Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
SARS-CoV-2 RdRp uses NDPs as a substrate and is able to incorporate NHC into RNA from diphosphate form molnupiravir

Maofeng Wang, Cancan Wu, Nan Liu, Fengyu Zhang, Hongjie Dong, Shuai Wang, Min Chen, Xiaqiong Jiang, Kundi Zhang, Lichuan Gu

State Key Laboratory of Microbial Technology, Shandong University, 72 Binhai Road, Qingdao 266237, PR China

ABSTRACT

The coronavirus disease 2019 has been ravaging throughout the world for three years and has severely impaired both human health and the economy. The causative agent, severe acute respiratory syndrome coronavirus 2 employs the viral RNA dependent RNA polymerase (RdRp) complex for genome replication and transcription, making RdRp an appealing target for antiviral drug development. Systematic characterization of RdRp will undoubtedly aid in the development of antiviral drugs targeting RdRp. Here, our research reveals that RdRp can recognize and utilize nucleoside diphosphates as a substrate to synthesize RNA with an efficiency of about two thirds of using nucleoside triphosphates as a substrate. Nucleoside diphosphates incorporation is also template-specific and has high fidelity. Moreover, RdRp can incorporate β-D-N4-hydroxycytidine into RNA while using diphosphate form molnupiravir as a substrate. This incorporation results in genome mutation and virus death. It is also observed that diphosphate form molnupiravir is a better substrate for RdRp than the triphosphate form molnupiravir, presenting a new strategy for drug design.

1. Introduction

The coronavirus disease 2019 (COVID-19) pandemic has severely impacted global human health and economy since the outbreak in late 2019 [1,2]. The causative agent, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is homologous to 2002 severe acute respiratory syndrome coronavirus (SARS-CoV) with a genome sequence similarity of 79 % [3,4], forming a sister clade to SARS-CoV and considered a newly β-coronavirus [5]. As of the date of writing, >491.75 million infections and >6.17 million deaths have been confirmed (https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports). Although a variety of vaccines have been developed to contain the pandemic, the rapidly spreading mutant strains presses need to develop effective medication or treatments to counter the virus [6–9].

SARS-CoV-2 is a positive-sense single-stranded RNA virus [5] with a large genome of approximately 30 kb organized in 14 open reading frames [4,10–12]. Replication of the genome and transcription of genes are all dependent on a protein complex known as RNA dependent RNA polymerase (RdRp). The RdRp of SARS-CoV-2 contains a catalytic subunit (non-structural protein 12, nsp12) and two accessory subunits (nsp7 and nsp8) [13,14]. Processivity is enhanced by the presence of nsp7, nsp8, and bound to nsp12 in a 1:2:1 stoichiometry [15,16]. Due to the significant role of RdRp and the lack of homologues in humans, RdRp has become the most appealing target for anti-coronavirus drug development [17,18].

To date, the most promising RdRp inhibitors are nucleotide/nucleoside analogs (NAs). NA prodrugs are designed to be metabolized to the active 5′-triphosphate form (5′-TP) once inside cells [29], thus they can compete with endogenous nucleotides for incorporation into newly synthesized viral RNA, resulting in viral genomes lethal mutation or inactivation [30]. Remdesivir is a broad antiviral drug approved for the treatment of SARS-CoV-2, which can be incorporated into RNA to sterically hinder RdRp, thereby blocking viral RNA synthesis [31]. Other NAs (favipiravir, ribavirin) are integrated into nascent viral RNA at high rates but are not recognized as endogenous nucleotides in subsequent rounds of replication, increasing the mutation rate and resulting in an inviable genome, a process known as “lethal mutagenesis” [32–34]. The recently approved NA antiviral drug molnupiravir significantly increases the frequency of viral RNA mutations and impairs SARS-CoV-2 in both animal models and humans [35,36]. However, the application of NAs also bring adverse reactions and side effects, including elevated uric acid and hemolytic anemia [37]. Reported cases had also suggest
that monopilavir can be incorporated into host DNA with the same mutagenesis activity that affects viral replication, putting the host at risk of mutagenesis [38].

By far none of these RdRp targeting drugs is effective enough to fulfill the need for treatment of COVID-19. To develop a better RdRp targeting drug we need to better understand the properties of RdRp and how these inhibitors function. Previous studies showed that some DNA polymerases [39,40] and HIV reverse transcriptase [41] can use deoxy-nucleoside diphosphates (dNDPs) as a substrate to synthesize DNA, and Escherichia coli RNA polymerase [42] use nucleoside diphosphates (NDPs) to synthesize RNA. Although both dNDPs and NDPs showed very weak activity in these experiments, the data prompted us to test if NDPs are an active substrate for RdRp. Our purpose is to find if RdRp has some previously unknown properties which can be taken advantage of. New discoveries about the RdRp function may not only broaden our understanding of SARS-CoV-2 but also yield new strategies for antiviral drug design. This would undoubtedly aid in COVID-19 treatment. In this study, we found that SARS-CoV-2 RdRp can recognize and utilize NDPs as a substrate to synthesize RNA with an efficiency of about two-thirds of using nucleoside triphosphates (NTPs) as a substrate. In addition, we proved that monopilavir is not only fully functional in its diphosphate form but also more active than in triphosphate form. This could imply a new strategy for designing COVID-19 treatment NA drugs with more efficacy and fewer side effects.

2. Materials and methods

2.1. Protein expression and purification

Preparation of SARS-CoV-2 RdRp, composed of nsp12, nsp7 and two copies of the nsp8 subunits, was carried out as described in ref. [43]. The SARS-CoV-2 nsp12 and nsp7 gene were cloned into a modified PET-21b vector with the C-terminus possessing a 6×His-tag. The nsp8 gene was cloned into the modified PET-32a vector with the N-terminus possessing a trxA-His6-tag and PreScission Protease (PPase) site. The nsp12-pET-21b and nsp8-pET-32a plasmids were transformed into E. coli BL21 (DE3) and the transformed cells were cultured at 37 °C in LB with a final concentration of 100 μg/ml ampicillin. The nsp7-pET-21b plasmid was transformed into E. coli Rosetta-gami2 (DE3) and the transformed cells were cultured at 37 °C in LB containing a final concentration of 100 μg/ml ampicillin and 25 μg/ml chloramphenicol.

Bacterial cultures were incubated by shaking at 200 rpm, 37 °C to an OD600 of 0.8, then the temperature was lowered to 16 °C and a final concentration of 0.2 mM of isopropyl β-D-thiogalactoside (IPTG) was added to induce protein expression for 20 h. Subsequently, the cells were harvested by centrifugation at 5000 × g for 15 min at 4 °C. The pellet was resuspended in lysis buffer containing 25 mM TrisHCl pH 8.0, 250 mM NaCl, 4 mM MgCl2, 10 % Glycerol and homogenized with a high-pressure cell dismuter at 4 °C. The lysate was centrifuged at 37,000 × g for 50 min at 4 °C, and the supernatant was then loaded onto Ni-NTA affinity chromatography column for purification. Nsp12, nsp7 were eluted with elution buffer (25 mM TrisHCl pH 8.0, 200 mM NaCl, 4 mM MgCl2, 250 mM Imidazole). Nsp8 was subject to on-column tag cleavage by PPase (purified in our laboratory) and then eluted by the lysis buffer. Next, all proteins were purified by ion exchange chromatography (Source 15Q HR 16/10, GE Healthcare, Boston, MA, USA) and size exclusion chromatography (Superdex 200 10/300GL, GE Healthcare, Boston, MA, USA) with 25 mM TrisHCl pH 8.0, 200 mM NaCl, 4 mM MgCl2.

For nsp12-nsp7-nsp8 complex assembly, protein samples were mixed with the molar ratio of nsp7:nsp8:nsp12 = 2:2:1 at 4 °C overnight. The incubated RdRp complex was then concentrated with a 100kDa molecular weight cut-off centrifugal filter unit (Millipore Corporation, Billerica, MA, USA) and then further purified by size exclusion chromatography using a Superdex 200 10/300 GL column in 25 mM TrisHCl pH 8.0, 200 mM NaCl, 4 mM MgCl2. SARS-CoV RdRp was purified by the same procedure. The non-structural 5 (NS5) RdRp domain of Zika Virus (ZIKV) was purified as previously described [44].

2.2. Preparation of fluorescently labeled RNA for polymerase activity

RNA template-product duplex was designed according to the published SARS-CoV-2 RNA extension assays [43]. FAM labeled 13 nt oligonucleotide with the sequence of FAM-CCUAAUUGGAUAU and the 23 nt complementary RNA strand with the sequence AGUUAAUUCUCAUUAGC were synthesized for polymerase activity assay. Quality of the 13 + 23 nt RNA was checked three months later to make sure no degradation occur the throughout the experiment (Supplementary Fig. S3). A 26 nt RNA strand with the sequence UAGCUUGCCUUAAUUCUCAUUAGC was synthesized for substrate efficiency assay. And a 29 nt RNA strand with the sequence AGUUAAUUGGAUAUUCUCAUUAGC were used for the detection of mutation rates due to the incorporation of β-d-N4-hydroxycytidine (NHC) triphosphate (MTP) and NHC diphosphate (MDP) into RNA. All un-modified and 5’ FAM-labeled RNA oligonucleotides and DNA primers used in other experiments were purchased from Tsingke Biotechnology Co., Ltd. (Beijing, China). The RNA strands were mixed in equal molar ratio in DEPC water, annealed by heating it to 95 °C for 10 min and gradually cooling to room temperature to make the RNA duplexes.

2.3. RdRp polymerase activity assays

DEPC-treated water was used in the preparation of all solutions. RdRp at final concentration of 2 μM was incubated with 200 mM 13 + 23 nt RNA duplex and 50 μM NTPs (or NDPs/nucleoside monophosphates (NMPs)) (All nucleotides are purchased from Sigma-Aldrich, Shanghai, China) in a 20 μl reaction buffer containing 20 mM TrisHCl pH 8.0, 10 mM KCl, 10 mM MgCl2, 0.01 % Triton-X100, 1 mM DTT for 1 min at 37 °C [43], and the reactions were stopped with 2 × stop buffer (10 M urea, 50 mM EDTA). Eventually, glycerol was added to a concentration of 6.5 %. 10 μl RNA product for each reaction was resolved on 20 % denaturing polyacrylamide-urea gels (2.5 g urea, 1.2 ml s × 3 TBE, 3 ml 40 % acrylamide, 40 μl 0.1 g/ml ammonium persulfate, 10 μl TEMED) and imaged with a Tanon-5200 Multi Fluorescence Imager. The experiments described were performed in triplet, unless specified otherwise.

2.4. RNA extension assays with NDPs as substrate

RdRp of different viruses at final concentration of 2 μM was incubated with 200 mM 13 + 23 nt RNA duplex and 50 μM NTPs (or NDPs/NMPs) adenosine diphosphate (ADP)/guanosine diphosphate (GDP)/cytidine diphosphate (CDP)/uridine diphosphate (UDP) in a 20 μl reaction buffer containing 20 mM TrisHCl pH 8.0, 10 mM KCl, 10 mM MgCl2, 0.01 % Triton-X100, 1 mM DTT for 1 min at 37 °C.

2.5. Sanger sequencing

The RNA product, primer 1 (CCCGTCGAGCCGGAGTAACTT) and deoxy-ribonucleoside triphosphate (dNTPs) were gently mixed at 65 °C for 10 min. The mixture was cooled down to 24 °C then 10 × MuMv buffer and MuMv reverse transcriptase (purified in our laboratory) were added and incubated for 2 h to obtain cDNA. Thereafter, the reverse transcriptase was inactivated at 70 °C for 15 min. Primer 2 (CCGGGATCCGGCGGTATGGA) was then added, and the product was amplified by PCR catalyzed by the high-fidelity DNA polymerase pfu-Phusion (purified in our laboratory). The PCR product was digested with Xbol and BamHI (Purchased from Takara Bio, Tianjin, China) and ligated into pET-15b vector, and transformed into E. coli DH5α to obtain clones. Plasmids from positive clones were then sequenced with T7 universal primers.
2.6. High-throughput sequencing

The dsRNA products were mixed with primer 1 (cagatgcaagctct-ggacaaaaagtgacaccaacaCGGCTGAGGCGAGTTAACC) and dNTPs at 65 °C for 10 min and then cooled down to 24 °C for primer annealing. 10 × MMulV buffer, RNase inhibitor, and MMulV reverse transcriptase were added to the above system at 24 °C for 2 h to obtain cDNA, and then the enzymes were inactivated at 70 °C for 15 min. Subsequently, primer 2 (cagattaatataaatcttacgagggagtctacaCGGCGGATCGCG GCTATGTGAGA) and high-fidelity DNA polymerase pfu-Phusion were added to perform PCR amplification. PCR products were then purified by agarose gel electrophoresis for high-throughput sequencing (Sangon Biotech, Shanghai, China).

2.7. Substrate efficiency assay by agarose gel electrophoresis for high-throughput sequencing (Sangon GCTATGTGAGA) and high-fidelity DNA polymerase pfu-Phusion were added to perform PCR amplification. PCR products were then purified by agarose gel electrophoresis for high-throughput sequencing (Sangon Biotech, Shanghai, China).

2.7. Substrate efficiency assay

RNA extension was performed by incubating 2 μM RdRp, 250 nM 13 + 23 nt RNA duplex, 50 μM substrate (which are NTPs, NDPs) in a 20 μl reaction system at 37 °C for 0, 10, 20, 30, 40, 50, 60 s. RdRp catalyzed reactions each containing three NTPs and one NDP (ATP, guanosine triphosphate (GTP), cytidine triphosphate (CTP), uridine triphosphate (UTP); adenosine triphosphate (ATP), GDP, CTP, UTP; ATP, GTP, CDP, UTP and ATP, GTP, CTP, UDP) and 13 + 26 nt RNA duplex was performed for 1 min. The reactions were stopped with 2 × stop buffer. Grayscale analysis was then carried out using ImageJ and illustrated by using OriginPro 8 (OriginLab, Northampton, Massachusetts, USA).

2.8. RNA extension assays with MDP as substrate

A 13 + 26 nt RNA duplex was used in this assay (Fig. 2c). The 5′ end of the RNA product strand was labeled with a 6-carboxyfluorescein (FAM) group, which allows us to monitor RNA elongation. RdRp can incorporate NHC monophosphate into RNA at the position where the nucleoside in the template is G or A. MDP and MTP were purchased from MedChemExpress (Shanghai, China).

For RNA extension RdRp at final concentration of 2 μM was incubated with 200 nM 13 + 26 nt RNA duplex and 50 μM each of ATP, GTP, UTP or 50 μM each of ATP, GTP, CTP with or without 50 μM MTP/MDP in 20 μl reaction buffer for 1 min at 37 °C. In the experiment to detect the specificity of other viral RdRp using MTP and MDP as substrate, because ZIKV NS5 had a poor polymerization effect on 13 + 23 nt RNA, 13 + 23 nt RNA was used as the template. The reactions were stopped with 2 × stop buffer. 10 μl RNA product for each reaction was resolved on 20 % denaturing polyacrylamide-urea gels and imaged with a Tanon-5200 Multi Fluorescence Imager to confirm the extension. Grayscale analysis was then carried out using ImageJ and illustrated by using OriginPro 8.

2.9. Incorporation mutation rate detection

RdRp at final concentration of 2 μM was incubated with 200 nM 13 + 29 nt RNA duplex and 50 μM each of ATP, GTP, UTP; 0 or 800 nM of CTP; 50 μM of MTP/MDP in 20 μl reaction buffer for 1 min at 37 °C. Mutation rates were determined by High-throughput sequencing as mentioned above. Raw reads were filtered according to three steps: 1) Removing adaptor sequence if reads contains by cutadapt (v 1.2.1); 2) Removing low quality bases from reads 3′ to 5′ (Q < 20) by PRINSEQ-lite (v 0.20.3); 3) Removing chimeras sequence by usearch software (v11.0.667) with de novo mode by default parameter. And the remaining clean data were used for further analysis. A frequency of 0.001 was used as a cutoff for variants. The absolute number and type of mutations in three consecutive C bases of 13 + 29 nt RNA are reported. The percentage of the total mutations for each specific mutation type was calculated using these numbers. The difference in percentage for each class of mutation compared with NTPs control is referred to as the relative proportion of these mutations.

2.10. Fluorescence assay for T7 RNA polymerase (RNAP) transcription

Transcription of T7 RNAP was measured by a fluorescence assay which employs a pair of nucleic acid probes with fluorescent and quenching groups respectively to detect mRNA production. When the probes pair with the mRNA FAM and quenching group gets closer and the fluorescence produced by FAM will be quenched. The change of the fluorescent signal is proportional to the mRNA production.

3. Results

3.1. RdRp synthesizes RNA by using both NTPs and NDPs as a substrate

Since RdRp is by far one of the most attractive targets for antiviral drug development, a thorough characterization of RdRp polymerase activity will, without doubt, greatly contribute to discovering the effective strategy for drug screening [17,18]. The RdRp of SARS-CoV-2 was purified by assembling nsp12, nsp7 and nsp8 (Fig. 1a). Based on previous research, SARS-CoV-2 RdRp polymerase activity was measured by conducting a conventional RNA elongation assay [43]. When NTPs was added, elongation catalyzed by RdRp complex occurred on the RNA duplex, resulting in an intact double-stranded RNA product (Fig. 1c).

To test if RdRp uses NDPs or NMPs as a substrate, we performed an RNA extension assay by using NTPs, NDPs or NMPs as substrate, respectively. The result revealed that RdRp synthesized RNA products of the same length regardless of NTPs or NDPs as substrate (Fig. 1d). To examine the template specificity of this RNA synthesis, 13 + 23 nt RNA was extended in the presence of ADP, GDP, CDP, and UDP, respectively (Fig. 1b). Only the cognate ADP promoted RNA extension, indicating that NDPs incorporation is template-specific (Fig. 1e).

This result also raised the question of whether the RNA made from NDPs has the right sequence. Hence, two experiments were performed: single-clonal sequencing of ligated plasmids, and high-throughput sequencing of PCR products. First, the cDNA was synthesized through reverse transcription of the RNA products, and then amplified by 10 cycles of PCR with primers containing the restriction sites of XhoI and BamHI. PCR products were then ligated into the vector pET-15b and used to transform E. coli DH5α strain. Plasmids extracted from the positive clones were sequenced by Sanger sequencing. The results gave a DNA sequence that correlated with the sequence of the nascent RNA, which is the complementary sequence of the RNA template (Fig. 1f). Second, the cDNA amplified by 10 cycles of PCR was sequenced by high-throughput sequencing (Fig. 1g). At least 50,000 sequences were obtained for each product (NTPs is Supplementary File SI.H1, NDPs is Supplementary File SI.H2). The same result was obtained whether NDPs or NTPs was used as substrate. These two experiments clearly indicated confirm that the RNA synthesized from NDPs has the right sequence. The ADP, GDP, CDP and UDP used in these experiments were characterized by mass spectrometry respectively to make sure these NDPs were not contaminated by NTPs (Supplementary Fig. S1).

3.2. NDPs are nearly two thirds as efficient as NTPs as a substrate of RdRp

Previous studies indicated that although some DNA polymerases and RNA polymerases are able to use diphosphate form substrate, the activity of the diphosphate form substrate, however, is much lower than that of the triphosphate form substrate [40–42]. To determine if this is also the case for RdRp we performed a kinetics study by using NTPs and NDPs as a substrate respectively. The 13 + 23 nt RNA duplex was used as a template and the elongation of product over time was observed (Fig. 2a). Kinetics curves indicated that the relative incorporation efficiency of NDPs is about two thirds (67.94 % ± 2.66 %) of that of NTPs (Fig. 2b). These data suggested that the incorporation rate for NDPs is slower than that of NTPs. Subsequently, we decided to test if all four
Fig. 1. RdRp uses both NTPs and NDPs as a substrate. a Size-exclusion chromatogram of the SARS-CoV-2 nsp12-nsp7-nsp8 (RdRp) complex. RdRp was also characterized by SDS-PAGE. b The 13 + 23 nt RNA template-product duplex. The direction of RNA extension is shown. The colour of the depicted circles indicates the experimental design: blue, RNA template strand; red, RNA product strand. The 5′ end of the RNA product contains a FAM fluorescent label. c Incubation of the RdRp with RNA duplex and NTPs leads to RNA extension. Nsp12 alone has almost no polymerase activity, and requires nsp7 and nsp8 to form an RdRp complex to have polymerase activity. d RdRp was incubated with the 13 + 23 nt duplex in the presence of NMPs, NDPs or NTPs respectively. These are showing duplicate experiments performed with different batches RdRp. The positions of the template RNA and the full-length extension product are indicated on denaturing gel. e Electrophoretic separation of reaction products of single nucleoside diphosphates as substrates. The positions of the original RNA and extension products are indicated. f Part of the Sanger sequencing chromatogram of the RT-PCR products corresponding to the extended 10 nt RNA on the 13 + 23 nt RNA duplex with NDPs as substrate. The dash line circle represents the 10 nt extension. g Schematic of high-throughput sequencing of RdRp reaction products.
NDPs equally contribute to the slower incorporation rate. A 13 + 26 nt RNA duplex was synthesized for the assay testing the incorporation efficiency for each NDP (Fig. 2c). Each reaction system contains three NTPs and the NDP of interest (Fig. 2d). Unexpectedly, our data showed that substitution of any single NTP by NDP has no observable effect on the efficiency of RNA synthesis (Fig. 2e). This may imply that the speed of RNA synthesis is only affected by continuous NDP incorporation but not the intermittent insertion.

3.3. SARS-CoV-2 RdRp incorporates NHC into RNA by using diphosphate form molnupiravir as a substrate

Since RdRp has become one of the most important targets for the development of antiviral drugs, many NA drugs have been constructed to target RdRp for COVID-19 treatment [34,45–47]. Among all these chemicals reported by far, molnupiravir has been regarded as a promising drug candidate. The molecular mechanism of molnupiravir was determined recently. Once ingested by patients, molnupiravir undergoes stepwise phosphorylation to yield the active nucleoside triphosphate (MTP). The RdRp then uses MTP as substrate and incorporates NHC monophosphate into RNA at the position where the nucleoside in the template is G or A, resulting in “error catastrophe” and virus death [36,38,48] (Fig. 3a).

Although the mechanism of molnupiravir has been well established, our finding of RdRp’s ability to utilize NDPs as a substrate still raises a question: is MDP also an active form and recognized by RdRp as a substrate? To address this question, the 13 + 26 nt RNA duplex was used to perform the RNA extension assay with NDPs and MDP as the substrates in comparison with the extension using NTPs and MTP as the substrates. The 13 + 26 nt RNA duplex allows ten nucleotides (nt) of extension. The 13 + 26 nt RNA duplex contains two adjacent G and two adjacent A, allowing NHC incorporation into the RNA (Fig. 3b). In this case, when CTP or UTP is replaced by MDP, the incorporation of NHC does not hinder the incorporation of the next subsequent nucleotide, and RdRp still completes the extension reaction. Similar results were obtained using MTP as the substrate (Fig. 3c). Therefore, it is reasonable to speculate that both MTP and MDP are active forms of monulpiravir in humans.

Molnupiravir is an isopropylester prodrug of NHC. We speculate that when the molnupiravir prodrug enters the cells, it is sequentially converted to NHC-monophosphate, MDP and MTP. Since RdRp of SARS-CoV-2 can use both MDP and MTP as a substrate instead of CTP or UTP, the MDP and MTP could be the active forms of the prodrug in the cells.

Fig. 2. a Primer-extension reactions on 13 + 23 nt RNA templates sampled at indicated time points. b Quantification of elongation products for (a). The experiment was performed in triplet. The 13 + 26 nt RNA template-product duplex. The direction of RNA extension is shown. The S' end of the RNA product was labeled by FAM. d RdRp catalyzed reactions containing different substrates. Lane 1: no substrate; lane 2: NTPs; lane 3: ATP, GTP, CTP; lane 4: ATP, GTP, CTP, UDP; lane 5: GTP, CTP, UTP; lane 6: ATP, CTP, UTP, ADP; lane 7: ATP, CTP, UTP, GDP; lane 8: ATP, GTP, UTP; lane 9: ATP, GTP, CDP. The positions of the 13 + 26 nt RNA and extension products are indicated. e Quantification of elongation products for (d). Data are shown as means ± standard and analyzed using unpaired t-test. There are not statistically significant.
UTP, mutations in the genomic (gRNA) and subgenomic RNA will accumulate. Over time, lethal mutations may occur and the viruses inside the cells are eliminated.

3.4. MDP is a better substrate than MTP

Since both MTP and MDP are active substrates of RdRp we performed an assay to figure out which one is the better substrate. By observing their efficiency in reaction, we found that when the template is G, the relative incorporation efficiency of MTP is 0.78, while MDP is 0.87 in comparison to CTP. When the template is A, the relative incorporation efficiency of MTP is 0.58, while MDP is 0.77 in comparison to UTP (Fig. 3d). In order to obtain a reliable result, the experiment was repeated six times with two batches of RdRp. The results indicated MDP is the better substrate for RdRp in comparison to MTP. MTP and MDP were characterized by HPLC and mass spectrometry to make sure they are the right pure compounds (Supplementary Fig. S2).

Next, we tried to determine if MDP is also more efficient in introducing mutations than MTP. The products of RdRp-catalyzed extension reactions containing MTP or MDP using 13 + 29 nt RNA as a template (Fig. 3c) were reverse-transcribed and PCR-amplified, and the number of mutations was counted by high-throughput sequencing. The statistical results were shown in Fig. 4 and Supplementary File SI.H3. Further analysis of the types of mutations introduced by NHC revealed an increase in the number of C > T mutations when MDP or MTP was added (Fig. 4a-c, f-g). The relative proportions of C > T transitions among all observed mutations were increased by 13 % (MDP) and 12 % (MTP) in the absence of CTP (Fig. 4d and e), by 11 % (MDP) and 12 % (MTP) in the presence of 800 nM CTP compared to the NTPs control (Fig. 4h and i). Together, similar mutation ratios were found when the same concentrations of MTP and MDP were added.

3.5. Comparison of substrate specificity of RNA polymerases from different viruses

To test if other virus RNA polymerases also use NDPs as a substrate, we purified SARS-CoV RdRp, NS5 protein of ZIKV and T7 RNAP and measured their activities when NDPs were used as a substrate. Results indicated that of all the RNA polymerases tested SARS-CoV-2 RdRp had the highest activity when using NDPs as substrate. SARS-CoV RdRp could also use NDPs as a substrate, but the product band seem a double RNA strand shorter than expectation implying stalled extension. NS5 of
ZIKV did not produce a well-defined product band when using NDPs as a substrate (Fig. 5a). The transcription of T7 RNAP was measured based on fluorescence technology. We found that except for the case in which ATP was replaced by ADP and weak transcription occurred, any other NTP replaced by NDP would inhibit the transcription completely (Fig. 5b).

Substrate specificity of these enzymes was also tested using MTP or MDP as a substrate. As expected, SARS-CoV RdRp could also effectively use MTP and MDP as a substrate. When UTP is replaced with MTP as a substrate, the relative incorporation efficiency reached 0.77 ± 0.01. Efficiency was even higher when MTP was used as a substrate (0.70 ± 0.02). NS5 of ZIKV weakly incorporates NHC into RNA when MTP or MDP was used to replace UTP, giving the relative incorporation an efficiency of 0.08 ± 0.03 for MTP and 0.06 ± 0.02 for MDP (Fig. 5c,d). As the only DNA dependent RNA polymerase in this experiment, T7 RNAP cannot use MTP and MDP as a substrate (Fig. 5b).

4. Discussion

Since its outbreak in late 2019, the COVID-19 pandemic has been plaguing people for three years and still poses a threat to human health [3,11]. Although many kinds of vaccines have been developed to suppress the pandemic, they are not sufficient to withhold the spread of SARS-CoV-2 [49]. The development of potent antivirals against SARS-CoV-2 is still an urgent need [6,50]. Since RdRp is critical for viral genome replication and transcription and conserved in RNA virus species, RdRp has become one of the most appealing targets for antivirals development [17,43,50,51]. Thoroughly characterization of RdRp both structurally and biochemically would no doubt be beneficial to drug development.

The substrate usage of RdRp is astonishing. To the best of our knowledge, most RNA polymerases, including DNA-dependent and RNA-dependent RNA polymerases, use NTPs as a substrate to synthesize RNA under the guidance of a template strand. We found that SARS-CoV-2 RdRp synthesizes RNA with comparable efficiency using NTPs and NDPs as a substrate, and other RNA viruses that can utilize NDPs as a substrate have also been found. We speculate that the ability to use NDPs as a substrate may be an advantage of RNA virus during infection. During later stages of infection, host cells may not be able to produce enough ATP and other NTPs. By using NDPs as a substrate, SARS-CoV-2 can continue the process of genome replication and assembly of progeny viruses, while most metabolic activities stop.

It has long been a general knowledge that to become active against virus NA drugs must undergo stepwise addition of phosphate groups to become the triphosphate form. In order to produce this outcome, this type of antivirals must be recognized by three kinds of kinases [52]. This raises a concern that the triphosphate form of antivirals may be incorporated into host mRNA [48]. Even worse, mutagenic ribonucleoside analogs could be reduced into the 2′-deoxyribonucleotide form by host ribonucleotide reductase and then incorporated into DNA [38,48]. Our finding that RdRp also uses MDP as a substrate could largely resolve these concerns. Since the diphosphate form is also active, there would be no issue regarding whether the prodrugs can become triphosphate form inside host cells. Furthermore, nucleoside diphosphate analogues, which
Fig. 5. Comparison of substrate specificity of RNA polymerases from different viruses. a RdRp of SARS-CoV-2 and SARS-CoV and NS5 of ZIKV were incubated with the 13 + 23 nt RNA duplex in the presence of NTPs or NDPs respectively. The positions of the template RNA and the full-length extension product are indicated. b The transcription of T7 RNAP was measured by monitoring the signal of a fluorescent probe. Fluorescent signal continuously changed when transcription occurred, indicating the presence of RNA product. However, when NDPs were used as a substrate, the fluorescent signal remained unchanged, indicating the lack of RNA product. Substitution of ATP by ADP led to weak transcription, any other NTP replaced by NDP or MTP as well as UTP replaced by MDP would inhibit the transcription completely. c RdRp of SARS-CoV-2, SARS-CoV and NS5 of ZIKV were incubated with the 13 + 23 nt RNA duplex in the presence of NTPs; ATP, GTP, CTP; ATP, GTP, CTP, MTP; ATP, GTP, CTP, MDP respectively. The positions of the template RNA and the full-length extension product are indicated. d Quantification of elongation products for (c). Statistical significance is indicated as compared with NTPs as a substrate using a t-test. All of the values shown represented the mean ± standard deviation of the results from three independent experiments (n.s., not significant; *P < 0.05, **P < 0.01, ***P < 0.001).

Fig. 6. Nucleoside diphosphate analog prodrugs designed by our strategy would have less potential to cause side effects. The terminal phosphate group (β-phosphate) is modified by lipophilic masks to give the prodrug more membrane permeability. Inside cell the masks are removed by enzymatic or chemical reactions. The ribose moiety and the base group should be designed in such a way that the diphosphate form drug can evade NDPK thus eliminating the risk of being incorporated into host mRNA or DNA (Note: NDPs and dNDPs are phosphorylated by the same NDPK).
can evade nucleoside diphosphate kinase (NDPK), would cease to have the possibility of being incorporated into mRNA or DNA. The best strategy is to use the membrane-permeable nucleoside diphosphate analog for COVID-19 treatment [52,53]. The advantage of this strategy is that the drug is not only able to bypass all the phosphorylation steps and has no risk of incorporation into host mRNA or DNA (Fig. 6).

Previous studies have indicated that the error rate (#mutations/10,000 bases) of viral genes treated with 10 μM NHC increased from 0.015 to 0.09 [54]. The G > A and C > U transition mutations are detectable at low frequencies across the genome [30]. This suggests that incorporation of NHC is discontinuously. The ability of RdRp to incorporate NHC into a short template implies the ability to incorporate NHC into the long genome at any position. Since RdRp is the core of the replication complex, its substrate specificity is unlikely determined by other proteins. However, our results obtained from in vitro data are preliminary. They may be further validated by additional tests in vivo SARS-CoV-2 infections.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijbiomac.2022.12.112.

CRediT authorship contribution statement

Maofeng Wang: Methodology, Experimental verification, Investigation, Data curation, Writing-manuscript, Visualization. Cancan Wu: Experimental verification, Writing - Review & Editing. Nan Liu: Experimental verification, Writing - Review & Editing. Fengyu Zhang: Experimental verification. Hongjie Dong: Formal analysis. Shuai Wang: Experimental verification, Writing - Review & Editing. Lichuan Gu: Conceptualization, Methodology, Writing - review & editing, Supervision, Funding acquisition.

Declaration of competing interest

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

Data availability

All data generated or analyzed during this study are included in this published article (and its Supporting information).

Acknowledgments

The authors thank Yan Zhang at Zhejiang University for offering SARS-CoV-2 RdRp gene, Peihui Wang at Cheeloo College of Medicine, and pathogenesis, J. Med. Virol. 92 (2020) 418–423, https://doi.org/10.1002/jmv.25681.

References

[1] Y. Chen, Q. Liu, D. Guo, Emerging coronaviruses: genome structure, replication, and pathogenesis, J. Med. Virol. 92 (2020) 418–423, https://doi.org/10.1002/jmv.25681.

[2] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G.F. Gao, W. Tan, China Novel Coronavirus Investigating and Research Team, a novel coronavirus from patients with pneumonia in China, 2019, N. Engl. J. Med. 382 (2020) 727–733, https://doi.org/10.1056/NEJMc2001017.

[3] Rui Xu, Zhao Zhao, Li Pi, Piu Niu, Bi Yang, Hu Wu, Wang Song, B Huang, N Zhu, Yi Bi, Xi Xu, Mu, F Zhang, L Wang, T Hu, H Zhou, Z Hu, Wu Zhou, L Zhao, J Chen, Ying Meng, J Wang, Y Lin, J Yuan, Z Xie, J Ma, WJ Liu, D Wang, W Xu, E
