REVIEW ARTICLE

Structure genomics of SARS-CoV-2 and its Omicron variant: drug design templates for COVID-19

Can-rong Wu[1], Wan-chao Yin[1], Yi Jiang[1] and H. Eric Xu[1,2]

Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has brought an unprecedented public health crisis and persistently threatens to humanity. With tireless efforts from scientists around the world, understanding of the biology of coronavirus has been greatly enhanced over the past 2 years. Structural biology has demonstrated its powerful impact on uncovering structures and functions for the vast majority of SARS-CoV-2 proteins and guided the development of drugs and vaccines against COVID-19. In this review, we summarize current progress in the structural biology of SARS-CoV-2 and discuss important biological issues that remain to be addressed. We present the examples of structure-based design of Pfizer's novel anti-SARS-CoV-2 drug PF-07321332 (Paxlovid), Merck's nucleotide inhibitor molnupiravir (Lagevrio), and VV116, an oral drug candidate for COVID-19. These examples highlight the importance of structure in drug discovery to combat COVID-19. We also discussed the recent variants of Omicron and its implication in immunity escape from existing vaccines and antibody therapies.

Keywords: COVID-19; SARS-CoV-2; structural biology; omicron; drug design; vaccine development

INTRODUCTION

Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has led to more than 274 million infections and over 5.35 million deaths by December 20, 2021 [1–5]. SARS-CoV-2, an enveloped, single-stranded positive-sense RNA virus, belongs to beta-coronavirus and is related to two highly pathogenic coronaviruses, SARS-CoV and Middle East respiratory syndrome coronavirus (MERS-CoV), which have caused over 8000 and 2500 confirmed cases, with ~10% and 35% fatality rates, respectively [5–7]. Compared with SARS-CoV and MERS-CoV, SARS-CoV-2 is spreading more rapidly and causes much higher number of deaths. This emerging virus has promoted massive hospitalizations, lockouts, financial loss, unemployment, the closure of schools in nearly all countries [8].

The SARS-CoV-2 genome is a non-segmented large positive-sense stranded RNA with a length of about 30 kb. It contains a 5′-cap structure and a 3′-poly-A tail [7, 9]. The viral genome encodes 29 proteins, including 25 putative non-structural and accessory proteins and four structural proteins [10]. Non-structural proteins (NSPs) play crucial roles in viral RNA replication and immune evasion, while accessory proteins carry out multiple functions that aid with viral infection, survival, and transmission in the host cells [8, 11, 12]. Structural proteins are responsible for viral assembly and make up the mature viral particles. Elucidating the structure and function of these SARS-CoV-2 encoded proteins will deepen the understanding of the viral infection cycle and offer a new opportunity to develop effective vaccines and drugs to combat this global pandemic.

Not long after the first SARS-CoV-2 genome sequence was published [9], Rao et al. deposited the first structure of SARS-CoV-2, Main protease (M[pro]), into Protein Data Bank [13]. With the tireless efforts of scientists worldwide, structures of SARS-CoV-2 proteins have burst and reached 1250 in the current moment. These structures cover more than 90% of the SARS-CoV-2 coding amino acids. The rapidly expanding repertoire of SARS-CoV-2 structures has provided new insights into the viral life cycle and facilitates drug and vaccine development.

In this review, we provide an overview of the SARS-CoV-2 genome and its coded proteins. We then describe the structures of SARS-CoV-2 proteins and mainly focus on RNA-dependent RNA polymerase and S protein. We provide insights into in situ structures of SARS-CoV-2 that have been solved by cryo-electron tomography (cryo-ET). Finally, we discuss the remaining challenges and future perspectives on the structural biology of SARS-CoV-2 and highlight the recent success of the development of anti-COVID-19 drugs.

THE GENOME OF THE SARS-COV-2 AND ITS CODED PROTEINS

The 5′-proximal two-thirds of the coronavirus genome contains the replicase gene, which encodes two open reading frames, ORF1a and ORF1b (Fig. 1a) [10]. The other one-third of the genome at the 3′-end encodes several ORFs. Among them, four ORFs encode coronavirus structural proteins, which are spike glycoprotein (S), membrane (M), envelope (E), and nucleocapsid (N) proteins, where S, E, and M present on virion membrane surfaces, with the N protein is involved in the binding and packing

[1]The CAS Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China and [2]School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China
Correspondence: Can-rong Wu (wucanrong@simm.ac.cn) or H. Eric Xu (eric.xu@simm.ac.cn)

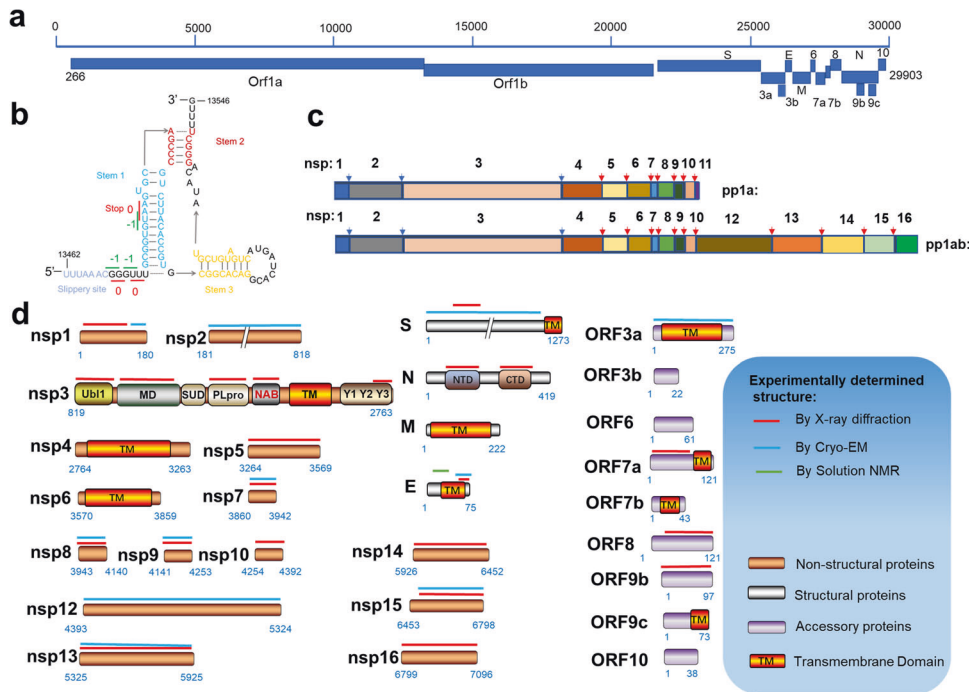Structure genomics of SARS-CoV-2: drug design for COVID-19
CR Wu et al.

3022

Fig. 1 The genome of SARS-CoV-2 and its coded proteins. a The organization of SARS-CoV-2 genome. **b** Schematic illustrations of the secondary structure of the frameshift stimulation element −1 programmed ribosomal frameshifting, with different functional regions labeled and colored accordingly. **c** The rectangle depicts the nsps derived from processing of the pp1a and pp1ab polyproteins. Labels indicate protein names. The position of blue arrow indicates cleavage site of PL^pro, and position of red arrow indicates the site of M^pro. **d** Domain architectures of SARS-CoV-2 genome coded proteins and summary of structural characterization of individual proteins. Bars above domain architectures indicate regions of the proteins for which high-resolution structures area are available. Ubl ubiquitin-like domain, MD macrodomain, SUD SARS unique domain, PL^pro papain-like protease, NAB nucleic acid-binding domain, TM transmembrane domain, Y Y region, NTD N-terminal domain, CTD C-terminal domain.

of the RNA genome. Other ORFs encode several accessory proteins [14, 15] (Fig. 1a), with their functions less well-defined.

ORF1a translation yields a replicase polyprotein 1a (ppla), and a −1 ribosomal frameshift at the 3′-end of ORF1a, which facilitates the translation of ORF1b to produce the replicase polyprotein 1ab (pp1ab) [16–18]. The −1 programmed ribosomal frameshifting mechanism is stimulated by a three-stem pseudoknot at the 3′ of slippery sequence -UUUAAAC- named the frameshift stimulation element (FSE) (Fig. 1b) [19]. When the frameshift does not occur, the stop codon at the stem 1 induces translation termination (Fig. 1b). Conversely, when the −1 ribosomal frameshift occurs, translation termination does not occur, which results in the generation of a protein with additional ~2700 amino acids. The ribosomal frameshifting occurs at a frequency of 0.25–0.75 at this site [19]. The structures of FSE alone and FSE complexed with ribosomes were recently reported [19, 20].

In SARS-CoV-2, the polyprotein pp1a is proteolytically cleaved into 11 functional NSPs, whereas pp1ab is cleaved into 15 NSPs (Fig. 1c). Each NSP plays specific or multifaceted roles in the viral life cycle. Nsp1 interacts with the host ribosome to inhibit the synthesis of host proteins. Nsp2 may be involved in coupling viral transcription with translation by interacting with both ribosome and replication-transcription complexes (RTC) [21]. The papain-like protease (PL^pro) from nsp3 cleaves at three sites forming mature nsps 1–3, while the nsp5, also named main protease (M^pro), cleaves pp1a and pp1ab at 11 sites, releasing the mature nsps 4–16 [22, 23] (Fig. 1c). Considering their vital roles in NSPs maturation, M^pro and PL^pro have been considered attractive therapeutic targets for COVID-19 [24, 25]. In addition to the PL^pro domain, nsp3 contains several other domains: ubiquitin-like domain, macrodomain (also named "X domain"), nucleic acid-binding domain (NAB), SARS coronavirus-unique domain, transmembrane domain and the Y1-3 domain [26]. The transmembrane

domain of nsp3 together with nsp4 and nsp6 plays a key role in rearranging endoplasmic reticulum (ER) membranes, leading to curvature of the ER membrane, which is adapted to form double-membrane vesicles (DMVs) and is essential for virus replication [27]. Nsp7 and nsp8 are assembled with nsp12 to enhance the RNA polymerase activity.

Nsp7-nsp16 are involved in the formation of a large RTC to regulate the replication and transcription of SARS-CoV-2. Nsp9 inhibits the nucleotidyltransferase activity of nsp12 [28]; nsp10 is a cofactor to nsp14 and nsp16 [29]; nsp12 functions as a RNA-dependent RNA polymerase and nucleotidyltransferase [30]; nsp13 acts as a helicase; nsp14 is exoribonuclease and N7-guaninemethyltransferase [31]; nsp15 has uridine-specific endoribonuclease activity [32]; nsp16 has 2′-O-methyltransferase activity that mediates mRNA cap process [33].

Among four structural proteins, S, M and E are located on the viral membrane. S plays an essential role in the host receptor-binding and membrane fusion [34, 35]. M protein is associated with N protein and other viral structural proteins to facilitate the viral assembly and is involved in the pathogenesis process [36]. E protein forms an ion channel, which promotes virus assembly and pathogenesis [37, 38]. N protein participates in viral genome RNA packaging and promotes virion assembly [39, 40].

Among nine accessory proteins, ORF3a, ORF7a and ORF7b are transmembrane proteins. ORF3a forms a homo-dimer, with its transmembrane domain forming an ion channel in the host cell membrane. It participates in virus release, pathogenesis and host cell apoptosis induced by the virus [41]. As a type I transmembrane protein, ORF7a is an immunomodulatory factor for immune cell binding and is involved in dramatic inflammatory responses. ORF7b is a small, integrative membrane protein found in the host cell Golgi, which may increase the virulence of SARS-CoV-2 [42]. It is coded by the same mRNA as ORF7a and expressed according to

Structure genomics of SARS-CoV-2: drug design for COVID-19
CR Wu et al.

3023

the leaky-scanning mechanism of ribosomes [42]. ORF8 is one of the newest genes and has low homology with SARS-CoV due to deletion [43]. This protein interacts with major histocompatibility complexes I, thus modulating their degradation in cell cultures, thereby contributing to immune evasion [44]. ORF9b mediates interferon response suppression by inhibiting the interaction between Hsp90 and TOM70 [45]. ORF10 is predicted to encode a small protein. Pancer et al. suggest that this protein is not essential for the viral life cycle in humans [46]. To date, only the structures of ORF3b, ORF6, ORF9c and ORF10 remain unsolved (Fig. 1d).

THE STRUCTURES OF RNA-DEPENDENT RNA POLYMERASE AND REPLICATION-TRANSCRIPTION COMPLEX
Nsp12 functions as an RNA-dependent RNA polymerase and is a core component of the replication-transcription machinery [47]. Auxiliary factors nsp7 and nsp8 are essential for enhancing the binding of nsp12 to the RNA template and its enzymatic activity [48]. The association of the nsp7–nsp8–nsp12 and template-primer RNA constitutes the minimum components of a functional holo-RdRp, which is the main component of the SARS-CoV-2 RTC [49]. Several structures of holo-RdRp or holo-RdRp complexed with inhibitors have been determined by cryo-electron micro-scopy (cryo-EM) [30, 47, 50–53]. These structures provide insights into the molecular architecture of the holo-RdRp components. The nsp12 subunit mainly contains several highly conserved domains in coronavirus, including a nidovirus-specific extension domain (NiRAN), a C-terminal polymerase domain and an interface domain in the middle (Fig. 2a). The β-hairpin (residues V31 to K50), a unique conserved domain located upstream of the NiRAN domain, is sandwiched by the NiRAN and palm subdomain in the polymerase domain and stabilizes the overall structure (Fig. 2a). Yan et al. disclosed that NiRAN catalyzes the formation of cap core structure (GpppA) [28]. The C-terminal polymerase core domain of nsp12 adopts a classic cupped right-handed conformation, consisting of the finger, palm, and thumb subdomains and several conserved structural motifs. The SARS-CoV-2 nsp7-nsp8 heterodimer shares significant structural similarity to that of SARS-CoV. The nsp7-nsp8 heterodimer binds above the thumb subdomain and sandwiches the extended finger loops to stabilize nsp12 (Fig. 2a). Nsp7 contributes to most of the binding of the heterodimer to nsp12, while nsp8 sparsely contacts nsp12. The second nsp8 (nsp8-2) attaches to the top of the finger subdomain and forms additional interactions with the interface domain. The active sites of SARS-CoV-2 RdRp and RNA recognition have been summarized and discussed in detail in the previous review [29].

To date, several antiviral drugs such as Remdesivir and Favipiravir that target the viral RdRp have been repurposed for COVID-19 therapy [54–59] (Fig. 2b). In April 2020, the first cryo-EM structure of holo-RdRp in complexed with Remdesivir was reported. Biochemical studies showed that 1 mM Remdesivir triphosphate (RTP) displayed a complete inhibition of RdRp polymerization activity in the presence of 10 mM ATP, which mimicked the physiological concentrations of ATP. While 100 µM RTP displayed a delayed chain termination mechanism [50]. Kinetic studies on RdRp demonstrated that the polymerase can still add 2 or 3 nt after incorporating Remdesivir monophosphate (RDV-MP) [60, 61]. This delayed chain termination model was further supported by recently solved structures of the SARS-CoV-2 polymerase-RNA complex in its catalytic states incorporated with RDV-MP [62, 63]. Upon incorporation of Remdesivir, RDV-MP moved to position −4 of the RNA product strand results in a steric clash between the C1 cyano group of RDV-MP and the side chain of nsp12 Ser861 (Fig. 2b), which may cause the delayed chain termination of Remdesivir on RdRp. Because Remdesivir is intravenously administered nucleotide prodrug that is inconvenient to patients, VV116, an orally available Remdesivir derivative,

was developed and showed excellent efficacy in inhibiting SARS-CoV-2 replication in cell-based and animal models [64] (Fig. 2b).

Favipiravir, an orally administered nucleoside analog, also showed a potential therapeutic effect on COVID-19 [65]. Unlike Remdesivir, Favipiravir showed a poor RNA replication termination effect in the presence of natural nucleotides [66]. It mainly displays as a mutagenic agent toward viral genomic RNA in vivo [67, 68], different from Remdesivir, which impairs the elongation of RNA products. The mechanisms of binding of Favipiravir to RdRp and the low efficiency of incorporation of Favipiravir-RTP are revealed by the cryo-EM structure of the Favipiravir-bound SARS-CoV-2 RdRp-RNA complex (Fig. 2b) [69]. A recent study reported similar findings that Favipiravir-RTP complexed with SARS-CoV-2 poly-merase in the pre-catalytic state [70].

Molnupiravir, a novel RdRp inhibitor, has recently received much attention, as it showed very promising phase III study results [https://www.drugs.com/history/molnupiravir.htm]. Patients with mild or moderate COVID-19 had a reduced risk of hospitalization or death by about 50% compared with placebo in a positive interim analysis of the oral antiviral drug Molnupiravir in a phase III study [71]. When Molnupiravir was incorporated into RNA template by the RdRp, Molnupiravir directs incorporation of A or G, resulting in mutation of the viral RNA genome [59]. Recently, the first structure of SARS-CoV-2 RdRp complexed with the non-nucleotide inhibitor, Suramin, was reported [30]. The RdRp harbors two independent Suramin binding sites. One suramin molecule binds to the template RNA binding site, directly blocking the substrate binding. The other one binds to the site near the catalytic site, thus blocking RNA primer binding by steric hindrance (Fig. 2b). This unique binding pattern clarifies that Suramin is at least 20-fold potent than the triphosphate form of Remdesivir (RDV-TP) in polymerase activity inhibition assays in vitro.

Promisingly, during the revision of this manuscript, Molnupiravir (Lagevrio) was approved by the Medicines and Healthcare products Regulatory Agency (MHRA) and FDA for the treatment of established infections of COVID-19 [72]. Benefiting from recent progress in structural biology, the precise interaction between drugs or drug candidates and holo-RdRp has been elucidated, and the mechanism of these drugs block virus replication by targeting RdRp has been revealed. The structural information will greatly promote drug design through the structure-based approach and facilitate the development of the next generation of RdRp inhibitors.

Holo-RdRp is located in the central of the SARS-CoV-2 RTC [29]. In addition, a number of conserved accessory factors, which are thought to coordinate with Holo-RdRp, are also required during the viral life cycle [18]. One of these accessory proteins, nsp13, acts as an RNA helicase to unwind the RNA double strand. Chen et al. reported the cryo-EM structure of two molecules of nsp13 in complex with nsp7-nsp8-nsp12 containing the RNA template products [73]. One of those interacted with nsp8-1 named nsp13-1, the other named nsp13-2. The nsp13-1-ZBD interacts with nsp12 thumb domain and nsp8-1. Meanwhile, nsp13-1 RecA1 domain abuts upon the nsp7 and the nsp8-1 head (Fig. 2c), while nsp13-2 binds less tightly in the complex than does nsp13-1. Aside from the interactions between the nsp13.2-ZBD and nsp8-2, nsp13-2 only interacts with nsp13-1. The RdRp translocates in the 3′ → 5′ direction on the template RNA strand. Meanwhile, nsp13 translocates on this strand in the 5′ → 3′direction direction. Integrating structural analysis, unbiased molecular dynamics simulations and RNA–protein cross-linking studies, Malone et al. proposed the backtracking mechanism [74]. The nsp13 facilitates reverse RNA translocation, resulting in a single-stranded 3′ seg-ment of the product RNA extruding through the NTP entry tunnel. This backtracking manner is similar to cellular DNA-dependent RNA polymerases. Thus, nsp13 may play an important role in enhancing the fidelity of RdRp [75]. More recently, Yan et al. have
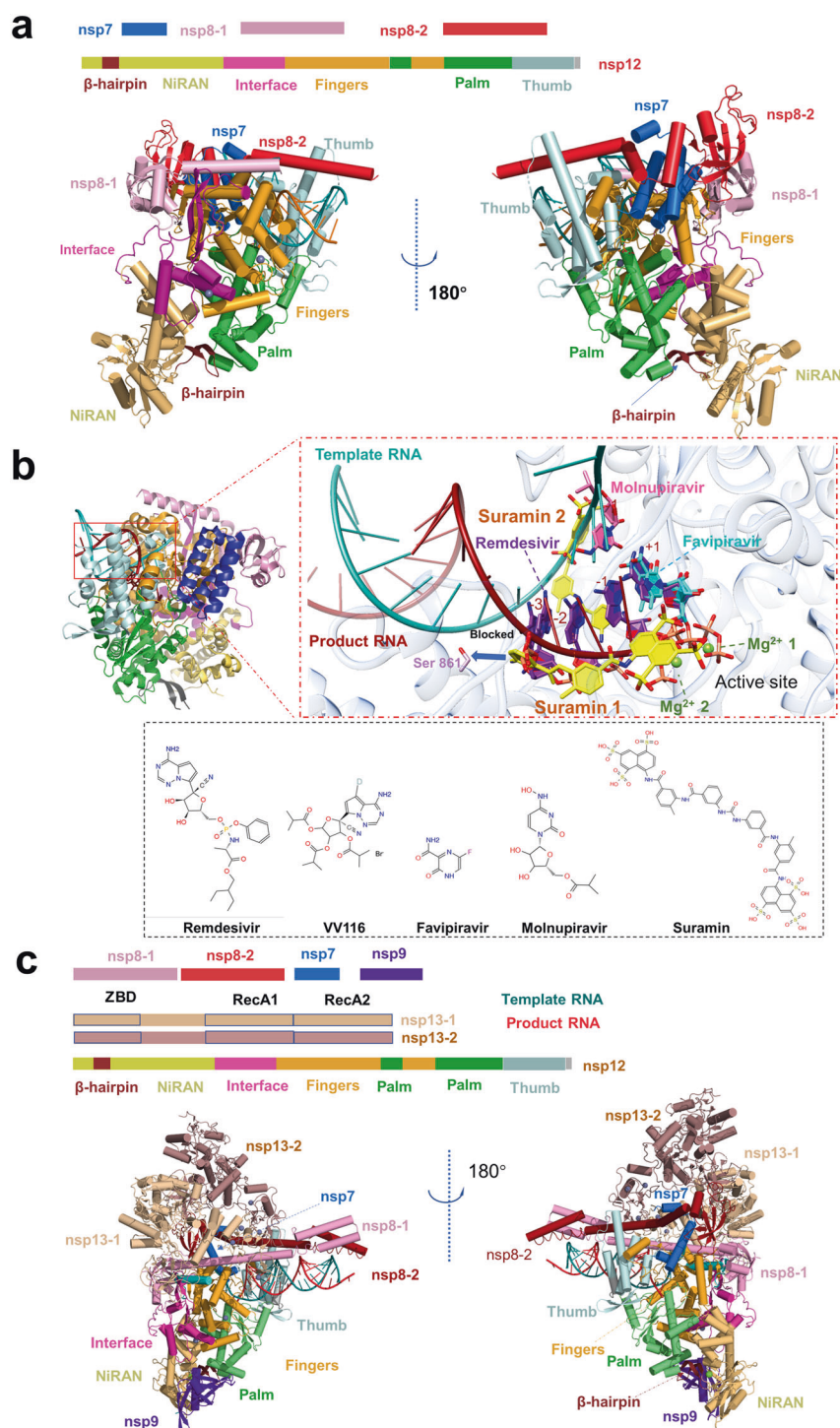
Structure genomics of SARS-CoV-2: drug design for COVID-19
CR Wu et al.

3024

Fig. 2 Structures of the SARS-CoV-2 holo-RdRp and RTX. a The schematic diagram for the domain organization of the holo-RdRp, containing nsp12, nsp7, nsp8-1 and nsp8-2 (PDB: 7C2K). Two views of the cartoon model of the cryo-EM structure of holo-RdRp. The subdomains and components of the holo-RdRp are colored as follows: β-hairpin, chocolate; NiRAN, light orange; Interface, light magenta; Fingers, bright orange; Palm, lime; Thumb, pale cyan; nsp7, marine; nsp8-1, pink; nsp8- 2, fire brick; template RNA, red; product RNA, deepteal. **b** Zoom in views of holo-RdRp bound to inhibitors, Favipiravir at +1 (PDB: 7AAP, 7CTT), Remedisivir at +1 (PDB: 7BV2), Remedisivir at −1 (PDB: 7C2K), Remedisivir at −3 (PDB: 7B3B), Remedisivir at +1, −1, −2, −3 (PDB: 7L1F), Suramin (PDB: 7D4F), Molnupiravir at template RNA, the four inhibitors are indicated in different colors. Proteins and RNA are shown in cartoon representation, and inhibitors are shown as sticks. Nsp12 is shown in transparent gray. Ligand atom color code: O atoms, red; N atoms, blue; P atoms, salmon; Mg^{2+} ion, green. 2D structures of five drugs are presented. **c** The schematic diagram for the domain organization of the RTC, containing nsp12, nsp7, nsp8-1, nsp8-2, nsp9, nsp13-1 and nsp13-2 (PDB: 7CYQ). Two views of the cartoon model of the cryo-EM structure of RTC. The subdomains and components of the holo-RdRp are colored as follows: the subdomains and components of the RTC are colored as follows: β-hairpin, chocolate; NiRAN, light orange; Interface, light magenta; Fingers, bright orange; Palm, lime; Thumb, pale cyan; nsp7, marine; nsp8-1, pink; nsp8-2, fire brick; nsp9, purple blue; nsp13-1, wheat; nsp13-2, dark salmon; template RNA, red; product RNA, deep-teal.

Structure genomics of SARS-CoV-2: drug design for COVID-19
CR Wu et al.

3025

presented a cryo-EM structure of the SARS-CoV-2 RTC that includes an nsp9 [28]. The structure reveals that nsp9 binds tightly to the catalytic center of nsp12 NiRAN and inhibits the GTPase activity of nsp12 NiRAN (Fig. 2c). How the complete RTC executes a series of complicated and fine processes in genome replication, such as the mRNA $^{7Me}GpppA_{2'OMe}$ cap process and the mismatch repair process, remains to be addressed by continuous and in-depth structural studies.

THE STRUCTURES OF S PROTEIN

The SARS-CoV-2 S protein belongs to the type I viral fusion protein and mainly comprises two functional subunits, S1 and S2 [76]. The S1 subunit is mainly responsible for receptor-binding, and the S2 subunit mediates membrane fusion following proteolytic activation [77–79]. The S1 subunit can be further subdivided into the N-terminal domain (S1-NTD), two subdomains (SD1 and SD2) and C-terminal domain (S1-CTD), which is also known as the receptor-binding domain (RBD) [79] (Fig. 3a). The SARS-CoV-2 S RBD is responsible for recognizing host receptor ACE2 and contains two subdomains, a receptor-binding motif (RBM) and a core structure (Fig. 3a). The S2 subunit consists of multi-structural components, including hydrophobic fusion peptide (FP), heptad repeats (HR1 and HR2), transmembrane domain (TM). Unlike SARS-CoV and other beta-CoVs in subtype B, SARS-CoV-2 S protein contains a furin cleavage site between S1 and S2 subunits [80] (Fig. 3a). In host cells, S protein is proteolytically cleaved at the PRRAR motif (residues 681–685) by furin into S1 and S2 to form a S1/S2 heterodimer, which was further assembled into the final trimeric spike complex [76, 81]. The S2 subunit contains an S' cleavage site, which is cleaved by host proteases like transmembrane protease serine 2 (TMPRSS2) to expose the fusion peptide for membrane fusion [79]. To date, numerous SARS-CoV-2 S protein structures in both prefusion and post-fusion forms have been reported [77, 78, 80, 82–84]. In the prefusion state, the SARS-CoV-2 trimeric S protein primarily adopts two conformations, a close form with all RBD domains in "down" conformation and an open form with one or multiple RBDs in the "up" conformation (Fig. 3b, c). In a closed form, the receptor-binding site was enshrouded by S1-NTD, thereby impeding the binding of the receptor ACE2, whereas, in the open conformation, one or multiple RBDs expose the receptor-binding site for ACE2 binding [83]. Ke et al. determined the S protein structure of intact SARS-CoV-2 virion by cryo-EM and cryo-ET and found that among all prefusion state S protein, fully closed trimers account for 31%, one RBD opened or two RBDs opened trimers account for 55% and 14%, respectively [83]. ACE2 binding destabilizes the prefusion structure and facilitates the S1 subunit dissociates, thereby enhancing the prefusion to the post-fusion transition of the S protein [85]. Fan et al. solved the cryo-EM structure of the post-fusion state of SARS-CoV-2 S and showed that S2 forms a tightly bound six-helix bundle by rotating HR1-HR2 and the linker region at the upstream of the HR2 motif [86] (Fig. 3d). The mechanism of S protein-mediated coronavirus attachment and fusion has been described in detail in previous reviews [87–90].

SARS-CoV-2 displays higher transmissibility than other human coronaviruses, mainly attributing to the following factors: a, higher receptor-binding affinity; b, more susceptibility to protease; c, existing more alternative host cell receptors. Walls et al. found that SARS-CoV-2 S is bound to ACE2 with significantly higher affinity compared to SARS-CoV S [76]. Sequence alignment showed that four of five key residues at the interface between SARS-CoV-S RBD and host ACE2 are non-conserved, including $Y422_{SARS-CoV}/L455_{SARS-CoV-2}$, $L472_{SARS-CoV}/F486_{SARS-CoV-2}$, $N479_{SARS-CoV}/Q493_{SARS-CoV-2}$, $T487_{SARS-CoV}/N501_{SARS-CoV-2}$ and $Y491_{SARS-CoV}/Y505_{SARS-CoV-2}$ [84, 90–92]. Shang et al. demonstrated that the N501 and Q493 contribute to the enhanced ACE2-binding affinity of SARS-CoV-2 relative to SAR-CoV by forming additional hydrogen bonds with RBM main chain, as well

as K31 and E35 side chains, respectively [77, 78]. Intriguingly, the B.1.1.7 variant with N501Y substitution, which was originated in the UK in September of 2020, displays higher transmissibility and rapid spread globally, which may be caused by the enhanced ACE2 affinity to the S variant of N501Y compared with the wild-type S protein [93–95].

As mentioned above, the SARS-CoV-2 S protein is proteolytically cleaved by furin during posttranslational maturation at ER to Golgi steps, leading to significantly decreased dependency of membrane protease. The non-covalently linked S protein is more flexible, which facilitates conformation change to the open state for receptor-binding and exposes the S2' cleavage site to protease like TMPRSS2 [81, 96]. Cleavage of the S protein at the S2' site by cell surface protease facilitates membrane fusion for viral cell entry. Mutations promoting the cleavage efficiency on S1/S2 or S2' site lead to higher transmissibility. By early April of 2020, a SARS-CoV-2 variant containing a spike D614G mutation is spread rapidly and become the dominant form in the pandemic. Korber et al. found that the D614G variant is closely related to greater infectivity and increased viral load [97]. Yan et al. reported the cryo-EM structure of the D614G variant S protein and found that this mutation makes S protein more flexible and is susceptible to protease cleavage [83]. Another study demonstrated that the D614G mutation disrupted a salt bridge between D614 and K854, thereby attenuating the interaction between S1 and S2 subunits [84]. In the spring of 2021, the B.1.617 lineage (Delta) containing a P681R mutation in S protein emerged from India and became the dominant strain globally [98]. Saito et al. suggested that P681R mutation in the spike protein facilitates spike protein cleavage and enhances viral fusogenicity [99].

In addition to higher affinity to ACE2 and more accessible to be digested by host proteases, S protein binds to several other host receptors or co-receptors to facilitate the entry of SARS-CoV-2 into host cells. The N-terminal domain of S protein specifically interacts with tyrosine-protein kinase receptor UFO (AXL) in the human cell surface [100]. It was reported that other proteins like neuropilin-1 and asialoglycoprotein receptors also facilitate SARS-CoV-2 cell entry and infectivity [101, 102].

The S protein determines the viral infection efficiency of human cells and is pursued as a prominent anti-virus drug target. Due to the large interaction interface between S protein and ACE2, it is hard to block their interaction by small-molecule drugs. SARS-CoV-2 S neutralizing antibodies (nAbs) may be the preferred therapeutics. Recently, a number of broad and potent SARS-CoV-2 nAbs have been identified, including nAbs which were also identified from convalescent COVID-19 patients and nanobodies screened in the surface-displayed library. A number of structures of the SARS-CoV-2 S protein or RBD in complexed with antibody have been determined by cryo-EM and X-ray crystallography [83, 95, 103–110]. Most nAbs bind to the RBD, which is in the "up" conformation resembling the conformation when interacting with ACE2. Conversely, only a few nAbs can bind to the "down" conformation and thus impede the conformational switching required for viral entry [105, 111] (Fig. 3f). Besides anti-RBD antibodies, an antibody named 4A8 responds to S1-NTD, which has also been identified from convalescent COVID-19 patients [112](Fig. 3f). Recently, two reviews have conducted a comprehensive analysis on epitopes on S recognized by various antibodies [113, 114]. Currently, more than 50 nAbs-related clinical trials have been conducted for patients with COVID-19. Among them, eight RBD-specific nAbs have been authorized by the Food and Drug Administration (FDA) for emergency use [115].

Unfortunately, the continuously evolving SARS-CoV-2 has led to vaccine and nAbs resistance. Several variants of SARS-CoV-2 with certain mutations have been designated as a variant of concern (VOC) by the World Health Organization (WHO) [98]. On 26 November 2021, the WHO designated a new variant B1.1.1.529 (Omicron) as a VOC only 2 days after it was reported [98]. The
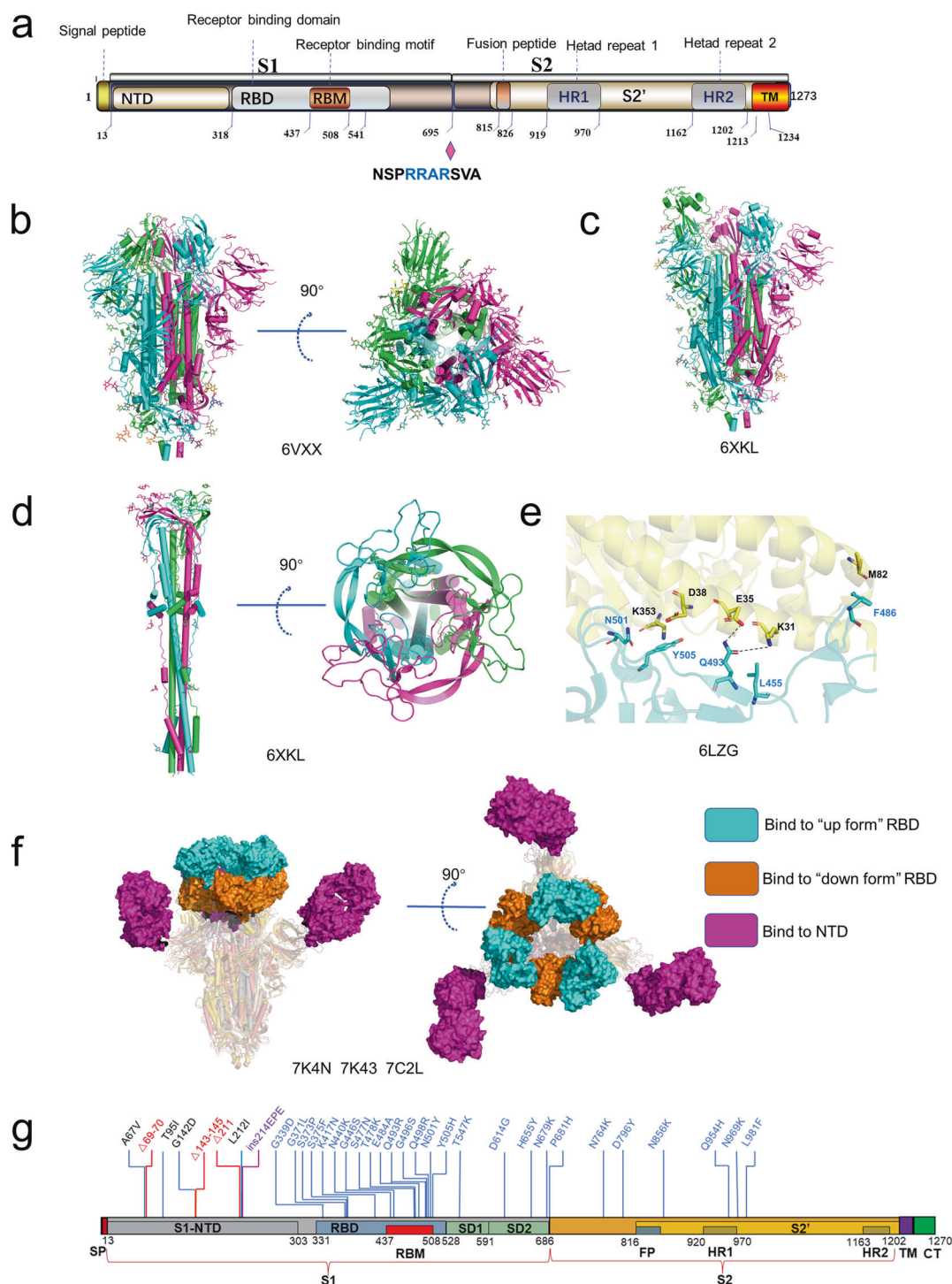
Structure genomics of SARS-CoV-2: drug design for COVID-19
CR Wu et al.

3026

Fig. 3 The structures of S protein. a Schematic of SARS-CoV-2 S protein domain architecture. The S1 and S2 subunits are indicated, with diamond representing the locations of furin cleavage site. SP signal peptide, RBD receptor-binding domain, RBM receptor-binding motif, SD1 subdomain 1, SD2 subdomain 2, FP fusion peptide, HR1 heptad repeat 1, HR2 heptad repeat 2, TM transmembrane region, CT cytoplasmic tail. **b** Side and top views of the prefusion structure of the SARS-CoV-2 S protein with all RBD in the "down" conformation (PDB: 6VXX). Protein is shown in cartoon representation, and glycosyls are shown as sticks. **c** Side view of the prefusion structure of the SARS-CoV-2 S protein with one RBD in the "up" conformation (PDB: 6XKL). **d** Side and top views of the post-fusion structure of the SARS-CoV-2 S protein (PDB: 6M3W). **e** Structure of the SARS-CoV-2 RBD complexed with ACE2 (PDB: 6LZG). ACE2 is shown in yellow. The RBD is shown in cyan. Key contacting residues are shown as sticks at the SARS-CoV-2 RBD–ACE2 interfaces. **f** Side and top views of superimposed three cryo-EM structures of SARS-CoV-2 S in complex with nAbs. S2E12 (represented as a cyan surface) binds to the "up" conformation of SARS-CoV-2 S RBD (PDB: 7K4N); S2M11 (represented as a brown surface) binds to the "down" conformation of SARS-CoV-2 S RBD (PDB: 7K43); 4A8 (represented as a magento surface) binds to the NTD of SARS-CoV-2 S (PDB: 7C2L). **g** Amino acids mutated in the Omicron variant S protein.
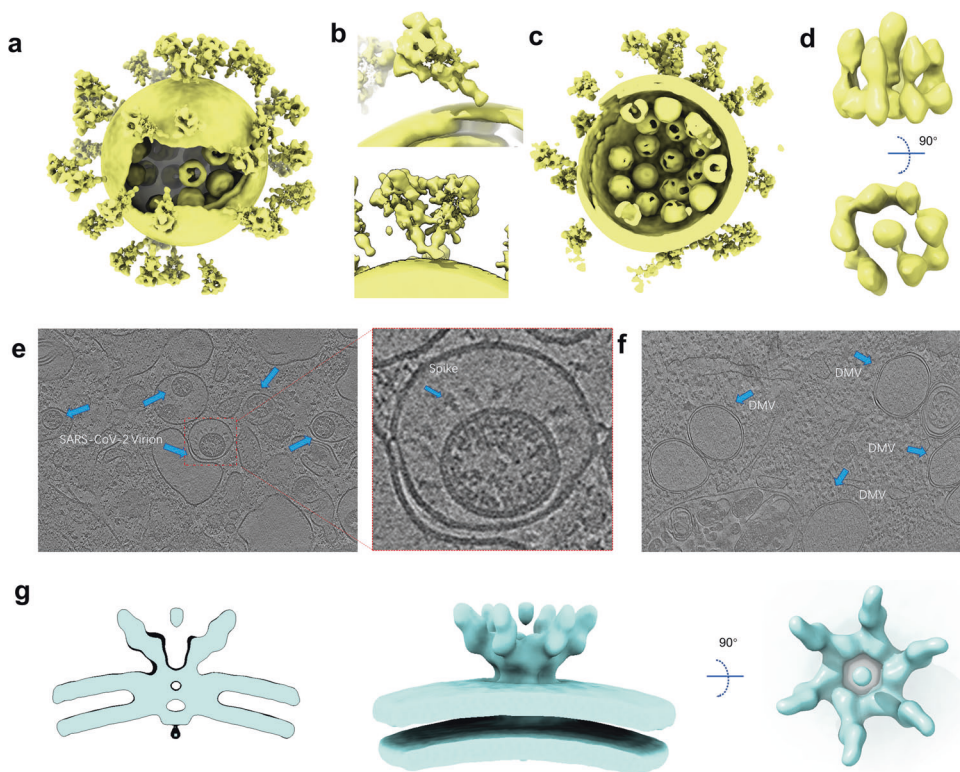
Structure genomics of SARS-CoV-2: drug design for COVID-19
CR Wu et al.

3027

Fig. 4 In situ structures of SARS-CoV-2 virions. a Cryo-EM map of SARS-CoV-2 virion structure. The S proteins with different conformations are distributed over the virion surface and can be tilted to different directions (EMD-30430). **b** A tilted conformation of S trimer is presented in the upper plots, while a Y-shaped spike pairs having two heads and one combined stem are presented in the lower plots. **c** Side view of cryo-EM map of SARS-CoV-2 virion structure, displaying RNPs assembled in the virus envelope. **d** Side and top views of in situ structure of SARS-CoV-2 RNP (EMD-30429). **e** Tomograms showing SARS-CoV-2 virions in VeroE6 cells (EMD-11865). **f** Tomograms showing DMVs in SARS-CoV-2 affected VeroE6 cells (EMD-11866). **g** Different views of the EM structure of the MHV-induced pore complex embedded in the DMVs membranes (EMD-11514).

Omicron variant accumulated a total of 60 mutations compared to the original SARS-CoV-2 variant [98]. There are 15 mutations on the RBD of S protein, which has more than 30 mutations in total (Fig. 3g). Recently, Xie et al. determined the RBD escaping mutation profiles for a total of 247 anti-RBD NAbs by yeast display screening [116]. Among all these tested nAbs, more than 85% are escaped by Omicron. For instance, nAbs whose epitope included in the RBM are escaped mainly by six mutations (K417N, N440K, G446S, E484A, Q493K, and G496S). Combination therapy may greatly reduce the probability of immune escape of mutant viruses caused by single-antibody treatment. High-resolution structures of nAbs and S protein are of vital importance in revealing their epitope distribution and understanding the neutralization mechanism. The structural information of neutralizing antibodies to different epitopes will be critical to accelerate the development of broadly protective nAb cocktail therapies for COVID-19. Moreover, structure-based design of nAb mutants help to generate novel nAbs with improved potency and efficacy [117].

IN SITU STRUCTURE OF SARS-COV-2 VIRIONS

Coronavirus belongs to the envelope virus. Its shape is relatively variable, and each virus has a unique structure. Therefore, it cannot be solved by the single-particle cryo-EM, which limits the acquisition of high-resolution entire virus structure. With the improvement of the cryo-EM facility and the optimization of the image processing algorithm in recent years, it is possible to obtain the in situ structure with a near-atomic resolution by cryo-ET. In the autumn of 2020, three research groups reported the overall structure of purified SARS-CoV-2 virions, which presents an approximately spherical conformation. Ke et al. measured the

average virion diameter of 91 ± 11 nm [118]. The virions may become less spherical after being concentrated by ultracentrifugation through a sucrose cushion [118]. Yao et al. measured the virion diameters for the short, medium, and the long axis of the envelope are 64.8 ± 11.8, 85.9 ± 9.4, and 96.6 ± 11.8 nm, respectively [119]. The S trimers are randomly distributed on the surface of the envelope (Fig. 4a). In three independent studies, the average number of S trimers resided on the virion surface was calculated to be between 20–40, with the reported statistical results being 40 [120], 24 ± 9 [118] and 26 ± 15 [119], respectively. Surprisingly, the S protein in the virus is not perpendicular to the viral membrane but tilted at different angles, which is significantly different from other enveloped viruses possessing class I fusion proteins (Fig. 4b). A small part of the S trimer even tilted by over 90° toward the membrane [118]. The tilt of S protein is thought to be attributed to the flexibility of the hinge region in proximity to the membrane. Interestingly, besides tilting from the membrane, S trimers can also move freely on the viral envelope. Yao et al. even observed two S trimers are combined together with heads and stalks forming a Y-shaped spike [119] (Fig. 4b). The flexibility of S protein both in direction and location may facilitate the virus sensing and binding to ACE2, allowing one S trimer to bind with two or three ACE2s, or two S proteins with one ACE2 dimer. Moreover, compared with recombinant S proteins, the enriched N-linked glycans present on the native spike are more complicated. These enriched glycosylation shields S from host protease digestion and antibody recognition.

The coronavirus has the largest genome across all RNA viruses but a smaller size (Ø = ~90 nm) compared with other RNA viruses like human immunodeficiency virus (Ø = 120–170 nm) [121], and human respiratory syncytial virus (Ø = 150–250 nm) [122].

Structure genomics of SARS-CoV-2: drug design for COVID-19
CR Wu et al.

3028

Considering that coronaviruses pack their large genomes to form a supercoiled dense structure into a relatively small viral particle [123], it seems that the RNPs of SARS-CoV-2 inside the envelope are more tightly packed relative to other coronaviruses.

The crystal structures of the nuclear acid-binding domain (NTD and CTD) of nucleocapsid protein have recently been solved [124, 125]. However, the lack of the native structure of the ribonucleoprotein complex limits our understanding of the assembly and function of coronavirus RNP. Yao et al. find that each SARS-CoV-2 virion contains an average of more than 26 RNPs, of which the membrane-proximal RNPs assemble as "hexon" and the membrane-free assemble as "tetrahedron" (Fig. 4c). They proposed an RNP assembly model in which the native RNPs interact with the RNA in a "beads on a string" pattern, similar to the mechanism of the chromatin forming in the eukaryotic cell [119]. Through the sub-tomogram averaging method, Yao et al. solved a 13.1 Å resolution structure of RNPs, which revealed their reverse G-shaped architecture, with its diameter 15 nm and height 16 nm (Fig. 4d). This is similar to the 30 Å-resolution RNPs structure in the intracellular virions, which is solved by Klein et al. [126]. This size is different from the reported architecture of SARS-CoV N protein, which is assembled by crystal packing of 24-mer CTD domain [127, 128]. The shape is also distinct from the released MHV RNP structure [129]. Limited by the resolution, it is hard to clarify the precise mechanism of RNP assembly. This structure also cannot provide more information on the interaction between N protein and other structural proteins like M and E [127]. Recently, a 4.3 Å cryo-EM map of full-length SARS-CoV-2 N protein has been released in the Electron Microscopy Data Bank, but the corresponding structural model is still unavailable [130]. This N protein structure of moderate resolution may provide valuable insights into RNP assembly of SARS-CoV-2.

To better understand the life cycle of the SARS-CoV-2 virus in host cells, Steffen et al. used the cryo-ET method to structurally characterize the near-native state of virus assembly, DMVs morphology, and extracellular virions [126] (Fig. 4e). DMVs inside the SARS-CoV-2 infected cells show an average diameter of 338 nm, which approaches the size of DMVs in SARS-CoV-1-infected VeroE6 cells [131]. The inner and outer membranes of DMVs are separated by 5–10 nm but are clamped together in several sites (Fig. 4f). These connected sites may be the proteinaceous pore complex formed by NSPs. This proteinaceous pore complex is analogous to the hexameric assembly crown-shaped complex, which is disclosed in murine hepatitis coronavirus-induced DMVs in infected cells [132]. This DMVs pore complex is a well-developed pathway for the transport of coronaviral RNA products out of the DMVs [132] (Fig. 4g). However, the exact component and the function of this molecular pore remain to be elucidated by high-resolution structures.

SUMMARY AND PERSPECTIVES

As a result of the rapid development of structural biology techniques and the efficient investment of global researchers over recent years, remarkable progress has been made on structural studies of SARS-CoV-2. In addition to the RTC and S protein, the structures of most SARS-CoV-2 proteins have been resolved in the past 2 years. These proteins include (1) NSPs M^{pro} [13], PL^{pro} [22, 133], nsp1 [134, 135], nsp9 [136], nsp10-nsp14 complex [31], nsp10-nsp16 complex [33] and nsp15 [32], (2) structural protein N [125, 137] and E proteins [138], (3) and accessory proteins such as orf8 [43] (Fig. 5). Recently, the structures of ORF3a [41], ORF7a [139], ORF9b [45], and Nsp2 [21] have also been reported (Fig. 5).

A better understanding of these viral protein structures has accelerated the structure-based development of novel drugs, including anti-influenza drugs Zanamivir [140] and Oseltamivir [141]. Based on the structure of SARS-CoV-2 M^{pro} in complex with inhibitors, Pfizer has developed the second-generation orally SARS-CoV-2 M^{pro} inhibitor, PF-07321332. It showed positive phase III results for the treatment of COVID-19 in combination with Ritonavir, which maintains higher circulating concentrations of PF-07321332 by inhibiting cytochrome enzymes [142].

PF-07321332 shows an overall structural similarity to the anti-HCV drug Boceprevir, which harbored a rigid P2 dimethylcyclo-propylproline residue and a hydrophobic P3 residue (Fig. 6a) [143]. The P2 residue fits well to the S2 subsite and hydrophobically interacts with Met149 and Asp187, while the P3 residue interacts with Met165 in S3 subsite (Fig. 6b). The P1 c-Bua residue of Boceprevir does not directly contact the SARS-CoV-2 M^{pro} S1 subsite, thus limiting its affinity [144]. The P1 residue of PF-07321332 is a glutamine surrogate that mimics the equivalent residue of GC376. As a broad-spectrum cysteine protease covalent inhibitor, GC376 has shown promise in treating cats with fatal feline infectious peritonitis (FIP) caused by FIPV and is being investigated as a treatment for COVID-19 [144]. The glutamine surrogate ring of both GC376 and PF-07321332 fits into the S1 pocket and forms three stable hydrogen bonds with Phe140, His163, and Glu166 (Fig. 6c). Unlike other M^{pro} inhibitors, PF-07321332 contains a nitrile warhead in P1' residue, forming a reversible covalent thioimidate adduct with the catalytic Cys145. A trifluoroacetyl capping group was applied to the P4 residue to improve the oral delivery efficiency. As shown in the SARS-CoV-2 M^{pro}-PF-07321332 complex structure, the trifluoromethyl group forms hydrogen bonds with Gln192 and two ordered water molecules, providing stronger interactions to the S4 subsite relative to Boceprevir (Fig. 6b). Collectively, through the structure-based approach, the affinity and physicochemical properties were optimized from initial leads in a short time. On November 16, an application was submitted to the U.S. FDA for PF-07321332 in combination with Ritonavir (brand name: Paxlovid) for emergency treatment of COVID-19. The advent of this promising new agent has offered much hope to people worldwide. In addition to SARS-CoV-2 M^{pro}, RdRp and spike protein, several other promising therapeutic targets have been identified in the laboratory, such as PL^{pro} and helicase; structure-based antiviral drugs development targeting these proteins are ongoing [145, 146].

So far, there are only a few SARS-CoV-2 proteins whose structures have not been characterized, and most of them are transmembrane proteins (Table 1). These studies shed light on the biological functions of SARS-CoV-2 proteins and offer new opportunities to develop vaccines and drugs. Nevertheless, there are still some important issues waiting for elucidation from a structural perspective.

At present, the largest available RTC structure only contains nsp7-8-9-12-13 [28]. Although the structures of other potential components such as nsp10, nsp14, nsp15 and nsp16 have been solved, it remains unclear how these subunits participate in the assembly of the complete RTC. As we know, coronavirus contains the largest single-stranded RNA genome among all RNA viruses. How the complete RTC executes a series of complicated and fine processes in genome replication, including the mRNA ^{7MeG}pppA$_{2'}$OMe cap process and the mismatch repair process, remains to be revealed by the structural study. For biosafety reasons, the current structural studies of RTC are carried out through in vitro recombinant expression and assembly, which hampers the acquisition of stable RTC. Overcoming the biosafety issues of SARS-CoV-2, or alternatively to culture low pathogenicity corona-viruses like MHV to isolate and purify natural RTC may be a feasible strategy to solve the complete RTC structure.

Although the high-resolution crystal structures of NTD and CTD of N protein have been solved, and a low resolution full-length N protein structure has been available recently [124, 125, 130], the structural basis for the assembly of RNPs has not been clarified. It has been proposed that the assembly of RNPs enables efficient genome packaging in a nucleosome-like manner [126].
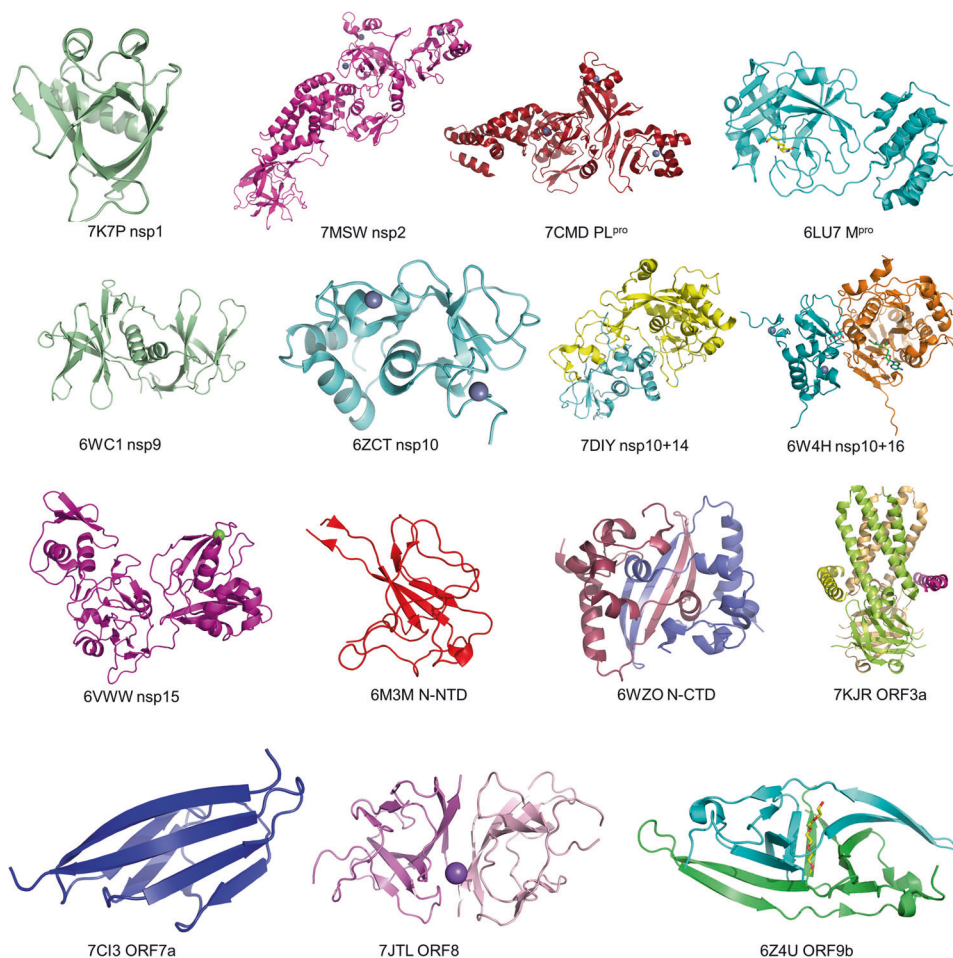
Structure genomics of SARS-CoV-2: drug design for COVID-19
CR Wu et al.

3029

Fig. 5 High-resolution structures of SARS-CoV-2 proteins. Crystal structure of SARS-CoV-2 nsp1 (PDB: 7K7P); Cryo-EM structure of the SARS-CoV-2 nsp2 (PDB: 7MSW); crystal structure of the SARS-CoV-2 PLᵖʳᵒ with GRL0617 (PDB: 7CMD); crystal structure of the SARS-CoV-2 Mᵖʳᵒ in complex with an inhibitor N3 (PDB: 6LU7); crystal structure of the SARS-CoV-2 Nsp9 in complex with a peptide (PDB: 6WC1); crystal structure of the SARS-CoV-2 nsp10 (PDB: 6ZCT); crystal structure of SARS-CoV-2 nsp10 bound to nsp14-exoribonuclease domain (PDB: 7DIY); crystal structure of the SARS-CoV-2 nsp10 in complex with nsp16 (PDB: 6W4H); crystal structure of the SARS-CoV-2 nsp15 endoribonuclease (PDB: 6VWW); crystal structure of SARS-CoV-2 nucleocapsid protein N-terminal RNA binding domain (PDB: 6M3M); crystal structure of SARS-CoV-2 nucleocapsid protein C-terminal RNA binding domain (PDB: 6WZO); Cryo-EM structure of SARS-CoV-2 ORF3a (PDB: 7KJR); the crystal structure of the SARS-CoV-2 ORF7a ectodomain (PDB: 7CI3); the crystal structure of SARS-CoV-2 ORF8 accessory protein (PDB: 7JTL); the crystal structure of SARS-CoV-2 ORF9b accessory protein (PDB: 6Z4U).

High-resolution of assembled RNP structure is needed to explain how many N proteins assembled in the natural state and how the RNA wounds around them. N protein plays a very important role in the process of virus assembly. Relevant studies have proven that N protein can interact not only with RNA but also with M and E proteins [147]. As described in the cryo-ET virion structure, the RNP is close to the inner membrane of the virus capsule, indicating that it is feasible to interact with membrane proteins like M and E proteins in terms of space proximity [119]. Elucidating the interaction mode of N protein with M protein and E protein would help understand the viral assembly process and provide a new opportunity for designing antiviral drugs that interfere with viral assembly.

Structurally, N proteins consist of two ordered domains: the N-terminal domain (NTD), and C-terminal domain (CTD), connected by flexible LKR linker regions and two disordered regions, N-arm and C-tail. These disordered regions of the N protein make it extremely difficult for high-resolution structure determination. In the SARS-CoV-2 virions, the N protein may be stabilized by interacting with several molecules, such as viral genome RNA, M protein, E protein and S protein, with an unknown assembly ratio. Besides, the N protein binds nonspecifically with the RNA phosphate backbone,

making it difficult for in vitro assembly of the recombinant N protein complex. Alternatively, virus-like particles (VLPs), which can be spontaneously assembled by co-expressing these structural proteins, could be a potential resource of the N-M-E protein complex samples, since VLPs are viral genome-free and are considered biologically safe [147]. Nevertheless, obtaining the homogeneous, soluble N-M-E protein complexes required for detailed structural studies is still a huge challenge, because detergents can hamper the assembly of such membrane protein complexes. Extensive efforts to improve the stability of these complexes, such as developing specific Fabs or nanobodies, should be made for high-resolution structure determination.

The mechanism of DMVs formation in the life process of coronavirus has not been clearly explained. Previous studies have shown that nsp3, nsp4 are mainly involved in the DMVs formation process [148]. Interestingly, a hexameric assembly pore complex that spans both membranes of the MHV-induced DMVs was visualized by the cryo-ET method [132]. Nsp3 was identified as a major constituent of the pore complex, and nsp3, nsp4, nsp6 and some other unknown proteins may also be involved in this process. It is speculated that the pore complex can export synthesized RNA to the cytosol. In a sense, DMVs are somewhat similar to the nucleus
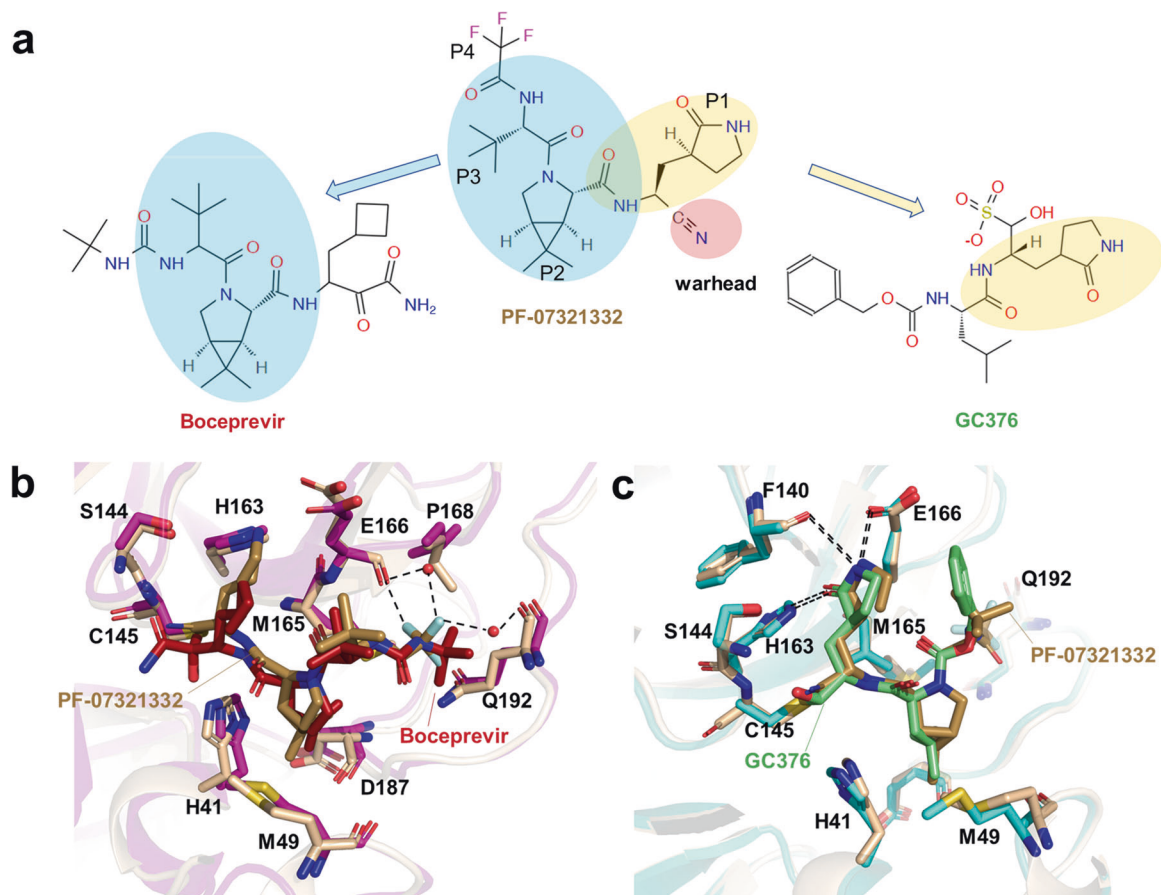
Structure genomics of SARS-CoV-2: drug design for COVID-19
CR Wu et al.

3030

Fig. 6 The structures of SARS-CoV-2 M^pro in complex with inhibitors. a Chemical structure of PF-07321332, Boceprevir and GC376. **b** The comparison of SARS-CoV-2 M^pro-PF-07321332 complex structure with that of SARS-CoV-2 M^pro-boceprevir. **c** The comparison of SARS-CoV-2 M^pro-PF-07321332 complex structure with that of SARS-CoV-2 M^pro-GC376.

Table 1. A summary of SARS-CoV-2 protein structures that have not been resolved.

Proteins name	Sequence range	Transmembrane protein or not
Nsp3	379–742, 1410–1842	Yes
Nsp4	1–500	Yes
Nsp6	1–290	Yes
S	1209–1273	Yes
M	1–222	Yes
N	1–43, 181–254, 364–419	No
ORF3b	1–22	No
ORF6	1–61	No
ORF7a	83–121	Yes
ORF7b	1–43	Yes
ORF9c	1–73	Yes
ORF10	1–38	No

of eukaryotic cells. As an independent platform of mRNA and genome replication, it is continuously exported to the cytoplasm. Similarly, the novel pore complex is analogous to a nuclear pore and also spans double membranes. The pore complex may finely control the inlet and outlet of raw materials and products in the DMVs. The structure of the pore complex is important for understanding the mechanism of DMVs formation and material transport. Moreover, it can also provide a structural basis for the design of antiviral drugs through intervening viral replication.

With the advances in recent cryo-EM methods, it is practicable to solve the structure of protein-free RNA. Recently, Wah Chiu group and Rhiju Das group developed a set of methods for RNA structure study by cryo-EM [149]. These two groups solved several near-atomic resolution RNA structures [150, 151], including a 3.1 Å cryo-EM structure of the full-length Tetrahymena ribozyme. To date, 3D structures of key RNA elements in the SARS-CoV-2 genome remain poorly characterized. Only the structure of FSE was solved [20]. There remain dozens of highly conserved RNA segments in the SARS-CoV-2 genome. Targeting these segments by antisense RNA is a good strategy for therapeutic agent development [152–154]. Applying the cryo-EM method to study these RNA structures is of great significance for the elucidation of SARS-CoV-2 RNA biology, hopefully, which would facilitate the development of genome-disrupting antiviral agents.

ADDITIONAL INFORMATION
Competing interests: The authors declare no competing interests.

Structure genomics of SARS-CoV-2: drug design for COVID-19
CR Wu et al.

3031

REFERENCES

1. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet. 2020;395:507–13.
2. Chan JF, Yuan S, Kok KH, To KK, Chu H, Yang J, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet. 2020;395:514–23.
3. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel Coronavirus-infected pneumonia. N Engl J Med. 2020;382:1199–207.
4. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis. 2020;20:533–4.
5. Johns Hopkins Coronavirus Resource Center. COVID-19 dashboard. JHU. 2021. https://coronavirus.jhu.edu/map.html.
6. WHO | Middle East respiratory syndrome coronavirus (MERS-CoV). 2021. https://www.who.int/emergencies/mers-cov/en/.
7. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J.China Novel Coronavirus Investigating and Research Team, et al. A novel Coronavirus from patients with pneumonia in China, 2019. N Engl J Med. 2020;382:727–33.
8. Thomas S. Mapping the nonstructural transmembrane proteins of severe acute respiratory syndrome Coronavirus 2. J Comput Biol. 2021;28:909–21.
9. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. Nature. 2020;579:265–9.
10. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020;579:270–3.
11. Yoshimoto FK. The proteins of severe acute respiratory syndrome coronavirus-2 (SARS CoV-2 or n-COV19), the cause of COVID-19. Protein J. 2020;39:198–216.
12. Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, White KM, et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature. 2020;583:459–68.
13. Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, et al. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. Nature. 2020;582:289–93.
14. Poran A, Harjanto D, Malloy M, Arieta CM, Rothenberg DA, Lenkala D, et al. Sequence-based prediction of SARS-CoV-2 vaccine targets using a mass spectrometry-based bioinformatics predictor identifies immunogenic T cell epitopes. Genome Med. 2020;12:70.
15. Lowery SA, Sariol A, Perlman S. Innate immune and inflammatory responses to SARS-CoV-2: implications for COVID-19. Cell Host Microbe. 2021;29:1052–62.
16. Jungreis I, Sealfon R, Kellis M. SARS-CoV-2 gene content and COVID-19 mutation impact by comparing 44 Sarbecovirus genomes. Nat Commun. 2021;12:2642.
17. Chiara M, D'Erchia AM, Gissi C, Manzari C, Parisi A, Resta N, et al. Next generation sequencing of SARS-CoV-2 genomes: challenges, applications and opportunities. Brief Bioinform. 2021;22:616–30.
18. Sola I, Almazán F, Zúñiga S, Enjuanes L. Continuous and discontinuous RNA synthesis in coronaviruses. Annu Rev Virol. 2015;2:265–88.
19. Bhatt PR, Scaiola A, Loughran G, Leibundgut M, Kratzel A, Meurs R, et al. Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome. Science. 2021;372:1306–13.
20. Zhang K, Zheludev IN, Hagey RJ, Haslecker R, Hou YJ, Kretsch R, et al. Cryo-EM and antisense targeting of the 28-kDa frameshift stimulation element from the SARS-CoV-2 RNA genome. Nat Struct Mol Biol. 2021;28:747–54.
21. Gupta M, Azumaya CM, Moritz M, Pourmal S, Diallo A, Merz GE, et al. CryoEM and AI reveal a structure of SARS-CoV-2 Nsp2, a multifunctional protein involved in key host processes. bioRxiv [Preprint]. 2021. Available from: https://doi.org/10.1101/2021.05.10.443524.
22. Shin D, Mukherjee R, Grewe D, Bojkova D, Baek K, Bhattacharya A, et al. Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. Nature. 2020;587:657–62.
23. Dai W, Zhang B, Jiang XM, Su H, Li J, Zhao Y, et al. Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. Science. 2020;368:1331–5.
24. Yang H, Xie W, Xue X, Yang K, Ma J, Liang W, et al. Design of wide-spectrum inhibitors targeting coronavirus main proteases. PLoS Biol. 2005;3:e324.
25. Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. Science. 2003;300:1763–7.
26. Lei J, Kusov Y, Hilgenfeld R. Nsp3 of coronaviruses: structures and functions of a large multi-domain protein. Antivir Res. 2018;149:58–74.
27. Oostra M, Hagemeijer MC, van Gent M, Bekker CP, te Lintelo EG, Rottier PJ, et al. Topology and membrane anchoring of the coronavirus replication complex: not all hydrophobic domains of nsp3 and nsp6 are membrane spanning. J Virol. 2008;82:12392–405.
28. Yan L, Ge J, Zheng L, Zhang Y, Gao Y, Wang T, et al. Cryo-EM structure of an extended SARS-CoV-2 replication and transcription complex reveals an intermediate state in cap synthesis. Cell. 2021;184:184–93.
29. Jiang Y, Yin W, Xu HE. RNA-dependent RNA polymerase: structure, mechanism, and drug discovery for COVID-19. Biochem Biophys Res Commun. 2021;538:47–53.
30. Yin W, Luan X, Li Z, Zhou Z, Wang Q, Gao M, et al. Structural basis for inhibition of the SARS-CoV-2 RNA polymerase by suramin. Nat Struct Mol Biol. 2021;28:319–25.
31. Lin S, Chen H, Chen Z, Yang F, Ye F, Zheng Y, et al. Crystal structure of SARS-CoV-2 nsp10 bound to nsp14-ExoN domain reveals an exoribonuclease with both structural and functional integrity. Nucleic Acids Res. 2021;49:5382–92.
32. Kim Y, Jedrzejczak R, Maltseva NI, Wilamowski M, Endres M, Godzik A, et al. Crystal structure of Nsp15 endoribonuclease NendoU from SARS-CoV-2. Protein Sci. 2020;29:1596–605.
33. Rosas-Lemus M, Minasov G, Shuvalova L, Inniss NL, Kiryukhina O, Brunzelle J, et al. High-resolution structures of the SARS-CoV-2 2'-O-methyltransferase reveal strategies for structure-based inhibitor design. Sci Signal. 2020;13:eabe1202.
34. Moreira RA, Guzman HV, Boopathi S, Baker JL, Poma AB. Characterization of structural and energetic differences between conformations of the SARS-CoV-2 spike protein. Materials. 2020;13:5362.
35. Huang Y, Yang C, Xu XF, Xu W, Liu SW. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. Acta Pharmacol Sin. 2020;41:1141–9.
36. Fu YZ, Wang SY, Zheng ZQ, Yi H, Li WW, Xu ZS, et al. SARS-CoV-2 membrane glycoprotein M antagonizes the MAVS-mediated innate antiviral response. Cell Mol Immunol. 2021;18:613–20.
37. Nieto-Torres JL, DeDiego ML, Verdiá-Báguena C, Jimenez-Guardeño JM, Regla-Nava JA, Fernandez-Delgado R, et al. Severe acute respiratory syndrome coronavirus envelope protein ion channel activity promotes virus fitness and pathogenesis. PLoS Pathog. 2014;10:e1004077.
38. Schoeman D, Fielding BC. Coronavirus envelope protein: current knowledge. Virol J. 2019;16:69.
39. Sheikh A, Al-Taher A, Al-Nazawi M, Al-Mubarak AI, Kandeel M. Analysis of preferred codon usage in the coronavirus N genes and their implications for genome evolution and vaccine design. J Virol Methods. 2020;277:113806.
40. Zeng W, Liu G, Ma H, Zhao D, Yang Y, Liu M, et al. Biochemical characterization of SARS-CoV-2 nucleocapsid protein. Biochem Biophys Res Commun. 2020;527:618–23.
41. Kern DM, Sorum B, Mali SS, Hoel CM, Sridharan S, Remis JP, et al. Cryo-EM structure of SARS-CoV-2 ORF3a in lipid nanodiscs. Nat Struct Mol Biol. 2021;28:573–82.
42. McBride R, Fielding BC. The role of severe acute respiratory syndrome (SARS)-coronavirus accessory proteins in virus pathogenesis. Viruses. 2012;4:2902–23.
43. Flower TG, Buffalo CZ, Hooy RM, Allaire M, Ren X, Hurley JH. Structure of SARS-CoV-2 ORF8, a rapidly evolving immune evasion protein. Proc Natl Acad Sci USA. 2021;118:e2021785118.
44. Zhang Y, Chen Y, Li Y, Huang F, Luo B, Yuan Y, et al. The ORF8 protein of SARS-CoV-2 mediates immune evasion through down-regulating MHC-I. Proc Natl Acad Sci USA. 2021;118:e2024202118.
45. Gao X, Zhu K, Qin B, Olieric V, Wang M, Cui S. Crystal structure of SARS-CoV-2 Orf9b in complex with human TOM70 suggests unusual virus-host interactions. Nat Commun. 2021;12:2843.
46. Pancer K, Milewska A, Owczarek K, Dabrowska A, Kowalski M, Łabaj PP, et al. The SARS-CoV-2 ORF10 is not essential in vitro or in vivo in humans. PLoS Pathog. 2020;16:e1008959.
47. Hillen HS, Kokic G, Farnung L, Dienemann C, Tegunov D, Cramer P. Structure of replicating SARS-CoV-2 polymerase. Nature. 2020;584:154–6.
48. Kirchdoerfer RN, Ward AB. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. Nat Commun. 2019;10:2342.
49. Subissi L, Posthuma CC, Collet A, Zevenhoven-Dobbe JC, Gorbalenya AE, Decroly E, et al. One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. Proc Natl Acad Sci USA. 2014;111:E3900–9.
50. Yin W, Mao C, Luan X, Shen DD, Shen Q, Su H, et al. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. Science. 2020;368:1499–504.
51. Wang Q, Wu J, Wang H, Gao Y, Liu Q, Mu A, et al. Structural basis for RNA replication by the SARS-CoV-2 polymerase. Cell. 2020;182:417–28.e13.
52. Gao Y, Yan L, Huang Y, Liu F, Zhao Y, Cao L, et al. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. Science. 2020;368:779–82.
53. Kabinger F, Stiller C, Schmitzová J, Dienemann C, Kokic G, Hillen HS, et al. Mechanism of molnupiravir-induced SARS-CoV-2 mutagenesis. Nat Struct Mol Biol. 2021;28:740–6.

Structure genomics of SARS-CoV-2: drug design for COVID-19
CR Wu et al.

3032

54. Hashemian SM, Farhadi T, Velayati AA. A review on favipiravir: the properties, function, and usefulness to treat COVID-19. Expert Rev Anti Infect Ther. 2021;19:1029–37.

55. Rochwerg B, Agarwal A, Zeng L, Leo YS, Appiah JA, Agoritsas T, et al. Remdesivir for severe covid-19: a clinical practice guideline. BMJ. 2020;370:m2924.

56. Spinner CD, Gottlieb RL, Criner GJ, Arribas López JR, Cattelan AM, Soriano Viladomiu A, et al. Effect of Remdesivir vs standard care on clinical status at 11 days in patients with moderate COVID-19: a randomized clinical trial. JAMA. 2020;324:1048–57.

57. Grein J, Ohmagari N, Shin D, Diaz G, Asperges E, Castagna A, et al. Compassionate use of Remdesivir for patients with severe Covid-19. N Engl J Med. 2020;382:2327–36.

58. Wang Y, Zhang D, Du G, Du R, Zhao J, Jin Y, et al. Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial. Lancet. 2020;395:1569–78.

59. Bosaeed M, Mahmoud E, Hussein M, Alharbi A, Alsaedy A, Alothman A, et al. A trial of favipiravir and hydroxychloroquine combination in adults hospitalized with moderate and severe Covid-19: a structured summary of a study protocol for a randomised controlled trial. Trials. 2020;21:904.

60. Gordon CJ, Tchesnokov EP, Woolner E, Perry JK, Feng JY, Porter DP, et al. Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency. J Biol Chem. 2020;295:6785–97.

61. Dangerfield TL, Huang NZ, Johnson KA. Remdesivir is effective in combating COVID-19 because it is a better substrate than ATP for the viral RNA-dependent RNA polymerase. iScience. 2020;23:101849.

62. Bravo JPK, Dangerfield TL, Taylor DW, Johnson KA. Remdesivir is a delayed translocation inhibitor of SARS-CoV-2 replication. Mol Cell. 2021;81:1548–52.

63. Kokic G, Hillen HS, Tegunov D, Dienemann C, Seitz F, Schmitzova J, et al. Mechanism of SARS-CoV-2 polymerase stalling by remdesivir. Nat Commun. 2021;12:279.

64. Xie Y, Yin W, Zhang Y, Shang W, Wang Z, Luan X, et al. Design and development of an oral remdesivir derivative VV116 against SARS-CoV-2. Cell Res. 2021;31:1212–4.

65. Ghasemnejad-Berenji M, Pashapour S. Favipiravir and COVID-19: a simplified summary. Drug Res. 2021;71:166–70.

66. Shannon A, Selisko B, Le NT, Huchting J, Touret F, Piorkowski G, et al. Rapid incorporation of Favipiravir by the fast and permissive viral RNA polymerase complex results in SARS-CoV-2 lethal mutagenesis. Nat Commun. 2020;11:4682.

67. Arias A, Thorne L, Goodfellow I. Favipiravir elicits antiviral mutagenesis during virus replication in vivo. Elife. 2014;3:e03679.

68. Sangawa H, Komeno T, Nishikawa H, Yoshida A, Takahashi K, Nomura N, et al. Mechanism of action of T-705 ribosyl triphosphate against influenza virus RNA polymerase. Antimicrob Agents Chemother. 2013;57:5202–8.

69. Naydenova K, Muir KW, Wu LF, Zhang Z, Coscia F, Peet MJ, et al. Structure of the SARS-CoV-2 RNA-dependent RNA polymerase in the presence of favipiravir-RTP. Proc Natl Acad Sci USA. 2021;118:e2021946118.

70. Peng Q, Peng R, Yuan B, Wang M, Zhao J, Fu L, et al. Structural basis of SARS-CoV-2 polymerase inhibition by favipiravir. Innovation. 2021;2:100080.

71. Jayk Bernal A, Gomes da Silva MM, Musungaie DB, Kovalchuk E, Gonzalez A, Delos Reyes V, et al. Molnupiravir for oral treatment of Covid-19 in non-hospitalized patients. N Engl J Med. 2021;NEJMoa2116044. https://doi.org/10.1056/NEJMoa2116044.

72. Parums DV. Editorial: Current status of oral antiviral drug treatments for SARS-CoV-2 infection in non-hospitalized patients. Med Sci Monit. 2022;28:e935952.

73. Chen J, Malone B, Llewellyn E, Grasso M, Shelton PMM, Olinares PDB, et al. Structural basis for helicase-polymerase coupling in the SARS-CoV-2 replication-transcription complex. Cell. 2020;182:1560–73.

74. Malone B, Chen J, Wang Q, Llewellyn E, Choi YJ, Olinares PDB, et al. Structural basis for backtracking by the SARS-CoV-2 replication-transcription complex. Proc Natl Acad Sci USA. 2021;118:e2102516118.

75. Kuljić-Kapulica N, Budisin A. Korona-virusi. Srp Arh Celok Lek. 1992;120:215–8.

76. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell. 2020;181:281–92.

77. Wang Q, Zhang Y, Wu L, Niu S, Song C, Zhang Z, et al. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. Cell. 2020;181:894–904.

78. Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H, et al. Structural basis of receptor recognition by SARS-CoV-2. Nature. 2020;581:221–4.

79. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature. 2020;581:215–20.

80. McCallum M, Walls AC, Bowen JE, Corti D, Veesler D. Structure-guided covalent stabilization of coronavirus spike glycoprotein trimers in the closed conformation. Nat Struct Mol Biol. 2020;27:942–9.

81. Wu C, Zheng M, Yang Y, Gu X, Yang K, Li M, et al. Furin: a potential therapeutic target for COVID-19. iScience. 2020;23:101642.

82. Xiao T, Lu J, Zhang J, Johnson RI, McKay LGA, Storm N, et al. A trimeric human angiotensin-converting enzyme 2 as an anti-SARS-CoV-2 agent. Nat Struct Mol Biol. 2021;28:202–9.

83. Yan R, Zhang Y, Li Y, Ye F, Guo Y, Xia L, et al. Structural basis for the different states of the spike protein of SARS-CoV-2 in complex with ACE2. Cell Res. 2021;31:717–9.

84. Benton DJ, Wrobel AG, Xu P, Roustan C, Martin SR, Rosenthal PB, et al. Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. Nature. 2020;588:327–30.

85. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science. 2020;367:1260–3.

86. Fan X, Cao D, Kong L, Zhang X. Cryo-EM analysis of the post-fusion structure of the SARS-CoV spike glycoprotein. Nat Commun. 2020;11:3618.

87. Wang X, Xia S, Zhu Y, Lu L, Jiang S. Pan-coronavirus fusion inhibitors as the hope for today and tomorrow. Protein Cell. 2021;12:84–8.

88. Li F. Structure, function, and evolution of coronavirus spike proteins. Annu Rev Virol. 2016;3:237–61.

89. Sun G, Xue L, He Q, Zhao Y, Xu W, Wang Z. Structural insights into SARS-CoV-2 infection and therapeutics development. Stem Cell Res. 2021;52:102219.

90. Ge J, Zhang S, Zhang L, Wang X. Structural basis of severe acute respiratory syndrome coronavirus 2 infection. Curr Opin HIV Aids. 2021;16:74–81.

91. Ge XY, Li JL, Yang XL, Chmura AA, Zhu G, Epstein JH, et al. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. Nature. 2013;503:535–8.

92. Wu C, Liu Y, Yang Y, Zhang P, Zhong W, Wang Y, et al. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. Acta Pharm Sin B. 2020;10:766–88.

93. Wang P, Casner RG, Nair MS, Wang M, Yu J, Cerutti G, et al. Increased resistance of SARS-CoV-2 variant P.1 to antibody neutralization. Cell Host Microbe. 2021;29:747–51.

94. Gobeil SM, Janowska K, McDowell S, Mansouri K, Parks R, Stalls V, et al. Effect of natural mutations of SARS-CoV-2 on spike structure, conformation, and antigenicity. Science. 2021;373:eabi6226.

95. Wang R, Zhang Q, Ge J, Ren W, Zhang R, Lan J, et al. Analysis of SARS-CoV-2 variant mutations reveals neutralization escape mechanisms and the ability to use ACE2 receptors from additional species. Immunity. 2021;54:1611–21.

96. Wrobel AG, Benton DJ, Xu P, Roustan C, Martin SR, Rosenthal PB, et al. SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. Nat Struct Mol Biol. 2020;27:763–7.

97. Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. Cell. 2020;182:812–27.

98. WHO. Tracking SARS-CoV-2 variants. 2021. https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/.

99. Saito A, Irie T, Suzuki R, Maemura T, Nasser H, Uriu K, et al. Enhanced fusogenicity and pathogenicity of SARS-CoV-2 Delta P681R mutation. Nature. 2021. https://doi.org/10.1038/s41586-021-04266-9.

100. Wang S, Qiu Z, Hou Y, Deng X, Xu W, Zheng T, et al. AXL is a candidate receptor for SARS-CoV-2 that promotes infection of pulmonary and bronchial epithelial cells. Cell Res. 2021;31:126–40.

101. Cantuti-Castelvetri L, Ojha R, Pedro LD, Djannatian M, Franz J, Kuivanen S, et al. Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. Science. 2020;370:856–60.

102. Collins DP, Steer CJ. Binding of the SARS-CoV-2 spike protein to the asialoglycoprotein receptor on human primary hepatocytes and immortalized hepatocyte-like cells by confocal analysis. Hepat Med. 2021;13:37–44.

103. Huo J, Zhao Y, Ren J, Zhou D, Duyvesteyn HME, Ginn HM, Carrique L, et al. Neutralization of SARS-CoV-2 by destruction of the prefusion spike. Cell Host Microbe. 2020;28:445–54.

104. Yuan M, Wu NC, Zhu X, Lee CD, So RTY, Lv H, et al. A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. Science. 2020;368:630–3.

105. Schoof M, Faust B, Saunders RA, Sangwan S, Rezelj V, Hoppe N, et al. An ultrapotent synthetic nanobody neutralizes SARS-CoV-2 by stabilizing inactive Spike. Science. 2020;370:1473–9.

106. Koenig PA, Das H, Liu H, Kümmerer BM, Gohr FN, Jenster LM, et al. Structure-guided multivalent nanobodies block SARS-CoV-2 infection and suppress mutational escape. Science. 2021;371:eabe6230.

107. Barnes CO, West AP Jr, Huey-Tubman KE, Hoffmann MAG, Sharaf NG, Hoffman PR, et al. Structures of human antibodies bound to SARS-CoV-2 spike reveal common epitopes and recurrent features of antibodies. Cell. 2020;182:828–42.

Structure genomics of SARS-CoV-2: drug design for COVID-19
CR Wu et al.

3033

108. Zhou D, Duyvesteyn HME, Chen CP, Huang CG, Chen TH, Shih SR, et al. Structural basis for the neutralization of SARS-CoV-2 by an antibody from a convalescent patient. Nat Struct Mol Biol. 2020;27:950–8.

109. Wang N, Sun Y, Feng R, Wang Y, Guo Y, Zhang L, et al. Structure-based development of human antibody cocktails against SARS-CoV-2. Cell Res. 2021;31:101–3.

110. Asarnow D, Wang B, Lee WH, Hu Y, Huang CW, Faust B, et al. Structural insight into SARS-CoV-2 neutralizing antibodies and modulation of syncytia. Cell. 2021;184:3192–204.

111. Tortorici MA, Beltramello M, Lempp FA, Pinto D, Dang HV, Rosen LE, et al. Ultrapotent human antibodies protect against SARS-CoV-2 challenge via multiple mechanisms. Science. 2020;370:950–7.

112. Chi X, Yan R, Zhang J, Zhang G, Zhang Y, Hao M, et al. A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. Science. 2020;369:650–5.

113. Niu L, Wittrock KN, Clabaugh GC, Srivastava V, Cho MW. A structural landscape of neutralizing antibodies against SARS-CoV-2 receptor binding domain. Front Immunol. 2021;12:647934.

114. Yuan M, Liu H, Wu NC, Wilson IA. Recognition of the SARS-CoV-2 receptor binding domain by neutralizing antibodies. Biochem Biophys Res Commun. 2021;538:192–203.

115. Raybould MIJ, Kovaltsuk A, Marks C, Deane CM. CoV-AbDab: the coronavirus antibody database. Bioinformatics. 2021;37:734–5.

116. Cao Y, Wang J, Jian F, Xiao T, Song W, Yisimayi A, et al. Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. Nature. 2021. https://doi.org/10.1038/s41586-021-04385-3.

117. Li T, Cai H, Yao H, Zhou B, Zhang N, van Vlissingen MF, et al. A synthetic nanobody targeting RBD protects hamsters from SARS-CoV-2 infection. Nat Commun. 2021;12:4635.

118. Ke Z, Oton J, Qu K, Cortese M, Zila V, McKeane L, et al. Structures and distributions of SARS-CoV-2 spike proteins on intact virions. Nature. 2020;588:498–502.

119. Yao H, Song Y, Chen Y, Wu N, Xu J, Sun C, et al. Molecular architecture of the SARS-CoV-2 virus. Cell. 2020;183:730–8.

120. Turoňová B, Sikora M, Schürmann C, Hagen WJH, Welsch S, Blanc FEC, et al. In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. Science. 2020;370:203–8.

121. Briggs JA, Wilk T, Welker R, Kräusslich HG, Fuller SD. Structural organization of authentic, mature HIV-1 virions and cores. EMBO J. 2003;22:1707–15.

122. Mejias A, Ramilo O. New options in the treatment of respiratory syncytial virus disease. J Infect. 2015;71:S80–7.

123. Bárcena M, Oostergetel GT, Bartelink W, Faas FG, Verkleij A, Rottier PJ, et al. Cryo-electron tomography of mouse hepatitis virus: insights into the structure of the coronavirion. Proc Natl Acad Sci USA. 2009;106:582–7.

124. Yang M, He S, Chen X, Huang Z, Zhou Z, Zhou Z, et al. Structural insight into the SARS-CoV-2 nucleocapsid protein C-terminal domain reveals a novel recognition mechanism for viral transcriptional regulatory sequences. Front Chem. 2021;8:624765.

125. Kang S, Yang M, Hong Z, Zhang L, Huang Z, Chen X, et al. Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites. Acta Pharm Sin B. 2020;10:1228–38.

126. Klein S, Cortese M, Winter SL, Wachsmuth-Melm M, Neufeldt CJ, Cerikan B, et al. SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. Nat Commun. 2020;11:5885.

127. Chang CK, Hou MH, Chang CF, Hsiao CD, Huang TH. The SARS coronavirus nucleocapsid protein–forms and functions. Antivir Res. 2014;103:39–50.

128. Chen CY, Chang CK, Chang YW, Sue SC, Bai HI, Riang L, et al. Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA. J Mol Biol. 2007;368:1075–86.

129. Gui M, Liu X, Guo D, Zhang Z, Yin CC, Chen Y, et al. Electron microscopy studies of the coronavirus ribonucleoprotein complex. Protein Cell. 2017;8:219–24.

130. Casasanta MA, Jonaid GM, Kaylor L, Luqiu WY, Solares MJ, Schroen ML, et al. Microchip-based structure determination of low-molecular weight proteins using cryo-electron microscopy. Nanoscale. 2021;13:7285–93.

131. Knoops K, Kikkert M, Worm SH, Zevenhoven-Dobbe JC, van der Meer Y, Koster AJ, et al. SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. PLoS Biol. 2008;6:e226.

132. Wolff G, Limpens RWAL, Zevenhoven-Dobbe JC, Laugks U, Zheng S, de Jong AWM, et al. A molecular pore spans the double membrane of the coronavirus replication organelle. Science. 2020;369:1395–8.

133. Fu Z, Huang B, Tang J, Liu S, Liu M, Ye Y, et al. The complex structure of GRL0617 and SARS-CoV-2 PLpro reveals a hot spot for antiviral drug discovery. Nat Commun. 2021;12:488.

134. Thoms M, Buschauer R, Ameismeier M, Koepke L, Denk T, Hirschenberger M, et al. Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. Science. 2020;369:1249–55.

135. Clark LK, Green TJ, Petit CM. Structure of nonstructural protein 1 from SARS-CoV-2. J Virol. 2021;95:e02019–20.

136. Littler DR, Gully BS, Colson RN, Rossjohn J. Crystal structure of the SARS-CoV-2 non-structural protein 9 Nsp9. iScience. 2020;23:101258.

137. Ye Q, West AMV, Silletti S, Corbett KD. Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. Protein Sci. 2020;29:1890–901.

138. Surya W, Li Y, Torres J. Structural model of the SARS coronavirus E channel in LMPG micelles. Biochim Biophys Acta Biomembr. 2018;1860:1309–17.

139. Zhou Z, Huang C, Zhou Z, Huang Z, Su L, Kang S, et al. Structural insight reveals SARS-CoV-2 ORF7a as an immunomodulating factor for human CD14$^+$ monocytes. iScience. 2021;24:102187.

140. von Itzstein M, Wu WY, Kok GB, Pegg MS, Dyason JC, Jin B, et al. Rational design of potent sialidase-based inhibitors of influenza virus replication. Nature. 1993;363:418–23.

141. Kim CU, Lew W, Williams MA, Liu H, Zhang L, Swaminathan S, et al. Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. J Am Chem Soc. 1997;119:681–90.

142. Owen DR, Allerton CMN, Anderson AS, Aschenbrenner L, Avery M, Berritt S, et al. An oral SARS-CoV-2 Mpro inhibitor clinical candidate for the treatment of COVID-19. Science. 2021;2:eabl4784.

143. Zhao Y, Fang C, Zhang Q, Zhang R, Zhao X, Duan Y, et al. Crystal structure of SARS-CoV-2 main protease in complex with protease inhibitor PF-07321332. Protein Cell. 2021;22:1–5.

144. Fu L, Ye F, Feng Y, Yu F, Wang Q, Wu Y, et al. Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease. Nat Commun. 2020;11:4417.

145. Patchett S, Lv Z, Rut W, Békés M, Drag M, Olsen SK, et al. A molecular sensor determines the ubiquitin substrate specificity of SARS-CoV-2 papain-like protease. Cell Rep. 2021;36:109754.

146. Newman JA, Douangamath A, Yadzani S, Yosaatmadja Y, Aimon A, Brandão-Neto J, et al. Structure, mechanism and crystallographic fragment screening of the SARS-CoV-2 NSP13 helicase. Nat Commun. 2021;12:4848.

147. Siu YL, Teoh KT, Lo J, Chan CM, Kien F, Escriou N, et al. The M, E, and N structural proteins of the severe acute respiratory syndrome coronavirus are required for efficient assembly, trafficking, and release of virus-like particles. J Virol. 2008;82:11318–30.

148. Angelini MM, Akhlaghpour M, Neuman BW, Buchmeier MJ. Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles. mBio. 2013;4:e00524–13.

149. Kappel K, Zhang K, Su Z, Watkins AM, Kladwang W, Li S, et al. Accelerated cryo-EM-guided determination of three-dimensional RNA-only structures. Nat Methods. 2020;17:699–707.

150. Zhang K, Li S, Kappel K, Pintilie G, Su Z, Mou TC, et al. Cryo-EM structure of a 40 kDa SAM-IV riboswitch RNA at 3.7 Å resolution. Nat Commun. 2019;10:5511.

151. Su Z, Zhang K, Kappel K, Li S, Palo MZ, Pintilie GD, et al. Cryo-EM structures of full-length Tetrahymena ribozyme at 3.1 Å resolution. Nature. 2021;596:603–7.

152. Jiang X, Jiang L, Shan A, Su Y, Cheng Y, Song D, et al. Targeting hepatic miR-221/222 for therapeutic intervention of nonalcoholic steatohepatitis in mice. EBioMedicine. 2018;37:307–21.

153. Lim KRQ, Maruyama R, Echigoya Y, Nguyen Q, Zhang A, Khawaja H, et al. Inhibition of DUX4 expression with antisense LNA gapmers as a therapy for facioscapulohumeral muscular dystrophy. Proc Natl Acad Sci USA. 2020;117:16509–15.

154. Di Martino MT, Gullà A, Gallo Cantafio ME, Altomare E, Amodio N, Leone E, et al. In vitro and in vivo activity of a novel locked nucleic acid (LNA)-inhibitor-miR-221 against multiple myeloma cells. PLoS One. 2014;9:e89659.