (Ullrich and Nitsche, 2020). More specifically, 3CLpro (nonstructural protein 5), following its autocleavage, processes the aforementioned polyproteins at their respective cleavage sites (Yu et al., 2004). The 3CLpro monomer contains three domains, with a long loop connecting the second and third domains. The enzyme’s active site is located between the first and the second domain, containing a catalytic pair of cysteine and histidine. The histidine’s imidazole attracts the side-chain proton of the cysteine, creating a thiolate nucleophile, which then attacks the amide bond on the enzyme’s substrate. Through proton abstraction from histidine, the N-terminal peptide product is released, while the C-terminal product is released through the hydrolysis of the thioester, allowing the subsequent cleavage and formation of the substrate-binding site (Zhang et al., 2020a). The 3CLpro monomer contains three domains, with a long loop connecting the second and third domains. The enzyme’s active site is located between the first and the second domain, containing a catalytic pair of cysteine and histidine. The histidine’s imidazole attracts the side-chain proton of the cysteine, creating a thiolate nucleophile, which then attacks the amide bond on the enzyme’s substrate. Through proton abstraction from histidine, the N-terminal peptide product is released, while the C-terminal product is released through the hydrolysis of the thioester, allowing the subsequent cleavage and formation of the substrate-binding site (Zhang et al., 2020a).

The mediation of nonstructural viral proteins’ maturation by 3CLpro makes it a very attractive target for the development of anti-coronavirus drugs. It is important to note that 3CLpro executes cleavage exclusively of the polypeptide sequences that follow a glutamine residue. This ability offers a significant advantage, given that there are no known host-cell protease enzymes that share this particular specificity of substrate. In essence, inhibitors of the main protease would be unlikely to have off-target effects (Ullrich and Nitsche, 2020). SARS-CoV inhibitors have so far steered towards peptide inhibitors and small-molecule inhibitors (Wu et al., 2020a). Studies conducted on the design and screening of 3CLpro inhibitors usually focus on cleavage sites, substrate-binding sites, the catalytic pair at the enzyme’s active center, as well as various reported residues of critical importance for enzyme function. Significant conservation of the catalytic site has been observed between the 3CLpro structures of the SARS-CoV, MERS-CoV, and SARS-CoV-2 viruses. This discovery allows the examination of already established inhibitors of the other coronaviruses’ 3CLpro enzyme, under the possibility that they may prove effective against the 3CLpro of the novel SARS-CoV-2 (He et al., 2020). Furthermore, research groups have been establishing screening processes for the possible repurposing of various drugs. Among those, anti-bacterial drugs, anti-hypertensive drugs, as well as natural compounds with antiviral properties have shown a high binding affinity to 3CLpro, making them attractive candidates for the inhibition of the viral 3CLpro towards the treatment of COVID-19 (Wu et al., 2020a).

### 3.2. NSP13 helicase

Another key enzyme necessary for the viral replication cycle that has been characterized as potent therapeutic target is the viral helicase (Vlachakis et al., 2013c). SARS-CoV-2 NSP13 has been shown to possess NTPase and RNA helicase activity; it catalyzes the unwinding of double stranded RNA in an NTP-dependent manner (Shu et al., 2020). Due to its sequence conservation across coronaviruses, targeted inhibition of its enzymatic activity is recognized as a promising strategy for antiviral therapy (Habsemair et al., 2020).

SARS-CoV-2 NSP13 has a high structural and functional similarity with MERS and SARS NSP13, belonging to the SF1 helicase family. It consists of 5 domains that form a triangular pyramid shape; the two ‘RecA-like’ domains, domain 1A and 2A, along with domain 1B form the triangular base, while the Zinc Binding Domain (ZBD) and stalk domain, that connects 1B and ZBD domains, are located at the apex of the pyramid (Hao et al., 2017; Jia et al., 2019). Domains 1A, 1B and 2A are involved in NTP and nucleic acid binding, while it has been shown that the NSP13 enzymatic activity is enhanced by direct interaction of NSP12 at the ZBD-1A interaction site (Jia et al., 2019; Romano et al., 2020).

A number of compounds have been reported as competitive inhibitors against SARS coronavirus (SCV) NSP13 activity that could pave the ground for clinical research for SARS-CoV-2 inhibition. Myricetin and scutellarein have been identified to strongly interrupt ATPase but not helicase activity, possibly by direct binding at the ATPase domain (Yu et al., 2012). Aryl diketoacids (ADK) and dihydroxychromone derivatives demonstrated selective inhibition against the SCV NSP13 duplex DNA-unwinding activity in multiple binding sites in the target enzyme, while 2-arylmethoxyloxy-6-(3-chlorobenzyloxy)-5-hydroxychromones were shown to inhibit both NTPase and helicase activities (Lee et al., 2009a, 2009b; Kim et al., 2011). Another class of antiviral compounds, bananins, have been shown to inhibit viral transcription and replication by significantly reducing the viral RNA levels whereas they did not affect viral entry. The compounds were tested against SCV NTPase/helicase activity and a number of bananin derivatives were found to potent ATPase and helicase inhibitors at concentrations significantly below toxicity levels (Tanner et al., 2005). 3-[(2-Nitrophenyl)sulfanyl)methyl]-4-prop-2-enyl-1H-1,2,4-triazole-5-thione, SSYA10-001, was recognized by Adedeji and coworkers as a novel inhibitor of SCV that interferes with the dsRNA/dsDNA activity of NSP13 without competitive inhibition of NTP and nucleic acid binding, but possibly by inducing conformational changes of the enzyme that disrupt the unwinding and translocation mechanisms (Adedeji et al., 2012).

Bismuth salts have been also characterized as strong inhibitors of both ATPase and duplex-unwinding activities through binding to the zinc binding domain (ZBD) of SCV NSP13 at micromolar concentrations (Yang et al., 2007a, 2007b). Similarly, a recent study on SARS-CoV-2 identified inhibitory effects of bismuth salts against NSP13, and specifically, bismuth potassium citrate (BPC) and ranitidine bismuth citrate (RBC) exhibited effective inhibition of NTPase and helicase activity with the same binding mode, interrupting the SARS-CoV-2 replication cycle (Shu et al., 2020).

In silico methods have been proven to be a crucial step for the rapid, cost-efficient identification of potent compounds with inhibitory effects in antiviral research (Vlachakis and Kossida, 2013; Loukatou et al., 2015; Papageorgiou et al., 2016). As such, computational approaches employing molecular modelling and virtual screening against known antivirals, active compounds and/or FDA approved drugs for drug repurposing, have been applied for the identification of candidate SARS-CoV-2 helicase inhibitors (Loukatou et al., 2015). Iftikhar et al. identified two compounds, meclonazepam and oxiphenisatin, that bind at β19–β20 loop on the 1A domain of SARS-CoV-2 helicase, that is directly involved in unwinding double stranded nucleic acids (iftikhar et al., 2020). Vapreotide, an AIDS-related diarrhea drug, was found to exhibit the lowest binding free energy amongst 23 approved antivirals in a computer-aided drug design pipeline (Borgio et al., 2020). Similar approaches have identified cmp1, cmp3a, cmp11 and cmp15 to competitively bind at the NTP binding site (Mirza and Froeyen, 2020), and elbasvir, approved for HCV treatment, to inhibit not only helicase activity by blocking ssDNA binding, but also RNA binding to RdRNP (Balasubramaniam and Reis, 2020).

### 3.3. NSP12 polymerase

RNA viruses encode RNA-dependent RNA Polymerase (RdRP), an essential enzyme for viral replication since it catalyzes RNA synthesis, governing the initiation and elongation phase (Ng et al., 2008). RdRPs share a similar core structure resembling the shape of a right hand with three subdomains, namely the palm, finger and thumb subdomains (Vlachakis et al., 2017; Venkataraman et al., 2018). The catalytic site of RdRPs is formed by seven conserved motifs (Motif A-G), five of which are located within the most conserved palm subdomain (A-E), while Motifs F and G are found in the fingers (Vlachakis et al., 2013a, 2013b; Shu and Gong, 2016). SARS-CoV-2 polymerase, or NSP12, recently
Several clinical trials have been registered and are ongoing (Cai et al., 2020). Due to their high significance in viral life cycle and their conserved mode of function across viral families, RdRPs have long been the target of antiviral research with a number of approved therapies targeting viral polymerases (Tsai et al., 2006; Ng et al., 2008; Reznik and Ashby, 2017; Yao et al., 2018; Peersen, 2019). Consequently, during the current COVID-19 outbreak, NSP12 is considered as a promising target for antiviral therapeutics. Identified nucleoside analogues that were originally developed to target viral RdRPs and novel compounds are evaluated as antiviral drug candidates against COVID-19 (Neogi et al., 2020).

Remdesivir is the most promising drug candidate to date. It is a prodrug of an adenosine nucleotide analog that is converted to active nucleoside triphosphate upon cell entry and competes with ATP in the RdRP binding site, leading to chain termination (Warren et al., 2016; Yin et al., 2020). Remdesivir was developed for the treatment of Ebola virus disease (EVD), but it has not been approved after showing low efficacy in a phase III clinical trial for EVD (Warren et al., 2016). However, in vitro and in vivo studies of remdesivir against MERS-CoV and SARS-CoV and recent studies against SARS-CoV-2 showed inhibitory effects and antiviral activity with low cytotoxicity levels (Sheahan et al., 2017; Wang et al., 2020). Additionally, remdesivir was tested in a compassionate clinical trial for severe COVID-19 patients, and after 10-day treatment an 86% clinical improvement was observed (Grein et al., 2020). Several clinical trials have been reported to evaluate remdesivir safety and efficacy, while a number of patents have already been granted (Eastman et al., 2020).

Favipiravir, a guanine analog developed for the treatment of influenza viruses, currently approved in China and Japan, is also an anti-COVID19 drug candidate (Furuta et al., 2017). Favipiravir inhibits polymerase activity, yet the exact mode of action has not been elucidated. Two possible mechanisms for chain termination have been proposed, either by inhibition of the ATP catalytic site, or by misincorporation in a nascent viral RNA (Furuta et al., 2005). Clinical trials for favipiravir have shown efficient treatment for moderate COVID-19 patients and several clinical trials have been registered and are ongoing (Cai et al., 2020; Du and Chen, 2020; Li and Clerq, 2020).

Besides remdesivir and favipiravir, several nucleotide analogs have been reported as potential antiviral agents. Ribavirin, a ribonucleoside analog that has been tested against several RNA viruses but exhibits antiviral activity with low cytotoxicity levels (Furuta et al., 2005). Additionally, remdesivir was tested in a compassionate clinical trial for severe COVID-19 patients and after 10-day treatment an 86% clinical improvement was observed (Grein et al., 2020). Several clinical trials have been reported to evaluate remdesivir safety and efficacy, while a number of patents have already been granted (Eastman et al., 2020).

Favipiravir, a guanine analog developed for the treatment of influenza viruses, currently approved in China and Japan, is also an anti-COVID19 drug candidate (Furuta et al., 2017). Favipiravir inhibits polymerase activity, yet the exact mode of action has not been elucidated. Two possible mechanisms for chain termination have been proposed, either by inhibition of the ATP catalytic site, or by misincorporation in a nascent viral RNA (Furuta et al., 2005). Clinical trials for favipiravir have shown efficient treatment for moderate COVID-19 patients and several clinical trials have been registered and are ongoing (Cai et al., 2020; Du and Chen, 2020; Li and Clerq, 2020).

Besides remdesivir and favipiravir, several nucleotide analogs have been reported as potential antiviral agents. Ribavirin, a ribonucleoside analog that has been tested against several RNA viruses but exhibits higher levels of cytotoxicity and side-effects, is also on ongoing clinical trial for COVID-19 treatment combined with interferon-α2b (Falzarano et al., 2013). Jockusch et al. have shown the inhibitory effects of sofosbuvir, alavudine, vidovudine, tenofovir alafenamide and emtricitabine and termination of polymerase RNA synthesis through in vitro assays (Jockusch et al., 2020). Sibilin is predicted to have a dual activity against SARS-CoV-2 infection; sibilin can potentially reduce viral replication activity by targeting NSP12 as a remdesivir-like inhibitor, and modulate inflammatory responses by direct inhibition of STAT3 (Bosch-Barrera et al., 2020). STAT3 showcases both pro- and anti-inflammatory abilities; it mediates interleukin 10 (IL-10) activity, a known anti-inflammatory cytokine. On the other hand, in the case of cancer, persistent activation of STAT3 mediates tumor-promoting inflammation (Hillmer et al., 2016; Yu et al., 2009).

As in any therapeutic approach concerning human disease, toxicity is a critical component in the development of safe and effective drugs against SARS-CoV-2. ACE2, as reviewed earlier, is a promising target for the development of therapeutic strategies. A recombinant form of the human ACE2 protein was synthesized as a therapeutic treatment for COVID-19, functioning as a decoy for SARS-CoV-2 and essentially preventing the virus from binding to the cell surface ACE2 (Schuster et al., 2010). The approach of recombinant proteins carries the risk of off-target binding, which is the binding of the recombinant protein to other cell-receptors, leading to unwanted signaling cascades. Furthermore, use of recombinant proteins may lead to the activation of the immune system, which may in turn result in the production of anti-drug antibodies and their deposition in the form of immune complexes, potentially hindering the function of important organs (De GROOT and Scott, 2007). Lastly, this sort of treatment may lead to serum sickness, in the case that the immune system may recognize the recombinant protein as an antigen (Schellekens, 2005).

In the case of viral replication inhibitors, instances of toxicity reported for other nucleoside analogs can provide insight into the possible toxicity of analogs such as favipiravir or remdesivir, which are being evaluated for use against SARS-CoV-2. While inhibiting the viral polymerase enzyme, nucleoside analogs additionally inhibit mitochondrial DNA polymerase-gamma, resulting in reduction of mitochondrial protein synthesis (Moyle, 2000). Aspects of toxicity include myopathy and pancreatitis (Johnson et al., 2001), severe metabolic acidosis paired with elevated lactate, with significant observed mortality (Falco et al., 2002) and bone marrow suppression in patients treated with 3′-azido-3′, 3′-dideoxythymidine (AZT) (Moyle, 2000). As far as protease inhibitors are concerned, drugs such as ritonavir, an inhibitor of cytochrome CYP3A4, can have an effect on the metabolism of drugs-substrates of CYP3A4, which in turn may result in unpredictable drug interactions (Rock et al., 2014). As summarized by Chary and colleagues, the use of humanized antibodies as a therapeutic method harbors dangers of hypersensitivity reactions, immunostimulation, which results in chills, fever and other reactions, and finally, immunosuppression, which may often become a window for opportunistic infections (Chary et al., 2020).

To date, efforts are being made by the scientific community to develop a vaccine against SARS-CoV-2. Vaccines that have been developed and are in the stage of clinical trials are based on one hand on the virus genome and on the other hand, on the structure of the viral S-protein (WH Chen et al., 2020; Li et al., 2020; Tu et al., 2020; Tu et al., 2020). One of the vaccines that started developing in early 2020 is Moderna’s mRNA-1273. This vaccine is a synthetic close of mRNA that encodes the viral S-protein. The goal of the vaccine is to provoke an immune response specific for S-protein of SARS-CoV-2. Another characteristic of this vaccine is that the virus is not required for its development, as the synthetic mRNA is encapsulated in lipid nanoparticles (Tu et al., 2020).

The ChAdOx1 nCoV-19 vaccine was developed by the University of Oxford and is currently in clinical phase III. This vaccine consists of a...
non-replicable adenovirus vector and the S protein sequence of the SARS-CoV-2 protein. Characteristic of this vaccine is that due to the adenovirus-based vector, it can cover both the respiratory and gastrointestinal epithelium, which are the two main sites expressing the SARS-CoV-2 ACE-2 receptor (Folegatti et al., 2020). Finally, the University of Queensland is developing a vaccine that is equally based on the molecular structure of the virus SARS-CoV-2. As SARS-CoV-2 is an enveloped virus, the fusion of its viral membrane with the host cell membrane is required for infection. For this purpose, the structure of the glycoproteins of the viral envelope alter from the pre-fusion form, which is more unstable, to the post-fusion form. Thus, this vaccine is a stabilized subunit vaccine based on molecular clamp technology, such as the vaccines against influenza virus and Ebola virus, and allows the recombinant viral proteins to remain stable in the pre-fusion form (Coleman et al., 2014; Tu et al., 2020).

CRediT authorship contribution statement

**Dimitrios Vlachakis:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Eleni Papakonstantinou:** Methodology, Writing – original draft, Writing – review & editing. **Thanasis Mitis:** Writing – original draft, Writing – review & editing. **Katerina Pieroulou:** Writing – review & editing. **Io Diakovou:** Writing – review & editing. **George Chrousos:** Supervision, Writing – review & editing. **Flora Bacopoulos:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Adedeji, A.O., Singh, K., Calcuttara, N.E., DeDiego, M.L., Enjuanes, L., et al., 2012. SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature. https://doi.org/10.1038/nature19887.

Almandil, N.B., et al., 2020. Declaration of competing interest.

Balasubramian, M., Reis, R.J.S., 2020. Computational target-based drug repurposing evaluation and treatment coronavirus (COVID-19). StatPearls. StatPearls Publishing LLC, Treasure Island (FL). https://doi.org/10.1097/JVL.000060-11.

Borgio, J.F., Alsuwat, H.S., Al Otaibi, W.M., Ibrahim, A.M., Almandil, N.B., et al., 2020. Purified coronavirus spike protein nanoparticles induce coronavirus neutralizing antibodies in mice. Vaccine 32 (26), 3169–3174. https://doi.org/10.1016/j.vaccine.2020.04.016.

Bouhaddou, M., Xu, J., Obernier, K., et al., 2020. The spike protein of the new coronavirus 2019-CoV contains a furin-like cleavage site absent in CoV of the same clade. Antivir. Res. 176, 104742. https://doi.org/10.1016/j.antiviral.2020.104742.

Cai, Q., Yang, M., Liu, D., Chen, J., Shu, D., et al., 2020. Experimental treatment with favipiravir for COVID-19: an open-label control study. Engineering (Beijing). https://doi.org/10.1021/acscentsci.0c00489.

Cai, Q., Yang, M., Liu, D., Chen, J., Shu, D., et al., 2020. Experimental treatment with favipiravir for COVID-19: an open-label control study. Engineering (Beijing). https://doi.org/10.1021/acscentsci.0c00489.

Chang, C.K., Hou, M.H., Chang, C.F., Hsiao, C.D., Huang, T.H., 2014. The SARS coronavirus nucleoside protein–forms and functions. Antivir. Res. 103, 39–50. https://doi.org/10.1016/j.antiviral.2014.11.005.

Chen, C.Y., Chang, C.K., Chang, Y.W., Sue, S.C., Bai, H.L., et al., 2004. Nucleoside protein as early diagnostic marker for SARS. Emerg. Infect. Dis. 10 (11), 1947–1949. https://doi.org/10.3201/eid1011.0402206.

Chen, C.Y., Chang, C.K., Chang, Y.W., Sue, S.C., Bai, H.L., et al., 2004. Nucleoside protein as early diagnostic marker for SARS. Emerg. Infect. Dis. 10 (11), 1947–1949. https://doi.org/10.3201/eid1011.0402206.

Chen, Y., Liu, Q., Guo, D., 2020. Emerging coronaviruses: genome structure, replication, and pathogenesis. J. Med. Virol. 92 (4), 418–423. https://doi.org/10.1002/jmv.25861.

Cheng, J.H., Strych, U., Hotes, P.J., Bottazzi, M.E., 2020. The SARS-CoV-2 vaccine pipeline: an overview. Current Tropical Medicine Reports 7 (2), 61–64. https://doi.org/10.1007/s40475-020-00201-6.

Cohen, J.R., Lin, L.D., Machamer, C.E., 2011. Identification of a Golgi complex-targeting signal in the cytoplasmic tail of the severe acute respiratory coronavirus envelope protein. J. Virol. 85 (12), 5794–5803. https://doi.org/10.1128/JVI.00600-11.

Cortaud, B., Vallé, C., de Lamballerie, X., Camard, B., Seilah, N.G., et al., 2020. The spike protein of SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature. https://doi.org/10.1038/s41586-020-2286-9.

Definiens: a review of its discovery and development leading to emergency use authorization for treatment of COVID-19. ACS Cent. Sci. 6 (5), 672–683. https://doi.org/10.1021/acscente.0c00489.

Estrada, E., 2020. Topological Analysis of SARS CoV-2 Main Protease. https://doi.org/10.1002/jbfa.4728.

Falcon, V., Rodríguez, D., Ribera, E., Martínez, E., Miro, J.M., et al., 2002. Severe nucleoside-associated lactic acidosis in human immunodeficiency virus-infected patients: report of 12 cases and review of the literature. Clin. Infect. Dis. 34 (6), 838–846. https://doi.org/10.1086/339041.

Falcón, V., de Wit, E., Martelaro, C., Callison, J., Munster, V.J., et al., 2013. Inhibition of novel beta coronavirus replication by a combination of interferon-alpha2b and ribavirin. Sci. Rep. 3, 1688. https://doi.org/10.1038/srep01686.

Fang, X., Gao, J., Zheng, H., Li, B., Kong, L., et al., 2007. The membrane protein of SARS-CoV suppresses NF-kappaB activation. J. Med. Virol. 79 (10), 1431–1439. https://doi.org/10.1002/jmv.20955.

Folegatti, P.M., Ewer, K.J., Aley, P.K., Angus, B., Becker, S., Belij-Rammerstorfer, S., et al., 2020. Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against COVID-19: a preliminary report of a phase 1/2, single-blind, randomised controlled trial. Lancet 396 (10249), 467–478. https://doi.org/10.1016/S0140-6736(20)31604-4.

Freitas, B.T., Durie, I.A., Murray, J., Longo, J.E., Miller, H.C., et al., 2020. Return of the coronavirus: 2019-nCoV. Viruses 12 (4), 4389089. https://doi.org/10.1128/JVI.02001-08.

Gralinski, L.E., Menachery, V.D., 2020. The SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA. J. Mol. Biol. 368 (4), 1075–1086. https://doi.org/10.1016/j.jmb.2007.02.069.

Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., et al., 2020. A SARS-CoV-2 membrane binding proteins of coronaviruses. Future Virol. 14 (4), 275–286. https://doi.org/10.2174/16612822.478.

Grau, G.E., de la Torre, J.C., 2003. Neuraminidase of paramyxoviruses: a mechanism for helical packaging of viral RNA. J. Mol. Biol. 368 (4), 1075–1086. https://doi.org/10.1016/j.jmb.2007.02.069.

Grifﬁths, W.G., 1999. Silencing of the human immunodeﬁciency virus: state of the art 2000. Trends Virol. 20 (3), 108–114. https://doi.org/10.1016/S0959-4479(99)01201-5.

Gupta, M.K., Vemula, S., Donde, R., Gouda, G., Behera, L., et al., 2020. In-silico approaches to detect inhibitors of the human severe acute respiratory syndrome}
Food and Chemical Toxicology 146 (2020) 111805
