


Effects of natural RNA modifications on the activity of SARS-CoV-2 RNA-dependent RNA polymerase

Ivan Petushkov, Daria Esyunina and Andrey Kulbachinskiy 

Institute of Molecular Genetics, National Research Center “Kurchatov Institute”, Moscow, Russia

Keywords

pyrophosphorolysis; RNA-dependent RNA polymerase; RNA modifications; RNA proofreading; SARS-CoV-2

Correspondence

A. Kulbachinskiy, Institute of Molecular Genetics, National Research Center “Kurchatov Institute”, Moscow 123182, Russia
Tel: +74991960015
E-mail: avkulb@yandex.ru

(Received 7 June 2022, revised 17 July 2022, accepted 1 August 2022)

doi:10.1111/febs.16587

RNA-dependent RNA polymerase (RdRp) plays a key role in the replication of RNA viruses, including SARS-CoV-2. Processive RNA synthesis by RdRp is crucial for successful genome replication and expression, especially in the case of very long coronaviral genomes. Here, we analysed the activity of SARS-CoV-2 RdRp (the nsp12–nsp7–nsp8 complex) on synthetic primer–templates of various structures, including substrates with mismatched primers or template RNA modifications. It has been shown that RdRp cannot efficiently extend RNA primers containing mismatches and has no intrinsic RNA cleavage activity to remove the primer 3'-end, thus necessitating the action of exoribonuclease for proofreading. Similar to DNA-dependent RNA polymerases, RdRp can perform processive pyrophosphorolysis of the nascent RNA product but this reaction is also blocked in the presence of mismatches. Furthermore, we have demonstrated that several natural post-transcriptional modifications in the RNA template, which do not prevent complementary interactions (N6-methyladenosine, 5-methylcytosine, inosine and pseudouridine), do not change RdRp processivity. At the same time, certain modifications of RNA bases and ribose residues strongly block RNA synthesis, either prior to nucleotide incorporation (3-methyluridine and 1-methylguanosine) or immediately after it (2'-O-methylation). The results demonstrate that the activity of SARS-CoV-2 RdRp can be strongly inhibited by common modifications of the RNA template suggesting a way to design novel antiviral compounds.

Introduction

RNA-dependent RNA polymerase (RdRp) is the principal enzyme involved in the replication of viral RNA genomes and the most conserved component of the viral replication machinery [1–3]. Due to its high conservation and its critical role in RNA replication, RdRp can serve as a universal target for antiviral compounds with broad specificity. At the same time, RdRps from various viral families can significantly differ in their properties and sensitivity to existing antivirals, as a result of lineage-specific differences in the basal replication machinery and the action of

accessory factors [1]. Coronaviruses have the largest genomes among RNA viruses (~ 30 kilobases), which pose a significant threat to genome integrity and its faithful replication [3–6]. Structural analysis of coronaviral RdRp, first from SARS-CoV [7] and then from SARS-CoV-2 [8–10], demonstrated that it has a typical right-hand structure composed of the fingers, thumb and palm domains (Fig. 1). Acidic residues from the palm domain bind divalent metal cations required for catalysis, while several conserved motifs of the active site coordinate the reacting substrates (Fig. 1B,C)

Abbreviations

1-meG, 1-methylguanosine; 2'-O-meG, 2'-O-methylguanosine; 3-meU, 3-methyluridine; 5-meC, 5-methylcytosine; HCV, hepatitis C virus; Ino, inosine; N6-meA, N6-methyladenosine; RdRp, RNA-dependent RNA polymerase; SARS-CoV-2, severe acute respiratory syndrome-related Coronavirus 2; Ψ, pseudouridine.

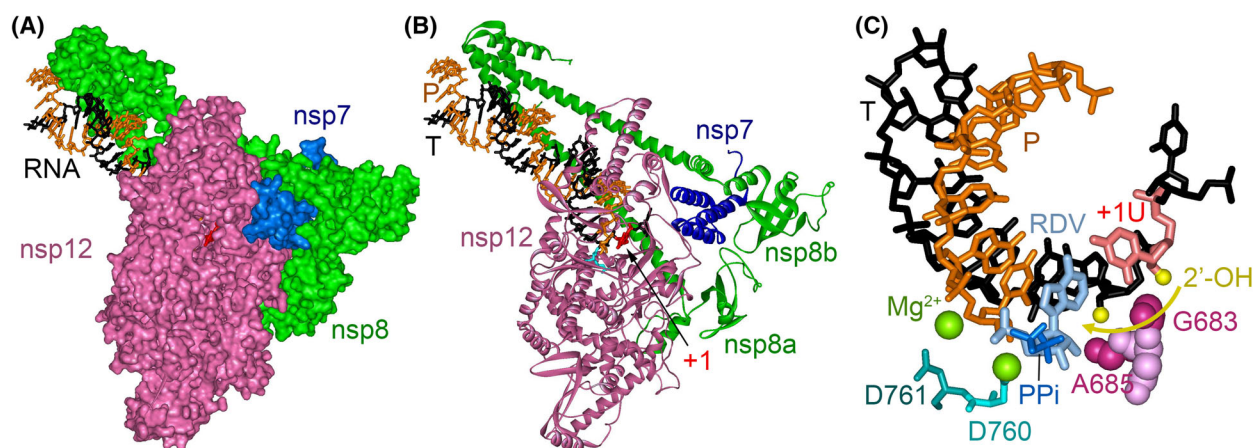


Fig. 1. Structure of the catalytic complex of SARS-CoV-2 RdRp and the mechanism of RNA synthesis. (A) Structure of the elongation complex of RdRp (nsp12, violet; nsp8, green; nsp7, blue) with an RNA substrate (PDB: 6YYT [12]). (B) The same structure showing interactions of RdRp with the RNA product (P, orange) and RNA template (T, black). The +1 template nucleotide is shown in red. (C) Structure of the active site of RdRp in complex with product-template RNA after incorporation of the triphosphate form of remdesivir (PDB: 7BV2 [59]). The template nucleotide in the +1 position (+1 U) is pink; the complementary remdesivir residue is blue and the pyrophosphate moiety is light blue. The catalytic aspartate residues of nsp12 (D760 and D761) are turquoise and catalytic magnesium ions are light green. The hydroxyl group of template nucleotide at the +1 position and the direction of its translocation are indicated. Nsp12 residues G683 and A685 that may block the translocation of 2'-modified templates are shown. The images were generated by ViewerLite 4.2 (Accelrys Ltd., Cambridge, UK).

[2,3]. In holoenzyme RdRp, the polymerase subunit nsp12 interacts with one nsp7 and two nsp8 subunits, which are required for efficient replication. Structural analysis of a replicative complex of SARS-CoV-2 RdRp demonstrated that the upstream RNA duplex interacts with N-terminal α -helical extensions of the nsp8 subunits (Fig. 1A,B), suggesting that these contacts may be important for stabilization of the contacts of RdRp with RNA [11,12].

The full replicative complex of SARS-CoV-2 also contains several accessory subunits, including the nsp13 helicase, the nsp14 3'-5' exonuclease/N⁷-methyltransferase with its cofactor nsp10, and the nsp16 2'-O-methyltransferase [2,3]. The replicative complex binds two molecules of the nsp13 helicase, one of which interacts with the RNA template and another likely has a regulatory role [13,14]. Intriguingly, nsp13 moves in the opposite 5'-3' polarity on the RNA template relative to RdRp, suggesting that it may act in RNA proofreading rather than in RNA synthesis. In particular, it was proposed that nsp13 may displace RdRp backward on the RNA template after nucleotide misincorporation, thus allowing the action of the nsp14 exonuclease [15]. The mechanism of coordination of the polymerization activity of RdRp and exonuclease activity of nsp14 remains largely unknown but recent structural analysis suggested that nsp14 may access the RNA 3'-end *in trans*, after dimerization of replicative complexes [16]. The

exonuclease activity of nsp14 is likely responsible for RNA proofreading after nucleotide misincorporation [17–19], as well as for the resistance of coronaviral RdRp to various nucleotide analogues, which are active against other viruses but can be removed after their incorporation into the RNA product during replication of coronaviruses [20,21].

Despite intensive studies of the replication machinery of coronaviruses, many structural details and activities of the replicative complex of RdRp remain unknown. In particular, the exact mechanism of RNA proofreading following nucleotide misincorporation remains to be elucidated. Furthermore, while many synthetic compounds inhibiting RdRp have been tested *in vitro* and *in vivo* (see Discussion), it remains unknown how natural modifications in the RNA template can affect RdRp activities. In this study, we have analysed the effects of primer and template modifications on the activity of recombinant RdRp from SARS-CoV-2. It has been shown that the efficiency of RNA extension can be strongly affected by the presence of mismatched nucleotides in the RNA primer. Furthermore, we have demonstrated that RdRp can perform nascent RNA cleavage in the reaction of pyrophosphorolysis, which is blocked by mismatches in the RNA 3'-end, but does not possess intrinsic RNA cleavage activity. Finally, we have shown that RdRp can sense naturally occurring nucleotide modifications and lesions in the template

RNA strand, resulting in strong inhibition of RNA synthesis.

Results

Activity and substrate specificity of recombinant SARS-CoV-2 RdRp

To test the activity and substrate specificity of recombinant RdRp consisting of the nsp12, nsp7 and nsp8 proteins, we analysed primer extension using primer–template substrates of various structures, with either RNA or DNA primer and template strands and with various lengths of the upstream duplex (Fig. 2). First, we measured RNA extension using substrates with long upstream duplex (35 bp) to allow the formation of the full set of contacts in the replication complex. It was shown that RdRp can extend RNA primer on both RNA and DNA templates, however, the efficiency of RNA synthesis on the DNA template was lower and it was stalled after the incorporation of several nucleotides (Fig. 2B, lanes 1–4). In contrast,

RdRp was completely inactive with DNA primer on either RNA or DNA templates (lanes 5–8). RdRp was strictly selective for NTPs and could not incorporate dNTPs even in the case of the optimal RNA–RNA primer–template substrate (lanes 9–11).

We then compared the activity of RdRp on RNA substrates containing RNA primers of various lengths (Fig. 2A). To reveal the role of upstream RdRp–RNA duplex interactions in the stabilization of the replication complex, the experiment was performed in reaction buffers containing various concentrations of monovalent salts, from 10 to 175 mM (Fig. 2C). In the case of the long upstream duplex (35 bp), the activity of RdRp was gradually decreased with increasing ionic strength and was decreased 20-fold at 175 mM KCl (Fig. 2C, left). Furthermore, the activity of RdRp depended on the primer length. Even at the lowest ionic strength (10 mM KCl), RNA synthesis was decreased ~3–4 fold when using 15 and 10 nt primers in comparison with the 35 nt primers (Fig. 2C, blue bars). This effect became much stronger at 30 or 100 mM KCl concentrations (6- and 16-fold decrease,

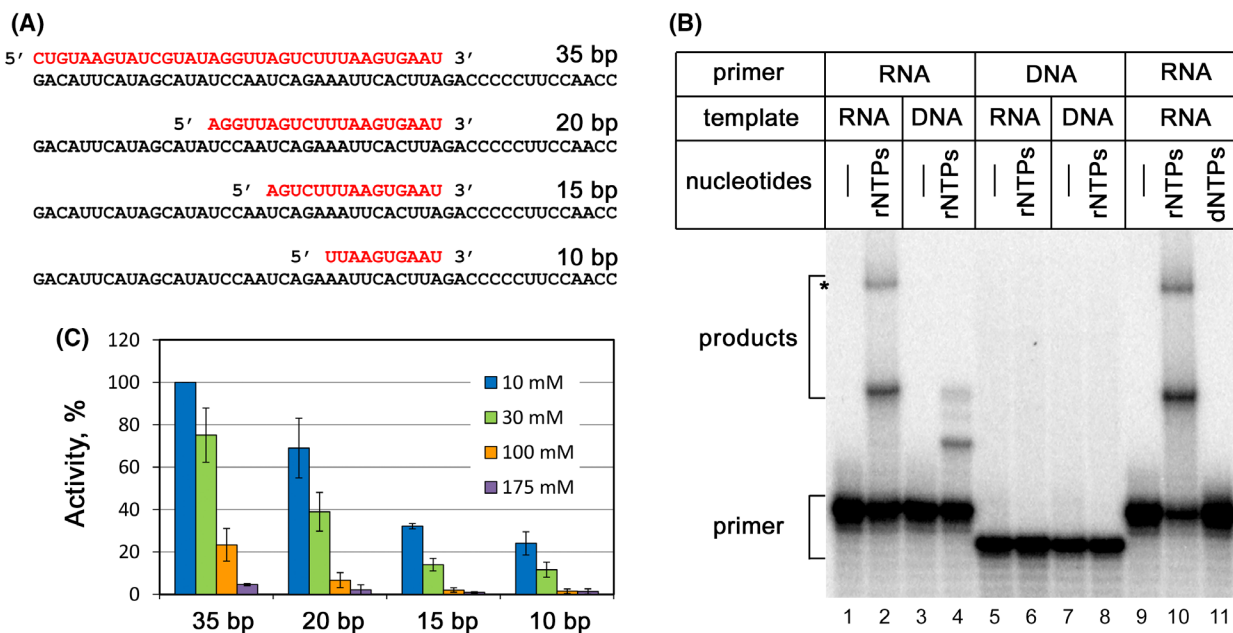


Fig. 2. Primer–template specificity of SARS-CoV-2 RdRp. (A) Structures of RNA substrates with various lengths of RNA primers. DNA primer and template oligonucleotides of the same sequences as in the top substrate were also used in the experiment shown in panel (B). (B) Extension of RNA or DNA primers on RNA or DNA templates by RdRp (the sequences of the 50 nt template and 35 nt primer oligonucleotides correspond to the upper substrate in panel A). The reactions were performed for 10 min at 30 °C with 100 μ M NTPs or dNTPs. Positions of the starting 5'-labelled primers and the extended products are indicated. The upper band indicated with an asterisk corresponds to the RNA product–template duplex that was not completely denatured during PAGE analysis. A representative gel from two independent experiments is shown. (C) Relative efficiency of full-length RNA synthesis by RdRp with primers of different lengths (forming 35, 20, 15 or 10 bp duplexes with the RNA template, see panel A) at increasing ionic strength (10, 30, 100 or 175 mM NaCl). For all reactions, the activity was measured as the sum of extended RNA products and normalized by the maximal activity observed with the 35 nt primer at the lowest ionic strength. Means and standard deviations from three independent experiments are shown.

respectively, for 10 nt primer, green and orange bars). At 175 mM KCl, only trace activity could be detected with the short primers. We, therefore, conclude that the upstream RNA duplex plays an important role in the formation of stable replication complexes by SARS-CoV-2 RdRp.

RdRp activities with mismatched primers and the reaction of pyrophosphorolysis

Previous studies demonstrated that replication complexes of SARS-CoV-2 RdRp with noncomplementary 3'-nucleotides in the RNA product can adopt a backtracked conformation, in which the mismatched 3'-end enters into the RdRp NTP entry tunnel [15]. It was proposed that RdRp backtracking is required for RNA proofreading through the nucleolytic activity of the nsp14 exoribonuclease. Some previously studied cellular and viral RNA polymerases can remove nascent RNA 3'-ends through intrinsic endonucleolytic or exonucleolytic RNA cleavage or pyrophosphorolysis; these reactions can be stimulated at increased pH or in the presence of noncomplementary nucleotides (see Discussion). However, the activity of SARS-CoV-2 RdRp with mismatched RNA substrates and its intrinsic ability to process mismatched RNA 3'-end were not studied.

To test the activities of SARS-CoV-2 RdRp on mismatched RNA substrates, we used RNA substrates containing 1, 2 or 3 mismatched nucleotides at the primer 3'-end (Fig. 3A). Analysis of their interactions with RdRp by electrophoretic mobility shift assay (EMSA) demonstrated that the presence of mismatches slightly decreases the efficiency of their binding to RdRp (Fig. 3B). Previously, RNA substrates having different nucleotide sequence and containing 3 or 5 mismatched nucleotides at the primer 5'-end were shown to be less efficiently bound by RdRp than fully-matched RNA, however, the activity of RdRp on such substrates was not tested [15]. Analysis of RNA synthesis in the presence of NTPs in complexes of SARS-CoV-2 RdRp with mismatched substrates demonstrated that the efficiency of RNA extension is greatly decreased in comparison with the fully matched RNA primer (30–50 fold for 1 nt and 2 nt mismatches; Fig. 3C, compare lanes 4–12 with lanes 2–3). However, some level of RNA extension was still observed with all primers, and it was even higher in the case of RNA primer with three mismatched nucleotides (Fig. 3C, lanes 11–12). The addition of heparin, a competitive inhibitor of RNA binding, had no major effects on RNA extension, suggesting that RdRp forms sufficiently stable complexes with all tested substrates.

These results suggested that RdRp may either accommodate and slowly extend the mismatched RNA 3'-end in the active site or first remove the mismatched nucleotides through exo- or endonucleolytic cleavage followed by extension of the newly formed 3'-end.

To reveal whether RdRp can perform RNA cleavage in the case of fully matched or mismatched RNA primers, we incubated the complexes of RdRp with various types of RNA substrates in the presence of divalent metal cations, which are required for catalysis, but in the absence of NTP substrates (Fig. 3D). No RNA shortening was observed with either fully complementary or mismatched RNA primers in the presence of 2 mM or 20 mM Mg^{2+} or 2 mM Mn^{2+} (Fig. 3D). The addition of a noncomplementary nucleotide (100 μ M GTP) also did not stimulate RNA cleavage (Fig. 3D, lanes 4, 9, 14 and 19). Therefore, we conclude that SARS-CoV-2 RdRp lacks intrinsic endonucleolytic or exonucleolytic activity, at least under the conditions of our experiments.

We further analysed the reaction of pyrophosphorolysis in the same complexes, by adding inorganic pyrophosphate at various concentrations (Fig. 3E). Efficient shortening of the 5'-labelled RNA primer was observed in the case of fully complementary RNA substrate (lanes 1–3). In contrast, no reaction products were detected in the case of mismatched RNA substrates, even in the case of a single nucleotide mismatch (lanes 4–12). To reveal whether the reaction of pyrophosphorolysis is processive, we compared the kinetics of this reaction in the absence and presence of heparin (Fig. 3F). The patterns of RNA products were identical in both cases, indicating that processive RNA shortening can occur in the same catalytic complex by stepwise nucleotide removal without dissociation of RdRp (compare lanes 2–6 and 7–11). From these experiments, we can conclude that SARS-CoV-2 RdRp can perform pyrophosphorolysis of complementary RNA products but this reaction is strongly inhibited in the presence of 3'-mismatches.

Effects of RNA modifications on the activity and fidelity of RdRp

Previous studies identified several nucleotide analogues acting as efficient inhibitors of the RdRp activity. Several of them, including remdesivir, favipiravir and molnupiravir (β -D-N4-hydroxycytidine), which are clinically approved for COVID-19 treatment, can be incorporated into nascent RNA by SARS-CoV-2 RdRp. This leads to inhibition of coronaviral replication by either blocking RdRp translocation after their incorporation into the RNA product (for remdesivir

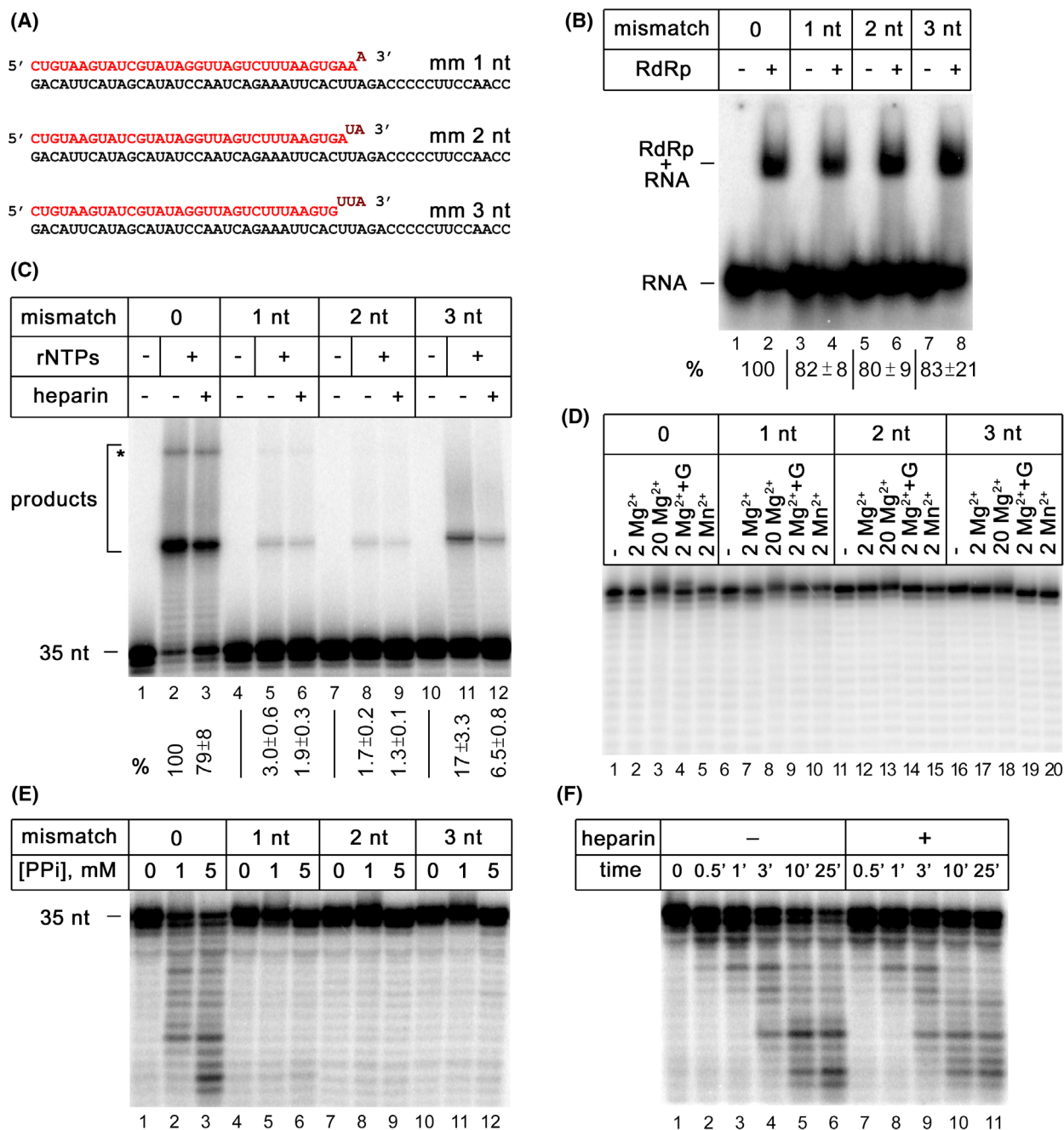


Fig. 3. Activities of SARS-CoV-2 RdRp with matched and mismatched primer–template substrates. (A) Structures of RNA substrates with mismatched 3'-ends. (B) Electrophoretic mobility shift assay with mismatched RNA substrates. The binding reactions were performed with 500 nM RdRp and 50 nM RNA. The percentages of bound RNA substrates relative to the fully matched primer–template RNA are shown below the gel (means and standard deviations from four independent experiments). (C) Extension of fully matched and mismatched RNA primers (0, 1, 2 or 3 nt mismatches) by RdRp in the absence and presence of heparin. Positions of extended RNA products are indicated. The percentages of RNA extension relative to the RdRp activity measured with fully matched RNA in the absence of heparin are shown below the gel (means and standard deviations from three independent experiments). (D) Analysis of RNA cleavage by RdRp in fully matched and mismatched complexes. Preformed RNA–RdRp complexes were incubated in the absence of divalent cations or presence of Mg²⁺ (2 or 20 mM) or Mn²⁺ (2 mM) for 15 min at 30 °C. GTP ('G') was added when indicated. A representative gel from two independent experiments is shown. (E) Pyrophosphorolysis of primer RNA by RdRp. The reactions were performed for 15 min at 30 °C at indicated concentrations of PPI. (F) Kinetics of pyrophosphorolysis in fully matched complexes in the absence or presence of heparin. The reactions were performed with 1 mM PPI for indicated time intervals. Panels E and F show representative gels from two independent experiments, which produced identical results.

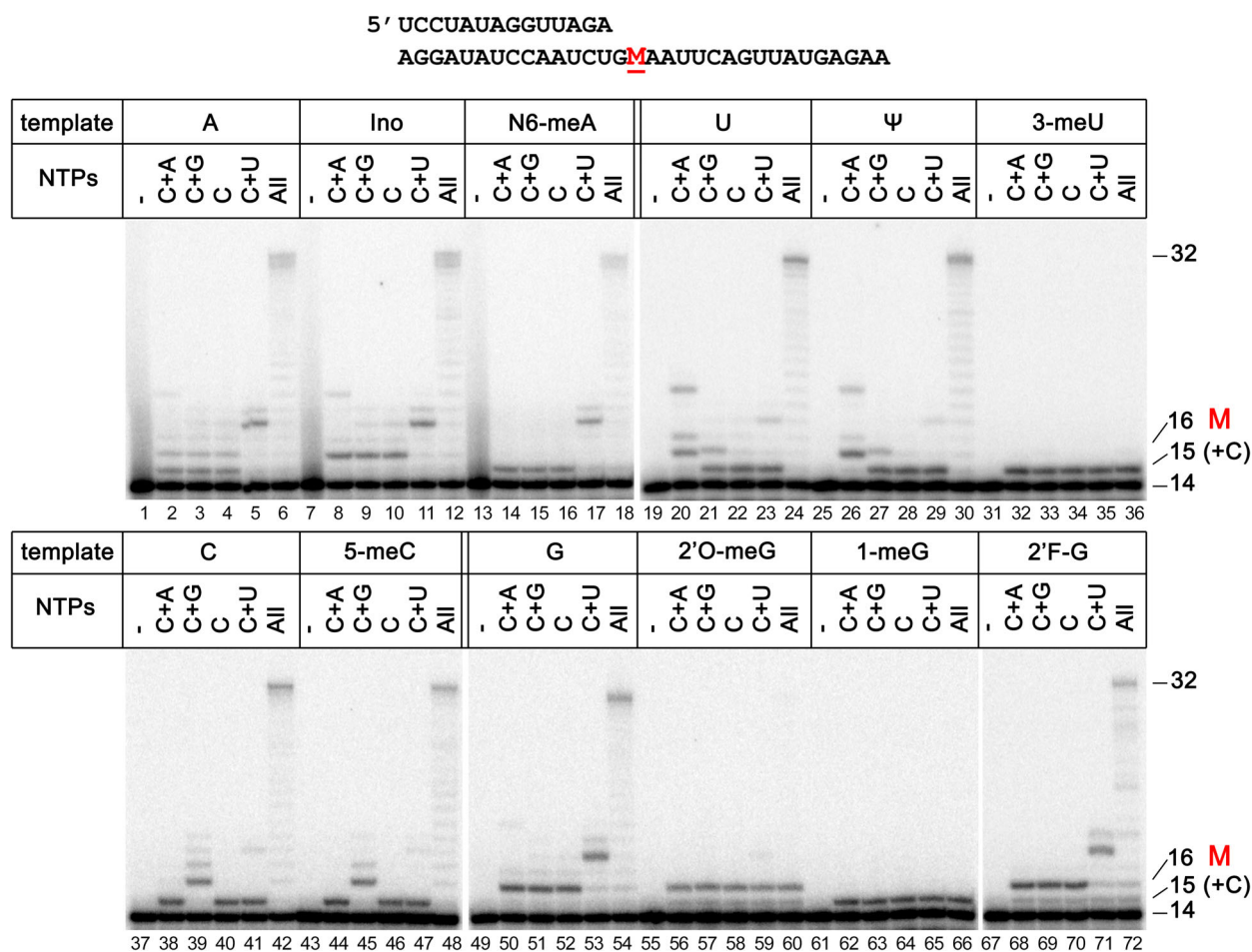


Fig. 5. Analysis of nucleotide incorporation by SARS-CoV-2 RdRp on modified RNA templates. The general structure of the RNA substrates containing site-specific modifications ('M') is shown on the top. The reactions were performed with modified RNA templates or corresponding control templates for 2 min at 30 °C in the presence of various combinations of NTPs. CTP was present in all reactions to allow one-nucleotide RNA extension opposite template G prior to the modified nucleotide. Positions of the starting RNA primer (14 nt), full-length RNA (32 nt) and the products of RNA extension stalled before (15 nt) or opposite the lesion (16 nt) are indicated.

could incorporate an additional nucleotide opposite the modification and stalled after it (with the total efficiency of stalling also approaching 90%; lane 11).

To better understand the effects of RNA modifications on the activity and fidelity of RdRp, we tested the incorporation of individual NTPs on all modified templates in comparison with corresponding unmodified RNAs (Fig. 5). In all cases, we added combinations of single NTPs together with CTP (to enable primer extension down to the site of modification). Control reactions contained all four NTPs (last lanes for each template in Fig. 5).

For all unmodified templates, the patterns of nucleotide incorporation were fully consistent with the template sequence and the primer was extended by one or more nucleotides depending on added NTP combinations. For example, the major extension

product on template A in the presence of CTP alone, CTP + ATP or CTP + GTP corresponded to the incorporation of a single C (with some readthrough as a result of misincorporation of C opposite next template A; Fig. 5, lanes 2–4), while in the presence of CTP + UTP RNA was extended by four nucleotides as a result of consecutive incorporation of C and three Us (lanes 5). In comparison, when inosine was present in place of template A, two nucleotides were successfully added to the primer 3'-end in the presence of CTP (as well as CTP + ATP or CTP + GTP, lanes 8–10), indicating that inosine promotes the incorporation of C instead of U. In the case of the N6-meA template, only correct C was added upstream of the site of modification, and no misincorporation of C was observed after opposite the modified base (lanes 14–16), indicating that

methylation of adenosine at the N6 position prevents mispairing of A with C.

For the pseudouridine template, the pattern of nucleotide addition was identical to the control template containing unmodified U (compare lanes 26–29 with 20–23). Similarly, 5'-methylcytosine had no effect on the fidelity of nucleotide incorporation in comparison with the control template C (compare lanes 44–47 with lanes 38–41). In contrast, in the case of 3-methyluridine and 1-methylguanine, only single nucleotide incorporation upstream of the modified nucleotide was observed with all tested nucleotide combinations (lanes 32–36 and 62–66), indicating that these lesions completely block RNA extension. For 2'-O-methylguanosine, RNA primer was extended by two nucleotides with all combinations of NTPs, as a result of the incorporation of two consecutive Cs, followed by strong RdRp stalling (lanes 56–60). This indicated that 2'-O-methyl ribose modification does not affect the fidelity of nucleotide incorporation but prevents further RdRp translocation. We, therefore, tested whether another 2'-ribose modification, 2'-fluoro ribose, can also affect RNA synthesis by RdRp. The pattern of nucleotide incorporation in the case of the 2'-fluoroguanosine (2'F-G) template was similar to the control unmodified template G, indicating that this modification does not strongly affect transcription (lanes 67–72). From these experiments, we can conclude that modifications in the RNA template can have highly varying effects on the activity of RdRp, depending on the type and position of modification.

Discussion

Since coronaviral RdRp plays a central role in genomic replication and transcription, it can serve as a promising target for the development of novel drugs inhibiting viral reproduction [1,3]. While most efforts have been focused on finding nucleoside analogues or non-nucleoside inhibitors targeting the catalytic cycle of RdRp [32–37], we have demonstrated that structural variations in the template and primer RNA can also strongly modulate its activity.

Structural studies revealed extensive protein–RNA contacts within the replicative complex of SARS-CoV-2, including an extended interface between the two nsp8 subunits of the replicase and the upstream RNA duplex (Fig. 1A,B) [11,12]. In accordance with previous studies [38], we have found that the activity of RdRp strongly depends on ionic strength, indicating that the replication complex is destabilized at high salt concentrations. However, we have also found that RdRp activity and its salt resistance are increased

when using longer RNA primers, suggesting that interactions of the RdRp replicase with the upstream product–template duplex are important for RdRp processivity. Furthermore, we have shown that SARS-CoV-2 RdRp can extend RNA on complementary template DNA, although with a decreased efficiency. Previously, SARS-CoV nsp12 was also shown to be active on a DNA template [39]. This ability may potentially be used for the development of DNA-based assays for testing RdRp activity [39]. At the same time, SARS-CoV-2 RdRp cannot extend DNA primers and is also strongly selective for NTP substrates. Structural analysis suggested that active-site residues of nsp12 (in particular, N691) recognize the 2'-OH group of the NTP, thereby enabling preferential incorporation of NTPs over dNTPs [12,40], but it remains to be established whether there are additional mechanisms of dNTP exclusion.

Cellular DNA-dependent RNA polymerases can remove misincorporated nucleotides from the RNA 3'-end through their intrinsic endonucleolytic RNA cleavage activity, which is increased at high pH and can be stimulated by dedicated RNA cleavage factors, Gre proteins in bacteria and TFIIS in eukaryotic Pol II [41–43]. This reaction requires backtracking of the transcription complex to allow attack of the active site of RNA polymerase on the internal phosphodiester bond [44–46]. In addition, noncomplementary NTPs were shown to induce exonucleolytic cleavage of a 3'-terminal nucleotide from the RNA product by bacterial DNA-dependent RNA polymerase [47] and also by hepatitis C virus (HCV) RNA-dependent RNA polymerase [48].

While no such activities have been reported for coronaviral RdRps, recent structural studies revealed that SARS-CoV-2 RdRp can form backtracked complexes with mismatched RNA substrates [15]. However, we could not detect either endonucleolytic or exonucleolytic RNA cleavage by RdRp under any tested conditions, including various concentrations of divalent metal cofactors (Mg^{2+} or Mn^{2+}), the presence of noncomplementary nucleotides or increased pH (Fig. 3 and unpublished observations). At the same time, we have demonstrated that SARS-CoV-2 RdRp can extend mismatched RNA primers, albeit with a low efficiency, which may result in error-prone genome replication. This likely explains the strong requirement of the proofreading exonuclease nsp14, which can remove mismatched nucleotides [16,17], for error-free replication and maintaining the genome integrity in coronaviruses [18,19].

Another potential way of the 3'-terminal RNA processing is pyrophosphorolysis, the reverse reaction for

RNA synthesis that removes the 3'-terminal residues and generates triphosphate nucleotides. Pyrophosphorolysis of nascent RNA can be performed by DNA-dependent RNA polymerases [49] and some previously studied RdRps including HCV RdRp [48] and hepatitis B virus RdRp [50]. We have demonstrated that coronaviral RdRp can also perform pyrophosphorolytic cleavage of the nascent RNA product, thus providing a way for 3'-nucleotide removal. However, pyrophosphorolysis is strongly blocked in the presence of 3'-mismatches, indicating that correct 3'-nucleotide positioning in the RdRp active site is critical for this reaction. Despite pyrophosphorolysis is unlikely to play a functional role in RNA processing by RdRp due to low pyrophosphate concentrations *in vivo*, it can be used as a tool to study the catalytic mechanism and the translocation state of RdRp.

Finally, we have tested the activity of SARS-CoV-2 RdRp on RNA templates containing various types of nucleotide modifications. Previously, it was reported that synthetic nucleotide analogues remdesivir and molnupiravir can inhibit RNA synthesis by SARS-CoV-2 RdRp when incorporated into the RNA template [25,27,28]. Analysis of RdRp activity on RNA templates containing a remdesivir residue demonstrated that RNA synthesis is strongly delayed before the modified nucleotide, as well as after incorporation of U opposite remdesivir [27,28]. Template molnupiravir modestly inhibits RNA synthesis by RdRp after nucleotide incorporation, but this inhibition can be overcome at high NTP concentrations. In addition, template molnupiravir induces misincorporation of A in addition to correct G and dramatically increases the level of mutagenesis [25]. Similarly, favipiravir incorporated into genomic RNA disrupts coronaviral replication by base pairing with both cytidine and uridine [26]. Structural studies demonstrated that template remdesivir clashes with residue A558 of nsp12, and mutations of adjacent V557 help RdRp bypass the site of modification [28]. Inhibition of RNA synthesis by template remdesivir is also alleviated by the V792I mutation in nsp12, which does not contact template RNA but may affect the dynamics of NTP incorporation [27,51]. It is likely that this and other substitutions in nsp12 may also modulate the effects of other types of RNA template modifications (see below) on the activity of RdRp.

We have demonstrated that natural RNA modifications can also modulate the activity of RdRp. It has been found that post-transcriptional modifications N6-meA, 5-meC and Ψ , which do not change base pairing, also do not affect primer RNA extension by RdRp. Interestingly, N6-meA may even increase the fidelity of

nucleotide incorporation by RdRp. At the same time, inosine, which occurs in RNA as a result of deamination of adenosine, promoted the misincorporation of C instead of U. Other tested RNA modifications, including 3-meU, 1-meG and 2'-O-meG, severely block RNA synthesis by RdRp. The strong effects of 3-meU and 1-meG on nucleotide incorporation by RdRp are likely explained by the disruption of complementary base pairing by these modifications. In contrast, 2'-O-meG contains a methyl group at the ribose moiety that does not affect base pairing. In comparison with 3-meU and 1-meG, 2'-O-methyl modification does not prevent nucleotide incorporation opposite the modified nucleotide but prevents further RNA extension. Previously, a strong inhibitory effect of the 2'-O-methyl RNA modification was also observed for reverse transcriptases [52]. Analysis of available structures of RdRp suggests that residues G683 and A685 of nsp12 can sterically interfere with translocation of the 2'-O-methyl group after nucleotide incorporation (Fig. 1C), while a smaller 2'-fluoro modification does not prevent RNA synthesis (Fig. 5).

Three of the analysed modifications, N6-methyladenosine, pseudouridine and 2'-O-methylribose, are naturally found in the SARS-CoV-2 genome [30]. N6-meA is introduced at about one or two dozen sites in coronaviral RNA by cellular RNA methyltransferases, and it was shown that inhibition of their activity delays coronaviral replication [30,53,54]. However, the exact functional role of this modification in the coronaviral life cycle remains to be established. Ψ is found in multiple positions in the SARS-CoV-2 genome and it was hypothesized that it may help the virus avoid the host immune response [55]. Both N6-meA and Ψ are easily bypassed by SARS-CoV-2 RdRp, without interfering with viral replication.

The 2'-O-methyl modification is introduced in the 5'-adenosine by the nsp16 methyltransferase and is a part of the 5'-cap that protects coronaviral RNA from degradation by host nucleases [29,56]. It is therefore essential that RdRp is stalled by this modification only after nucleotide incorporation (Fig. 4C), thus allowing complete genome replication. Intriguingly, multiple 2'-O-methyl modifications were found in coronaviral RNA in host cells, enriched in 5'- and 3'-untranslated regions [57]. Since the presence of these modifications in template RNA would interfere with genomic RNA synthesis, modified RNAs may represent a fraction of coronaviral genomes not participating in replication but potentially involved in translation. Indeed, only a fraction of each target site in coronaviral RNA is 2'-O-methylated [57]. Further studies are needed to establish whether there is any kind of functional

specialization between individual copies of the coronaviral genome, depending on their 2'-O-methylation.

An important goal of future studies is to test whether other natural RNA modifications or chemically induced RNA lesions can efficiently inhibit RdRps from SARS-CoV2 and other RNA viruses or change the fidelity of nucleotide incorporation by these polymerases. Further analysis of the effects of various classes of RNA modifications on coronaviral replication may lead to the discovery of highly potent template-specific RdRp inhibitors.

Materials and methods

Protein expression and purification

The nsp12, nsp7 and nsp8 genes of SARS-CoV-2 were codon optimized for expression in *E. coli* using OPTIMIZER (<http://genomes.urv.es/OPTIMIZER/>) and obtained by chemical synthesis. To purify the nsp12–nsp7–nsp8 complex, we co-expressed nsp12 containing a C-terminal His8 tag and an nsp7–His6–nsp8 polypeptide, in which the C terminus of nsp7 was fused to the N terminus of nsp8 through a hexahistidine linker. As was previously demonstrated, this approach allows to obtain highly active RdRp preparations [58]. Nsp12 and fused nsp7–His6–nsp8 were cloned into the pET-28 vector using HiFi DNA assembly master mix (New England Biolabs, Ipswich, MA, USA) under the control of T7 RNA polymerase promoter (T7-promoter_nsp12–His8_nsp7–His6–nsp8). *E. coli* BL21 (DE3) carrying the expression plasmid were grown at 37 °C until OD₆₀₀ 0.45 and chilled on ice for 30 min. IPTG was added to 0.1 mM and the cells were grown for 16 h at 37 °C. The cells were collected by centrifugation, resuspended in a 10-fold excess of lysis buffer (50 mM Tris–HCl, pH 7.9, 500 mM NaCl and 100 µg·mL⁻¹ PMSF) and lysed with a cell disruptor CF (Constant Systems Ltd, Daventry, UK). The lysate was centrifuged twice for 20 min at 45 000 g. The RdRp complex was purified from the cleared lysate by Ni-affinity chromatography using a 1 mL HiTrap TALON crude column (GE Healthcare, Chicago, Illinois, USA), pre-equilibrated with buffer containing 40 mM Tris–HCl, pH 7.9 and 500 mM NaCl. The sample was loaded at 1 mL·min⁻¹ and the column was washed with the same buffer containing 0, 30 and 100 mM imidazole. RdRp was eluted by the same buffer containing 300 mM imidazole. The eluted proteins were dialysed against buffer containing 50 mM Tris–HCl, pH 7.9, 50 mM NaCl, 5% glycerol, 0.1 mM DTT and 0.5 mM EDTA overnight. The sample was loaded on a HiScreenQ HP column (GE Healthcare) at the rate of 0.3 mL·min⁻¹. The column was washed with a linear salt gradient from 50 mM to 1 M NaCl in the same buffer for 150 min at the rate of 0.5 mL·min⁻¹. The complex of nsp12 with the nsp7–nsp8 fusion was eluted at ~ 300 mM NaCl. The fractions were concentrated using

Amicon Ultra-4 centrifugal filters Ultracel-50 K; NaCl, DTT and glycerol were added to 250 mM, 1 mM and 50%, respectively; the samples were aliquoted, frozen in liquid nitrogen and stored at –70 °C.

In vitro transcription

Analysis of substrate specificity of SARS-CoV-2 RdRp was performed using synthetic RNA and/or DNA oligonucleotide substrates corresponding to the template and primer strands (see Fig. 2A for oligonucleotide sequences). Primer oligonucleotides were 5'-labelled with γ -[P³²]-ATP and T4 polynucleotide kinase and mixed with template oligonucleotides in transcription buffer containing 10 mM Tris–HCl, pH 7.9, 10 mM KCl (or 30, 100 and 175 mM KCl when indicated), 2 mM MgCl₂ and 1 mM DTT at final concentrations of 2 and 4 µM respectively. The samples were incubated for 3 min at 95 °C, cooled down to 85 °C for 2 min and then cooled down to 25 °C at ~ 0.5 °C·min⁻¹. The assembled primer-template substrate was diluted with the same buffer (50 nM final primer concentration), mixed with RdRp (500 nM final concentration) and incubated for 15 min at 30 °C. Transcription reactions were initiated by adding 100 µM NTPs (together with 50 µg·mL⁻¹ heparin when indicated) at 30 °C and stopped with 1.1 volume of formamide with 100 µg·mL⁻¹ heparin after 15 min. The samples were heated at 95 °C for 4 min and rapidly loaded onto pre-running 15% denaturing PAGE. The reaction products were visualized by phosphor imaging with a Typhoon 9500 scanner (GE Healthcare).

To analyse *in vitro* transcription of modified RNA templates, primer–template substrates were prepared in the same way using modified or control RNA oligonucleotides shown in Fig. 4A. The annealed substrates were diluted with transcription buffer to 50 nM, mixed with 2.5 µM RdRp and incubated for 15 min at 30 °C. Transcription reactions were initiated by adding all four NTPs or combinations of individual NTPs with CTP (10 µM each). The reactions were performed for 10 min (when measuring RdRp activity with all NTPs) or for 2 min (when measuring the fidelity of incorporation of individual NTPs) at 30 °C and stopped with 1.1 volume of formamide with 100 µg·mL⁻¹ heparin. The samples were heated at 95 °C and separated in 19% denaturing PAGE.

Electrophoretic mobility shift assay

Complexes of oligonucleotide primer–template substrates with RdRp were assembled as described above in a transcription buffer containing 10 mM Tris–HCl, pH 7.9, 10 mM KCl, 1 mM DTT, 0.1 mM EDTA and 100 µg·mL⁻¹ BSA. The samples were mixed with 1/5 volume of loading dye buffer containing 50% glycerol, 2.5 \times TBE and separated in 4.5% native gel (acrylamide:bisacrylamide 37.5 : 1, 0.5 \times TBE and 10 V·cm⁻¹) at 25 °C.

Analysis of RNA cleavage and pyrophosphorolysis

For analysis of intrinsic RNA cleavage, primer–template substrates were prepared from oligonucleotides containing 1, 2 or 3 mismatched nucleotides in the primer 3'-end (Fig. 3A) in transcription buffer containing 10 mM Tris–HCl, pH 7.9, 10 mM KCl, 1 mM DTT and 0.1 mM EDTA and assembled with RdRp as described above. The reactions were initiated by adding different concentrations of divalent cations (2 or 20 mM MgCl₂ or MnCl₂) and proceeded for 15 min at 30 °C. GTP was added to 100 μM, when indicated. For analysis of pyrophosphorolysis, the complexes of RdRp were assembled in the same way. The reaction was initiated by adding sodium pyrophosphate (PPi) to a final concentration of 1 or 5 mM together with 7 mM MgCl₂, heparin was added to 50 μg·mL⁻¹ when indicated. The reaction was stopped after indicated time intervals and the samples were processed as described above.

Acknowledgements

This study was supported by the Russian Foundation for Basic Research grant 20-04-60571. We thank A. Olina and N. Miropolskaya for help in figure preparation.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

IP, DE and AK conceptualized the study, IP performed experiments, DE developed the expression system for RdRp, IP and AK interpreted the results and AK wrote the manuscript.

Data availability statement

The data supporting the findings of this study are available in the figures within this article; data from replicate experiments are available from the corresponding author upon request.

References

- Lu L, Su S, Yang H, Jiang S. Antivirals with common targets against highly pathogenic viruses. *Cell*. 2021;**184**:1604–20.
- Hillen HS. Structure and function of SARS-CoV-2 polymerase. *Curr Opin Virol*. 2021;**48**:82–90.
- Malone B, Urakova N, Snijder EJ, Campbell EA. Structures and functions of coronavirus replication-transcription complexes and their relevance for SARS-CoV-2 drug design. *Nat Rev Mol Cell Biol*. 2022;**23**:21–39.
- Gorbalenya AE, Enjuanes L, Ziebuhr J, Snijder EJ. Nidovirales: evolving the largest RNA virus genome. *Virus Res*. 2006;**117**:17–37.
- Gulyaeva AA, Gorbalenya AE. A nidovirus perspective on SARS-CoV-2. *Biochem Biophys Res Commun*. 2021;**538**:24–34.
- Hartenian E, Nandakumar D, Lari A, Ly M, Tucker JM, Glaunsinger BA. The molecular virology of coronaviruses. *J Biol Chem*. 2020;**295**:12910–34.
- Kirchdoerfer RN, Ward AB. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nat Commun*. 2019;**10**:2342.
- Hillen HS, Kokic G, Farnung L, Dienemann C, Tegunov D, Cramer P. Structure of replicating SARS-CoV-2 polymerase. *Nature*. 2020;**584**:154–6.
- 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.
- 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.
- 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.e13.
- Hillen HS, Kokic G, Farnung L, Dienemann C, Tegunov D, Cramer P. Structure of replicating SARS-CoV-2 polymerase. *bioRxiv*. 2020. <https://doi.org/10.1101/2020.04.27.063180>
- 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.e10.
- Yan L, Zhang Y, Ge J, Zheng L, Gao Y, Wang T, et al. Architecture of a SARS-CoV-2 mini replication and transcription complex. *Nat Commun*. 2020;**11**:5874.
- 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.
- Yan L, Yang Y, Li M, Zhang Y, Zheng L, Ge J, et al. Coupling of N7-methyltransferase and 3'-5' exoribonuclease with SARS-CoV-2 polymerase reveals mechanisms for capping and proofreading. *Cell*. 2021;**184**:3474–85.e11.
- Robson F, Khan KS, Le TK, Paris C, Demirbag S, Barfuss P, et al. Coronavirus RNA proofreading: molecular basis and therapeutic targeting. *Mol Cell*. 2020;**79**:710–27.
- Ogando NS, Zevenhoven-Dobbe JC, van der Meer Y, Bredenbeek PJ, Posthuma CC, Snijder EJ. The

- enzymatic activity of the nsp14 exoribonuclease is critical for replication of MERS-CoV and SARS-CoV-2. *J Virol.* 2020;**94**:e01246-20.
- 19 Smith EC, Blanc H, Surdel MC, Vignuzzi M, Denison MR. Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics. *PLoS Pathog.* 2013;**9**:e1003565.
 - 20 Ferron F, Subissi L, Silveira De Moraes AT, Le NTT, Sevajol M, Gluais L, et al. Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. *Proc Natl Acad Sci USA.* 2018;**115**:E162–71.
 - 21 Liu C, Shi W, Becker ST, Schatz DG, Liu B, Yang Y. Structural basis of mismatch recognition by a SARS-CoV-2 proofreading enzyme. *Science.* 2021;**373**:1142–6.
 - 22 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.
 - 23 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.
 - 24 Shannon A, Le NT, Selisko B, Eydoux C, Alvarez K, Guillemot JC, et al. Remdesivir and SARS-CoV-2: structural requirements at both nsp12 RdRp and nsp14 exonuclease active-sites. *Antiviral Res.* 2020;**178**:104793.
 - 25 Gordon CJ, Tchesnokov EP, Schinazi RF, Gotte M. Molnupiravir promotes SARS-CoV-2 mutagenesis via the RNA template. *J Biol Chem.* 2021;**297**:100770.
 - 26 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.
 - 27 Gordon CJ, Lee HW, Tchesnokov EP, Perry JK, Feng JY, Bilello JP, et al. Efficient incorporation and template-dependent polymerase inhibition are major determinants for the broad-spectrum antiviral activity of remdesivir. *J Biol Chem.* 2022;**298**:101529.
 - 28 Tchesnokov EP, Gordon CJ, Woolner E, Kocinkova D, Perry JK, Feng JY, et al. Template-dependent inhibition of coronavirus RNA-dependent RNA polymerase by remdesivir reveals a second mechanism of action. *J Biol Chem.* 2020;**295**:16156–65.
 - 29 Nencka R, Silhan J, Klima M, Otava T, Kocek H, Krafcikova P, et al. Coronaviral RNA-methyltransferases: function, structure and inhibition. *Nucleic Acids Res.* 2022;**50**:635–50.
 - 30 Izadpanah A, Rappaport J, Datta PK. Epitranscriptomics of SARS-CoV-2 infection. *Front Cell Dev Biol.* 2022;**10**:849298.
 - 31 Eyler DE, Franco MK, Batool Z, Wu MZ, Dubuke ML, Dobosz-Bartoszek M, et al. Pseudouridylation of mRNA coding sequences alters translation. *Proc Natl Acad Sci USA.* 2019;**116**:23068–74.
 - 32 Chien M, Anderson TK, Jockusch S, Tao C, Li X, Kumar S, et al. Nucleotide analogues as inhibitors of SARS-CoV-2 polymerase, a key drug target for COVID-19. *J Proteome Res.* 2020;**19**:4690–7.
 - 33 Jockusch S, Tao C, Li X, Anderson TK, Chien M, Kumar S, et al. A library of nucleotide analogues terminate RNA synthesis catalyzed by polymerases of coronaviruses that cause SARS and COVID-19. *Antiviral Res.* 2020;**180**:104857.
 - 34 Seifert M, Bera SC, van Nies P, Kirchdoerfer RN, Shannon A, Le TT, et al. Inhibition of SARS-CoV-2 polymerase by nucleotide analogs from a single-molecule perspective. *Elife.* 2021;**10**:e70968.
 - 35 Ju J, Li X, Kumar S, Jockusch S, Chien M, Tao C, et al. Nucleotide analogues as inhibitors of SARS-CoV polymerase. *Pharmacol Res Perspect.* 2020;**8**:e00674.
 - 36 Tian L, Qiang T, Liang C, Ren X, Jia M, Zhang J, et al. RNA-dependent RNA polymerase (RdRp) inhibitors: the current landscape and repurposing for the COVID-19 pandemic. *Eur J Med Chem.* 2021;**213**:113201.
 - 37 Yin W, Luan, X, Li, Z, Zhou, Z, Wang, Q, Gao, M, Wang, X, Zhou, F, Shi, J, You, E, Liu, M, Wang, Q, Jiang, Y, Jiang, H, Xiao, G, Zhang, L, Yu, X, Zhang, S & Eric Xu, H (2021) Structural basis for inhibition of the SARS-CoV-2 RNA polymerase by suramin, *Nat Struct Mol Biol* **28**, 319–325.
 - 38 Wang B, Svetlov V, Wolf YI, Koonin EV, Nudler E, Artsimovitch I. Allosteric activation of SARS-CoV-2 RNA-dependent RNA polymerase by Remdesivir triphosphate and other phosphorylated nucleotides. *MBio.* 2021;**12**:e0142321.
 - 39 te Velthuis AJ, Arnold JJ, Cameron CE, van den Worm SH, Snijder EJ. The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. *Nucleic Acids Res.* 2010;**38**:203–14.
 - 40 Malone B, Perry JK, Olinares PDB, Chen J, Appelby TK, Feng JY, et al. Structural basis for substrate selection by the SARS-CoV-2 replicase. *bioRxiv.* 2022. <https://doi.org/10.1101/2022.05.20.492815>
 - 41 Borukhov S, Sagitov V, Goldfarb A. Transcript cleavage factors from *E. coli*. *Cell.* 1993;**72**:459–66.
 - 42 Orlova M, Newlands J, Das A, Goldfarb A, Borukhov S. Intrinsic transcript cleavage activity of RNA polymerase. *Proc Natl Acad Sci USA.* 1995;**92**:4596–600.
 - 43 Izban MG, Luse DS. SII-facilitated transcript cleavage in RNA polymerase II complexes stalled early after initiation occurs in primarily dinucleotide increments. *J Biol Chem.* 1993;**268**:12864–73.
 - 44 Komissarova N, Kashlev M. Transcriptional arrest: Escherichia coli RNA polymerase translocates

- backward, leaving the 3' end of the RNA intact and extruded. *Proc Natl Acad Sci USA*. 1997;**94**:1755–60.
- 45 Nudler E, Mustaev A, Lukhtanov E, Goldfarb A. The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell*. 1997;**89**:33–41.
- 46 Nudler E. RNA polymerase backtracking in gene regulation and genome instability. *Cell*. 2012;**149**:1438–45.
- 47 Sosunov V, Sosunova E, Mustaev A, Bass I, Nikiforov V, Goldfarb A. Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *EMBO J*. 2003;**22**:2234–44.
- 48 Jin Z, Leveque V, Ma H, Johnson KA, Klumpp K. NTP-mediated nucleotide excision activity of hepatitis C virus RNA-dependent RNA polymerase. *Proc Natl Acad Sci USA*. 2013;**110**:E348–57.
- 49 Rozovskaya TA, Chenchik AA, Beabealashvili R. Processive pyrophosphorolysis of RNA by *Escherichia coli* RNA polymerase. *FEBS Lett*. 1982;**137**:100–4.
- 50 Urban S, Urban S, Fischer KP, Tyrrell DL. Efficient pyrophosphorolysis by a hepatitis B virus polymerase may be a primer-unblocking mechanism. *Proc Natl Acad Sci USA*. 2001;**98**:4984–9.
- 51 Stevens LJ, Puijssers AJ, Lee HW, Gordon CJ, Tchesnokov EP, Gribble J, et al. Mutations in the SARS-CoV-2 RNA dependent RNA polymerase confer resistance to remdesivir by distinct mechanisms. *Sci Transl Med*. 2022;**14**:eabo0718.
- 52 Houlihan G, Arangundy-Franklin S, Porebski BT, Subramanian N, Taylor AI, Holliger P. Discovery and evolution of RNA and XNA reverse transcriptase function and fidelity. *Nat Chem*. 2020;**12**:683–90.
- 53 Burgess HM, Depledge DP, Thompson L, Srinivas KP, Grande RC, Vink EI, et al. Targeting the m(6)a RNA modification pathway blocks SARS-CoV-2 and HCoV-OC43 replication. *Genes Dev*. 2021;**35**:1005–19.
- 54 Zhang X, Hao H, Ma L, Zhang Y, Hu X, Chen Z, et al. Methyltransferase-like 3 modulates severe acute respiratory syndrome Coronavirus-2 RNA N6-Methyladenosine modification and replication. *MBio*. 2021;**12**:e0106721.
- 55 Fleming AM, Mathewson NJ, Howpay Manage SA, Burrows CJ. Nanopore dwell time analysis permits sequencing and conformational assignment of Pseudouridine in SARS-CoV-2. *ACS Cent Sci*. 2021;**7**:1707–17.
- 56 Wilamowski M, Sherrell DA, Minasov G, Kim Y, Shuvalova L, Lavens A, et al. 2'-O methylation of RNA cap in SARS-CoV-2 captured by serial crystallography. *Proc Natl Acad Sci USA*. 2021;**118**:e2100170118.
- 57 Yang SL, DeFalco L, Anderson DE, Zhang Y, Aw JGA, Lim SY, et al. Comprehensive mapping of SARS-CoV-2 interactions in vivo reveals functional virus-host interactions. *Nat Commun*. 2021;**12**:5113.
- 58 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.
- 59 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.