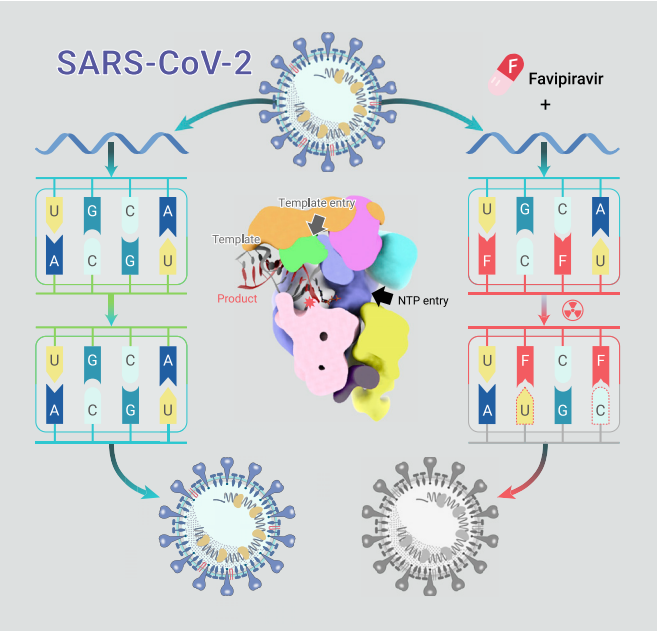
The Innovation

Report

Structural Basis of SARS-CoV-2 Polymerase Inhibition by Favipiravir

Qi Peng,[1,6] Ruchao Peng,[1,6] Bin Yuan,[1,2,6] Min Wang,[1] Jingru Zhao,[1,2] Lifeng Fu,[1,3] Jianxun Qi,[1,2,3] and Yi Shi[1,2,3,4,5,*]
*Correspondence: shiyi@im.ac.cn
Received: November 5, 2020; Accepted: January 11, 2021; Published Online: January 18, 2021; https://doi.org/10.1016/j.xinn.2021.100080

GRAPHICAL ABSTRACT



PUBLIC SUMMARY

- Favipiravir inhibits viral replication mainly by inducing mutations
- Structure of Favipiravir bound to SARS-CoV-2 polymerase in the pre-catalytic state
- The structure reveals the base-pairing pattern of Favipiravir to pyrimidine residues

Structural Basis of SARS-CoV-2 Polymerase Inhibition by Favipiravir

Qi Peng,[1,6] Ruchao Peng,[1,6] Bin Yuan,[1,2,6] Min Wang,[1] Jingru Zhao,[1,2] Lifeng Fu,[1,3] Jianxun Qi,[1,2,3] and Yi Shi[1,2,3,4,5,*]

[1]CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China

[2]Savaid Medical School, University of Chinese Academy of Sciences, Beijing, China

[3]Center for Influenza Research and Early-warning (CASCIRE), CAS-TWAS Center of Excellence for Emerging Infectious Disease (CEEID), Chinese Academy of Sciences, Beijing, China

[4]Chongqing Key Laboratory of Neurodegenerative Diseases, Chongqing General Hospital, University of Chinese Academy of Sciences, Chongqing, China

[5]College of Basic Medicine, Jilin University, Changchun, China

[6]These authors contributed equally

*Correspondence: shiyi@im.ac.cn

Received: November 5, 2020; Accepted: January 11, 2021; Published Online: January 18, 2021; https://doi.org/10.1016/j.xinn.2021.100080

Citation: Peng Q., Peng R., Yuan B., et al. (2020). Structural basis of SARS-CoV-2 polymerase inhibition by Favipiravir. The Innovation **2(1)**, 100080.

The outbreak of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has developed into an unprecedented global pandemic. Nucleoside analogs, such as Remdesivir and Favipiravir, can serve as the first-line broad-spectrum antiviral drugs by targeting the viral polymerases. However, the underlying mechanisms for the antiviral efficacies of these drugs are far from well understood. Here, we reveal that Favipiravir, as a pyrazine derivative, could be incorporated into the viral RNA products by mimicking both adenine and guanine nucleotides. This drug thus inhibits viral replication mainly by inducing mutations in progeny RNAs, different from Remdesivir or other RNA-terminating nucleoside analogs that impair the elongation of RNA products. We further determined the cryo-EM structure of Favipiravir bound to the replicating polymerase complex of SARS-CoV-2 in the pre-catalytic state. This structure provides a missing snapshot for visualizing the catalysis dynamics of coronavirus polymerase, and reveals an unexpected base-pairing pattern between Favipiravir and pyrimidine residues that may explain its capacity for mimicking both adenine and guanine nucleotides. These findings shed light on the mechanism of coronavirus polymerase catalysis and provide a rational basis for developing antiviral drugs to combat the SARS-CoV-2 pandemic.

KEYWORDS: SARS-CoV-2; Favipiravir; cryo-EM; polymerase; pre-catalytic state

INTRODUCTION

The *Coronaviridae* family includes many life-threatening human pathogens, such as SARS-CoV,[1] Middle East respiratory syndrome coronavirus (MERS-CoV),[2] and the cause of the ongoing pandemic (SARS-CoV-2).[3,4] Coronaviruses harbor a non-segmented positive-sense RNA genome with a 5′ cap and a 3′ poly(A) tail, which can directly serve as mRNA to guide the production of viral proteins.[5] A total of 4 structural proteins and at least 16 non-structural proteins (nsps) as well as 9 accessory proteins are encoded by the coronavirus genome. Among them, nsp12 is the core catalytic subunit of viral RNA-dependent RNA polymerase (RdRp) complex, which executes transcription and replication of the viral genomic RNA.[6] Two cofactor subunits, nsp7 and nsp8, associate with nsp12 to constitute an obligatory core polymerase complex to confer processivity for RNA elongation.[7] To achieve complete transcription and replication, a panel of other nsps are involved to accomplish other enzymatic functions, including the nsp14-nsp10 exonuclease for proofreading,[8–10] the nsp13 helicase for RNA unwinding,[11–14] the nsp14 N7-methyltransferase and the nsp16-nsp10 2′-O-methyltransferase for capping.[15–19] Due to the key roles of polymerase complex for viral replication, it has long been thought of as a promising antiviral drug target.[20] Recently, we and other groups have determined the structures of the SARS-CoV-2 core polymerase complex in both apo and RNA-bound states,[21–25] providing important information for structure-based antiviral drug design.

In the efforts to combat the SARS-CoV-2 pandemic, Remdesivir was initially expected to be a highly competent drug candidate for disease treatment. However, the recently disclosed outcomes of clinical trials revealed controversial efficacies as well as some cases of side effects.[26,27] The clinical improvement rate of patients receiving Remdesivir within 7 days was only 3%, and ~66% subjects developed adverse effects.[26,27] In contrast, another broad-spectrum antiviral Favipiravir showed good promise in that 40.1%–73.8% patients with different symptoms revealed significant clinical improvement at day 7 of treatment, and 31.90% subjects were reported with side effect manifestations.[28–30] Both of them are administered as pro-drugs and should be processed into triphosphate forms as nucleotide mimicries to interfere with RNA synthesis. The active form of Remdesivir has a similar base moiety to adenine nucleotide, which can be incorporated into the growing strand of the RNA product using a uracil nucleotide as the template.[25] Besides, the cyano group in the ribose ring may cause a steric clash with polymerase residues during elongation, resulting in aberrant termination of RNA synthesis.[24,31] It is yet unclear how Favipiravir could be recognized by the polymerase and disturb the faithful process of RNA production.

In this work, we present evidence for *in vitro* incorporation of Favipiravir into RNA products by SARS-CoV-2 polymerase, and demonstrate that this drug can mimic both A and G nucleotides to interfere with RNA synthesis by inducing mutations in product strands. Also, we provide the structural basis of Favipiravir recognition by the SARS-CoV-2 polymerase complex and discuss the mechanism of catalysis by this particular polymerase.

RESULTS

Favipiravir Mimics both A and G Nucleotides for RNA Synthesis

To investigate the mechanism for the antiviral efficacy of Favipiravir, we performed *in vitro* primer extension assays using a template derived from the 3′ untranslated region of the authentic viral genome. The +1 catalytic position was inserted with different template residues accordingly to allow only one nucleotide to be incorporated in the presence of each individual nucleotide triphosphate (NTP) substrate (Figure 1). Even though the product strand was supposed to grow by only one nucleotide, some larger RNA products with two or three nucleotide extensions were generated for each specific NTP substrate (Figures 1C and 1D). This phenomenon suggests that the SARS-CoV-2 polymerase is prone to mis-incorporate unwanted nucleotides into the product RNA, resulting in low fidelity for transferring the genomic information during transcription and replication. Similar observations were also reported recently that SARS-CoV-2 polymerase is more tolerant for mismatches between template and product residues than other viral RdRps,[32] which further highlights the requirement for the proofreading nuclease nsp14 to maintain the integrity of viral genome. Interestingly, Favipiravir could be incorporated into the RNA product with similar efficiencies to those of ATP or GTP substrates guided by U or C template residues, respectively (Figure 1C). In contrast, Remdesivir could only be incorporated with a U in the

Report

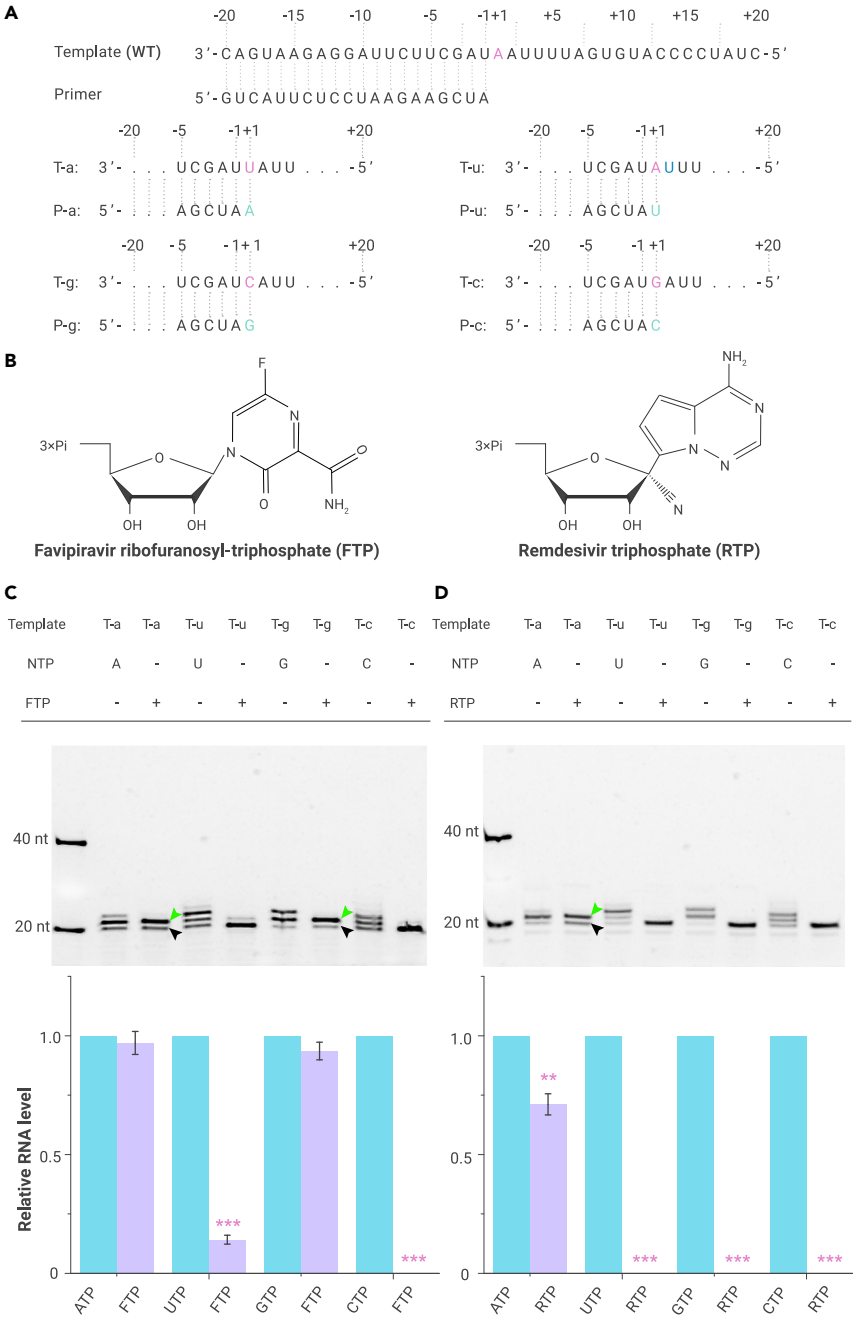


Figure 1. Incorporation of Favipiravir and Remdesivir into RNA products (A) RNA template and primer sequences engineered for single nucleotide extension. The +1 template residues are highlighted in red color and the nucleotide to grow is indicated by green.
(B) Structures of the metabolically active form of Favipiravir and Remdesivir.
(C and D) Single nucleotide extension assays for Favipiravir (C) and Remdesivir (D). The remaining 20-nt primer and the 21-nt single nucleotide extension products are indicated by black and green arrowheads, respectively. Different extents of extra-extension products are also generated in reactions for each individual NTP substrate, which may result from mismatches between template residues and incoming nucleotides. The products were quantified by integrating the intensity of each band. The incorporation efficiency of Favipiravir for each specific template was compared with the corresponding NTP substrate. The differences were estimated by two-tailed Student's t test with results from three independent experiments using different protein preparations. *p < 0.05, **p < 0.01, ***p < 0.001.

template strand (Figure 1D). This evidence demonstrates that Favipiravir is a universal mimicry for purine nucleotides, different from other specific nucleotide analogs, such as Remdesivir[24,25] and Sofosbuvir.[33]

Even though Favipiravir could be efficiently incorporated into RNA products to impair the fidelity of RNA synthesis, it did not significantly inhibit RNA production *in vitro* (Figure 2). In contrast, Remdesivir effectively inhibited the RNA elongation by SARS-CoV-2 polymerase (Figure 2A). The presence of Favipiravir in the RNA product did not immediately terminate the extension of the growing strand except when repeat pyrimidine residues were encountered in the template. The RNA synthesis seemed prone to stall if multiple consecutive incorporations of Favipiravir were supposed to take place (Figure 2C), consistent with the observations reported by other groups recently.[32] This evidence suggests that the presence of repetitive Favipiravir residues might distort the configuration of the template-product RNA duplex to prevent its further extension. However, the consecutive incorporation of Favipiravir might take place with extremely low probabilities in the cellular environment due to the competition of ATP/GTP substrates. Thus, Favipiravir is more likely to be discretely incorporated during virus replication and induce mutations in progeny RNA copies. This hypothesis was supported by a recently reported virus-based inhibition assay, which revealed that Favipiravir was able to escape the proofreading mechanism of the SARS-CoV-2 replication complex and led to mutations in the progeny viral genome.[32]

Structure of Favipiravir Bound to SARS-CoV-2 Polymerase Before Catalysis

To uncover the structural basis of Favipiravir recognition by the polymerase, we determined the structure of the SARS-CoV-2 nsp12-nsp7-nsp8 core polymerase complex in the presence of a template-product partial duplex RNA and Favipiravir at 3.2 Å global resolution by cryoelectron microscopy (cryo-EM) reconstruction (Figures 3 and S1). The density map reveals clear

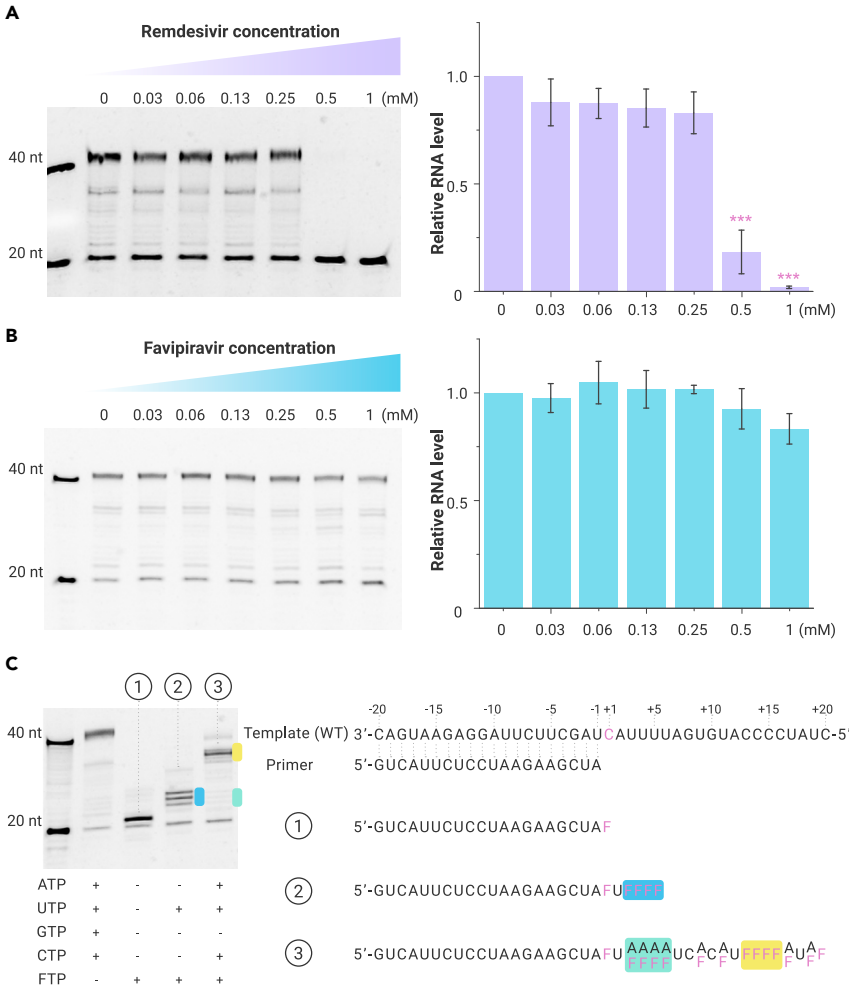


Figure 2. Incorporation of Favipiravir into RNA products by SARS-CoV-2 Polymerase *In vitro* Inhibition of Remdesivir and Favipiravir Against SARS-CoV-2 Polymerase (A and B) *In vitro* inhibition of SARS-CoV-2 polymerase activity by Remdesivir (A) and Favipiravir (B) in serially diluted concentrations. No significant inhibition for RNA synthesis was observed.
(C) RNA extension in the presence of FTP and different NTP substrates. The RNA synthesis tends to stall at the repeat sequence where consecutive incorporation of FTP is supposed to take place. The potential stalled regions are highlighted in blue, green, and yellow backgrounds as indicated in the sequences of RNA products.

features for most amino acid side chains and nucleotide base moieties, allowing the detailed analysis for protein-RNA interactions and Favipiravir recognition (Figures S2 and S3). In the structure, the nsp8 subunit (nsp8.1) in the nsp7-nsp8 heterodimer is mostly unresolved, similar to the recently reported structure of Remdesivir-bound complex in the pre-translocation conformation[25] (Figure 3A). Also, the N-terminal long helices of nsp8 subunits could not be visualized, which are supposed to form a sliding platform for template-product duplex elongation.[22,24] This might result from the insufficient length of the RNA duplex to establish extensive interactions with the helix track. This structure resolved 18 nucleotide residues in the template strand and 16 residues in the primer strand. Importantly, Favipiravir was captured in the triphosphate form (FTP) before catalysis which bound at the +1 position and paired with a C template residue (Figures 3B and 3C). The α-phosphate of FTP is located in vicinity of the 3′-hydroxyl group of the A^{-1} residue with a distance of ~3.3 Å, and no density for a phosphate-di-ester bond could be observed (Figure 3C). Moreover, the density of FTP could only be visualized at lower contour levels than the other residues, suggesting the low occupancy of Favipiravir in the complex, which might be related to the pre-catalytic state without stable covalent interactions with the product RNA. Thus, this structure provides a missing snapshot for the pre-catalytic conformation of SARS-CoV-2 polymerase, an earlier stage before the previously reported post-catalysis/pre-translocation and post-translocation conformations.[22,24,25]

As a mimicry of nucleotide substrate, FTP is recognized as an incoming NTP by conserved catalytic residues (Figure 4A). Residues D623 (motif A) and N691 (motif B) stabilize the ribose ring of FTP, and residue S682 (motif B) may potentially interact with both the ribose and pyrazine branch moieties. The A^{-1} residue of product RNA potentially further contacts with the α-phosphate and ribose of FTP^{+1} to drag the two residues in close proximity for catalysis. The amide group of Favipiravir pairs with the template C residue and potentially forms three hydrogen bonds for base-pairing (Figures 4A and S4). When paired with a U template residue, there might be two hydrogen bonds in between (Figure S4). To capture the pre-catalytic conformation, we did not include catalytic metals in the buffer solutions for protein purification and assembled the complex by incubation on ice to prevent catalysis (Figure S5). Yet, we observed the density for a single metal ion anchored by D761 (motif C) of nsp12, which is offset from the catalytic position and may arise from the cellular environment during expression (Figures 4A and S2). A similar conformation was also observed in the structure of enterovirus 71 (EV71) polymerase (PDB: 5F8I).[34] To enable catalysis, two metal ions are required to neutralize the negative charges of phosphate groups in the NTP substrate and stabilize the transition intermediate.[34,35] Based on the previous structure of post-catalytic conformation in the presence of Remdesivir,[25] we modeled two magnesium ions into our structure to analyze the potential catalytic state of this polymerase. The two metal atoms are coordinated by D618 (motif A) and D761 (motif C) and further bridge the three phosphate groups of the FTP substrate. The terminal γ-phosphate is further sequestered by a salt bridge contributed by K798 (Figure 4A). Besides, residues R553 and R555 from motif F may also interact with the β- and γ-phosphate groups. These interactions together stabilize the incoming nucleotide and 3′ terminus of product RNA in close vicinity to facilitate the nucleophilic attack to α-phosphate by the 3′-hydroxyl oxygen of the −1 nucleotide.

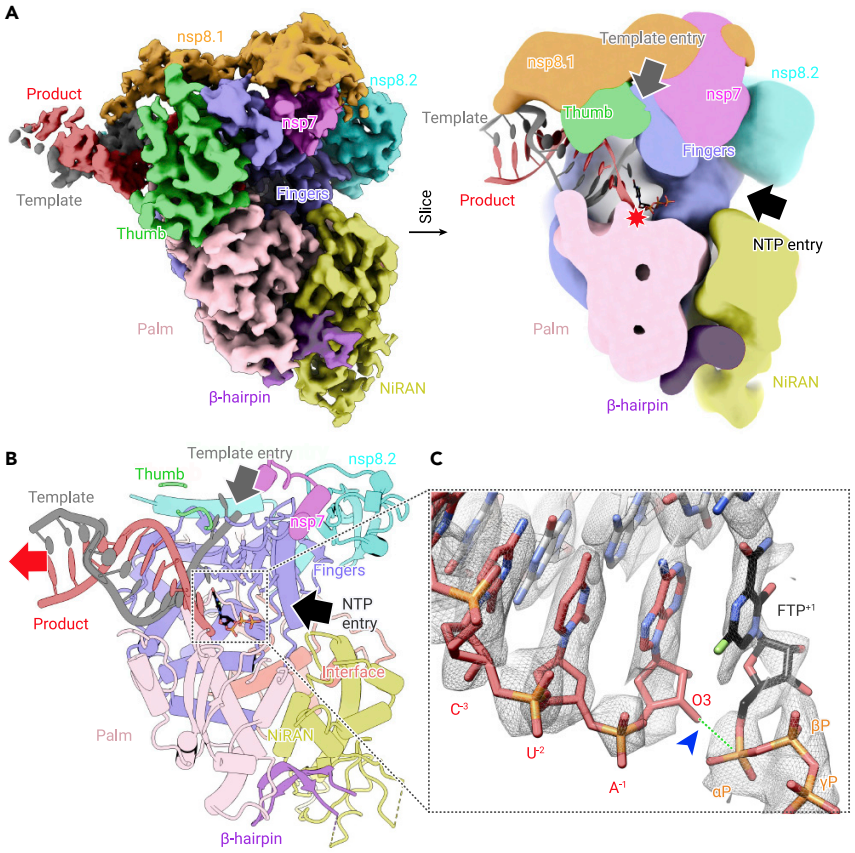


Figure 3. Cryo-EM Structure of Favipiravir Bound to Replicating Core Polymerase Complex of SARS-CoV-2 (A) Overall density map (left) of SARS-CoV-2 core polymerase complex bound to template and product RNA strands and FTP. The nsp12 polymerase subunit is colored by domains. The nsp7-nsp8 cofactors and RNA strands are colored by chains. The central section (right) of the density map is shown to reveal the inner tunnels within the polymerase. The RNA strands are shown in cartoons and colored by chains. The entrances for template and NTP substrate are indicated by arrows. The catalytic site is highlighted by a red star.
(B) Atomic model of the Favipiravir bound polymerase complex of SARS-CoV-2. The protein and RNA are shown in cartoons with the same color code as in (A). The FTP is shown as a stick model and highlighted in black.
(C) Close-up view of FTP and the growing terminus of product RNA strand. No density for a covalent bond is observed between O3 of A^{-1} and α-P of FTP, highlighted by a blue arrowhead. The two atoms are linked by a green dashed line to indicate the position for catalysis.

Conserved Mechanism of Substrate Recognition by Viral RdRps

Since Favipiravir could mimic both adenine and guanine nucleotides for RNA synthesis, we modeledGTP into the structure to compare the potential different recognition patterns for different NTP substrates (Figure 4). Basically, the GTP reveals a highly similar interaction network with the polymerase residues to that for FTP except that an additional residue K545 from motif F may also be involved in interactions with the carbonyl group of the guanine base (Figure 4B). The ATP substrate could also be accommodated similarly with the amino group of adenine base interacting with K545. For pyrimidine nucleotides, however, the base moiety is too far away from K545 to form such interactions (Figure 4D). This residue is highly conserved in all viral RdRps and may serve as a signature residue for discriminating purine and pyrimidine substrates (Figure S6). Comparing the RdRps from different viruses, including positively/negatively sensed and segmented/non-segmented RNA viruses, we found that the residues for stabilizing the ribose, base, and catalytic metals are highly conserved across all viral families, whereas the residues for accommodating the phosphate groups reveal obvious diversity and may also involve residues outside the seven canonical catalytic motifs of RdRp, e.g., residue R48 in HCV polymerase[33] (Figure 4C).

Catalysis Cycle of SARS-CoV-2 Polymerase

With the available structures of SARS-CoV-2 polymerase before and after catalysis,[22,24,25] we were able to assemble a complete scenario of the catalytic cycle to analyze the potential conformational changes of the polymerase during catalysis (Figure 5). It has been established in flavivirus and enterovirus RdRps that the binding of incoming NTP substrates would induce active site closure and relocation of metal ions to facilitate catalysis.[33–35] A similar scenario may also exist for coronavirus polymerase, and potentially other viral RdRps. The motif A loop in apo polymerase adopts an open conformation with the active site fully accessible for incoming NTP substrates (Figures 5A and 5B). One of the catalytic metal ions (metal A) is anchored by D761 offset from the catalytic site, whereas the other one (metal B) might be less stable and frequently exchange between the polymerase and the solvent. Therefore, the metal A atom could be visualized in several viral polymerases without the template RNA and NTP substrate in the active site, including bunyavirus (PDB: 6Z6G)[36] and arenavirus (PDB: 6KLC and 6KLD)[37] polymerases. With the NTP substrate coming in, the motif A loop flips inward to close the active site and the side chain of D761 (motif C) rotates to relocate metal A into the catalytic site. Besides, residue K798 (motif D) would also be re-oriented to stabilize the γ-phosphate of NTP (Figures 5C and 5D). After catalysis, the motif A loop and residue K798 would move back to the open conformation, allowing the pyrophosphoric acid fragment to be released. This process is accompanied by the translocation of the template-product RNA duplex, which resumes the polymerase to get ready for the next round of catalysis (Figures 5E and 5F).

DISCUSSION

Viral polymerases are very important targets for antiviral drug design. Compared with the viral envelop proteins, the polymerase machinery of different viruses is more stringently conserved in evolution, which offers the possibility of developing broad-spectrum antiviral drugs against different viral groups.[20] This unprecedented SARS-CoV-2 pandemic again highlights the potential of viral polymerase as a critical target for antiviral therapy as some nucleoside analog drugs have shown promising efficacies in pre-clinical trials, including Remdesivir and Favipiravir.[26–29] The underlying mechanisms of these nucleoside analogs reveal quite obvious diversity, many of which have not been well understood. Elucidating the molecular basis of interactions between viral polymerase and these drugs is of particular importance for developing better antiviral drugs and minimizing the potential side effects.

We and other groups have previously reported the structures of the SARS-CoV-2 polymerase complex in the apo form or in complex with template/primer RNA strands in the post-catalytic states.[21–25] The pre-catalytic conformation we captured in this study fills an important knowledge gap for understanding the conformational dynamics during the complete

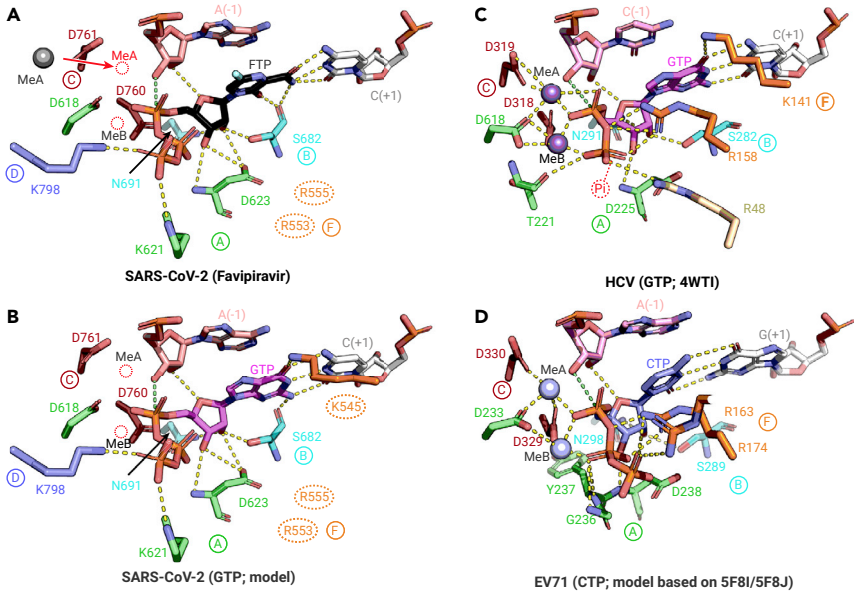


Figure 4. Recognition of Favipiravir and Comparison with Other NTP Substrates (A) Recognition of Favipiravir by SARS-CoV-2 polymerase in the pre-catalytic conformation. The key residues of nsp12 involved in FTP interaction are shown as sticks and colored by different catalytic motifs. Positions of the two catalytic metal ions are modeled based on the post-catalytic structure in complex with Remdesivir (PDB: 7BV2),[25] which are indicated by red dashed circles. The density for a metal ion is observed offset from the catalytic site, which presumably corresponds to the catalytic metal A (MeA) before being relocated to the active site, similar to the scenario observed for EV71 polymerase.[34] Two residues in motif F (R553 and R555) are potentially involved in interactions with FTP but are poorly resolved in the density map, represented by dashed ovals. The potential polar contacts are shown as yellow dashed lines. The O3 atom of A^{-1} and the αP atom of FTP are linked by a green dashed line to indicate the site of catalysis.
(B) Predicted model of GTP recognition by SARS-CoV-2 polymerase complex. An additional residue in motif F (K545) is potentially involved in interactions with the guanine base.
(C) Recognition of GTP substrate by HCV polymerase (PDB: 4WTI).[33] The γ-phosphate of GTP is not present in the structure, indicated by a dashed circle. A residue outside the seven conserved catalytic motifs of RdRp (R48) stabilizes the β-phosphate of GTP by a salt bridge, colored in wheat.
(D) Structural model of CTP bound to enterovirus 71 (EV71) polymerase in the closed conformation. This model was generated based on the pre-catalytic complex in open conformation (PDB: 5F8I) and the closed post-catalytic conformation (PDB: 5F8J) of the same complex.[34] All four panels represent the closed conformation in the pre-catalytic state of different structures to facilitate comparison.

catalytic cycle of this polymerase. It is interesting to note that the key residues for NTP substrate recognition are conserved across many RNA viruses from different families, and that Favipiravir is efficiently incorporated into RNA products similar to other typical NTP substrates. These observations suggest the potential broad-spectrum antiviral efficacies of this drug for a panel of different RNA viruses by similar mechanisms. The administration of Favipiravir has been shown to induce mutations in the progeny virion of SARS-CoV-2 in cell-based assays,[32] indicating that this drug can evade the proofreading mechanism of coronaviruses to exert antiviral efficacy via multiple rounds of RNA synthesis. This is in sharp contrast to other RNA-terminating nucleoside analog drugs that directly impair the elongation process to produce abortive RNA products.[24,31] These differences imply the diversity of the underlying mechanisms of various nucleoside analog antiviral drugs and offer clues for further engineering of these drugs to inhibit viral RNA synthesis at different stages of the catalytic cycle.

Apart from the small-molecule inhibitors, viral infection could also be effectively inhibited by antibodies that interfere with viral entry pathways. The envelope spike glycoprotein has been demonstrated to be a dominant antigen during coronavirus infection.[38,39] Numerous therapeutic antibodies have been developed for targeting this molecule to impair receptor recognition or membrane fusion.[40–44] Because of the pressure rendered by host immunity, the viral envelope proteins are prone to undergo antigenic drift, resulting in immune evasion.[45] In a short term of clinical usage, antibodies are highly efficient for blocking viral infections, which can achieve ideal specificity and minimize side effects. For long-term therapeutic usage and perspective preparation, small-molecule inhibitors targeting other conserved targets, such as viral polymerase and main protease, might be better options that bear less chance of developing resistance or evasion. In addition, these targets are better suited for developing broad-spectrum antiviral drugs.[46–48] Simultaneously, these broad-spectrum inhibitors are also likely to cause side effects due to non-specific targeting on cellular components. On the other hand, due to the relatively lower cost and convenience for drug delivery, the small-molecule inhibitors display remarkable advantages for large-scale application among human populations, especially in less-developed areas. Therefore, drug development targeting the viral polymerase would represent a highly promising and universal strategy against various emerging and re-emerging viruses.

In summary, we present a missing structural snapshot of coronavirus polymerase replication in the pre-catalytic conformation, facilitating the extrapolation of the dynamic catalytic cycle for RNA nucleotide polymerization. Importantly, we reveal the structural basis of Favipiravir incorporation by SARS-CoV-2, which suggests the feasibility of developing other nucleotide-mimicking antiviral drugs by utilizing non-base-derived molecular entities. Given the better record of Favipiravir than Remdesivir with regard to side effects,[28,29] it may represent a better option for clinical treatment of SARS-CoV-2 infections, as well as a panel of other human-infecting RNA viruses.[49] In addition, these findings provide an important basis for developing better broad-spectrum inhibitors with higher potency to combat infection by various emerging RNA viruses.

MATERIALS AND METHODS

Protein Expression and Purification

The SARS-CoV-2 nsp7, nsp8, and nsp7L8 fusion proteins were expressed in *E. coli*, and the nsp12 polymerase subunit was expressed with the Bac-to-Bac system (Invitrogen) as described previously.[23] All these proteins were purified by tandem affinity chromatography and size-exclusion chromatography (SEC) accordingly. To constitute the core polymerase complex, the purified nsp12, nsp8, and nsp7L8 proteins were incubated on ice overnight with a molar ratio of nsp12:nsp8:nsp7L8 = 1:3:3. The complex was then purified by SEC using a Superdex 200 increase column (GE Healthcare) equilibrated with a buffer consisting of 25 mM HEPES-NaOH (pH 7.5), 300 mM NaCl, and 2 mM tris(2-carboxyethyl) phosphine (TCEP). The fractions for nsp12-nsp8-nsp7L8 complex were pooled and concentrated to 4 mg/mL for subsequence experiments.

In Vitro Polymerase Activity Assay

The activity of the SARS-CoV-2 polymerase complex was tested as described previously[23] with slight modifications. In brief, 40-nt template RNA strands (sequences adapted for each specific NTP substrate as shown in Figure 1) were annealed to a complementary 20-nt primer containing a 5′-fluorescein label (5′-FAM- GUCAUUCUCCUAAGAAGCUA-3′, Takara). To perform the primer extension assay, 1 μM nsp12, 1 μM nsp7, and 2 μM nsp8 were incubated for 30 min at 30 °C with 1 μM annealed RNA and 0.5 mM individual NTP or drug in a reaction buffer containing 10 mM Tris-HCl (pH 8.0), 10 mM KCl, 1 mM TCEP, and 2 mM $MgCl_2$ (freshly added before use). The products were denatured by heating to 100 °C for 10 min in the presence of formamide and resolved by 20% PAGE containing 9 M urea, which was run with 0.5 × TBE buffer. Images were recorded with a Vilber Fusion imaging system. The RNA products were quantified by integrating the intensity of each band using ImageJ software. The significance of difference was estimated by the two-tailed Student's t test to calculate the p values for each experimental group.

Cryo-EM Sample Preparation and Data Collection

To prepare the Favipiravir bound polymerase complex, the purified nsp12-nsp8-nsp7L8 complex was mixed with annealed RNA duplex (template: 5′-CUAUCCCCAUGUGAUUUUACUAGCUUCUUAGGAGAAUGAC-3′; primer: 5′-GUCAUUCUCCUAAGAAGCUA-3′) with a molar ratio of nsp12:RNA = 1:1.5. The mixture was incubated on ice for

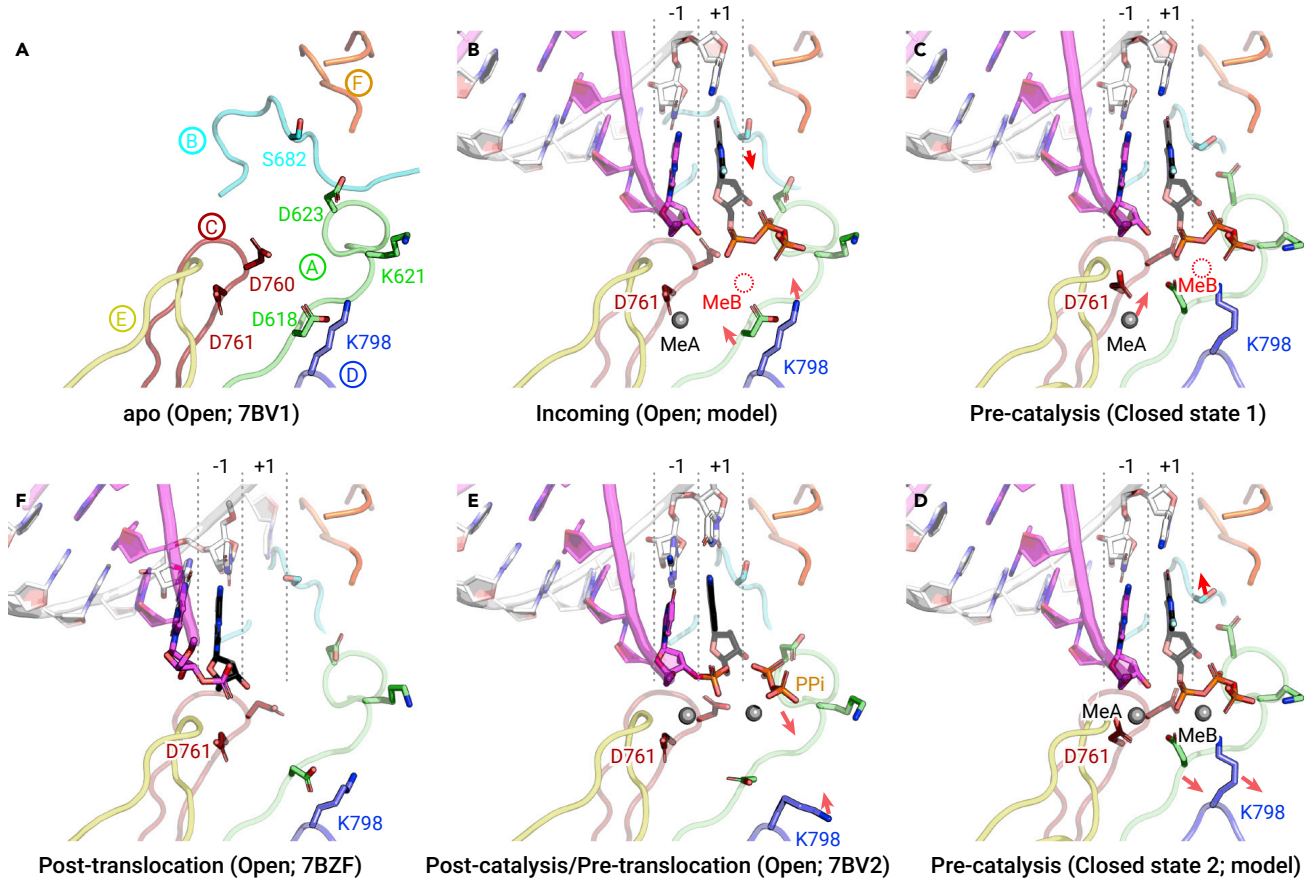


Figure 5. Catalytic Cycle of SARS-CoV-2 Polymerase (A) Catalytic motifs of the apo SARS-CoV-2 polymerase complex. Each motif is represented by a unique color. This structure represents the open conformation of the polymerase. (B) Predicted model of the SARS-CoV-2 polymerase with the NTP substrate coming in. This structure is generated by modeling the RNA and FTP substrate of the pre-catalytic complex into the apo polymerase complex (PDB: 7BV1).[25] The metal A atom is also modeled offset from the catalytic site, which interacts with residue D761. The movements of key residues or motifs for closing the active site are indicated by red arrows. (C) The early-stage pre-catalytic structure in the closed conformation (the reported structure in this study). The metal A is offset from the catalytic site and should be relocated as indicated to enable catalysis. The metal B is absent in the structure, indicated by a red dashed circle. (D) The predicted structure of closed conformation before catalysis. The two metal ions are modeled based on the post-catalytic structure (PDB: 7BV2).[25] After catalysis, the motifs A and B, and residue K798 would shift to re-open the active site, indicated by red arrows. (E) The post-catalytic structure of SARS-CoV-2 polymerase in the open conformation before translocation (PDB: 7BV2).[25] The pyrophosphoric acid fragment would be released and the side chain of K798 would resume as indicated by arrows. (F) The structure of SARS-CoV-2 polymerase in open conformation after translocation (PDB: 7BZF).[24]

2 h in a buffer containing 25 mM HEPES-NaOH (pH 7.5), 150 mM NaCl, and 1 mM TCEP, and supplemented with 0.5 mM FTP. An aliquot of 3 μL protein solution (0.4 mg/mL) was applied to a glow-discharged Quantifiol 1.2/1.3 holey carbon grid and blotted for 2.5 s in a humidity of 100% before plunge freezing with an FEI Vitrobot Mark IV. Cryo-EM samples were screened using an FEI Tecnai TF20 electron microscope and transferred to an FEI Talos Arctica microscope for data collection. The microscope was operated at 200 kV and equipped with a post-column BioQuantum energy filter (Gatan), which was used with a slit width of 20 eV. Cryo-EM data were automatically collected using SerialEM software (http://bio3d.colorado.edu/SerialEM/). Images were recorded with a Gatan K2 Summit camera in super-resolution counting mode with a calibrated pixel size of 1.0 Å at the specimen level. Each exposure was performed with a dose rate of 10 e^-/pixel/s (approximately 10 $e^-/Å^2$/s) and lasted for 6 s, resulting in an accumulative dose of ~60 $e^-/Å^2$, which was fractionated into 30 movie frames. The final defocus range of the dataset was approximately −1.7 to −3.4 μm.

Image Processing

The movie frames were aligned using MotionCor2 to correct beam-induced motion and anisotropic magnification.[50] Initial contrast transfer function (CTF) values were estimated with CTFFIND4.1[51] at the micrograph level. Particles were automatically picked with RELION-3.0[52] following the standard protocol. In total, approximately 2,895,000 particles were picked from ~5,600 micrographs. After 4 rounds of 2D classification, ~1,002,000 particles were selected for 3D classification with the density map of the SARS-CoV-2 polymerase replicating complex (EMDB-11007) low-pass filtered to 60 Å resolution as the reference. After two rounds of 3D classification, three different 3D classes were identified with clear features of secondary structural elements. These classes showed different extents of flexibility at the distal end of the RNA duplex but the main body of polymerase complex was highly similar. Therefore, these 3 classes were combined (including ~329,000 particles) and subjected to 3D refinement supplemented with per-particle CTF refinement and dose-weighting, which led to a reconstruction of 3.2 Å resolution estimated by the gold standard Fourier shell correlation (FSC) 0.143 cutoff value. In the density map, the long helices within the N terminus of nsp8 subunits were not observed. To better resolve this region, a local mask for the N-terminal region of nsp8 subunits and the RNA duplex was applied to perform 3D classification without particle alignment. However, no definable class with a long helix track could be identified, suggesting the flexibility of this region in the structure. This phenomenon might result from the shorter RNA duplex observed in other structures,[22,24] which would help to stabilize the long helices. In addition, this structure was restricted by some extent of preferred orientation of particles, which somehow limited the attainable resolution in certain views of the final reconstruction. Basically, the density map was sufficient to support faithful atomic modeling in most regions. The local resolution distribution of the final density map was calculated with ResMap.[53]

Model Building and Refinement

The structure of the Remdesivir-bound SARS-CoV-2 polymerase complex (PDB: 7BV2) was rigidly docked into the density map using CHIMERA.[54] The model was manually corrected for local fit in COOT.[55] The structure of FTP was built using the "Ligand builder" plug-in in COOT and manually fitted into the density. The initial model was refined in real space using PHENIX[56] with the secondary structural restraints and Ramachandran restrains applied. The model was further adjusted and refined iteratively for several rounds aided by stereochemical quality assessment using MolProbity.[57] The representative density and atomic models are shown in Figures S2 and S3. The statistics for image processing and model refinement are summarized in

Table S1. Structural figures were prepared using either CHIMERA[54] or PyMOL (https://pymol.org/).

Structure-Based Conservation Analysis

The structures of SARS-CoV-2, SARS-CoV (PDB: 6NUR),[58] EV71 (PDB: 5F8I),[34] poliovirus (PDB: 3OL8),[35] HCV (PDB: 4WTI),[33] Zika virus (PDB: 5TMH),[59] vesicular stomatitis virus (PDB: 5A22),[60] influenza virus (PDB: 4WRT),[61] Lassa virus (PDB: 6KLC), Machupo virus (PDB: 6KLD),[37] severe fever with thrombocytopenia syndrome virus (PDB: 6L42),[62] La Crosse virus (PDB: 6Z6G)[36] polymerases were aligned with MUSTANG.[63] The aligned sequences were then used for conservation analysis via the ConSurf[64] server. The figure was generated with PyMOL.

Data Availability

The cryo-EM density map and atomic coordinates have been deposited in the Electron Microscopy Data Bank (EMDB) and the Protein Data Bank (PDB) with the accession codes EMD-30469 and 7CTT, respectively. All other data are available from the authors on reasonable request.

REFERENCES

1. Zhong, N.S., Zheng, B.J., Li, Y.M., et al. (2003). Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. Lancet ***362***, 1353–1358.
2. Zumla, A., Hui, D.S., and Perlman, S. (2015). Middle East respiratory syndrome. Lancet ***386***, 995–1007.
3. Zhou, P., Yang, X.L., Wang, X.G., et al. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature ***579***, 270–273.
4. Zhu, N., Zhang, D., Wang, W., et al. (2020). A novel coronavirus from patients with pneumonia in China, 2019. N. Engl. J. Med. ***382***, 727–733.
5. Fehr, A.R., and Perlman, S. (2015). Coronaviruses: an overview of their replication and pathogenesis. Methods Mol. Biol. ***1282***, 1–23.
6. Ahn, D.G., Choi, J.K., Taylor, D.R., et al. (2012). Biochemical characterization of a recombinant SARS coronavirus nsp12 RNA-dependent RNA polymerase capable of copying viral RNA templates. Arch. Virol. ***157***, 2095–2104.
7. Subissi, L., Posthuma, C.C., Collet, A., et al. (2014). One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. Proc. Natl. Acad. Sci. U S A ***111***, E3900–E3909.
8. Ma, Y., Wu, L., Shaw, N., et al. (2015). Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. Proc. Natl. Acad. Sci. U S A ***112***, 9436–9441.
9. Ferron, F., Subissi, L., De Morais, A.T.S., et al. (2018). Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. Proc. Natl. Acad. Sci. U S A ***115***, E162–E171.
10. Ogando, N.S., Ferron, F., Decroly, E., et al. (2019). The curious case of the nidovirus exoribonuclease: its role in RNA synthesis and replication fidelity. Front. Microbiol. ***10***, 1813.
11. Adedeji, A.O., Marchand, B., te Velthuis, A.J.W., et al. (2012). Mechanism of nucleic acid unwinding by SARS-CoV helicase. PLoS One ***7***, 36521.
12. Shu, T., Huang, M.H., Wu, D., et al. (2020). SARS-coronavirus-2 Nsp13 possesses NTPase and RNA helicase activities that can be inhibited by bismuth salts. Virol. Sin. ***35***, 321–329.
13. Hao, W., Wojdyla, J.A., Zhao, R., et al. (2017). Crystal structure of Middle East respiratory syndrome coronavirus helicase. PLoS Pathog. ***13***, e1006474.
14. Jia, Z.H., Yan, L.M., Ren, Z.L., et al. (2019). Delicate structural coordination of the severe acute respiratory syndrome coronavirus Nsp13 upon ATP hydrolysis. Nucleic Acids Res. ***47***, 6538–6550.
15. Sevajol, M., Subissi, L., Decroly, E., et al. (2014). Insights into RNA synthesis, capping, and proofreading mechanisms of SARS-coronavirus. Virus Res. ***194***, 90–99.
16. Chen, Y., and Guo, D. (2016). Molecular mechanisms of coronavirus RNA capping and methylation. Virol. Sin. ***31***, 3–11.
17. Rosas-Lemus, M., Minasov, G., Shuvalova, L., et al. (2020). High-resolution structures of the SARS-CoV-2 2'-O-methyltransferase reveal strategies for structure-based inhibitor design. Sci. Signal. ***13***, eabe1202.
18. Chen, Y., Su, C.Y., Ke, M., et al. (2011). Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2'-O-methylation by nsp16/nsp10 protein complex. PLoS Pathog. ***7***, e1002294.
19. Viswanathan, T., Arya, S., Chan, S.H., et al. (2020). Structural basis of RNA cap modification by SARS-CoV-2. Nat. Commun. ***11***, 3718.
20. Shi, F., Xie, Y., Shi, L., et al. (2013). Viral RNA polymerase: a promising antiviral target for influenza A virus. Curr. Med. Chem. ***20***, 3923–3934.
21. Gao, Y., Yan, L., Huang, Y., et al. (2020). Structure of the RNA-dependent RNA polymerase from COVID-19 virus. Science ***368***, 779–782.
22. Hillen, H.S., Kokic, G., Farnung, L., et al. (2020). Structure of replicating SARS-CoV-2 polymerase. Nature ***584***, 154–156.
23. Peng, Q., Peng, R., Yuan, B., et al. (2020). Structural and biochemical characterization of the nsp12-nsp7-nsp8 core polymerase complex from SARS-CoV-2. Cell Rep. ***31***, 107774.
24. Wang, Q., Wu, J., Wang, H., et al. (2020). Structural basis for RNA replication by the SARS-CoV-2 polymerase. Cell ***182***, 417–428.e3.
25. Yin, W., Mao, C., Luan, X., et al. (2020). Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. Science ***368***, 1499–1504.
26. Beigel, J.H., Tomashek, K.M., Dodd, L.E., et al. (2020). Remdesivir for the treatment of covid-19—preliminary report. N. Engl. J. Med. ***383***, 994.
27. Wang, Y., Zhang, D., Du, G., et al. (2020). Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial. Lancet ***395***, 1569–1578.
28. Cai, Q., Yang, M., Liu, D., et al. (2020). Experimental treatment with favipiravir for COVID-19: an open-label control study. Engineering (Beijing) ***6***, 1192–1198.
29. Chen, C., Zhang, Y., Huang, J., et al. (2020). Favipiravir versus Arbidol for COVID-19: a randomized clinical trial. medRxiv. https://doi.org/10.1101/2020.2003.2017.20037432.
30. James Ives, M. (2020). Preliminary report of favipiravir observational study in Japan released [Internet]. News-Medical.net [cited 23 July 2020]. https://www.news-medical.net/news/20200602/Preliminary-report-of-Favipiravir-Observational-Study-in-Japan-released.aspx.
31. Gordon, C.J., Tchesnokov, E.P., Woolner, E., et al. (2020). 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. ***295***, 6785–6797.
32. Shannon, A., Selisko, B., Le, N., et al. (2020). Favipiravir strikes the SARS-CoV-2 at its Achilles heel, the RNA polymerase. bioRxiv. https://doi.org/10.1101/2020.2005.2015.098731.
33. Appleby, T.C., Perry, J.K., Murakami, E., et al. (2015). Viral replication. Structural basis for RNA replication by the hepatitis C virus polymerase. Science ***347***, 771–775.
34. Shu, B., and Gong, P. (2016). Structural basis of viral RNA-dependent RNA polymerase catalysis and translocation. Proc. Natl. Acad. Sci. U S A ***113***, E4005–E4014.
35. Gong, P., and Peersen, O.B. (2010). Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. Proc. Natl. Acad. Sci. U S A ***107***, 22505–22510.
36. Arragain, B., Effantin, G., Gerlach, P., et al. (2020). Pre-initiation and elongation structures of full-length La Crosse virus polymerase reveal functionally important conformational changes. Nat. Commun. ***11***, 3590.
37. Peng, R.C., Xu, X., Jing, J.M., et al. (2020). Structural insight into arenavirus replication machinery. Nature ***579***, 615.
38. Duan, L.W., Zheng, Q.Q., Zhang, H.X., et al. (2020). The SARS-CoV-2 spike glycoprotein biosynthesis, structure, function, and antigenicity: implications for the design of spike-based vaccine immunogens. Front. Immunol. ***11***, 576622.
39. Walls, A.C., Park, Y.J., Tortorici, M.A., et al. (2020). Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell ***181***, 281.
40. Shi, R., Shan, C., Duan, X., et al. (2020). A human neutralizing antibody targets the receptor-binding site of SARS-CoV-2. Nature ***584***, 120–124.
41. Wu, Y., Wang, F., Shen, C., et al. (2020). A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2. Science ***368***, 1274–1278.
42. Chi, X., Yan, R., Zhang, J., et al. (2020). A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. Science ***369***, 650–655.
43. Cao, Y., Su, B., Guo, X., et al. (2020). Potent neutralizing antibodies against SARS-CoV-2 identified by high-throughput single-cell sequencing of convalescent patients' B cells. Cell ***182***, 73–84.e16.
44. Lv, Z., Deng, Y.Q., Ye, Q., et al. (2020). Structural basis for neutralization of SARS-CoV-2 and SARS-CoV by a potent therapeutic antibody. Science ***369***, 1505–1509.
45. Singh, P.K., Kulsum, U., Rufai, S.B., et al. (2020). Mutations in SARS-CoV-2 leading to antigenic variations in spike protein: a challenge in vaccine development. J. Lab. Phys. ***12***, 154–160.
46. Furuta, Y., Komeno, T., and Nakamura, T. (2017). Favipiravir (T-705), a broad spectrum inhibitor of viral RNA polymerase. Proc. Jpn. Acad. Ser. B Phys. Biol. Sci. ***93***, 449–463.
47. Van Dycke, J., Arnoldi, F., Papa, G., et al. (2018). A single nucleoside viral polymerase inhibitor against norovirus, rotavirus, and sapovirus-induced diarrhea. J. Infect. Dis. ***218***, 1753–1758.
48. Yi, D., Li, Q., Pang, L., et al. (2020). Identification of a broad-spectrum viral inhibitor targeting a novel allosteric site in the RNA-dependent RNA polymerases of dengue virus and norovirus. Front. Microbiol. ***11***, 1440.
49. Shi, Y. (2020). New Virus, New Challenge. The Innovation ***1***, 10005.
50. Zheng, S.Q., Palovcak, E., Armache, J.P., et al. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods ***14***, 331–332.
51. Rohou, A., and Grigorieff, N. (2015). CTFFIND4: fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. ***192***, 216–221.
52. Zivanov, J., Nakane, T., Forsberg, B.O., et al. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. eLife ***7***, e42166.

53. Kucukelbir, A., Sigworth, F.J., and Tagare, H.D. (2014). Quantifying the local resolution of cryo-EM density maps. Nat. Methods ***11***, 63–65.
54. Pettersen, E.F., Goddard, T.D., Huang, C.C., et al. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. ***25***, 1605–1612.
55. Emsley, P., Lohkamp, B., Scott, W.G., et al. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. ***66***, 486–501.
56. Adams, P.D., Afonine, P.V., Bunkoczi, G., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. ***66***, 213–221.
57. Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., et al. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. ***66***, 12–21.
58. Kirchdoerfer, R.N., and Ward, A.B. (2019). Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. Nat. Commun. ***10***, 2342.
59. Wang, B., Tan, X.F., Thurmond, S., et al. (2017). The structure of Zika virus NS5 reveals a conserved domain conformation. Nat. Commun. ***8***, 14763.
60. Liang, B., Li, Z., Jenni, S., et al. (2015). Structure of the L protein of vesicular stomatitis virus from electron cryomicroscopy. Cell ***162***, 314–327.
61. Reich, S., Guilligay, D., Pflug, A., et al. (2014). Structural insight into cap-snatching and RNA synthesis by influenza polymerase. Nature ***516***, 361.
62. Wang, P., Liu, L., Liu, A., et al. (2020). Structure of severe fever with thrombocytopenia syndrome virus L protein elucidates the mechanisms of viral transcription initiation. Nat. Microbiol. ***5***, 864–871.
63. Konagurthu, A.S., Whisstock, J.C., Stuckey, P.J., et al. (2006). MUSTANG: a multiple structural alignment algorithm. Proteins ***64***, 559–574.
64. Landau, M., Mayrose, I., Rosenberg, Y., et al. (2005). ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Res. ***33***, W299–W302.

ACKNOWLEDGMENTS

We thank all staff members in the Center of Biological Imaging (CBI), the Institute of Biophysics (IBP), and the Chinese Academy of Sciences (CAS) for assistance with data collection. This study was supported by the Strategic Priority Research Program of CAS (XDB29010000), the National Science and Technology Major Project (2018ZX10101004), the National Key Research and Development Program of China (2020YFC0845900), the National Natural Science Foundation of China (NSFC) (82041016, 81871658, and 81802010), a grant from the Bill & Melinda Gates Foundation, and is partially supported by the Yanqi Lake Meeting organized by the Academic Divisions of CAS. M.W. is supported by the National Science and Technology Major Project (2018ZX09711003) and the National Natural Science Foundation of China (NSFC) (81802007). R.P. is supported by the Young Elite Scientist Sponsorship Program (YESS) by the China Association for Science and Technology (CAST) (2018QNRC001). Y.S. is also supported by the Youth Innovation Promotion Association of CAS.

AUTHOR CONTRIBUTIONS

Y.S., Q.P., and R.P. conceived the study. Q.P., B.Y., J.Z., and M.W. purified the protein samples and conducted biochemical studies. Q.P. and R.P. performed cryo-EM analysis. R.P. and J.Q. built the atomic model. Q.P., R.P., M.W., L.F., and Y.S. analyzed the data and wrote the manuscript. All authors participated in the discussion and manuscript editing. Q.P., R.P., and B.Y. contributed equally to this work.

DECLARATION OF INTERESTS

The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xinn.2021.100080.

LEAD CONTACT WEBSITE

http://www.im.cas.cn/jgsz2018/yjtx/zgkxybywswymyxzdsys/202003/t20200324_5518747.html.