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.
A mechanism for SARS-CoV-2 RNA capping and its inhibition by nucleotide analog inhibitors

Graphical abstract

Highlights

• SARS-CoV-2 polymerase NiRAN domain catalyzes nsp9 RNAylation

• The RNAylated nsp9 serves as an intermediate to form GpppA with GTP catalyzed by NiRAN

• Nucleotide analog inhibitors inhibit nsp9 RNAylation and GpppA formation

• An “induce-and-lock” mechanism for inhibitor design is proposed

Authors
Liming Yan, Yucen Huang, Ji Ge, ..., Yan Gao, Zihe Rao, Zhiyong Lou

Correspondence
gaoyan@shanghaitech.edu.cn (Y.G.), raozh@tsinghua.edu.cn (Z.R.), louzy@mail.tsinghua.edu.cn (Z.L.)

In brief
Structural analyses reveal how proteins from SARS-CoV-2 cooperate and use GTP to form the cap on viral mRNA and how this process is interrupted by nucleotide analogs that serve as antiviral drugs.
A mechanism for SARS-CoV-2 RNA capping and its inhibition by nucleotide analog inhibitors

Liming Yan,1,10 Yucen Huang,1,3,10 Ji Ge,1,2,10 Zhenyu Liu,1,4,10 Pengchi Lu,5 Bo Huang,6 Shan Gao,1 Junbo Wang,1 Liping Tan,1 Sihan Ye,1 Fengxi Yu,3 Weiqi Lu,4 Shiya Xu,1 Feng Zhou,6 Lei Shi,6 Luke W. Guddat,7 Yan Gao,4,* Zihe Rao,1,2,4,8,9,11,* and Zhiyong Lou1,*

1MOE Key Laboratory of Protein Science, School of Medicine, Tsinghua University, Beijing, China
2Innovation Center for Pathogen Research, Guangzhou Laboratory, Guangzhou, China
3Tsinghua University-Peking University Joint Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing, China
4Shanghai Institute for Advanced Immunochemical Studies and School of Life Science and Technology, ShanghaiTech University, Shanghai, China
5Division of Biosciences, University College London, London, London, UK
6Beijing StoneWise Technology Co Ltd., Beijing, China
7School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Australia
8State Key Laboratory of Medicinal Chemical Biology, College of Life Sciences and College of Pharmacy, Nankai University, Tianjin, China
9National Laboratory of Biomacromolecules, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China
10These authors contributed equally
11Lead contact
*Correspondence: gaoyan@shanghaitech.edu.cn (Y.G.), raozh@tsinghua.edu.cn (Z.R.), louzy@mail.tsinghua.edu.cn (Z.L.)
https://doi.org/10.1016/j.cell.2022.09.037

SUMMARY

Decoration of cap on viral RNA plays essential roles in SARS-CoV-2 proliferation. Here, we report a mechanism for SARS-CoV-2 RNA capping and document structural details at atomic resolution. The NiRAN domain in polymerase catalyzes the covalent link of RNA 5’ end to the first residue of nsp9 (termed as RNAylation), thus being an intermediate to form cap core (GpppA) with GTP catalyzed again by NiRAN. We also reveal that triphosphorylated nucleotide analog inhibitors can be bonded to nsp9 and fit into a previously unknown “Nuc-pocket” in NiRAN, thus inhibiting nsp9 RNAylation and formation of GpppA. S-loop (residues 50-KTN-52) in NiRAN presents a remarkable conformational shift observed in RTC bound with sofosbuvir monophosphate, reasoning an “induce-and-lock” mechanism to design inhibitors. These findings not only improve the understanding of SARS-CoV-2 RNA capping and the mode of action of NAIs but also provide a strategy to design antiviral drugs.

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the causative agent of coronavirus disease 2019 (COVID-19). To date, COVID-19 has caused over 585 million infections and over 6.4 million deaths worldwide (WHO, 2020). The genome of SARS-CoV-2 encodes 16 nonstructural proteins (nsps) to assemble the replication-transcription complex (RTC) for replicating and transcribing viral RNAs (Lou and Rao, 2022; Ziebuhr, 2005). The positive-sense RNAs (mRNA and genomic RNA) of SARS-CoV-2 bear a 5’ cap(1) structure (7MeGpppA2’OMe) (Bouvet et al., 2010; Chen et al., 2011; van Vliet et al., 2002). The co-transcriptional capping of SARS-CoV-2 RNAs has been previously shown to occur by a four-step sequential reaction in RTCs. First, a 5’ RNA triphosphatase (RTPase) activity in the viral helicase, nsp13, removes the γ-phosphate of the 5’-triphosphate end (pppA) of the nascent RNAs to generate the 5’-diphosphate end (ppA) (Ivanov et al., 2004; Ivanov and Ziebuhr, 2004). Second, the N-terminal nidovirus RdRp-associated nucleotidyltransferase (NiRAN) domain in viral polymerase, nsp12, acts as a guanylyltransferase (GTase) to transfer a GMP to ppA to yield the cap core (GpppA). The co-transcriptional capping of SARS-CoV-2 RNAs has been previously shown to occur by a four-step sequential reaction in RTCs. First, a 5’ RNA triphosphatase (RTPase) activity in the viral helicase, nsp13, removes the γ-phosphate of the 5’-triphosphate end (pppA) of the nascent RNAs to generate the 5’-diphosphate end (ppA) (Ivanov et al., 2004; Ivanov and Ziebuhr, 2004). Second, the N-terminal nidovirus RdRp-associated nucleotidyltransferase (NiRAN) domain in viral polymerase, nsp12, acts as a guanylyltransferase (GTase) to transfer a GMP to ppA to yield the cap core (GpppA) (Walker et al., 2021; Yan et al., 2021a). In the last two steps, an N7-methyltransferase (N7-MTase) in nsp14 and a 2’-O-methyltransferase (2’-O-MTase) in nsp16 sequentially methylates GpppA to cap(0) (7MeGpppA) and cap(1) (7MeGpppA2’OMe), thereby finalizing the capping actions (Bouvet et al., 2010; Chen et al., 2009; Chen et al., 2011; Decroly et al., 2008). Nsp9, a virally encoded nonstructural...
protein, has been shown to bind to nsp12 NiRAN and play essential roles in RNA capping (Yan et al., 2021a; Yan et al., 2021b). Nsp9 not only impacts on the enzymatic activity of nsp12 NiRAN but also plays a key role as an adaptor to recruit nsp10/14 or nsp10/16 into RTCs to constitute the co-transcriptional capping complexes (CCCs) for the third or final capping actions (Yan et al., 2021a; Yan et al., 2021b). Further studies have shown that the attenuation of the activity of nsp12 NiRAN by nsp9 is due to the additional N-terminal residues acquired from the expression plasmid, although these additional residues do not affect the binding of nsp9 to nsp12 NiRAN (Slanina et al., 2021; Wang et al., 2021). Moreover, nsp9 has also been reported to be NMPylated by nsp12 NiRAN, thus priming the synthesis of viral RNA (Shannon et al., 2022; Slanina et al., 2021; Wang et al., 2021).

Here, we report an alternative mechanism of RNA capping facilitated by nsp12 NiRAN and nsp9, contrasting with the mechanism proposed for the nsp13/nsp12/nsp14/nsp16 enzymes. Our structural studies show here that nsp12 NiRAN catalyzes a reaction to remove pyrophosphate (PPI) from the 5’-end of the RNA and subsequently forms a covalent bond between the remaining χ-phosphate with the first asparagine residue of nsp9 (nsp9N1), a step termed as RNAylation. The RNAylated nsp9 is further used as an intermediate to generate GpppA with a GTP molecule bound in the so-called “G-pocket,” again catalyzed by nsp12 NiRAN. We also show that nsp12 NiRAN covalently bonds nucleotide analog inhibitors (NAIs), remdesivir, sofosbuvir, molnupiravir, and AT-9010 (the produg is AT-527), to nsp9N1 and locate them in previously undescribed pocket we refer to as the “Nuc-pocket.” As a result, RNAylation of nsp9 and formation of GpppA is inhibited. The structures of RTC in complex with the RNAylated nsp9 and further with GMPNP or GTP, as well as in complex with GMPNP and nsp9 decorated by the monophosphorylated remdesivir (RMP), sofosbuvir (SMP), molnupiravir (MMP), or AT-9010 (ATMP), have also been determined. These structures not only provide insights to improve the understanding of SARS-CoV-2 RNA capping and the mechanism for NAI action but also suggest an “induce-and-lock” mechanism that can be utilized for the development of antiviral inhibitors targeting the process of RNA capping.

RESULTS

**Nsp12 catalyzes the covalent bonding of RNA to nsp9**

The incubation of an RdRp-defective nsp12 mutant (S759A-nsp12), with nsp9 and a 10-nt RNA derived from 5’ UTR of SARS-CoV-2 genome (5’-pppAUUAAAGGUU-3’), resulted in an upward migration of the apo nsp9 (apo-nsp9) band on SDS-PAGE in a time-dependent manner (Figures 1A and S1A). This band was isolated, and mass spectrometry showed it has a molecular weight (MW) of 15,609 Da (Figures 1B and S1B). Considering the MWs of the RNA (3,410 Da) and apo-nsp9 (12,378 Da), the change of mass (3,410 + 12,378 − 15,609 Da = 179 Da) shows the RNA is covalently bonded to nsp9 and with the loss of PPI (178 Da). We term this modification of nsp9 by the RNA as RNAylation. The RNAylation of nsp9 is strongly dependent on the presence of Mn$^{2+}$ (Figure 1A). In the presence of Mg$^{2+}$, only weak activity was observed, whereas no activity was detected in the presence of Ca$^{2+}$.

**Nsp12 catalyzes the formation of GpppA with the RNAylated nsp9 and GTP**

When the RNAylated nsp9 (RNA-nsp9) and S759A-nsp12 were incubated with [$\gamma$-32P]GTP in the presence of Mg$^{2+}$, a signal corresponding to the radioactively labeled GpppA-RNA was detected (Figures 1C, S1C, and S1D). The formation of GpppA-RNA is not dependent on Mn$^{2+}$, which is in contrast to the Mn$^{2+}$-dependency of the RNAylation of nsp9. Because the nsp12 used in these studies is an RdRp-defective mutant, the results indicate nsp12 NiRAN can catalyze the RNAylation of nsp9 and the formation of GpppA using RNA-nsp9 and GTP. Furthermore, the GpppA formed by RNA-nsp9 and GTP can be further modified by nsp14 and nsp16 to yield cap(0) (1$^{\prime}$MeGpppA) and cap(1) (1$^{\prime}$MeGpppA$^{2^{'}}$OMe), suggesting nsp9 RNAylation mediates a pathway for SARS-CoV-2 RNA capping (Figure 1D).

**Structure of RTC bound with the RNAylated nsp9**

To dissect the mechanisms for the RNAylation of nsp9 and the formation of GpppA, we purified RNA-nsp9 and determined the cryo-EM structures of the elongation RTC (E-RTC, composed by nsp7-nsp82-nsp12-nsp132) in complex with RNA-nsp9 and subsequently with GMPNP (a non-hydrolyzable GTP analog) at 2.75 and 2.66 Å, respectively (Figures 1E, 2, and Table S1). In the structure of E-RTC:RNA-nsp9, RNA-nsp9 binds to nsp12 NiRAN, which is in agreement with the previously reported structures of SARS-CoV-2 RTCs that also include apo-nsp9 (Yan et al., 2021a; Yan et al., 2021b). The N-terminal three residues of nsp9 insert into the catalytic center of nsp12 NiRAN (Figure 2A). The first two nucleotides of the RNA, A+1 and U+2, are identified in the density (Figure 2B). The other nucleotides are not observed, possibly because they are in a flexible conformation due to a lack of interaction with the RTC proteins.

The χ-phosphorus of the A+1 nucleotide is ~1.6 Å from the backbone nitrogen of the first asparagine residue of nsp9 (nsp9N1), indicating a phospho-amide covalent bond links the RNA and nsp9 (Figure 2C). The base and the ribose of A+1 bind in a cavity surrounded by the anti-parallel strands β2-β3 in nsp12 NiRAN and rest on a hydrophobic base consisting of the side chains of nsp12F35/I37/F48 (Figures 2D and 2A). The base of A+1 forms hydrogen bonds with nsp12N39 in nsp12 NiRAN and with nsp12N713 in nsp12 Palm. The side chain of nsp9N1 is also involved in the interaction with the base of A+1. Henceforth, we term this pocket as the “Nuc-pocket.” The second nucleotide, U+2, extends outward from the catalytic center of nsp12 NiRAN. The base of U+2 stacks parallel with the base of A+1. The side chains of nsp12K41 in nsp12 NiRAN and nsp12N713 in nsp12 Palm contact with the base of U+2, whereas nsp12K50 and nsp12T51/N52 bond with the χ-phosphate and the ribose of U+2. Two Mn$^{2+}$ ions are chelated by the side chains of nsp12D208/N209/D218 and coordinate with three solvent molecules in the catalytic center of nsp12 NiRAN (Figure 2E). Two water molecules form a bridge linking Mn(1) and the side chain of nsp12F1116. The distances of two Mn$^{2+}$ to the χ-phosphate of A+1 are both around 4.0 Å (Figure 2E, red dashes).
For mutagenesis studies, we selected five residues which have close contacts with Mn$^{2+}$ or the RNA and substituted them to alanine to verify their contributions to the RNAylation of nsp9 (Figure 2F). SDS-PAGE shows the substitutions of nsp12R116, nsp12D208, nsp12N209, and nsp12D218 to alanine abolishes the RNAylation of nsp9, confirming their essential roles in coordinating Mn$^{2+}$. The mutation of nsp12N39, which contacts with the base of A+1, also significantly attenuates nsp9 RNAylation, although a weak band corresponding to RNA-nsp9 could still be observed, indicating the stabilization of the RNA is also important for the RNAylation of nsp9.

Structure of RTC bound with GMPPNP and the RNAylated nsp9
In the structure of E-RTC:RNA-nsp9:GMPPNP, four nucleotides are identified (Figures 3A, 3B, S2B, S3A, and S3B). The first nucleotide, A+1, binds in the Nuc-pocket, similar to that observed in E-RTC:RNA-nsp9 (Figures 3C and S3C). The conformation shift of U+2 leads to the loss of contact of its uracil base with the side chain of nsp12N713, but it does form an interaction (distance 3.2 Å) with...
Figure 2. Structure of E-RTC:RNA-nsp9

(A) A close-up view of the catalytic center of nsp12 NiRAN in the structure of E-RTC:RNA-nsp9. Nsp12 Palm, nsp12 NiRAN, and nsp9 are shown as red, yellow, and purple cartoons, respectively. The first residue of nsp9 (nsp9N1), the catalytic residues (nsp12D208, nsp12N209, and nsp12D218) and the nucleotides of the RNA linked to nsp9 are displayed as colored sticks. Two Mn²⁺ ions bound at nsp12 NiRAN catalytic center are represented as light purple spheres, whereas the solvent molecules coordinated with Mn²⁺ ions are shown as red spheres.

(B and C) Cryo-EM densities covering nsp9N1, the nucleotides of RNA, Mn²⁺, and the coordinated solvent molecules are shown in gray mesh. The region of the phospho-amide bond is enlarged in (C).

(legend continued on next page)
the carbonyl oxygen of nsp12 L49 (Figure 3C). The hydroxyl group of nsp12 T51 side chain also participates in an interaction with the α-phosphate of U+2, together with nsp12 K50/N52, but does not contact with the ribose. The third and the fourth nucleotides, U+3 and A+4, run anti-parallel with respect to A+1 (Figure 3D). The base groups of U+3 and A+4 point toward the catalytic center. The base of U+3 is clamped by the base groups of A+1 and A+4. U+3 does not make contact with the RTC proteins. The base and the ribose of A+4 contact with nsp12 R74/T76/N79, whereas the base of A+4 is clamped by the side chain of nsp12 H75 and U+3.

A GMPPNP molecule binds to nsp12 NiRAN (Figures 3E, S2B, S3A, and S3B). The base and the ribose of GMPPNP are buried in a deep pocket, which is formed by the residues from β2-β3 and β4-β5 of nsp12 NiRAN. Henceforth, we term this region as the “G-pocket.” In the G-pocket, the carbonyl oxygens of nsp12 T120 and nsp12 K121 make interactions with the nucleobase of GMPPNP. The side chain nitrogen atom of nsp12 R55 also interacts with the hydroxyl of the base, whereas nsp12 V31/K50/C53/Y69/V71/Y122/T123/Y217 form a relatively hydrophobic environment to surround and stabilize GMPPNP.

The binding of AT-527 diphosphate (ATDP) and GTP in the G-pocket has been observed in two previous studies (Malone et al., 2022; Shannon et al., 2022). As the coordinates for GTP binding in G-pocket were not available when this work was submitted, we compared it with the binding of ATDP in the G-pocket (Figure S3D). The base, the ribose, and the α-β-phosphates of ATDP are in similar positions as GMPPNP, whereas a Mg2+ ligates with ATDP, occupying the position of the γ-phosphate in GMPPNP. The catalytic residues nsp12 D208/D218/N209/D218 are in the same conformation in both structures. Since AT-527/AT-9010 is a guanosine analog, the similarity with the binding of ATDP and GMPPNP in the G-pocket is expected. It is also noteworthy that nsp12 R55 and nsp12 Y217, two conserved residues across the α-β-coronavirus clades, have previously been suggested to provide base specificity for guanosine (Malone et al., 2022).

Three phosphate groups of GMPPNP bind at the catalytic center of nsp12 NiRAN in a space clamped by nsp12 Y217 and nsp12 D221 (Figures 3F and S2B). The α-phosphate of GMPPNP is stabilized by the positively charged side chains of nsp12 K73 and nsp12 R116. The β-phosphate forms interactions with nsp12 D208 and nsp12 N209, whereas the γ-phosphate contacts with the side chain of nsp12 D221 and the backbone nitrogen of nsp12 G220, and the side chain of nsp12 D218 is located in the central of three phosphates and forms interactions with them. A Mg2+ ligates (distances ~2.6 Å) with the phosphol oxygens of the α- and γ-phosphates of GMPPNP (Figure 3F, red dashes). An interesting observation is that the β-γ-phosphates of GMPPNP occupy the positions that bind the two Mn2+ in the structure of E-RTC:RNA-nsp9 (Figure S3E). In the context that the RNAylation of nsp9 is dependent on Mn2+, but the formation of GpppA is not, this structural observation suggests that the β-γ-phosphates of GTP may compete with the Mn2+ prior the subsequent catalytic reaction for GpppA formation.

We have also determined the structure of E-RTC in complex with RNA-nsp9 and GTP in the presence of Ca2+ at 3.39 Å (Table S1). Structural comparison showed that GTP binds in the G-pocket with a generally similar conformation as GMPPNP, but the orientation of β-γ-phosphates varies (Figure S4A). The densities for the nucleotides are largely not visible and only poor density corresponding to the ribose of A+1 remains (Figure S4B), suggesting the reaction still occurs. We assume the non-hydrolyzable P-N bond between β-γ-phosphates in GMPPNP helps to stabilize the conformation of RTC, thus resulting in the significantly higher resolution. Because of the higher resolution of E-RTC:RNA-nsp9:GMPPNP, as well as the similar conformation of GTP and GMPPNP in the G-pocket, we used GMPPNP for structural analysis in this study.

We next mutated five residues in nsp12 NiRAN to alanine and verified their importance on the formation of GpppA (Figure 3G). Residue nsp12 R55 is in G-pocket and interacts with the base of GMPPNP, nsp12 K73 contacts with the α-phosphate of GMPPNP, whereas nsp12 D208/N209/D218 are the catalytic residues. Among them, the mutation on R55 showed a significant decrease in the yield of GpppA, similar to the substitutions of the catalytic residues nsp12 D208 and nsp12 D218, suggesting the stabilization of the base of GMPPNP in the G-pocket is essential for reaction to occur. In contrast, the mutants, nsp12 K73A and nsp12 N209A, exhibit minimal effects on GpppA formation.

**NiRAN catalyzes the linking of nucleotide analog inhibitors to nsp9**

The incubation of nsp9 with the triphosphorylated NAIs (NAI-TP) and nsp12 resulted in a covalent bond between the NAIs and nsp9 (Figures S5A–S5D). We selected the triphosphorylated forms of remdesivir (RTD), sofosbuvir (STP), molnupiravir (MTP), and AT-9010 (ATTP) in this study, representing analogs of A/U/G/C, respectively. The structures show all four NAI-TPs bonded to nsp9 with catalysis due to the presence of nsp12. Mass spectrometry validated the formation of covalent bonding of the NAIs with nsp9. The MWs for the four reaction products are 12,732, 12,701, 12,710, and 12,739 Da, which correspond to the MW of apo-nsp9 (12,378 Da) with RTP (531 Da), STP (500 Da), MTP (499 Da), and ATTP (539 Da) added and a PPI (178 Da) subtracted. In summary, these results show that PPI is released from NAIs-TPs and their monophosphorylated forms (RMP, SMP, MMP, and ATMP) are covalently bonded to nsp9.

When the 10-nt RNA derived SARS-CoV-2 genome 5' UTR was incubated with nsp12 and the purified RMP/SMP/MMP/ATMP-nsp9, the RNA cannot be decorated to nsp9 by replacing RMP,
SMP, MMP, and ATMP (Figure S5E). Moreover, the incubation of RMP-/SMP-/MMP-/ATMP-nsp9 with nsp12 and GTP does not break the covalent bonds between NAI-MP and nsp9, prohibiting the removal of RMP/SMP/MMP/ATMP from the NAI-decorated nsp9 (Figure S5F). These results suggest that the covalent bonds between NAI-MP and nsp9N1 cannot be broken in nsp12 NiRAN and the N terminus of nsp9 cannot be released in the presence of the substrates for nsp9 RNAylation and GpppA formation, thus preventing RNAylation of nsp9 and subsequent GpppA formation.

Overall structures of RTC bound with NAI-MP-nsp9
To dissect the mechanism for the NAI-inhibition of nsp9 RNAylation and GpppA formation, we determined the structure of E-RTC in complex with purified RMP-nsp9, SMP-nsp9, MMP-nsp9, or ATMP-nsp9 in the presence of GMPPNP at resolutions of 2.72, 3.37, 2.64, and 3.38 Å, respectively (Figures 4A–4D and S2C–S2E; Table S2). In all four structures, GMPPNP is observed in the G-pocket, but the binding modes of the NAI-MPs in the Nuc-pocket is variable (Figures 4A–4H). In the Nuc-pocket, the densities for SMP and RMP are well defined, and the density for MMP is also visible. In sharp contrast, the density corresponding to ATMP linked with nsp9N1 is not observed, although mass spectrometry confirmed that the decoration of ATMP to nsp9 had occurred. We reason the lack of ATMP density is due to the loss of contact with RTC proteins. In addition, we also determined the structure of E-RTC:SMP-nsp9:GTP and have shown that GTP binds to the G-pocket in a similar mode as GMPPNP, as expected (Figure S4C, Table S2).

Conformational change of “S-loop”
The loop region constituted by residues 50–KNT–52 exhibits a conformational change with a downward rotation of ~180° in the structure of E-RTC:SMP-nsp9:GMPPNP, compared with structures with RMP-/MMP-/ATMP-nsp9 or RNA-nsp9 or apo-nsp9 (Yan et al., 2021a) bound (Figures 4A–4D, S4F, and S4G). As the fluoro modification and the hydroxyl group on the ribose of SMP leads to this conformational change.

Figure 3. Structure of E-RTC:RNA-nsp9:GMPPNP
(A) A close-up view of the catalytic center of nsp12 NiRAN in E-RTC:RNA-nsp9:GMPPNP. Nsp12 Palm, nsp12 NiRAN, and nsp9 are shown as red, yellow, and purple cartoons. The first residue of nsp9 (nsp9N1), the catalytic residues (nsp12D208, nsp12N209, and nsp12D218), the nucleotides of RNA linked to nsp9 and GMPPNP are displayed as colored sticks. A MgII ion bound with the 2'-phosphates of GMPPNP is displayed as a green sphere.
(B) Cryo-EM density covering nsp9N1, the nucleotides of RNA, and GMPPNP is shown as gray mesh. (C and D) The binding of RNA nucleotides. Enlarged view of U+3/A+4 nucleotides (C) and U+3/A+4 nucleotides (D) in nsp12 NiRAN are shown in (C) and (D), respectively. The polypeptides of nsp12 Palm, nsp12 NiRAN, and nsp9 are shown as semi-transparent colored cartoons. The nucleotides and the residues for RNA binding are represented in the same way as in (A). The black dashed lines denote the intermolecular interactions. For clarity, U+3/A+4 are not shown in (C), whereas U+2 and the interaction of A+1 are not shown in (D).
(E and F) Enlarged views of GMPPNP binding to G-pocket (D) and (E). The bound GMPPNP and the interacting residues in G-pocket and in the catalytic center are shown as colored sticks, whereas the MgII ion is displayed as a green sphere. The black dashed lines denote intermolecular interactions, whereas the red dashed lines in (F) denote the interactions of MgII with the 2'/3'-phosphates of GMPPNP. (G) The impact of nsp12 mutations on the formation of GpppA. The wt nsp12 and mutants used in the assays are labeled in the panel. The positions corresponding to 2'/3'P labeled GTP and GpppA-RNA are indicated. The result is a representative of three independent experiments. See also Figures S2, S3, and S4.

In the structure of E-RTC:UMP-nsp9, UMP is linked with nsp9N1 and the loop 50–KNT–52 in an upward conformation (Figures S4H and S4J), suggesting that the conformational change of this loop is not associated with the base and the ribose hydroxyl of the natural UMP. Moreover, in the structure of E-RTC:UMP-nsp9:GMPPNP, loop 50–KNT–52 also exhibits an upward conformation, indicating that the conformational change is not induced by simultaneous binding with UMP and GMPPNP (Figures S4I and S4K). Taken together, these results suggest that the conformational change of the loop 50–KNT–52 observed in the structure of E-RTC:SMP-nsp9:GMPPNP is induced by the interaction of nsp12 NiRAN with the fluoro modification on the ribose of SMP. Hence, we identify this sofosbuvir-specific loop as the “S-loop.” The downward movement of the S-loop induced by the binding of SMP results in notable changes in the binding of NAI-MPs and GMPPNP, as well as in the arrangement of residues in the catalytic center in the E-RTC:SMP-nsp9:GMPPNP complex (discussed below).

Interactions of NAI-MPs with RTC
RMP, SMP, and MMP linked with nsp9 bind in the Nuc-pocket (Figures 4I, S2F, and S2G). Although the three NAI-MPs bind in the Nuc-pocket similar to the binding of A+1, the chemical modifications on each NAI-MP, as well as conformational changes in the S-loop results in variations in their mode of binding. In E-RTC:SMP-nsp9:GMPPNP, the 2'-F and the 3'-hydroxyl group of the ribose of SMP form bonds with the carbonyl oxygen atoms of nsp12L49/K50 in the S-loop and the backbone carbon atom of nsp12T51. In contrast, because the S-loop is in an upward conformation, the nitrogen of the nsp12K50 side chain, instead of its backbone carbonyl oxygen, makes contacts with the ribose of RMP and MMP. Moreover, the carbonyl oxygen of nsp12L49 contacts with the hydroxyl group in the ribose of RMP and MMP, as well as the chemical modifications in the ribose of RMP, stabilizing the inhibitors. It is also noteworthy that the side chain of nsp12T51 contacts with the ribose of RMP but does not interact with MMP. Furthermore, the base of the bound SMP forms contacts with nsp12N39/V42/N713, but additional contacts of the nitrogen of nsp12K41 side chain with the base groups of the bound MMP and RMP are also observed.
The binding of GMPPNP in the G-pocket and the catalytic center of nsp12 NiRAN are generally similar in the four NAI structures and the E-RTC:RNA-nsp9:GMPPNP complex (Figures 4A-4D). However, as a result of the downward movement of the S-loop, the width of the channel from the catalytic center to the G-pocket is narrowed from ~8.7 Å to ~7.7 Å in E-RTC:SMP-nsp9:GMPPNP (Figures S4L and S4M). The ribose and the β-phosphate are pushed down with the largest distance being ~1.8 Å due to the downward movement of nsp12T51 (Figure S4N).

The binding mode of the base and the ribose of GMPPNP in the G-pocket in the four RTC complexes with NAI-nsp9 is generally similar to that observed in E-RTC:RNA-nsp9:GMPPNP (Figures 4J).
and $S2H$–$S2J$). In all of the structures, $nsp12T120$/$K121$/$R55$ form a set of hydrogen bonds with the base of GMPPNP. Residues $nsp12T31$/$K50$/$C53$/$Y69$/$V171$/$Y122$/$Y217$ also contribute to the stabilization of the base and the ribose of GMPPNP. As a result of S-loop movement, the side chain of $nsp12T51$ contacts with the ribose and the $\alpha$-phosphate upon SMP binding, whereas the side chain of $nsp12T$/$K50$ stacks parallel with GMPPNP in the RMP/MMP/ATMP complexes.

The arrangements of the catalytic center in RTC in complex with ATMP/MMP/RMP- $nsp9$ are also similar to those in E-RTC:RNA-$nsp9$:$G$MPPNP, where $nsp12K73$, $nsp12R116$, $nsp12D208$, $nsp12N209$, $nsp12D218$, $nsp12G220$, and $nsp12D221$ bind the three phosphates of GMPPNP. It is noteworthy that Mg$^{2+}$ binding with the phosphate groups cannot be clearly observed in the densities of RTC in complex with MPP-$nsp9$ and ATMP-$nsp9$, possibly due to the limitation of the cryo-EM data (Figures 4G and 4H). A significant exception is in the structure of RTC with SMP-$nsp9$ (Figures 4B, 4F, 4I, S4O, and S4P). This appears to be as a result of the downward movement of the S-loop in E-RTC:SMPP-$nsp9$:GMPPNP, which allows the side chain of $nsp12T51$ to form contacts with both the ribose and the $\alpha$-phosphate of SMP. Meanwhile, the distance between Mg$^{2+}$ and the $\alpha$-phosphate is lost, although the Mg$^{2+}$-$\gamma$-phosphate bond still exists. In addition, the side chain of $nsp12D208$ rotates $\sim$90° away from the phosphate groups, thus losing interactions with all three phosphate groups.

A “bond-breaking” water molecule

In the structures of E-RTC:RNA-$nsp9$ and E-RTC:RNA-$nsp9$:GMPPNP, a water molecule is observed coordinated by the side chain of $nsp12D208$ and the backbone carboxyl oxygen of $nsp12D36$ (Figures 5A and 5B). In E-RTC:RNA-$nsp9$, this water is 4.6 Å from the $\alpha$-phosphorus of A+1. However, in E-RTC:RNA-$nsp9$:GMPPNP, the distance between this water and the $\alpha$-phosphorus of A+1 is reduced to 3.8 Å. This is likely to be caused by the conformational change upon GMPPNP binding. In a previously reported structure of the deAMPylation FIC (filamentation induced by cAMP) enzymes, a water molecule coordinated by a negatively charged hydrophilic residue locates near to the phospho-ester bond at a distance of $\sim$3.6 Å and is proposed to play an essential role in breaking the phospho-ester bond (Veyron et al., 2019). The similar arrangement here allow us to also propose that $nsp12D208$ may attract a proton from the water molecule to activate the nucleophilic attack by its oxygen on the $\alpha$-phosphorus of A+1. This could only occur when the distance between the water molecule and $\alpha$-phosphorus of A+1 decreases to $\sim$3.8 Å. This action will lead to the breaking of the phospho-amide bond between $nsp9N1$ and A+1, providing a lone pair of electrons to the oxygen of the $\alpha$-phosphor. In the last step, the non-bridging oxygen from the remaining $\alpha$-phosphate group of A+1 attacks the $\beta$-phosphorus of GTP bound in the G-pocket, thus releasing Pi and generating the GpppA at 5′-RNA end.

When this manuscript was under revision, an independent study biochemically revealed the RNAylation of $nsp9$ and mediated GpppA formation (Park et al., 2022). In that work, GDP was reported to be a better substrate for deRNAylation than GTP. That work proposed that a GDP binds at the position, which is reported in the structure of E-RTC:apo-$nsp9$:GDP (PDB: 7CYQ) (Yan et al., 2021a) (termed as “base-out pose” position by Malone et al. (Malone et al., 2022)), represents the GDP subform to GpppA with the RNAylated $nsp9$ (Figures S3F–S3H). First, the high purity of the GTP samples used in this work shows that GDP is also a substrate for GpppA formation (Figures S1C and S1D). The structures of E-RTC:RNA-$nsp9$/SMPP-$nsp9$ in complex with GTP also demonstrate that GTP can bind in the G-pocket. Superimposition of the structures of E-RTC:apo-$nsp9$:GDP (PDB: 7CYQ) (Yan et al., 2021a) and E-RTC:RNA-$nsp9$:GMPPNP shows that the GDP in the “base-out pose” position extensively clashes with the nucleotides bonded to $nsp9N1$ (Figure S3), suggesting this model is structurally less likely. Third, previous structural studies have revealed that the “base-out pose” position not only binds GDP (PDB: 7CYQ) but also accommodates ADP (PDB: 6XEZ) (Chen et al., 2020) (Figure S3J), which is consistent with the structural observation that the bases of the bound nucleotides lack interaction with the $\alpha$-phosphorus of SMP is 5.5 Å, much longer than that observed in E-RTC:RNA-$nsp9$:GMPPNP (Figure 5D). A similar long distance between bb-wat with $\alpha$-phosphorus of the phospho-amide bond is observed in E-RTC:RMP-$nsp9$:GMPPNP (4.7 Å) and in E-RTC:MMP-$nsp9$:GMPPNP (4.2 Å) (Figures 5C and 5E). However, in the structure of E-RTC:ATMP-$nsp9$:GMPPNP, the density does not support a convincing bb-wat in the model (Figure 5F). In the three structures with the observed bb-wat, the distance from bb-wat to the $\alpha$-phosphorus of RMP/SMP ranges from 4.2 to 5.5 Å, even upon the binding of GMPPNP. This distance is 4.9 Å in the E-RTC:SPM-$nsp9$:GTP complex (Figure S4E). This may be an explanation as to why the covalent bond of NAI-MP with $nsp9$ cannot be broken with the incubation of $nsp12$ and GTP (Figure S5F).

**DISCUSSION**

A proposed mechanism for SARS-CoV-2 RNA capping

Combining the structural data and assay results, we propose a model for an alternative mechanism for the formation of GpppA catalyzed by $nsp12$ NiRAN in SARS-CoV-2 (Figures 6, S6A, and S6B). In the first step, the 5′ end of product RNA binds to the catalytic center of $nsp12$ NiRAN, possibly with its first adenosine nucleotide (A+1) binding at the Nuc-pocket. In the second step, the lone pair of electrons from the free amino group of $nsp9N1$ attack the $\alpha$-phosphorus of the A+1 nucleotide, thus releasing PPI and the RNAylated $nsp9$. In the third step, upon binding of GTP to the G-pocket, the bb-wat coordinated by $nsp12D208$ is nearby and in a position to break the phospho-amide bond between $nsp9N1$ and the $\alpha$-phosphor of A+1, providing a lone pair of electrons to the oxygen of the $\alpha$-phosphor. In the last step, the non-bridging oxygen from the remaining $\alpha$-phosphate group of A+1 attacks the $\beta$-phosphorus of GTP bound in the G-pocket, thus releasing Pi and generating the GpppA at 5′-RNA end.

Cell 185, 4347–4360, November 10, 2022 4355
Figure 5. The “bond-breaking” water
Close-up view of the structural details around bb-wat in E-RTC:RNA-nsp9 (A), E-RTC:RNA-nsp9:GMPPNP (B), E-RTC:RMP-nsp9:GMPPNP (C), E-RTC:SMP-nsp9:GMPPNP (D), E-RTC:MMP-nsp9:GMPPNP (E), and E-RTC:ATMP-nsp9:GMPPNP (F) are shown in the same orientation. The residues, the RNA, and the bound GMPPNP, as well as the cations and water molecules, are shown using the same scheme as in previous figure and are covered by the cryo-EM densities (gray meshes). The black dashed lines indicate the interactions of bb-wat with nsp12 residues, whereas the red dashed lines represent the distance from bb-wat to the α-phosphor of RNA nucleotide or NAI-MP covalently linking to nsp9N1.
See also Figures S2, S3, S4, and S6.
NiRAN residues (Chen et al., 2020; Yan et al., 2021a). This is contradictory to the specificity of GTP/GDP as the substrate for GpppA formation. Finally, the G-pocket (termed as “base-in pose” position by Malone et al. [Malone et al., 2022]) has been suggested to accommodate GTP, but not GDP (Malone et al., 2022). Moreover, R55 and Y217 in the G-pocket, two conserved residues across the $\alpha$-/b- coronavirus clades, were suggested to provide base specificity for guanosine in the G-pocket (Yan et al., 2021a). Notably, our results showed that the mutation of R55 significantly decreased the formation of GpppA (Figure 3G).

These results suggest the negative impact of the R55 mutation on GpppA formation is relevant to GTP binding. Very interestingly, a recent work found that ATDP, a GDP analog bearing a modification in the ribose group, can also bind to the G-pocket along with Mg$^{2+}$ which ligates to its $\beta$-phosphate to occupy the position of $\gamma$-phosphate of GTP (Shannon et al., 2022). That structure suggests the ligation of ions may help GDP bind to the G-pocket, whereby the $\alpha$-/b- phosphates are in similar positions as the GTP when bound in the same pocket. According to this evidence, we favor the idea that both GTP and GDP can be substrates to produce GpppA (Figure 3G).

Alternative capping pathways

In the conventional pathway, the co-transcriptional capping of SARS-CoV-2 RNAs has been known to occur by four sequential actions facilitated by nsp13, nsp12 NiRAN, nsp14 MTase, and nsp16 (Bouvet et al., 2010; Chen et al., 2009, 2011; Decroly et al., 2008; Ivanov et al., 2004; Ivanov and Ziebuhr, 2004; Walker et al., 2021; Yan et al., 2021a). In this pathway, nsp12 NiRAN functions as a GTase to transfer a GMP to ppA to produce GpppA (Lehmann et al., 2015; Walker et al., 2021; Yan et al., 2021a). The relevance and the efficiency of the conventional and the nsp9-RNAylation-mediated capping pathways raised our interest.

In our studies, we have shown that the two capping pathways have comparable in vitro efficiencies to form GpppA (Figure S5G). The attenuation of GpppA formation was observed by the mixed two pathways. Park et al. show that the RNA with 5'-diphosphate end (pp-RNA), which is the product catalyzed by nsp13 in the conventional capping pathway, cannot be used for nsp9 RNAylation and nsp9-RNAylation-mediated capping (Park et al., 2022). The attenuation of GpppA formation in the mixed system is therefore not surprising. We favor the idea that the conventional and the nsp9-RNAylation-mediated capping pathways can both exist, but there is no crosstalk between them.

In previously reported RTC structures, nsp13 and NiRAN are located at the two farthest ends of RTC. Furthermore, the distance between the 5' end of the product RNA to nsp12 NiRAN is over 120 Å in currently known RTC structures. This distance is too great to easily transfer the 5' end of the product RNA to nsp12 NiRAN for capping. These results indicate that either (1) an additional nsp13 is recruited into RTC adjacent to NiRAN or (2) the RTC would oligomerize to allow the close contact of nsp13 with nsp12 NiRAN to unwind the product-template RNA duplex and catalyze the removal of the $\gamma$-phosphate group from the 5' end of the product RNA to generate the 5'-diphosphate end for the conventional capping pathway. Although RTC with these properties have not been structurally characterized, the alternative nsp9-RNAylation-mediated capping pathway is consistent with the available structural data. In the nsp9-RNAylation-mediated capping pathway, nsp12 NiRAN catalyzes nsp9 RNAylation and subsequent GpppA formation, indicating NiRAN can initiate the capping actions in the absence of an additional nsp13. Such a model fits better with the currently known RTC structures. Therefore, we would like to propose that in the RTC in which no additional nsp13 is adjacent to nsp12 NiRAN, the nsp9-RNAylation-mediated RNA capping may play a major role.
role in capping. In a more complicated RTC in which an additional nsp13 would be adjacent to nsp12 NiRAN, nsp13 is present to allow pp-pRNA to initiate the conventional capping pathway.

Roles of nsp9 in SARS-CoV-2 replication/transcription
CoV-encoded nsp9 has been known to have RNA and DNA binding ability and to co-localize with NSPs, suggesting it is a key component of RTC (Bost et al., 2000; Brockway et al., 2003; Egloff et al., 2004; Sutton et al., 2004). A recent structural study showed that nsp9 tightly binds nsp12 NiRAN (Yan et al., 2021a). Upon binding with nsp12 NiRAN, nsp9 has been suggested to have multiple roles in SARS-CoV-2 replication/transcription. First, nsp9 is NMPylated by nsp12 NiRAN, thus priming the synthesis of viral RNA (Shannon et al., 2022; Slanina et al., 2021; Wang et al., 2021). Second, Park et al. and the present work show that nsp9 is RNAylated by nsp12 NiRAN, thus being an intermediate that contributes viral RNA capping (Park et al., 2022). The RNA binding ability of nsp9 may allow nsp9 to capture the free ppp-RNA unwound by helicase and carry the RNA together to the catalytic center of nsp12 NiRAN for capping. It is noteworthy that the RNAylation of nsp9 is reversible in the presence of PPI but not GTP (Wang et al., 2021); on the contrary, the RNAylation of nsp9 is reversible in the presence of GTP/GDP but cannot be broken by PPI (Park et al., 2022; Wang et al., 2021), suggesting these two functions of nsp9 would not cross-talk, and thus, their reaction mechanisms are likely different. Third, nsp9 acts as an adaptor to recruit nsp14, possibly as well as nsp16, to RTC for the last two capping actions (Yan et al., 2021b).

Mechanisms of nucleotide analog inhibitors that prevents SARS-CoV-2 proliferation
NAI is a major type of antiviral inhibitors against SARS-CoV-2. Conventionally, the active form of NAI is known to be incorpo-rated into the product RNA chain by the polymerase reaction, mimicking the natural NTPs. The incorporated NAI-NTPs can then terminate or stall the synthesis of RNA or result in lethal mutagenesis in the product RNA (Kabinger et al., 2021; Kokic et al., 2021; Wang et al., 2020; Yin et al., 2020).

In this work, our results show that NiRAN covalently binds to four NAI-NTP at nsp9. This is consistent with a previous result that shows remdesivir can be covalently linked to nsp9 (Wang et al., 2021). Our results also showed that the covalent bond of NAI-MP to nsp9 is not reversible in the presence of pppRNA, GTP, GMPPPNP, GDP, and GMP (Figures S5F and S5H), thus inhibiting subsequent processing. In addition, ATDP with the ligation of Mg²⁺ can bind to the G-pocket (Shannon et al., 2022), suggesting AT-527 may have an additional impact on GpppA formation by out competing GTP from the G-pocket. Since the G-pocket has the guanosine base specificity, it is conceivable that only NAI-NTPs derived from GTP/GDP may have such an additional impact.

It is noteworthy that the decoration of NAI-MP to nsp9 can be removed by PPI (Figure S5H), which is similar to the removal of nsp9 NMPylation by PPI (Wang et al., 2021). An interesting observation is that when we reacted the mixture of STP and ATP/UTP/CTP/GTP with nsp9 and nsp12, STP had optimal bonding efficiency to nsp9 (Figure S5I). Meanwhile, the decoration of SMP and UMP to nsp9 cannot be competed by each other (Figures S5J and S5K), suggesting the NAI-NTPs may also interfere nsp9 NMPylation and it primed RNA synthesis in virus life cycle. In summary, these data provide important insights into the understanding of the mechanism of action of the NAIs against SARS-CoV-2.

An “induce-and-lock” mechanism for NiRAN inhibitor design
The rapid emergence of the SARS-CoV-2 variants challenges the available vaccines and neutralizing antibodies, urging the demand to understand the molecular mechanisms underlying the life cycle of the virus and to develop broad-spectrum antiviral drugs (Micocohova et al., 2021; Planas et al., 2021). The high conservation among emerging SARS-CoV-2 variants, the lack of homologs in humans, as well as the essential role in viral RNA capping make NiRAN an ideal target to discover antiviral drugs. The binding of sofosbuvir monophosphate to the Nuc-pocket leads to a sharp downward movement of S-loop through formation of interactions between the NiRAN residues and the chemical modifications on the ribose ring of sofosbuvir. Such a movement of the S-loop narrows the entrance of the G-pocket, which blocks the turnover of GTP used for the formation of GpppA.

We also conducted a molecular dynamics (MDs) simulation to investigate the interactions of nucleotide with Nuc-pocket and S-loop. A coordinate system was defined to monitor the feasibility of the nucleophilic attack of nsp9 in free amino group on the x-phosphorus of NTP, in which the distance between N and P is defined as “d” and N-P<sub>x</sub>-O<sub>α</sub> is defined as “θ” (Figure S6C). The optimal geometry for initiating the nucleophilic attack is demonstrated by positioning the N in the direction of the reverse extension of P<sub>x</sub>-O<sub>α</sub>-P<sub>y</sub> (i.e., θ = 180°) (Gao and Yang, 2016). In a 200-ns MDs simulation, UTP presented better geometry alignment between its x-phosphorus and the free amino group of nsp9 in terms of facilitating nucleophilic attack, compared with other nucleotides (Figure S6D). Such a difference in positional alignment for different nucleotides is associated with the motion of the nucleobases (Figure S6E). The uracil base intends to shift in a relatively small region close to the S-loop, whereas the other three nucleobases were in more diverse locations, implying that the uracil base is favored in the Nuc-pocket. Furthermore, it is interesting to see that the interaction of nucleobase with S-loop orients the nucleotides in a position feasible for nucleophilic attack (Figure S6F).

These results allow us to propose an “induce-and-lock” mechanism that can be used to develop inhibitors for SARS-CoV-2 RNA capping in nsp12 NiRAN. An ideal inhibitor ruled by this mechanism would have three parts. (1) A GTP/GDP-like part that binds in the G-pocket. This would inhibit the binding of GTP and prevent GpppA from forming. (2) A SMP-like part binds in the Nuc-pocket and induces a conformational change to S-loop to provide additional assistance to lock a GTP/GDP-like part in the G-pocket, thus conceivably increasing the binding affinity of the inhibitor to NiRAN and advancing its ability to out compete GTP/GDP. The introduction of halogen atoms or hydrogen bond donors on the ribose group of the SMP-like part could enhance interactions with the main chain carbonyl groups. (3) A linker part is needed that fuses the GTP/GDP-like and the SMP-like parts. We believe that such compounds would be candidates as broad-spectrum antiviral drugs that target SARS-CoV-2 and variants.
Limitations of the study
This study has two limitations. First, the structure of a free ppp-RNA in complex with RTGs remains unclear. We have tried to collect a cryo-EM data set for a sample incubating free ppp-RNA with E-RTC:ns9 in the presence of Ca\textsuperscript{2+}, but no convincing density that supports building ppp-RNA can be observed. This is likely due to the relatively low binding affinity between free ppp-RNA and RTC. Second, according to the current known structural and assay results, the catalytic residues of NiRAN for the conventional pathway and the pathway reported overlap, making it difficult to assess the exclusive impact of the pathway reported here. Further investigations on this topic are warranted.

DECLARATION OF INTERESTS
The authors declare no competing interests.

REFERENCES

Afonine, P.V., Klahtolz, B.P., Moriarty, N.W., Poon, B.K., Sobolev, O.V., Terwilliger, T.C., Adams, P.D., and Urzhumtsev, A. (2018). New tools for the analysis and validation of cryo-EM maps and atomic models. Acta Crystallogr. D Struc. Biol. 74, 814–840.

Bost, A.G., Carnahan, R.H., Lu, X.T., and Denison, M.R. (2000). Four proteins processed from the replicase gene polyprotein of mouse hepatitis virus localize in the cell periphery and adjacent to sites of virion assembly. J. Virol. 74, 3379–3387.

Bouvet, M., Debonnet, C., Imbert, I., Seilsko, B., Snijder, E.J., Canard, B., and Decroly, E. (2010). In vitro reconstitution of SARS-coronavirus mRNA cap methylation. PLoS Pathog. 6, e1000863.

Brockway, S.M., Clay, C.T., Lu, X.T., and Denison, M.R. (2003). Characterization of the expression, intracellular localization, and replication complex association of the putative mouse hepatitis virus RNA-dependent RNA polymerase. J. Virol. 77, 10515–10527.

Bussi, G., Donadio, D., and Parrinello, M. (2007). Canonical sampling through velocity rescaling. J. Chem. Phys. 126, 014101.

Chen, J., Malone, B., Llewellyn, E., Grasso, M., Shelton, P.M.M., Oliareas, P.D.B., Maruthi, K., Eng, E.T., Vatandaslar, H., Chai, B.T., et al. (2020). Structural basis for helicase-polymerase coupling in the SARS-CoV-2 replication-transcription complex. Cell 182, 1560–1573.e13.

Chen, S., McMullan, G., Faruqi, A.R., Murshudov, G.N., Short, J.M., Scheres, S.H., and Henderson, R. (2013). High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. Ultramicroscopy 125, 24–35.

Chen, Y., Cai, H., Pan, J., Xiang, N., Tien, P., Ahola, T., and Guo, D. (2009). Functional screen reveals SARS coronavirus nonstructural protein ns9 as a novel cap N7 methyltransferase. Proc. Natl. Acad. Sci. USA 106, 3484–3489.

Chen, Y., Su, C., Ke, M., Jin, X., Xu, L., Zhang, Z., Wu, A., Sun, Y., Yang, Z., Tien, P., et al. (2011). Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2'-O-methylation by ns9/18ns19 protein complex. PLoS Pathog. 7, e1002294.

Daffis, S., Szerret, K.J., Schriewer, J., Lin, T.Y., Schneller, S., Zurst, R., Dong, H., et al. (2010). 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. Nature 468, 452–456.

Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: an N-\log(N) method for Ewald sums in large systems. J. Chem. Phys. 98, 10089–10092.

Daura, X., Gademann, K., Jaun, B., Seebach, D., van Gunsteren, W.F., and Mark, A.E. (1999). Peptide folding: when simulation meets experiment. Angew. Chem. Int. Ed. Engl. 38, 236–240.

Decroly, E., Imbert, I., Coutard, B., Bouvet, M., Seilsko, B., Alvarez, K., Gorbalenya, A.E., Snijder, E.J., and Canard, B. (2008). Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing (nucleoside-2'-O)-methyltransferase activity. J. Virol. 82, 8071–8084.

Egloff, M.P., Ferron, F., Campanacci, V., Longhi, S., Rancurel, C., Dutartre, H., Snijder, E.J., Gorbalenya, A.E., Cambillau, C., and Canard, B. (2004). The severe acute respiratory syndrome-coronavirus replicative protein ns9 is a molecular dynamic simulation. L.Y., Z. Lou, and Z.R. analyzed the data. Z.R., Z. Lou, L.Y., and L.W.G. wrote the manuscript. All authors discussed the experiments, read, and approved the manuscript.

SUPPLEMENTAL INFORMATION
Supplemental information can be found online at https://doi.org/10.1016/j.cell.2022.09.037.

ACKNOWLEDGMENTS
This work was supported by the National Natural Science Foundation of China (32188101 and U20A20135), the National Program on Key Research Project of China (2020YFA0707500 and 2017YFC0840300), and Tsinghua University Spring Breeze Fund. We thank Dr. Jin Li of Tsinghua University for their assistance for mass spectrum analysis and Dr. Hongjie Zhang (the Core Facility of Protein Sciences, Institute of Biophysics, CAS) for technical assistance. We also thank the Bio-Electron Microscopy Facility of ShanghaiTech University and the Tsinghua University Branch of the China National Center for Protein Sciences (Beijing) for the cryo-EM facility.

AUTHOR CONTRIBUTIONS
L.Y., Z.R., and Z. Lou conceived the project and designed the experiments. L.Y., Y.H., J.G., P.L., S.G., J.W., L.T., S.Y., and S.X. performed sample purification and RTG assembly. L.Y., Y.H., and J.G. for biochemical assays. W.L., Z. Liu, F.Y., and Y.G. collected cryo-EM data. B.H., F.Z., and L.S. performed molecular dynamic simulation. L.Y., Z. Lou, and Z.R. analyzed the data. Z.R., Z. Lou, L.Y., and L.W.G. wrote the manuscript. All authors discussed the experiments, read, and approved the manuscript.
single-stranded RNA-binding subunit unique in the RNA virus world. Proc. Natl. Acad. Sci. USA 101, 3792–3796.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501.

Gao, Y., and Yang, W. (2016). Capture of a third Mg\(^{2+}\) is essential for catalyzing DNA synthesis. Science 352, 1334–1337.

Ivanov, K.A., Thiel, V., Dobbe, J.C., van der Meer, Y., Snijder, E.J., and Ziebuhr, J. (2004). Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase. J. Virol. 78, 5619–5632.

Ivanov, K.A., and Ziebuhr, J. (2004). Human coronavirus 229E nonstructural protein 13: characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5′-triphosphatase activities. J. Virol. 78, 7833–7838.

Kabinger, F., Stiller, C., Schmitzova, J., Diemennan, C., Koci, G., Hillen, H.S., Hübartner, C., and Cramer, P. (2021). Mechanism of mpolurivir-induced SARS-CoV-2 mutagenesis. Nat. Struct. Mol. Biol. 28, 740–746.

Koci, G., Hillen, H.S., Tegunov, D., Diemennan, C., Seitz, F., Schmitzova, J., Farnung, L., Siewert, A., Hübartner, C., and Cramer, P. (2021). Mechanism of SARS-CoV-2 polymerase stalling by remdesivir. Nat. Commun. 12, 279.

Lehmann, K.C., Gulyaeva, A., Zevenhoven-Dobbe, J.C., Janssen, G.M.R., Ruben, M., Overkleeft, H.S., van Veelen, P.A., Samborskiy, D.V., Kravchenko, A.A., Leontovich, A.M., et al. (2015). Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses. Nucleic Acids Res. 43, 8416–8434.

Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J.L., Dror, R.O., and Shaw, D.E. (2010). Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins 78, 1950–1958.

Lou, Z., and Rao, Z. (2022). The life of SARS-CoV-2 inside cells: replication-transcription complex assembly and function. Annu. Rev. Biochem. 91, 381–401.

Malone, B.F., Perry, J.K., Olinares, P.D.B., Chen, J., Appelby, T.K., Feng, J.Y., Bielejo, J.P., Ng, H., Sotrini, J., Ebrahim, M., et al. (2022). Structural basis for substrate selection by the SARS-CoV-2 replicase. Preprint at bioRxiv. https://doi.org/10.1101/2022.05.20.492815.

Mark, P., and Nilsson, L. (2001). Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. J. Phys. Chem. A. 105, 9954–9960.

Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. 152, 486–501.

Park, G.J., Osinski, A., Hernandez, G., Eitson, J.L., Majumdar, A., Tonelli, M., Berlow, M., Overkleeft, H.S., van Veelen, P.A., Samborskiy, D.V., Kravchenko, A.A., Leontovich, A.M., et al. (2015). Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses. Nucleic Acids Res. 43, 8416–8434.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods 14, 290–296.

Rohou, A., and Grigorieff, N. (2015). CTFIND4: fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. 192, 216–221.

Rosenthal, P.B., and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J. Mol. Biol. 333, 721–745.

Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. 180, 519–530.

Sousa da Silva, A.W., and Vranken, W.F. (2012). ACPPYPE - AnteChamber Python parser interface. BMC Res. Notes 5, 367.

Sutton, G., Fry, E., Carter, L., Sainsbury, S., Walter, T., Nettleship, J., Berrow, N., Owens, R., Gilbert, R., Davidson, A., et al. (2004). The nsp9 replicase protein of SARS-coronavirus, structure and functional insights. Structure 12, 341–353.

ván Vliet, A.L., Smits, S.L., Rottier, P.J., and de Groot, R.J. (2002). Discontinuous and non-discontinuous subgenomic RNA transcription in a nidovirus. EMBO J. 21, 6571–6580.

Veyron, S., Oliva, G., Rolando, M., Buchrieser, C., Peyroche, G., and Cherfils, J. (2019). A Ca\(^{2+}\)-regulated deAMPylation switch in human and bacterial FIC proteins. Nat. Commun. 10, 1142.

Walker, A.P., Fan, H., Keown, J.R., Knight, M.L., Grimes, J.M., and Fodor, E. (2021). The SARS-CoV-2 RNA polymerase is a viral RNA capping enzyme. Nucleic Acids Res. 49, 13019–13030.

Wang, B., Svetlov, D., and Artsimovitch, I. (2021). NMPylation and de-NMPylation of SARS-CoV-2 nsp9 by the NIRAN-RdRp subunit. Nucleic Acids Res. 49, 8822–8835.

Wang, J., Wang, W., Kolman, P.A., and Case, D.A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. J. Mol. Graph. Model. 25, 247–260.

Zhao, W., Gao, M., et al. (2020). Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. Science 368, 3474–3485.e11.

WHO (2020). WHO coronavirus disease (COVID-19) dashboard. https://covid19.who.int/.
### STAR★METHODS

#### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| **Bacterial and virus strains** |        |            |
| E. coli BL21 (DE3)  | TIANGEN | Cat# CB105 |
| **Chemicals, peptides, and recombinant proteins** |        |            |
| n-Dodecyl β-D-maltoside (DDM) | INALCO | Cat# 17581350 |
| GDP                 | SIGMA  | Cat# G7127 |
| GTP                 | BBI     | Cat# i114KA3300 |
| NaF                 | SIGMA  | Cat# 201154 |
| Be2SO4              | ALADDIN | Cat# B106711 |
| Remdesivir-TP       | MedChemExpress | Cat# HY-126303 |
| Sofosbuvir-TP       | MedChemExpress | Cat# HY-15745 |
| Molnupiravir-TP     | MedChemExpress | Cat# HY-135853 |
| AT-9010             | MedChemExpress | Cat# HY-139165 |
| **Deposited data**  |        |            |
| E-RTC:RNA-nsp9 EM map | This paper | EMD: 34308 |
| E-RTC:RNA-nsp9 coordinate | This paper | PDB: 8GWB |
| E-RTC:RNA-nsp9:GMPPNP EM map | This paper | EMD: 34310 |
| E-RTC:RNA-nsp9:GMPPNP coordinate | This paper | PDB: 8GWE |
| E-RTC:RNA-nsp9:GTP EM map | This paper | EMD: 34311 |
| E-RTC:RNA-nsp9:GTP coordinate | This paper | PDB: 8GWF |
| E-RTC:RMP-nsp9:GMPPNP EM map | This paper | EMD: 34314 |
| E-RTC:RMP-nsp9:GMPPNP coordinate | This paper | PDB: 8GWK |
| E-RTC:SMC-nsp9:GMPPNP EM map | This paper | EMD: 34312 |
| E-RTC:SMC-nsp9:GMPPNP coordinate | This paper | PDB: 8GWG |
| E-RTC:MMP-nsp9:GMPPNP EM map | This paper | EMD: 34316 |
| E-RTC:MMP-nsp9:GMPPNP coordinate | This paper | PDB: 8GWM |
| E-RTC:ATMP-nsp9:GMPPNP EM map | This paper | EMD: 34317 |
| E-RTC:ATMP-nsp9:GMPPNP coordinate | This paper | PDB: 8GWN |
| E-RTC:SMC-nsp9:GTP EM map | This paper | EMD: 34313 |
| E-RTC:SMC-nsp9:GTP coordinate | This paper | PDB: 8GWI |
| E-RTC:UMP-nsp9 EM map | This paper | EMD: 34302 |
| E-RTC:UMP-nsp9 coordinate | This paper | PDB: 8GW1 |
| E-RTC:EMP-nsp9:GMPPNP EM map | This paper | EMD: 34318 |
| E-RTC:EMP-nsp9:GMPPNP coordinate | This paper | PDB: 8GWO |
| **Oligonucleotides** |        |            |
| Template RNA: 5’-CAUGCAUGCCUCUCUAAAUGUCAGCUUCCUCGACAGCUACUACCAGAGCAUG-3’ | Takara | N/A |
| Primer RNA: 5’-GCCGAGUAGCAUGCUAGGGGACG-3’ | Takara | N/A |
| 10-nt RNA of SARS-CoV-2 genome 5’ UTR : 5’-pppAUUAAAGGCUU-3’ | ChemGenes | N/A |
| **Recombinant DNA** |        |            |
| pET22b-nsp12       | This paper | N/A |
| pET22b-nsp7        | This paper | N/A |

(Continued on next page)
**RESOURCE AVAILABILITY**

**Lead contact**
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Zihe Rao (raozh@tsinghua.edu.cn).

**Materials availability**
This study did not generate new unique reagents.

**Data and code availability**
The cryo-EM density maps and the structures were deposited into the Electron Microscopy Data Bank (EMDB) and Protein Data Bank (PDB) with the accession numbers 34308 and 8GW8 for E-RTC:RNA-nsp9, 34310 and 8GW8 for E-RTC:RNA-nsp9:GMPPNP, 34311 and 8GW9 for E-RTC:RNA-nsp9:GTP, 34314 and 8GWK for E-RTC:RMP-nsp9:GMPPNP, 34317 and 8GWN for E-RTC:ATMP-nsp9:GMPPNP, 34316 and 8GEM for E-RTC:MMP-nsp9:GMPPNP, 34312 and 8GWG for E-RTC:SMP-nsp9:GMPPNP, 34313 and 8GIW for E-RTC:SMP-nsp9:GTP, 34302 and 8GWO for E-RTC:UMP-nsp9, 34318 for 8GW0 for E-RTC:UMP-nsp9:GMPPNP.

**EXPERIMENTAL MODEL AND SUBJECT DETAILS**

Proteins were obtained through recombinant expression in *E. coli* BL21 (DE3).

**METHODS DETAILS**

**Protein production and purification**
The gene encoding nsp12 of SARS-CoV-2 (GenBank: MN908947) was cloned into a modified pET-22b vector, with the C terminus possessing a 10 x His-tag. Protein was expressed in *E. coli* BL21 (DE3) as described previously (Yan et al., 2021a; Yan et al., 2021b). The transformed cells were cultured at 37 °C in LB media containing 100 mg/L ampicillin. After the OD<sub>600</sub> reached 0.6, the culture was

---

**Continued**

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| pET28b-SUMO-nsp8    | This paper | N/A |
| pET28a-nsp13        | This paper | N/A |
| pET28a-nsp9         | This paper | N/A |

**Software and algorithms**

- **SerialEM**
  - Mastronarde, 2005
  - http://bio3d.colorado.edu/SerialEM

- **MotionCor2** (Zheng et al., 2017)
  - https://emcore.ucsf.edu/ucsf-motioncor2

- **RELION 3.0**
  - Scheres, 2012
  - https://www2.mrc-imb.cam.ac.uk/relion

- **cryoSPARC**
  - Punjani et al., 2017
  - https://cryosparc.com/

- **UCSF Chimera**
  - Pettersen et al., 2004
  - https://www.cgl.ucsf.edu/chimera

- **Coot**
  - Emsley et al., 2010
  - https://www.cgl.ucsf.edu/chimera

- **PHENIX**
  - Afonine et al., 2018
  - https://www.phenix-online.org

- **PyMOL**
  - Schrodinger, LLC
  - Schrodinger

**Other**

- **Superdex-200 10/300 Increase**
  - GE Healthcare
  - Cat# 28990944

- **Superdex-75 10/300 Increase**
  - GE Healthcare
  - Cat# 29148722

- **Hitrap-Q HPGE Healthcare**
  - GE Healthcare
  - Cat# 17115401

- **Hitrap-SP HP**
  - GE Healthcare
  - Cat# 17115201

- **Mono-Q 5/50 GL**
  - GE Healthcare
  - Cat# 17516601

- **10 kDa cutoff concentrators**
  - Millipore
  - Cat# UFC01096

- **100 kDa cutoff concentrators**
  - Millipore
  - Cat# UFC91096

- **R0.6/1.0 200 mesh Cu holey carbon grids**
  - Quantifoil
  - Cat# Q250CR-06
cooled to 16°C and supplemented with 0.25 mM IPTG. After overnight induction, the cells were harvested through centrifugation, and the pellets were resuspended in lysis buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 4 mM MgCl₂, 10% glycerol) and then lysed by high-pressure homogenization and sonication at 4°C. The insoluble material was removed by centrifugation at 14,000 rpm for 50 min. The fusion protein was first purified by Ni-NTA (Novagen, USA) affinity chromatography and then further purified by passage through a Hitrap Q ion-exchange column (GE Healthcare, USA) with buffer A (20 mM Tris-HCl, pH 8.0, 4 mM MgCl₂, 10% glycerol, 4 mM DTT) and buffer B (20 mM Tris-HCl, pH 8.0, 1 M NaCl, 4 mM MgCl₂, 10% glycerol, 4 mM DTT). Next, it was loaded onto a Superdex 200 10/300 Increase column (GE Healthcare, USA) with DEPC-treated buffer (50 mM HEPES, pH 7.0, 100 mM NaCl, 4 mM MgCl₂, 4 mM DTT). Purified nsp12 was concentrated to 7 mg/mL and stored at 4°C.

SARS-CoV-2 nsp7 and nsp8 were co-expressed in E. coli BL21 (DE3) cells as a no-tagged protein and a 6x His-SUMO fusion protein, respectively (Yan et al., 2021b). The co-transformed cells were cultured at 37°C in LB media containing 100 mg/L ampicillin and 50 mg/L kanamycin. After the OD₆₀₀ reached 0.8, the culture was supplemented with 0.5 mM IPTG. After shaking at 26°C for 8 hours, the cells were harvested through centrifugation, and the pellets were resuspended in buffer 1 (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 4 mM MgCl₂) and homogenized with an ultra-high-pressure cell disrupter at 4°C. After purification by Ni-NTA (Novagen, USA) affinity chromatography, the nsp7-nsp8 complex was eluted through on-column tag cleavage by SUMO protease (ULP). The complex was further purified by using a Hitrap Q ion-exchange column (GE Healthcare, USA) and a Superdex 75 10/300 column (GE Healthcare, USA) in buffer 2 (50 mM HEPES, pH 7.0, 100 mM NaCl, 4 mM MgCl₂). Purified nsp7-nsp8 complex was concentrated to 20 mg/mL and stored at 4°C.

The SARS-CoV-2 nsp9 was cloned into a pET-28a vector with a N-terminal 6x His-tag and an ALVKQ sequence which has an enzymatic cleavage site for nsp5 to produce an unmodified N-terminal for nsp9. Nsp9 was expressed in E. coli strain BL21 (DE3). After harvesting by centrifugation, the pellets were resuspended in lysis buffer (20 mM Tris, pH 8.0, 150 mM NaCl, 4 mM MgCl₂). Next, homogenization was completed using a ultra-high-pressure cell disrupter at 4°C. The product was then centrifuged at 13000 rpm for 40 min. The fusion protein was purified by Ni-affinity chromatography and the ALVKQ sequence was removed by nsp5. The nsp9 with its native N-terminal was further purified by passage through a Hitrap SP ion-exchange column (GE Healthcare, USA) with buffer A (20 mM Tris, pH 7.5, 4 mM MgCl₂, 4 mM DTT) and buffer B (20 mM Tris, pH 7.5, 1 M NaCl, 4 mM MgCl₂, 4 mM DTT). The eluant was then loaded onto a Superdex 75 10/300 Increase column (GE Healthcare, USA) with buffer C (50 mM HEPES, pH 7.0, 100 mM NaCl, 2 mM MnCl₂, 4 mM DTT). The purified nsp9 was concentrated to 23 mg/mL and stored at 4°C.

Nsp13 was purified as described previously (Yan et al., 2021a). The nsp13 gene was inserted into the modified pET-28a vector with a 6x His tag and thrombin enzyme sites attached at its N-terminus, and the protein was expressed in E. coli strain BL21 (DE3) cells. Cells were harvested and resuspended in lysis buffer (20 mM HEPES, pH 7.0, 150 mM NaCl, 4 mM MgCl₂, 10% glycerol). The cells were centrifuged at 14,000 rpm for 40 min and then lysed by high-pressure homogenization and sonication. The fusion protein was first purified by Ni-NTA (Novagen, USA) affinity chromatography. Eluted nsp13 was incubated with thrombin protease to cleave the 6x His tag in digestion buffer (20 mM HEPES, pH 7.0, 500 mM NaCl, 4 mM MgCl₂, 10% (v/v) glycerol). Cleaved nsp13 was run through a Ni-affinity column for three times to remove undigested nsp13. The flow through was then further purified by ion exchange column Hitrap S (GE Healthcare, USA), and then the sample was loaded onto a Superdex 200 10/300 Increase column (GE Healthcare, USA) in a buffer (50 mM HEPES, pH 7.0, 100 mM NaCl, 2 mM MnCl₂). Purified nsp13 was concentrated to 5 mg/mL and stored at 4°C.

Preparation of the RNAylated nsp9
Nsp9 (50 μL, 23 mg/mL) was incubated with nsp12 (50 μL, 2 mg/mL) and pppRNA (80 μL, 3.43 mg/mL) and inorganic pyrophosphatase (10 μL, 100 U/mL) at 37°C for one hour in a buffer containing 50 mM HEPES, pH 7.0, 100 mM NaCl and 2 mM MnCl₂. The mixture was purified by Hitrap Q ion-exchange column (GE Healthcare, USA), with buffer A (50 mM Tris-HCl, pH 8.0, DEPC treated) and buffer B (50 mM Tris-HCl, pH 8.0, 1 M NaCl, DEPC treated), resulting in the RNA-nsp9 complex.

Purification of NAI-MP-nsp9
Nsp9 (50 μL, 23 mg/mL) was incubated with nsp12 (50 μL, 2 mg/mL) and NAI-TP (10 μL, 50 mg/mL) and inorganic pyrophosphatase (10 μL, 100 U/mL) at 37°C for one hour in a buffer containing 50 mM HEPES, pH 7.0, 100 mM NaCl and 2 mM MnCl₂. Then the mixture was loaded onto a Superdex-75 10/300 Increase column in a buffer (50 mM HEPES, pH 7.0, 100 mM NaCl and 2 mM MnCl₂), resulting in the formation of the NAI-MP-nsp9 complex (RMP-nsp9, or ATMP-nsp9, or MMP-nsp9, or SMP-nsp9). The UMP-nsp9 complex was purified in same way.

NAI-MP-nsp9 and RNA-nsp9 of LC-MS analysis
NAI-MP-nsp9 and RNA-nsp9 were separated by a 10 min gradient elution at a flow rate 0.5 mL/min with an ACQUITY UPLC system, which was directly interfaced with a SYNAPT-G2-Si mass spectrometer produced by Waters. Separation was achieved using a Protein BEH C4 silica capillary column (2.1 mm ID, 50 mm length; Made in Ireland) packed with C-4 resin (300 Å, 1.7 μm) purchased from Waters. Mobile phase A consisted of 0.1% formic acid aqueous solution, and mobile phase B consisted of 100% acetonitrile and 0.1% formic acid.

Aliquots of 1 μL samples were loaded into an autosampler for electrospray ionization. Samples were analyzed on a Q-TOF mass spectrometer (SYNAPT G2-Si, Waters company) instrument optimized for high-mass protein analysis. The measurements were
performed with a 3000–3500 V capillary. Data were collected over the m/z range of 500–2000. Having acquired the raw electrospray mass spectra, the data were deconvoluted using MaxEnt 1 (Waters) to generate a spectrum (relative intensity versus mass) where all the charge-state peaks of a single species have been collapsed into a single (zero-charge) peak.

Biochemical assays for GpppA-RNA synthesis
For these assays we used the Vaccinia capping system (except for SAM) (NewEngland Biolabs) and [α-32P]GTP to label the 5' terminus of the RNA (ChemGenes, USA) as a positive control. GpppA-RNA activity was tested as follows: 1 μg nsp12 or mutant nsp12 was mixed with 5 μg nsp9-RNA and [α-32P]GTP in a buffer consisting of 50 mM HEPES, pH 7.0, 6 mM KCl, 4 mM DTT, 2 mM MgCl2. After incubation at 37 °C for 60 min, a G-50 Sephadex column (Roche) was used to remove excess α-[32P]GTP. GpppA-RNA was extracted with phenol-chloroform and precipitated with ethanol. 5 μg nuclease P1 (NewEngland Biolabs) and 1 μM ZnCl2 were added to digest the RNA. Reaction products were spotted on polyethylenimine cellulose plates (Merck) to separate G'pppA. The plates were visualized using a PhosphorImager (PerkinElmer, USA).

Biochemical assays for MTase activity
MTase assay was performed in 40 μl volume: 2 μl nsp12 (2 μg/μl), 2 μl RNA-nsp9 (5 μg/μl), 2 μl GTP (0.5 mM), 1μl [α-32P]GTP (PerkinElmer, USA) in a buffer of 50 mM HEPES, pH 7.0, 6 mM KCl, 4 mM DTT, 2 mM MgCl2, 2 mM MnCl2. After incubation at 37 °C for 30 min, 2 μl nsp14 (2 μg/μl) and 1 μl SAM was added to the reactions and incubated at 37 °C for 30 min. For 2'-O methylation, 2 μl nsp10/16 (2 μg/μl) was added in the mixture. A G-50 Sephadex column (Roche, USA) was used to remove excess α-[32P]GTP.

Assembly of the RTC
Nsp12 was incubated with the nsp7/nsp8 complex at 4 °C for three hours at a molar ratio of 1:2.2 in a buffer containing 50 mM HEPES, pH 7.0, 100 mM NaCl and 2 mM MnCl2. Then the mixture was purified by mono Q 5/50 ion-exchange chromatography (GE Healthcare, USA), resulting in a stable nsp7-nsp8-nsp12 complex (C-RTC). C-RTC and nsp13 and RNA were mixed to form E-RTC at a 1:2.2:1 molar ratio, as described previously (Yan et al., 2020b). The E-RTC was incubated with RNA-nsp9 at a 1:1.2 molar ratio with or without 0.5 mM GMPPNP to assemble the E-RTC:RNA-nsp9:GMPPNP and E-RTC:RNA-nsp9. E-RTC:NAI-MP-nsp9:GMPPNP, E-RTC:UMP-nsp9:GMPPNP, and E-RTC:UMP-nsp9 were prepared in same way.

Cryo-EM sample preparation and data collection
In total, 3 μl of protein RNA complex at 3 mg/mL (added with 0.025% DDM) was applied onto a H2/O2 glow-discharged, 200-mesh Quantifoil R0.6/1.0 grid (Quantifoil, Micro Tools GmbH, Germany). The grid was then blotted for 3.0 s with a blot force of 0 at 8 °C and 100% humidity and plunge-frozen in liquid ethane using a Vitrobot (Thermo Fisher Scientific, USA). Cryo-EM data were collected with a 300 keV Titan Krios electron microscope (Thermo Fisher Scientific, USA) and a K3 direct electron detector (Gatan, USA). Images were recorded at 29000× magnification and calibrated at a super-resolution pixel size of 0.82 Å/pixel. The exposure time was set to 2 s with a total accumulated dose of 60 electrons per Å2. All images were automatically recorded using SerialEM (Mastronarde, 2005) and were collected with a defocus range from -2.0 μm to -1.0 μm. Statistics for data collection and refinement are provided in Tables 1A–1C. The methods for processing are described in Data S1.

Cryo-EM image processing
All dose-fractioned images were motion-corrected and dose-weighted by MotionCorr2 (Zheng et al., 2017) software and their contrast transfer functions were estimated by ctfind4 (Rohou and Grigorieff, 2015). Eight RTC samples particles were auto-picked using the model from SARS-CoV-2 Cap(+1')-RTC (PDB: 7CYQ) (Yan et al., 2021a) and extracted with a box size of 448 pixels in cryoSPARC (Punjani et al., 2017). The subsequent 2D, 3D classifications and refinements were all performed in cryoSPARC. The particles were selected after two rounds of 2D classification based on complex integrity. This particle set was used to perform Ab-Initio reconstruction in four classes, which were then used as 3D volume templates for heterogeneous refinement. Next, the particles were imported into RELION 3.03 (Scheres, 2012) to perform local classification to obtain one class particle. The methods are graphically described in Data S1.

Model building and refinement
To build the structure of the SARS-CoV-2 RTC complex, the coordinates of the SARS-CoV-2 nsp12 and nsp7-8 complex (PDB: 7BTF), SARS-CoV-2 nsp13 (PDB: 6ZSL), nsp9 (PDB: 6W9Q) were individually placed and fitted by rigid-body refinement into the cryo-EM map using UCSF Chimera (Petterson et al., 2004). The model was manually built with RNA-nsp9 or NAI-MP-nsp9 or UMP-nsp9 using coot (Emsley et al., 2010) with the guidance of the cryo-EM map, and in combination with real space refinement using Phenix (Afonine et al., 2018). The data validation statistics are shown in Tables S1 and S2.
Molecular dynamics simulation
To prepare the starting model for the simulation, nsp9 and nsp12 were extracted. UTP, ATP, GTP, and CTP were respectively put in Nuc-pocket, with the base, the ribose, and the \( \alpha \)-phosphate in the position of the nucleotide counterparts and the \( \gamma \)-phosphate extending towards R116 and K73. The amber99sb-star-ildn force field (Lindorff-Larsen et al., 2010) and TIP3P model (Mark and Nilsson, 2001) was used for the protein and water, respectively. The GAFF2 force field (Wang et al., 2006) parameters were generated by ACPYE (Sousa da Silva and Vranken, 2012) and used for ATP, CTP, GTP and UTP. 150 mM sodium chloride was added to make the whole system neutral with respect to charge. The energy minimization was performed using steepest descent method, followed by 400 ps NVT, 200 ps NPT simulation and production runs of 300 ns. The overall temperature of the system was kept constant, coupling independently for protein/ligand and solvents/ions at 300 K with a V-rescale thermostat (Bussi et al., 2007). The pressure was coupled to a Parrinello-Rahman barostat (Parrinello and Rahman, 1981) at 1 atm separately in every dimension. The temperature and pressure time constants of the coupling were 0.1 and 2 ps respectively. The integration of the equations of motion was performed by using a leapfrog algorithm with a time step of 2 fs. Periodic boundary conditions were implemented in all systems. A cutoff of 1.0 nm was implemented for the Lennard–Jones and the direct space part of the Ewald sum for Coulombic interactions. The Fourier space part of the Ewald splitting was computed by using the particle-mesh-Ewald method (Darden et al., 1993), with a grid length of 0.12 nm on the side and a cubic spline interpolation. The structures of the 100-300 ns trajectory were clustered using GROMOS algorithm as previously described (Daura et al., 1999).

QUANTIFICATION AND STATISTICAL ANALYSIS
In Data S1, the resolution estimations of cryo-EM density maps are based on the 0.143 Fourier Shell Correlation (FSC) criterion (Chen et al., 2013; Rosenthal and Henderson, 2003)
Figure S1. Purification of the modified nsp9 and GTP used in this study, related to Figure 1A

Time-dependence of nsp9 RNAylation. S759A-nsp12, apo-nsp9 and the 10-nt RNA were incubated under the condition described in the STAR Methods for 5, 10, 20, 40, 80 and 120 min, respectively. The reaction products were assessed by MOPS 4%–20% gradient SDS-PAGE. The molecular weights for the standard markers are displayed. The bands corresponding to nsp12, apo-nsp9 and RNA-nsp9 are indicated.

(B) Purification of RNA-nsp9. The product of nsp9 RNAylation reaction was loaded on to Hitrap Q ion-exchange column (GE Healthcare, USA) and was eluted with the gradient of NaCl from 50 to 370 mM. The peak for RNA-nsp9 is indicated in the figure.

(C) The purity of GTP used in this study was verified by HPLC analysis. Chromatographic separation of GDP and GTP were carried out using an Ultimate XB-C18 Column (Welch Technology Shanghai, China). The injection volume used was 10 μL and total chromatographic run time was 30.0 min. The mobile phase used consisted of 0.1% (v/v) trifluoroacetic acid in water (A) and 0.1% (v/v) acetonitrile (B). At a flow rate of 1 mL/min, the gradient conditions were from 0% (B) to 30% (B) in 30 min.

(D) The purity of [α-32P]GTP used in this study was checked by thin layer chromatography (TLC). As a control, [α-32P]GTP was treated by nsp13 to produce the hydrolyzed product [α-32P]GDP.

(E–I) Purifications of NAI-labeled or UMP labeled nsp9. The products of nsp12 and nsp9 incubation with remdesivir triphosphate (E), sofosbuvir triphosphate (F), monopiravir triphosphate (G), AT-9010 triphosphate (H) or UTP (I) were loaded on Superdex-75 column (GE Healthcare, USA). The peak positions for NAIs or UMP decorated nsp9 are indicated in each panel.
Figure S2. Interactions of ligands with SARS-CoV-2 RTC and molecular details for the Nuc-pocket in MMP-nsp9 and RMP-nsp9 bound RTC and the G-pocket in ATMP-nsp9, MMP-nsp9 and RMP-nsp9 bound RTC, related to Figures 2, 3, 4I, and 5A–5E.

(A–E) LIGPLOT diagrams show critical contacts between the RNA, GMPPNP, and NAIs with RTC in the structure of E-RTC:RNA-nsp9 (A), E-RTC:RNA-nsp9:GMPPNP (B), E-RTC:RMP-nsp9:GMPPNP (C), E-RTC:SMP-nsp9:GMPPNP (D) and E-RTC:MMP-nsp9:GMPPNP (E). The interacting components are depicted in ball-and-stick mode. Gray ball, carbon; blue ball, nitrogen; red ball, oxygen; cyan ball, fluorine.

(F and G) The polypeptides of nsp12 NiRAN and nsp9 are shown as semi-transparent colored cartoons. The bound RMP (F) or MMP (G), the interacting residues and GMPPNP are shown as colored sticks. The black dashed lines denote the intermolecular interactions.

(H–J) The polypeptides of nsp12 NiRAN and nsp9 are shown as semi-transparent colored cartoons. The bound RMP (H), MMP (I) or ATMP (J), as well as the interacting residues and GMPPNP, are shown as colored sticks.
Figure S3. Structural details of E-RTC:RNA-nsp9 and E-RTC:RNA-nsp9:GMPPNP and comparison with previously reported RTC structures, related to Figures 2, 3, and 5

(A) The binding of RNA and GMPPNP in nsp12 NiRAN. The molecules of RNA and GMPPNP are shown as colored sticks, while nsp12/nsp9 polypeptides are covered by an electrostatic potential surface. The positions of G-pocket and Nuc-pocket are indicated by arrows.

(B) A schematic representation of the interactions between RNA, GMPPNP and nsp12 NiRAN. The dashed lines indicate the intermolecular interactions.

(C) Two perpendicular views of the comparison of the binding of A+1/U+2 in the structures of E-RTC:RNA-nsp9 with or without GMPPNP. The residues of RTC and A+1/U+2 in E-RTC:RNA-nsp9 are shown as colored sticks, while they are colored in white in the structure of E-RTC:RNA-nsp9:GMPPNP. The distance for shifting and the angle for rotation are indicated as arrows.

(D) Comparison of GMPPNP and AT-9010-DP binding in the G-pocket. Structures of E-RTC:RNA-nsp9:GMPPNP and C-RTC:AT-9010-DP (Shannon et al., 2022) are aligned and are shown in the same orientation. The polypeptides of nsp12 and nsp9 are shown as a semi-transparent cartoon. A+1 nucleotide, nsp9N1, GMPPNP and three catalytic residues nsp12D208/N209/D218 are shown as colored sticks, while AT-9010-DP and nsp12D208/N209/D218 in the structure of C-RTC:AT-9010-DP are shown as gray sticks. One Mg$^{2+}$ ion in the structure of E-RTC:RNA-nsp9:GMPPNP and two Mg$^{2+}$ ions in the structure of C-RTC:AT-9010-DP are shown as green and gray spheres, respectively.

(E) Comparison of the cryo-EM densities for the catalytic center of nsp12 NiRAN in two structures. GMPPNP, Mg$^{2+}$ and Mn$^{2+}$ and the solvent molecules are shown as previously used. The cryo-EM density covering GMPPNP in the structure of E-RTC:RNA-nsp9:GMPPNP is shown as gray mesh, while the density covering Mn$^{2+}$ ions and the coordinated waters is shown as red mesh.

(F and G) Close-up views of the catalytic center of nsp12 NiRAN in the structure of E-RTC:RNA-nsp9 (F) and E-RTC:apo-nsp9:GDP (Yan et al., 2021a) (G). All representations are used as the same in Figure 2A.

(H) Comparison of A+1 nucleotide in E-RTC:RNA-nsp9 and GDP in E-RTC:apo-nsp9:GDP. Structures of E-RTC:RNA-nsp9 and E-RTC:apo-nsp9:GDP are aligned and shown the same three perpendicular views. The polypeptides are shown as semi-transparent cartoons. The carbon atoms for A+1 nucleotide, nsp9N1, the catalytic residues in the structure of E-RTC:RNA-nsp9 are colored as green, purple and yellow, respectively. Whereas, the carbon atoms for GDP, nsp9N1, the catalytic residues in the structure of E-RTC:apo-nsp9:GDP are colored in white. Mn$^{2+}$ and Mg$^{2+}$ are shown as gray and green spheres.

(I) Superimpose of E-RTC:apo-nsp9:GDP with E-RTC:RNA-nsp9:GMPPNP. The structural elements are shown as previously used. For clearer representation, GMPPNP in E-RTC:RNA-nsp9:GMPPNP is hidden, and only GDP in E-RTC:apo-nsp9:GDP is shown with C atom colored by white.

(J) Comparison of GDP, ADP and GMPPNP in RTC structures. The structures of E-RTC:apo-nsp9:GDP (PDB: 7CYQ), E-RTC:ADP (PDB: 6XEZ) and E-RTC:RNA-nsp9:GMPPNP are aligned. GDP, ADP and GMPPNP in them are shown as white, red and magenta sticks.
Figure S4. Structural comparison of GTP and GMPPNP bound in the G-pocket and sofosbuvir-specific conformational change of S-loop and structural arrangements following that, related to Figures 3, 4, and 5

(A) GTP in the structure of E-RTC:RNA-nsp9:GTP and GMPPNP in the structure of E-RTC:RNA-nsp9:GMPPNP were aligned and shown in the same orientation. GTP and nsp9N1 in E-RTC:RNA-nsp9:GTP were shown as colored sticks, while the structural elements in E-RTC:RNA-nsp9:GMPPNP were shown as white color with the exception that the nitrogen atom in P-N bond is colored as blue.

(B) The density for GTP in E-RTC:RNA-nsp9:GTP.

(C) GTP in the structure of E-RTC:SMP-nsp9:GTP and GMPPNP in the structure of E-RTC:SMP-nsp9:GMPPNP were aligned and shown in the same orientation.

(D) The density for GTP in E-RTC:SMP-nsp9:GTP.

(legend continued on next page)
(E) The density covering the environment surrounding bb-wat in E-RTC:SMP-nsp9:GTP. The representation scheme is the same as in Figure 5.

(F) The cryo-EM density covering the S-loop in the structure of E-RTC:SMP-nsp9:GMPPNP is shown as gray mesh, while the S-loop is displayed as colored sticks.

(G) Comparison of the S-loop in reported structures. The polypeptides spanning the S-loop in each structures indicated in the panel are shown as colored ribbons.

(H–K) Close-up views of the catalytic center of nsp12 NiRAN in E-RTC:UMP-nsp9 (H and J) and in E-RTC:UMP-nsp9:GMPPNP (I and K) are shown in the same representations as those in Figure 5.

(L and M) Close-up views of the catalytic center in E-RTC:RNA-nsp9:GMPPNP (L) and E-RTC:SMP-nsp9:GMPPNP (M). The polypeptides of nsp12 and nsp9 are covered by electrostatic potential surfaces. A+1 nucleotide, SMP, GMPPNP are shown as colored sticks. The positions of nsp12K50/N52 and nsp12T51 of the S-loop are indicated in (L) and (M) by arrows.

(N) Comparison of GMPPNP binding. GMPPNP molecules in the structures of RTC:RNA-nsp9:GMPPNP and E-RTC:SMP-nsp9:GMPPNP are shown as colored sticks, in which the carbon atoms are colored as magenta and white, respectively. The S-loop and nsp12T51 are also shown as colored cartoons and sticks.

(O and P) Comparison of the arrangement of the catalytic center. The polypeptides of nsp12 are shown as semi-transparent cartoon, while A+1 nucleotide, SMP, GMPPNP and the interacting residues are shown as colored sticks. It should be noted that only the interacting residues with significant conformational change are shown here, and the interacting residues with conserved positions are hidden for clear representation. The black dashed lines denote the intermolecular interactions, while the red dashed lines denote the close distance between Mg²⁺ and GMPPNP.
Figure S5. Covalent bond of NAI to ns9 and relevance of ns9-RNAylation-mediate capping pathway with currently known evidence, related to Figures 1 and 4A–4D

(A–D) SDS-PAGE and mass spectrum analysis of the products for the incubation of ns9, ns12 and NAI-TPs. The protein of ns12 was incubated with ns9 and remdesivir-TP (RTP) (A), sofosbuvir-TP (STP) (B), molnupiravir-TP (MTP) (C) and AT-9010-TP (ATTP) (D), all in the presence of Mn²⁺. The reaction products were subsequently analyzed using a MOPS 4%–20% gradient SDS-PAGE gel stained with R250. The molecular weights of the standard protein markers are shown in the panel. The bands corresponding to ns12, apo-ns9 and NAI-ns9 are indicated. The purified NAI-ns9 proteins were further analyzed by Q-TOF mass spectrometry.

(E) The purified RMP-/ATMP-/MMP-/SMP-ns9 or apo-ns9 proteins were incubated with ns12 and the 10-nt RNA in the presence of Mn²⁺ and the products were analyzed using a MOPS 4%–20% gradient SDS-PAGE gel stained with R250.

(F) The purified RMP-/ATMP-/MMP-/SMP-ns9 were incubated with ns12 and GTP in the presence of Mg²⁺. The products were analyzed by Q-TOF mass spectrometry. The peaks with molecular weights ranging from 11,000 to 14,000 Da are shown in each panel.

(G) The relative efficiency of the conventional pathway with ns9-RNAylation-mediate capping pathway. We used the vaccinia capping system (except for SAM) (NewEngland Biolabs, USA) and [α-³²P]-GTP to label the 5’ terminus of the RNA (ChemGenes, USA) as a positive control. (1) E-RTC (first pathway only): 2 μL nsp12 (2 μg/μL) and 1 μL nsp13 (2 μg/μL) was mixed with pppRNA and [α-³²P]-GTP in a buffer consisting of 50 mM HEPES, pH 7.0, 6 mM KCl, 4 mM DTT, 2 mM MgCl₂ and 2 mM MnCl₂. (2) E-RTC and apo-ns9 (both pathways): 2 μL nsp12 (2 μg/μL), 2 μL ns9 (10 μg/μL) and 1 μL nsp13 (2 μg/μL) was mixed with pppRNA and [α-³²P]-GTP in a buffer consisting of 50 mM HEPES, pH 7.0, 6 mM KCl, 4 mM DTT, 2 mM MgCl₂ and 2 mM MnCl₂. (3) RTC (with no nsp13) and apo-ns9 (second pathway only): 2 μL nsp12 (2 μg/μL), 2 μL ns9 (10 μg/μL) was mixed with pppRNA and [α-³²P]-GTP in a buffer consisting of 50 mM HEPES, pH 7.0, 6 mM KCl, 4 mM DTT 2 mM MnCl₂.

(H) Competition of SMP decoration by GTP, GDP, GMP, PPi and GMPPNP. The purified SMP-ns9 (20 μg/μL) was treated by GTP (50 μM), GDP (50 μM), GMP (50 μM), PPi (50 μM) and GMPPNP (50 μM) for 30 min at 30°C. The reaction products were analyzed by SDS-PAGE.

(I–K) Competition of SMP decoration by the native NTPs. In the (C), the purified apo-ns9 (20 μg/μL) was incubated with the equal concentration (100 μM) of STP, ATP, UTP, CTP, GTP for 60 min at 30°C. The reaction products were by Q-TOF mass spectrometry. Please note, because of the very similar molecular weights of UMP-ns9 and CMP-ns9, the peak at 12,683 Da cannot distinguish them. In the (D), the purified SMP-ns9 (5 μg/μL) was incubated with UTP (100 μM) for 30 min at 30°C. In the (E), the purified UMP-ns9 (5 μg/μL) was incubated with STP (100 μM) for 30 min at 30°C. The reaction products were by Q-TOF mass spectrometry.
Figure S6. Propose reaction mechanisms and molecular dynamics based analysis on the interaction of NTP with Nuc-pocket, related to Figures 5 and 6

(A and B) Propose reaction mechanisms for nsp9-RNAylation-mediated GpppA formation.

(C) Coordinate system of the nucleophilic attack of nsp9N1 free amino group on the $\alpha$-phosphorus of NTP. The distance between N and $\alpha$-P is defined as $d$, and N:$\alpha$-P:z$\beta$-O is defined as $\theta$. The optimal alignment of the free amino group of nsp9N1 and the $\alpha$-phosphorus of NTP for nucleophilic attack is demonstrated by assigning $\theta \approx 180^\circ$.

(D) The distribution of $\theta$ and $d$ within 200-ns molecular dynamics.

(E) Representative conformations of NTP observed in the molecular dynamics simulation. Clustering analysis was performed on molecular simulation trajectories to get the representative conformations.

(F) Sample frames showing residues interacting with the nitrogenous base of NTP. The red triangle indicates the position of the free amino group of nsp9N1. The pink circle indicates the optimal direction for nucleophilic attack on $\alpha$-P.