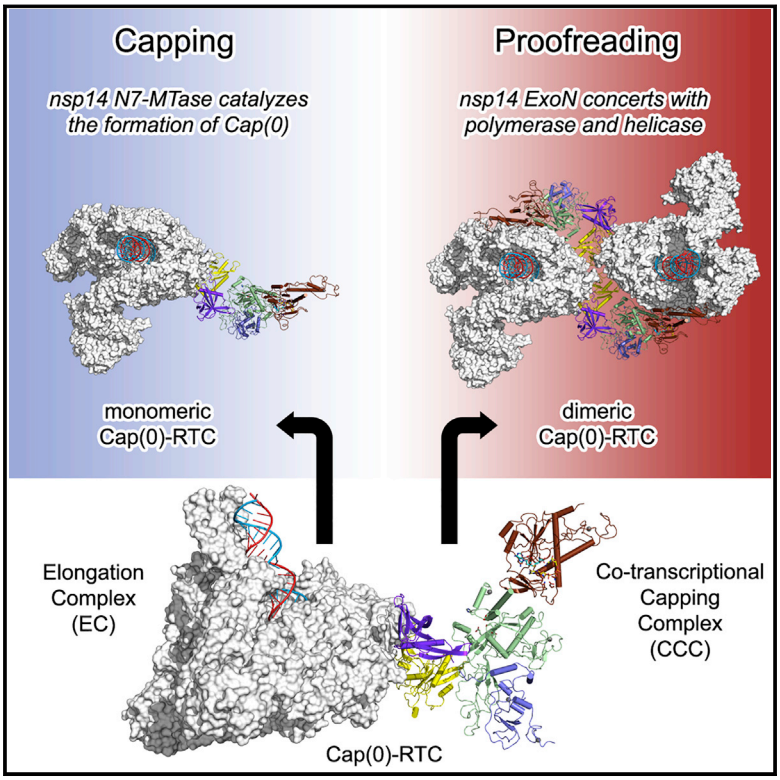
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Cell

Coupling of N7-methyltransferase and 3′-5′ exoribonuclease with SARS-CoV-2 polymerase reveals mechanisms for capping and proofreading

Graphical abstract



Highlights

- Structure of SARS-CoV-2 Cap(0)-RTC is determined
- Cap(0)-RTC couples a co-transcriptional capping complex bearing nsp10/nsp14
- The third action for mRNA capping occurs in Cap(0)-RTC
- The dimeric Cap(0)-RTC suggests an *in trans* backtracking proofreading mechanism

Authors

Liming Yan, Yunxiang Yang, Mingyu Li, ..., Yan Gao, Zihe Rao, Zhiyong Lou

Correspondence

yan-gao13@mails.tsinghua.edu.cn (Y.G.), raozh@tsinghua.edu.cn (Z.R.), louzy@mail.tsinghua.edu.cn (Z.L.)

In brief

Structures of the capping machinery and the polymerase super-complex of SARS-CoV-2 provides insights into virus replication, mRNA capping, and proofreading.

Yan et al., 2021, Cell *184*, 3474–3485
June 24, 2021
https://doi.org/10.1016/j.cell.2021.05.033

CellPress

CellPress

Cell

Article

Coupling of N7-methyltransferase and 3′-5′ exoribonuclease with SARS-CoV-2 polymerase reveals mechanisms for capping and proofreading

Liming Yan,[1,2,10] Yunxiang Yang,[1,4,10] Mingyu Li,[1,10] Ying Zhang,[1,10] Litao Zheng,[2,3,10] Ji Ge,[3,10] Yucen C. Huang,[1,10] Zhenyu Liu,[1,10] Tao Wang,[1,2] Shan Gao,[1] Ran Zhang,[3] Yuanyun Y. Huang,[5] Luke W. Guddat,[6] Yan Gao,[2,*] Zihe Rao,[1,2,3,7,8,9,11,*] and Zhiyong Lou[1,9,*]

[1]MOE Key Laboratory of Protein Science, School of Medicine, Tsinghua University, Beijing, China
[2]Shanghai Institute for Advanced Immunochemical Studies and School of Life Science and Technology, ShanghaiTech University, Shanghai, China
[3]School of Life Sciences, Tsinghua University, Beijing, China
[4]Tsinghua University-Peking University Joint Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing, China
[5]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA, USA
[6]School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, QLD, Australia
[7]State Key Laboratory of Medicinal Chemical Biology, College of Life Sciences and College of Pharmacy, Nankai University, Tianjin, China
[8]National Laboratory of Biomacromolecules, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China
[9]Guangzhou Laboratory, Guangzhou, China
[10]These authors contribute equally
[11]Lead contact
*Correspondence: yan-gao13@mails.tsinghua.edu.cn (Y.G.), raozh@tsinghua.edu.cn (Z.R.), louzy@mail.tsinghua.edu.cn (Z.L.)
https://doi.org/10.1016/j.cell.2021.05.033

SUMMARY

The capping of mRNA and the proofreading play essential roles in SARS-CoV-2 replication and transcription. Here, we present the cryo-EM structure of the SARS-CoV-2 replication-transcription complex (RTC) in a form identified as Cap(0)-RTC, which couples a co-transcriptional capping complex (CCC) composed of nsp12 NiRAN, nsp9, the bifunctional nsp14 possessing an N-terminal exoribonuclease (ExoN) and a C-terminal N7-methyltransferase (N7-MTase), and nsp10 as a cofactor of nsp14. Nsp9 and nsp12 NiRAN recruit nsp10/nsp14 into the Cap(0)-RTC, forming the N7-CCC to yield cap(0) (^{7Me}GpppA) at 5′ end of pre-mRNA. A dimeric form of Cap(0)-RTC observed by cryo-EM suggests an *in trans* backtracking mechanism for nsp14 ExoN to facilitate proofreading of the RNA in concert with polymerase nsp12. These results not only provide a structural basis for understanding co-transcriptional modification of SARS-CoV-2 mRNA but also shed light on how replication fidelity in SARS-CoV-2 is maintained.

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the causative agent of coronavirus disease 2019 (COVID-19). The genome of SARS-CoV-2 encodes 16 non-structural proteins (nsp1–nsp16) to assemble the replication-transcription complex (RTC) that plays key roles in the replication of genome-length viral RNAs and in the transcription of viral mRNAs (Wu et al., 2020). CoV mRNAs bear a 5′ cap(1) structure (^{7Me}GpppA$_{2′OMe}$) that plays an essential role in the life cycle of the virus (Daffis et al., 2010). In SARS-CoV-2 RTC, co-transcriptional capping of mRNA occurs after elongation initiation by four sequential actions in the co-transcriptional capping complex(s) (CCC) (Yan et al., 2021). The first and the second capping actions that sequentially generate the 5′-diphosphate end (ppA) and the cap core (GpppA) at the 5′ end of the nascent pre-mRNA by nsp13 and the nidovirus RdRp-associated nucleotidyltransferase (NiRAN) domain in nsp12. In the subsequent capping actions, an N7-methyltransferase (N7-MTase) in nsp14 methylates the first guanine of GpppA at the N7-position to produce the cap(0) (^{7Me}GpppA), being the substrate for the final capping action facilitated by a 2′-O-methyltransferase (2′-O-MTase) in nsp16 to yield the mature mRNA (Bouvet et al., 2010; Chen et al., 2009, 2011; Decroly et al., 2008; Ivanov et al., 2004; Ivanov and Ziebuhr, 2004; Yan et al., 2021).

CoV nsp14 is a bifunctional enzyme bearing an N-terminal exonuclease (ExoN) domain that has been implicated in proofreading to maintain replication fidelity (Eckerle et al., 2007, 2010; Minskaia et al., 2006) and a C-terminal N7-MTase domain that catalyzes the third mRNA capping action to yield the cap(0)(-Chen et al., 2009). The C-terminal N7-MTase adopts an atypical MTase fold with a conserved DxG S-adenosylmethionine (SAM) binding motif (Case et al., 2016; Ferron et al., 2018; Ma et al., 2015). Compared with the canonical MTase fold, an insertion of a three-stranded β sheet between the canonical MTase β5/β6 strands is observed in CoV nsp14. A similar insertion is also

observed in Vaccinia virus N7-MTase(Grimm et al., 2019). The binding site of SAM is in the space between the loops connecting strands β2′-β3′ and β3′-β4′ and is adjacent to the GpppA binding site (Ma et al., 2015). The N-terminal ExoN is a member of the DEDDh exonuclease superfamily and possesses a 3′-5′ exonuclease activity proposed to correct errors in replication by removing mis-incorporated nucleotides from the 3′ end of the nascent RNA strand (Ferron et al., 2018; Munster et al., 2020; Ogando et al., 2019). A feature of this exonuclease is that it has a proofreading mechanism to maintain the replication fidelity of the ~29.7 kb CoV genome. Nsp14 ExoN has also been shown to play roles in sub-genomic mRNA recombination and primary viral RNA synthesis (Gribble et al., 2021; Ogando et al., 2020). The proofreading is suggested to be a factor that impacts the inhibitory efficacy of nucleotide analog inhibitors (Robson et al., 2020). Escape from nsp14 ExoN proofreading is suggested to be associated with enhanced inhibitory efficacy of β-D-N4-hydroxycytidine (NHC; EIDD-1931) and derivatives against SARS-CoV-2 that has resistance mutations to remdesivir (Robson et al., 2020; Sheahan et al., 2020). Furthermore, the function of nsp14 requires its association with a virus-encoded cofactor nsp10 (Ma et al., 2015). Nsp10 exclusively binds with and confers structural integrity and stability to nsp14 ExoN and thus enhances the activity of nsp14 ExoN over 35-fold, preferring to use double-stranded RNA (dsRNA) with a 3′ mis-matched substrate that mimics an erroneous replication product (Bouvet et al., 2010, 2012). But nsp10 does not impact the structure and the activity of nsp14 N7-MTase (Bouvet et al., 2010; Ma et al., 2015).

In previous studies, the architectures of the central RTC (C-RTC) composed by nsp12 and cofactors nsp7 and nsp8 and the elongation RTC (E-RTC) composed by C-RTC and helicase (nsp13) have been revealed, and the structure of Cap(−1)′-RTC formed by E-RTC with a single-strand RNA binding protein nsp9 reveals an intermediate state toward mRNA capping (Chen et al., 2020; Gao et al., 2020; Wang et al., 2020; Yan et al., 2020, 2021). Interactions between nsp14 (both N7-MTase and ExoN) and the nsp12-nsp7-nsp8 complex (Subissi et al., 2014) indicate its involvement in the assembly of the RTC. However, the structural details as to how the nsp10/nsp14 complex assembles remain elusive, impeding our understanding of the mechanisms of the function of nsp14 in mRNA capping and proofreading. To dissect these mechanisms, we generated a complex [identified as Cap(0)-RTC] formed by Cap(−1)′-RTC (i.e., nsp7-nsp8$_2$-nsp9-nsp12-nsp13$_2$-RNA) and nsp10/nsp14 and determined its cryoelectron microscopy (cryo-EM) structure to 3.78 and 3.35 Å resolution in both monomeric and dimeric forms. The structure not only reveals the architecture of the capping machinery coupled to RTC that achieves the capping actions decorating the 5′ end of pre-mRNA but also provides a structural basis to understand the polymerase activity and the mechanism by which nsp14 ExoN is used for proofreading.

RESULTS

Preparation and structure determination of Cap(0)-RTC

The recombinant proteins used to assemble Cap(0)-RTC were expressed in *E. coli* cells and purified (Figure S1A). In a native electrophoretic mobility shift assay (EMSA), the addition of nsp10/nsp14 to Cap(−1)′-RTC resulted in a smeared upper-shifted band, indicating the formation of a potential complex, possibly in multiple forms (Figure S1B). With this sample, we collected a cryo-EM dataset containing 507 micrograph movies to determine whether a complex had formed. In total, 12% of the particles could be identified as belonging to Cap(−1)′-RTC with additional density, while the other particles had no clear features (Figure S2; Table S1). A reconstruction using 5,985 particles gave a density map at 7.9 Å. Although the resolution was limited due to the small amount of particles, a convincing model could be built. We proposed that additional density adjacent to nsp9 and nsp12 NiRAN likely corresponded to nsp10/nsp14 (Figure S3). We therefore hypothesized the association of nsp9/nsp10/nsp14 may mediate the coupling of nsp10/nsp14 into Cap(0)-RTC, but the small number of interactions between them may result in a low number of complexes being formed or there being inherent flexibility when they associate.

A previous study has shown that the proteolytic processing of nsp9-nsp10 in coronaviruses is not required for virus proliferation and fitness, indicating a natural connection between nsp10 and nsp9 exists in the CoV life cycle (Deming et al., 2007). A strategy where a linker is inserted to fuse two polypeptide components into one protein is an approach that can be used to improve the prospects of forming a homogeneous sample where two individual subunits may only weakly associate (Shabek et al., 2018; Soon et al., 2012; Wang et al., 2017). Using this approach, we fused the C terminus of nsp9 with the N terminus of nsp10 by using the natural polypeptide linker in the viral polyprotein connecting nsp9 and nsp10 and expressed this recombinant nsp9-nsp10 fusion protein together with nsp14 (Figures S1A and 1C). This nsp9-nsp10 fusion protein forms a stable complex with nsp14 as occurs for nsp10 alone (Figure 1C). This purified nsp9-nsp10/nsp14 complex was thus subsequently used to assemble Cap(0)-RTC by incubating with E-RTC [i.e., Cap(−1)′-RTC excludes nsp9].

After collecting and processing 12,704 cryo-EM micrograph movies, 5 major classes were identified: 9.0% belong to a monomeric Cap(0)-RTC [hereafter named as mCap(0)-RTC], 15.3% belongs to a previously unobserved large particle representing a dimeric form of Cap(0)-RTC [hereafter named as dCap(0)-RTC], and 48.4% belongs to E-RTC, and another 2 classes belong to small fragments (Figure S4). The classes representing mCap(0)-RTC (with 29,941 particles) and dCap(0)-RTC (with 47,981 particles) were reconstructed to resolutions of 3.35 and 3.78 Å, respectively (Figure S4; Table S1). The high-quality density clearly shows the position of nsp10/nsp14 adjacent to nsp9 and nsp12 NiRAN with a resolution of 3.0–4.5 Å in this region (Figures S3 and S5A). The local resolution for the density of nsp7-nsp8$_2$-nsp9-nsp12 ranges from 3.0 to 4.0 Å. The local resolution for nsp13-2, which anchors the unpaired 5′ extension of the RNA template, is significantly optimized compared to previously reported RTC structures (Yan et al., 2020, 2021).

Architecture of Cap(0)-RTC

In mCap(0)-RTC, one nsp10/nsp14 complex associates with Cap(−1)′-RTC composed of one nsp7, two nsp8, one nsp9, one nsp12, two nsp13, and a paired template-primer RNA, being

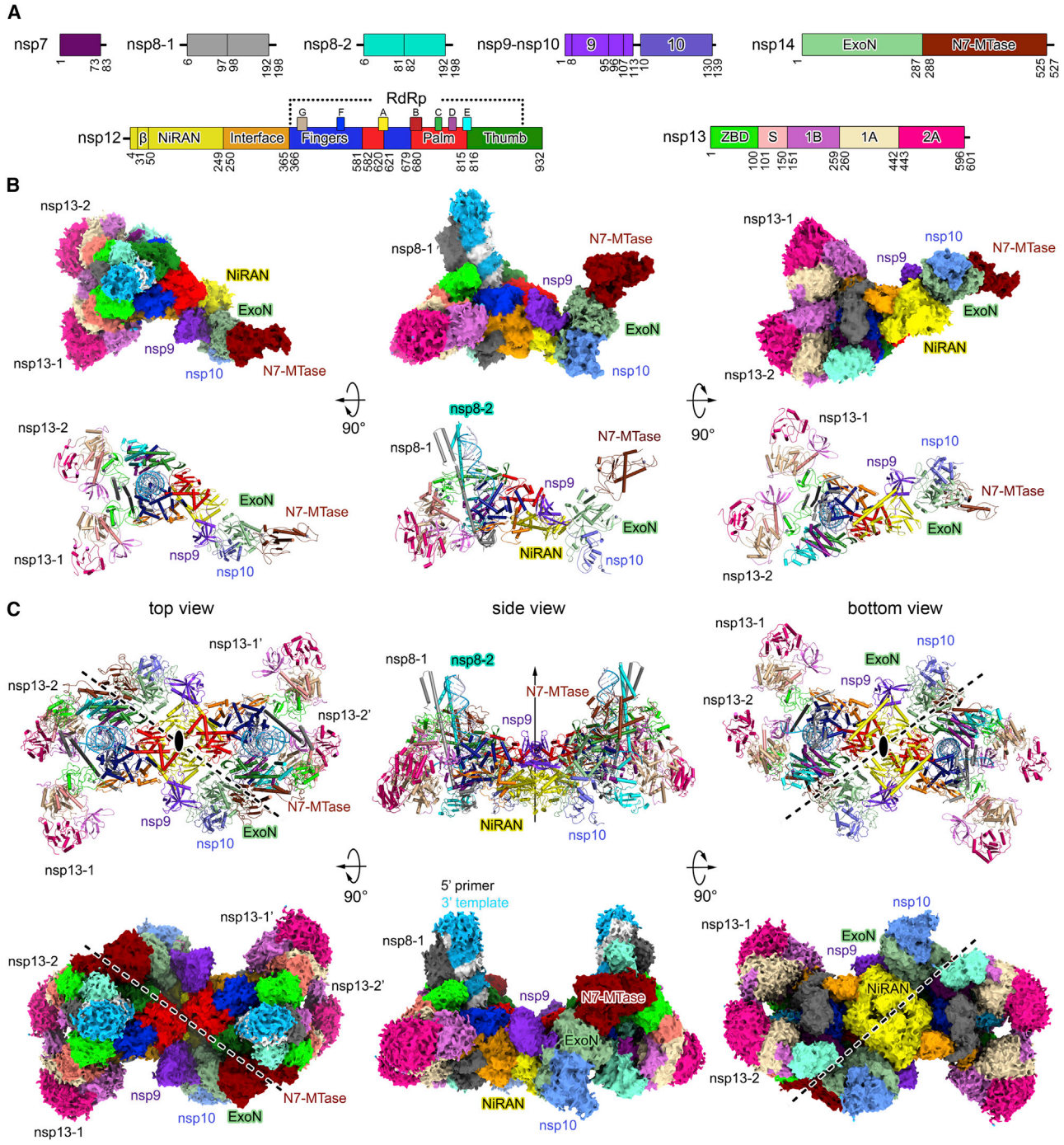


Figure 1. Overall structure

(A) Domain organization of each component in Cap(0)-RTC. The color scheme for each component in Cap(0)-RTC is generally similar to that used previously (Gao et al., 2020; Yan et al., 2020, 2021), with modifications. Nsp7, deep purple; nsp8-1, gray; nsp8-2, green cyan; nsp9, purple blue; nsp10, slate; nsp12 NiRAN, yellow; nsp12 Interface, orange; nsp12 fingers, blue; nsp12 palm, red; nsp12 thumb, forest; nsp13 ZBD, light green; nsp13 S, salmon; nsp13 1B, violet; nsp13 1A, sand; nsp13 2A, hot pink; nsp14 ExoN, pale green; nsp14 N7-MTase, brown.

(B and C) Structure of mCap(0)-RTC (B) and dCap(0)-RTC (C) in cartoon diagrams (top panels) and the cryo-EM densities (bottom panels) are shown in three perpendicular views. The 2-fold axis is indicated by ellipses (left and right panel) or an arrow (middle panel). The dashed lines roughly indicate the boundary of two Cap(0)-RTC protomers in dCap(0)-RTC (C).

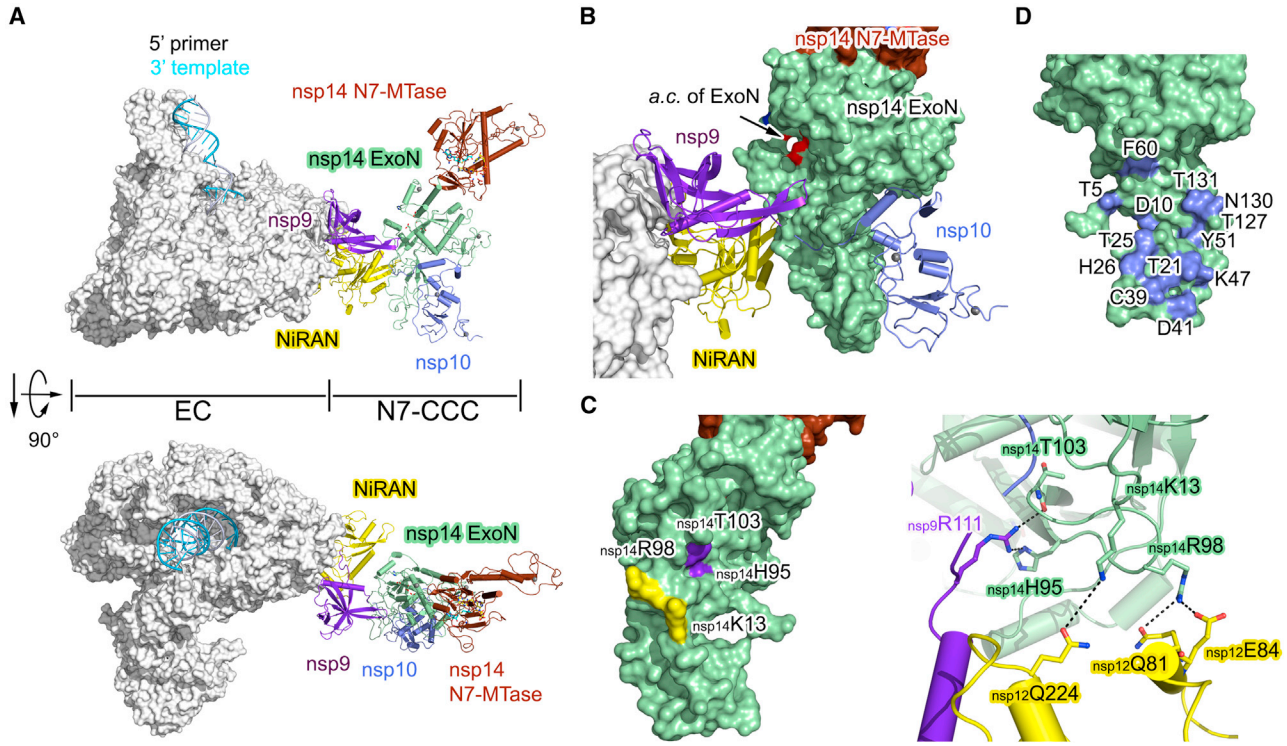


Figure 2. Architecture of N7-CCC

(A) N7-CCC in Cap(0)-RTC protomer is shown as colored cartoons from a side view (top panel) and a top view (bottom panel). For a clear representation, the components of EC in Cap(0)-RTC protomer is shown as a white molecular surface, whereas the primer-template RNAs are shown as cartoon diagrams.

(B–D) Inter-molecular interactions in N7-CCC. The polypeptides of nsp9, nsp10, and nsp12 NiRAN are shown as colored cartoons; nsp14 is covered by a colored surface. The interacting residues of nsp14 to contact with nsp9 and nsp12 are highlighted by the color of their interacting partners with labels (B and C). The inter-molecular interactions of nsp14 with nsp9/nsp12 are shown in detail (C, right panel). The interacting residues of nsp14 to contact with nsp10 are highlighted by the color as in nsp10 (D).

mediated by the contact of nsp9 and nsp12 NiRAN with nsp14 (Figures 1A and 1B). The architecture of Cap(−1)′-RTC is generally similar to a previously reported structure (Yan et al., 2021). Nsp9 and nsp12 NiRAN make contacts with nsp14 ExoN to stabilize the association of nsp10/nsp14 in Cap(0)-RTC but do not interact with nsp14 N7-MTase. According to a previous definition, nsp13-1 refers to the helicase molecule that does not bind with RNA, while nsp13-2 refers to the helicase molecule that binds the 5′ extension of template (Yan et al., 2020, 2021).

In dCap(0)-RTC, two Cap(0)-RTC protomers are oriented by a 2-fold axis to form a dimer (Figures 1A and 1C; Figure S5B). Nsp12 NiRAN and the palm of two Cap(0)-RTC protomers face each other and locate in the center of dCap(0)-RTC, while four nsp13s are at the two distal ends of the dimer. The binding of nsp9 and nsp12 NiRAN with nsp14 ExoN is similar in both mCap(0)-RTC and dCap(0)-RTC. Nsp14 N7-MTase, which is distant from other components in one Cap(0)-RTC protomer, closely contacts with the nsp13-2 zinc binding domain (ZBD)′ and nsp12 thumb′ in another Cap(0)-RTC protomer, playing a key role in dimer formation. Moreover, nsp13-2, which anchors the unpaired 5′ extension of template RNA, shows a significant conformational change in dCap(0)-RTC, compared to nsp13-1 in dCap(0)-RTC, nsp13-2 in mCap(0)-RTC and other reported CoV helicase structures (discussed below).

Structure of N7-CCC

In Cap(0)-RTC, nsp12 NiRAN, nsp9, nsp10, and nsp14 form a CCC catalyzing the methylation of the first guanine at the N7-position to produce the cap(0) (^{7Me}GpppA)(Chen et al., 2009) (hereafter referred to as N7-CCC), being boundaried by nsp12 NiRAN and nsp9 from the elongation complex (EC) composed of nsp7, nsp8, nsp12 Interface/Thumb/Palm/Fingers, and nsp13 (Figure 2A).

In N7-CCC, the rod-shaped nsp10/nsp14 complex vertically inserts into and is stabilized by a canyon region formed by nsp9 and nsp12 NiRAN (Figures 2A and 2B). Nsp9 and nsp10 flank nsp14 ExoN from two sides to stabilize the assembly. Nsp14 N7-MTase protrudes distal from the interface of nsp14 with nsp9 and nsp12 NiRAN. Although nsp10 is fused with nsp9 in Cap(0)-RTC, nsp10 does not contact with nsp9 except for the fusion residues, indicating again that the nsp9-nsp10 fusion does not affect the assembly of RTC. The catalytic center of nsp14 ExoN composed by residues D90/E92/E191/D273/H268 directly faces the central β sheet of nsp9. Moreover, the catalytic center of nsp14 N7-MTase is localized on the opposite side of nsp12 NiRAN.

Nsp14 ExoN contributes the majority of interactions for the association of nsp10/nsp14 into Cap(0)-RTC with a contacting surface of 2,396 Å^2 (Figures 2C and 2D; Table S2). The side-chain

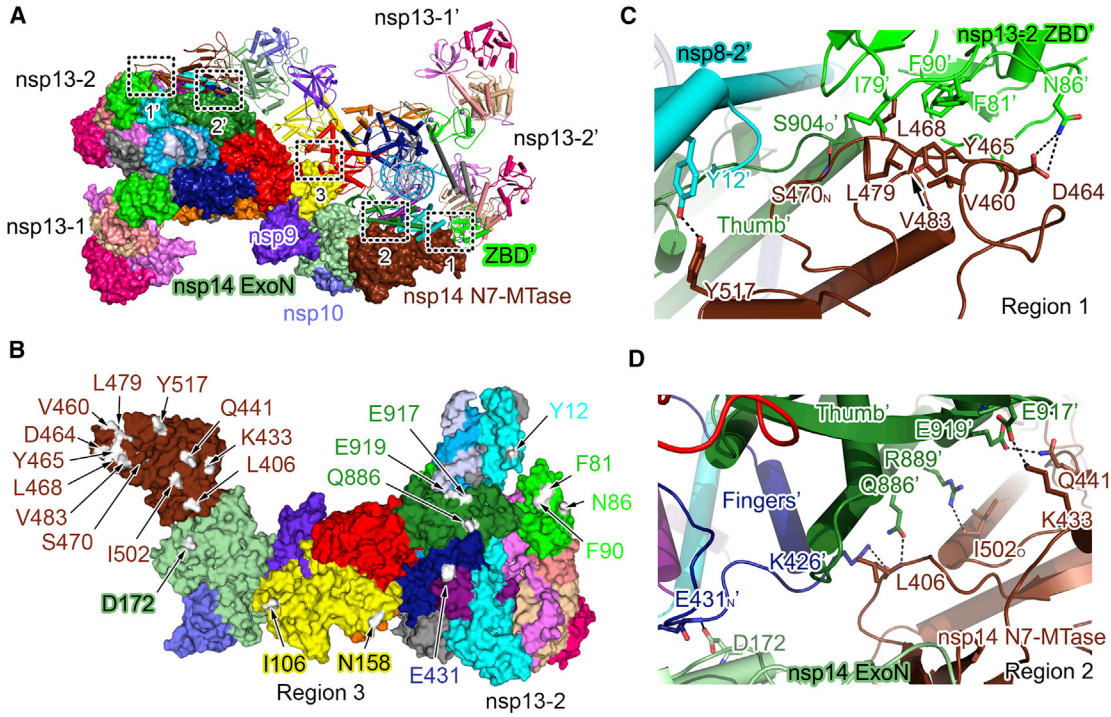


Figure 3. Inter-protomer interactions

(A) One Cap(0)-RTC protomer is shown as a cartoon diagram and the another is shown as a molecular surface. The inter-protomer interacting regions are indicated by the dashed frames.

(B) Key residues on the inter-protomer interface. The Cap(0)-RTC protomer is colored as in Figure 1. The interacting residues are highlighted by a white color on the surface with labels.

(C and D) The interaction details in regions 1 and 2. The interacting residues are displayed as colored sticks. Dashed lines indicate contact distances less than 3.5 Å.

nitrogen atoms of $_{nsp9}$R111 form hydrogen bonds with the imidazole group of $_{nsp14}$H95 and the carbonyl of $_{nsp14}$T103 and a side-chain nitrogen of $_{nsp14}$K13 forms a hydrogen bond with $_{nsp12}$Q224. The side-chain nitrogen atoms of $_{nsp14}$R98 contact with $_{nsp12}$Q81 and $_{nsp12}$E84. The overall structure of nsp10/nsp14 complex in Cap(0)-RTC shows strong similarities with that reported in the crystal structures of nsp10/nsp14 complex (Ma et al., 2015) (Figure S6A). The residues of nsp14 interacting with nsp10 includes $_{nsp14}$T5, $_{nsp14}$D10, $_{nsp14}$T21, $_{nsp14}$T25, $_{nsp14}$H26, $_{nsp14}$C39, $_{nsp14}$D41, $_{nsp14}$K47, $_{nsp14}$Y51, $_{nsp14}$F60, $_{nsp14}$T127, $_{nsp14}$N130, and $_{nsp14}$T131, which are consistent with the reported crystal structure (Ma et al., 2015). Sequence comparison shows that these key interacting residues are highly conserved in SARS-CoV-2, SARS-CoV, MERS-CoV, and RaTG13 (Figure S6B), suggesting a common mechanism for N7-CCC formation in CoVs. The N terminus of nsp9 inserts into the catalytic center of nsp12 NiRAN, which is similar to that observed in Cap(−1)′-RTC (Yan et al., 2021), but the bound GDP-Mg^{2+} in the nsp12 NiRAN catalytic center is missing. It is noteworthy that a recent study suggests that nsp9 is a substrate for nsp12 NiRAN NMPylation, but in another study it is debated that nsp7 is the substrate (Conti et al., 2020; Slanina et al., 2021). We hypothesize that the post-translation modification of nsp9 by nsp12 NiRAN might play a role in complex formation.

Inter-protomer interaction of Cap(0)-RTC

The inter-protomer interactions of dCap(0)-RTC are contributed by nsp7, nsp8-2, nsp12 NiRAN/thumb/palm/fingers, nsp13-2 ZBD, nsp14 ExoN, and N7-MTase at five separate regions (Figure 3A; Table S3).

In region 1, $_{nsp14}$V460, $_{nsp14}$Y465, $_{nsp14}$L479, and $_{nsp14}$V483 of nsp14 N7-MTase in protomer A forms a hydrophobic interface with the residues $_{nsp13\text{-}2}$I79, $_{nsp13\text{-}2}$F81, and $_{nsp13\text{-}2}$F90 of nsp13-2 ZBD′ in protomer B (Figure 3B). Three hydrogen bonds are observed between $_{nsp14}$D464, $_{nsp14}$S470, and $_{nsp14}$Y517 of nsp14 N7-MTase in protomer A and $_{nsp8\text{-}2}$Y12 (in nsp8-2′), $_{nsp12}$S904 (in nsp12 thumb′), and $_{nsp13\text{-}2}$N86 (in nsp13-2 ZBD′) in protomer B (Figure 3C). Notably, the loop region of nsp14 N7-MTase spanning residues $_{nsp14}$S454- $_{nsp14}$D464, which cannot be traced in the reported crystal structure, is well defined with unambiguous cryo-EM density (Figure S6A, right panel). We reason that the contact of this region with nsp13-2 ZBD′ in the other Cap(0)-RTC protomer helps to stabilize its conformation.

In region 2, the residues of nsp14 N7-MTase (including $_{nsp14}$L406, $_{nsp14}$K433, $_{nsp14}$Q441, and $_{nsp14}$I502) and ExoN (residue $_{nsp14}$D172) in protomer A constitute a set of inter-molecular interactions with the residues of nsp12 thumb′ (including $_{nsp12}$Q886, $_{nsp12}$R889, $_{nsp12}$E917, and $_{nsp12}$E919) and nsp12 fingers′ (including $_{nsp12}$K426 and $_{nsp12}$E431) in protomer B (Figure 3D). In

region 3, although nsp12 NiRAN and palm are located in the central parts of dCap(0)-RTC and face each other, only a single bond formed by $_{nsp12}$I106 in protomer A and $_{nsp12}$N158 in protomer B is observed (Figure 3A). It is noteworthy that the inter-protomer interface are located on the opposite side of the SAM and GpppA binding sites of nsp14 N7-MTase, allowing full solvent exposure of the nsp14 N7-MTase catalytic site in dCap(0)-RTC.

Nsp13-2 1B has a significant structural rearrangement

In Cap(0)-RTCs, two nsp13s are located on a plane formed by nsp12 and two nsp8s, which are in the similar position as they are in E-RTC and Cap(−1)′-RTC (Figures 1B and 1C). In mCap(0)-RTC, the folding of the two nsp13s is virtually identical. In dCap(0)-RTC, nsp13-1, which does not bind template RNA and was shown to play a role in the stimulation of the helicase activity of another nsp13 (Yan et al., 2020), exhibit structural similarity in dCap(0)-RTC and E-RTC with a root-mean-square deviation (RMSD) of 0.74 Å for all Cα atoms (Figure 4A, left panel). In contrast, nsp13-2, which binds to the 5′ extension of the unpaired template RNA, exhibits a remarkable conformational change within its 1B domain (Figure 4A, right panel). Nsp13-2 1B in dCap(0)-RTC exhibits a ~90° rotation toward nsp13-2 ZBD alongside the C terminus of nsp13-2 S and is approximately 25 Å from nsp13-2 2A (Figure 4A, right panel). In contrast, the 1B domain in all other reported structures of CoV nsp13 (either crystal or cryo-EM structures) is adjacent to the 2A domain. This previously unobserved position of nsp13-2 1B not only causes variation of the inter-helicase interactions but also results in significant differences in helicase-RNA interaction (Figures 4B–4L).

The movement of nsp13-2 1B results in variations of inter-molecular interactions

In mCap(0)-RTC, the loop region spanning residues $_{nsp13-2}$T214-$_{nsp13-2}$V221 of nsp13-2 1B makes contacts with the long loop region connecting nsp13-1 1B and 1A (Figures 4B and 4F). Similar inter-helicase interactions can be also found in the structures of E-RTC (Yan et al., 2020) and Cap(−1)′-RTC (Yan et al., 2021) (Figures 4D, 4E, 4H, and 4I). Following the movement of nsp13-2 1B in dCap(0)-RTC, the contacts between the 1B domains of the two nsp13 in RTC is completely abolished (Figures 4C and 4G). Moreover, the loop region spanning residues $_{nsp13-2}$G203-$_{nsp13-2}$A208 of nsp13-2 1B in dCap(0)-RTC contacts nsp13-1 ZBD and nsp8-1 (Figure 4J). In this region, the side chain of $_{nsp13-2}$D204 hydrogen bonds with $N^{\eta1}$ of $_{nsp8-2}$R75 in the central long helix of nsp8-2, and the carbonyl oxygen of $_{nsp13-2}$G206 contacts with the side chain of $_{nsp13-1}$Q11 in the ZBD. Previous studies on SARS-CoV-2 RTCs have shown that nsp8-1 participates in the assembly of C-RTC to arm the paired template-product RNA and in the recruitment of nsp13-1 into RTCs (Wang et al., 2020; Yan et al., 2020, 2021). The structural feature presented in dCap(0)-RTC suggests that nsp8-1 also contacts with nsp13-2 with a conformational change of the 1B domain, presumably to play a role in the regulation of the function of nsp13-2.

Variations in the binding of nsp13-2 with 5′ extension of the template RNA

The movement of nsp13-2 1B is associated with the variations in RNA interactions. Although we used the same template and primer RNAs in our investigations on RTCs in this and previous studies (Yan et al., 2020, 2021), the unpaired 5′ extension of the template RNA is not observed in the mCap(0)-RTC like in Cap(−1)′-RTC. We therefore only compared the binding of nsp13-2 with the 5′ extension of template RNA in dCap(0)-RTC and in E-RTC (Figures 4K and 4L).

In E-RTC, tU+7 to tU+12 in the unpaired 5′ extension of template RNA is anchored in a RNA-binding channel clamped by nsp13-2 1A, 2A and 1B domains (Yan et al., 2020) (Figures 4H and 4L). The positively charged side chains of $_{nsp13-2}$N177 and $_{nsp13-2}$R178 of nsp13-2 1B bond with tA+10 and tA+11, and the imidazole group of $_{nsp13-2}$H230 packs with the base of tA+8 (Figure 4L, right panel). In dCap(0)-RTC, the RNA-binding channel is disappears following the downward movement of nsp13-2 1B. As a result, the cryo-EM density shows there is no binding of RNA to nsp13-2 1A and 2A domains. However, the unambiguous cryo-EM density supports a model of template RNA continuously extended from tC+4 to tA+9 (Figures 4G and 4K). The side chains of $_{nsp13-2}$N177 and $_{nsp13-2}$R178 of nsp13-2 1B contact with tU+7 and tA+8 of the unpaired 5′ extension of template RNA, but $_{nsp13-2}$H230 does not interact with template RNA (Figure 4K). The nucleotides of tU+3 to tG+6 in E-RTC cannot be traced in this cryo-EM density map (Yan et al., 2020) but have good density in dCap(0)-RTC (Figure 4G). We reasoned that the contact of $_{nsp13-2}$N177 and $_{nsp13-2}$R178 with RNA in dCap(0)-RTC helps to stabilize the conformation of these four nucleotides. It is noteworthy that the densities of the first 17 nucleotides cannot be visualized in dCap(0)-RTC, though there is sufficient length of RNA to reach nsp13-2 1A/2A, suggesting 1B, but not 1A and 2A, plays the most important role in RNA binding of helicase.

Comparison of the helicase orientations in RTCs

The orientations of helicases show distinct features in different RTCs (Figures 4B-4E). In E-RTC, ZBDs of two nsp13 are placed on the long helical structure of nsp8-1 and nsp8-2 (Yan et al., 2020) (Figure 4D). Although the folding of two helicases are identical in Cap(−1)′-RTC, the orientation of nsp13-2 changes, with nsp13-2 ZBD inserting deeply into the second minor groove of the paired template-primer RNA (Yan et al., 2021) (Figure 4E), reasoning E-RTC is in a "Elongation" state and Cap(−1)′-RTC is in a "Suspending" state.

In mCap(0)-RTC, nsp13-2 has a similar orientation as in Cap(−1)′-RTC and inserts its ZBD into the paired template-primer RNA (Figure 4B), suggesting mCap(0)-RTC is also in a "Suspending" state. In contrast, despite nsp13-2 1B having a large conformational change in dCap(0)-RTC, nsp13-2 ZBD stays in a state as it appears in E-RTC. Both ZBD and 1B with conformational change of nsp13-2 makes no interaction with the paired template-primer RNA (Figure 4C), suggesting the movement of RNA is possible in dCap(0)-RTC as in E-RTC, rather in mCap(0)-RTC.

DISCUSSION

The coupling of capping machinery(s) in SARS-CoV-2 RTC is essential for virus mRNA maturation and gene transcription. In previous studies, the assembly of RTCs for RNA synthesis and

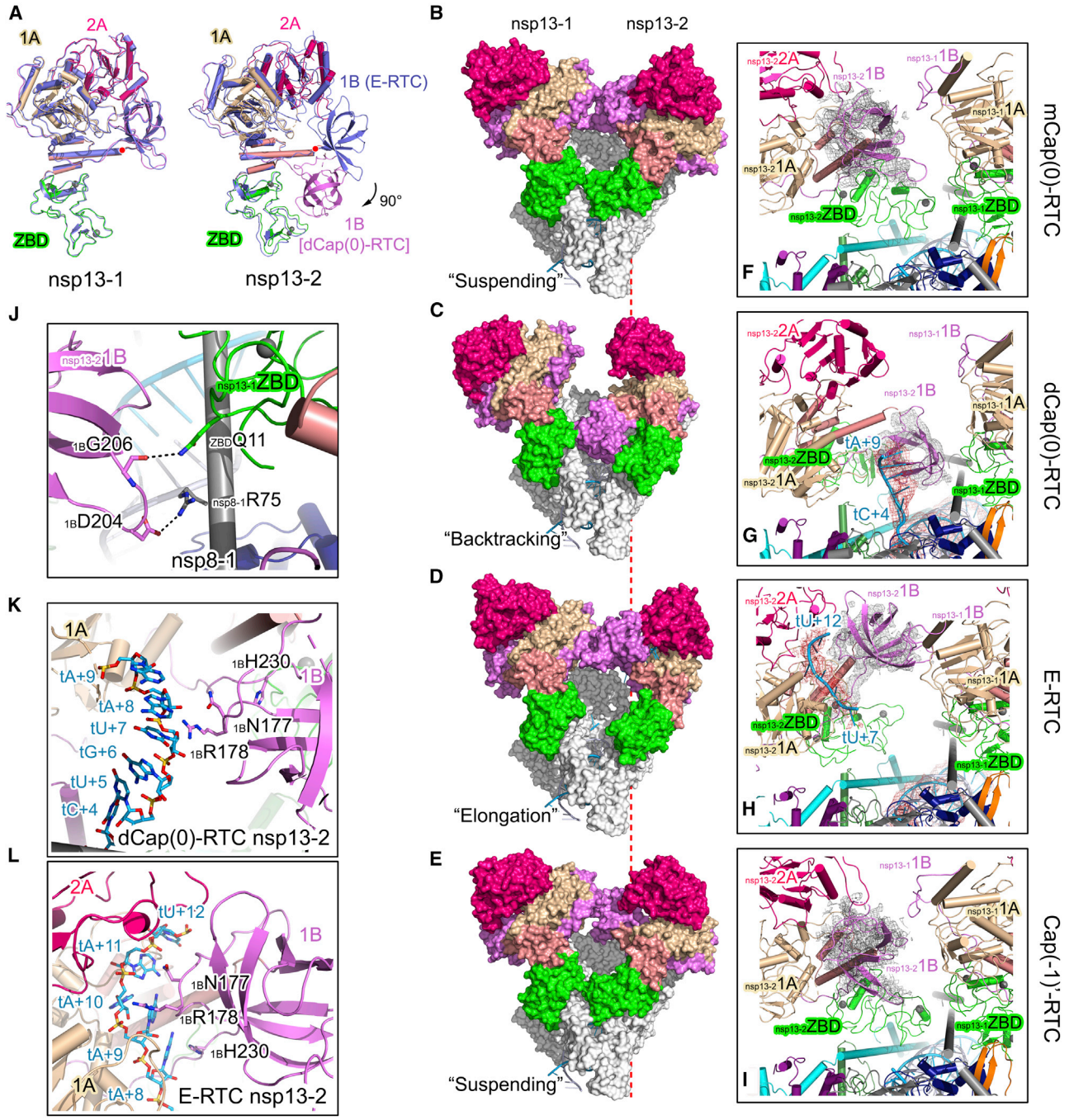


Figure 4. Conformational change of nsp13-2 1B

(A) Comparison of nsp13 in E-RTC and dCap(0)-RTC. The polypeptides of nsp13-1 and nsp13-2 in E-RTC are colored light blue, whereas these in dCap(0)-RTC are shown as the color scheme in Figure 1.

(B–E) Comparison of orientations of nsp13 in mCap(0)-RTC (B), dCap(0)-RTC (C), E-RTC (D), and Cap(−1)′-RTC (E). The helicases are shown as colored surfaces, while nsp7, nsp8, and nsp12 are shown as white surfaces. Template and primer RNAs are shown as colored cartoons. Four structures are aligned with the guidance of nsp12 and shown in the same orientation. The red dashed lines indicate the positions of the edge of nsp8-2 to compare the locations of subunits.

(F–I) Enlarged views of nsp13-2 1B in RTCs. The polypeptides in RTCs are shown as colored cartoons with the color scheme as in Figure 1. Nsp13-2 1B is covered by cryo-EM densities (gray mesh); the unpaired 5′ extension of template RNA bound with nsp13-2 in dCap(0)-RTC (G) and E-RTC (H) are covered by cryo-EM densities as red meshes.

(J) Interaction of nsp13-2 1B with nsp13-1 ZBD and nsp8-1 in dCap(0)-RTC. The interacting residues are displayed as colored sticks.

(K and L) Interaction of nsp13-2 1B with the unpaired 5′ extension of template RNA in dCap(0)-RTC (K) and E-RTC (L). Key interacting residues and RNA molecules are shown as colored sticks.

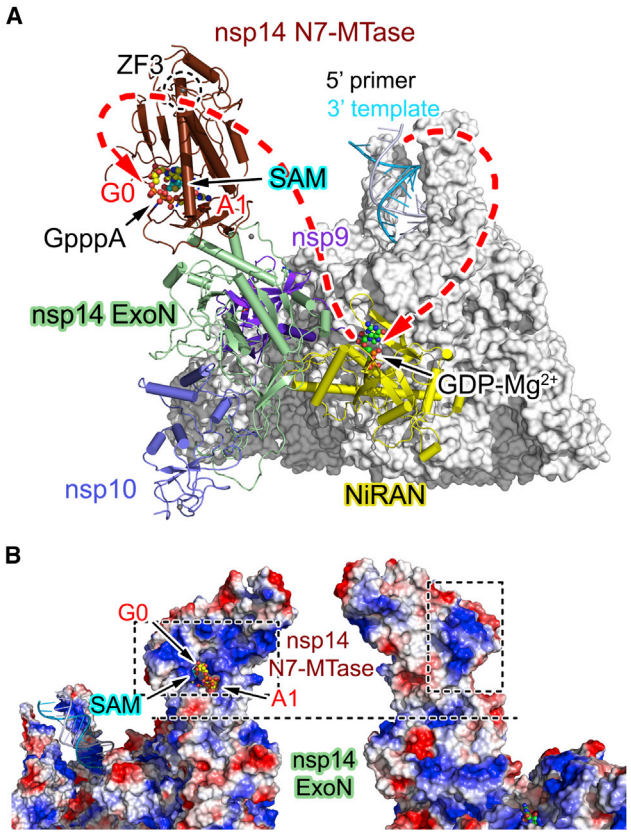


Figure 5. A potential transferring path for pre-mRNA in capping

(A) An overall view of N7-CCC. N7-CCC is shown as colored cartoon and the EC is covered by white molecular surface. Nsp9, nsp10, nsp12 NiRAN, nsp14 ExoN, and nsp14 N7-MTase are colored purple, blue, yellow, pale green, and brown, respectively. A GpppA and a SAM bound to nsp14 N7-MTase in the crystal structure (Ma et al., 2015) (PDB code: 5C8S) are shown as spheres with colors as following: C atoms for SAM, cyan; C atoms for GpppA, yellow; O atoms, red; N atoms, blue; P atoms, gold. GDP-Mg^{2+} bound in the catalytic center of nsp12 NiRAN in Cap(−1)′-RTC (Yan et al., 2021) (PDB code: 7CYQ) is shown as spheres with C atoms in green color to indicate the catalytic center of nsp12 NiRAN. A zinc finger (ZF3) in nsp14 N7-MTase is highlighted by a black dashed circle.

(B) The electrostatic potential surface of nsp14 in Cap(0)-RTC in views from the side of and the opposite side where the catalytic center of nsp12 NiRAN is located. Positively charged regions that potentially constitute pre-mRNA transferring path are highlighted by dashed frames.

its inhibition by nucleotide analog inhibitors (e.g., remdesivir) have been well characterized (Chen et al., 2020; Gao et al., 2020; Wang et al., 2020; Yan et al., 2020, 2021). However, how CCC couples in SARS-CoV-2 RTC remains unclear. Here, we have reconstituted SARS-CoV-2 Cap(0)-RTC and determined its structure, providing the first images for the coupling of N7-CCC in RTC and for dissecting the mechanism of the capping actions. In N7-CCC, the catalytic site of nsp12 NiRAN, which catalyzes the second capping action to yield GpppA, locates on the opposite side of the catalytic center of nsp14 N7-MTase, which facilitates the third capping action to generate the cap(0) (^{7Me}GpppA) (Figure 5A). A possible path for transferring the 5′ end of pre-mRNA with GpppA is reasoned that initiating from nsp12 NiRAN, passing through a positively charged region surrounded by $_{nsp14}$K433/K440/K512, and reaching the catalytic site of nsp14 N7-MTase (Figure 5B). Notably, a zinc finger (ZF3) has been reported at $_{nsp14}$C484/C452/C477/H487 in the crystal structure (Ma et al., 2015). This zinc finger is located on the potential path of transferring mRNA and may play a role in stabilizing mRNA for further modification. Moreover, A1 of GpppA is buried in a deep pocket in nsp14 N7-MTase, but G0 of GpppA is relatively exposed to solvent in the crystal structure (Ma et al., 2015) (Figure 5B, left panel). This conformation is not likely to allow the connection of A1 to the downstream sequence of mRNA, indicating the binding of the 5′ end of pre-mRNA with GpppA to the catalytic center of nsp14 N7-MTase may vary from that observed in the crystal structure. Moreover, the potential path for transferring mRNA on N7-CCC is occupied in dCap(0)-RTC by another Cap(0)-RTC protomer, indicating the capping actions may not occur in dCap(0)-RTC.

RNA viruses commonly exhibit high mutation rates, being attributed to the relatively low replication fidelity of their RNA-dependent RNA polymerase (RdRp) and the lack of proofreading mechanism (Ogando et al., 2020). Therefore, the expansion of RNA genome sizes is restricted to below 15 kb for most RNA virus families to balance replication fidelity, genome size, and complexity (Saberi et al., 2018). CoVs have the largest RNA genomes among RNA viruses and employ a proofreading mechanism to ensure replication fidelity. In proofreading, mis-incorporated nucleotides are excised through the 3′-5′ ExoN activity of nsp14 (Ferron et al., 2018; Munster et al., 2020; Ogando et al., 2019), but the concerted motion between nsp14 ExoN and polymerase remains unclear. In bacteria and eukaryotes, backtracking, which refers to the reversible sliding of RNA polymerase along DNA/RNA, has been implicated in many critical processes, including providing a basic mechanism for transcriptional proofreading (Nudler, 2012; Nudler et al., 1997; Sosunov et al., 2003). Any mis-match nucleotides in the RNA:DNA hybrid would induce immediate backtracking, thus resulting in the subsequent removal of the mis-incorporated nucleotides at 3′ end of RNA by exonuclease (Nudler, 2012; Wang et al., 2009). In SARS-CoV-2 mCap(0)-RTC, the catalytic center of nsp14 ExoN is distal from the polymerase reaction center of nsp12 and the exit of product RNA with an approximate distance of 80 and 120 Å (Figure 6A), indicating proofreading occurring in the monomeric Cap(0)-RTC is less likely. Interestingly, in dCap(0)-RTC, the catalytic center of nsp14 ExoN of one Cap(0)-RTC protomer faces to the polymerase reaction center of nsp12 at the 3′ end of primer chain at a distance of ~35 Å (Figure 6A). Considering the ~6 Å average length between two nucleotides in a single-stranded RNA, a backtracking of RNA with 6 nucleotides from the nsp12 polymerase reaction center in one Cap(0)-RTC protomer is plausible to reach the catalytic center of nsp14 ExoN in another Cap(0)-RTC protomer (Figures 6B and 6C). Based on these structural features, we propose a speculative model for an *in trans* backtracking mechanism for SARS-CoV-2 proofreading concerted by nsp12 polymerase, nsp13-2 helicase, and nsp14 ExoN (Figure 6D). In the elongation step, nsp13-2 functions as a helicase to unwind the high-ordered structure in the genome with the help of nsp13-1. At this stage, nsp13-2 1B domain is in position to constitute a RNA-binding groove together with

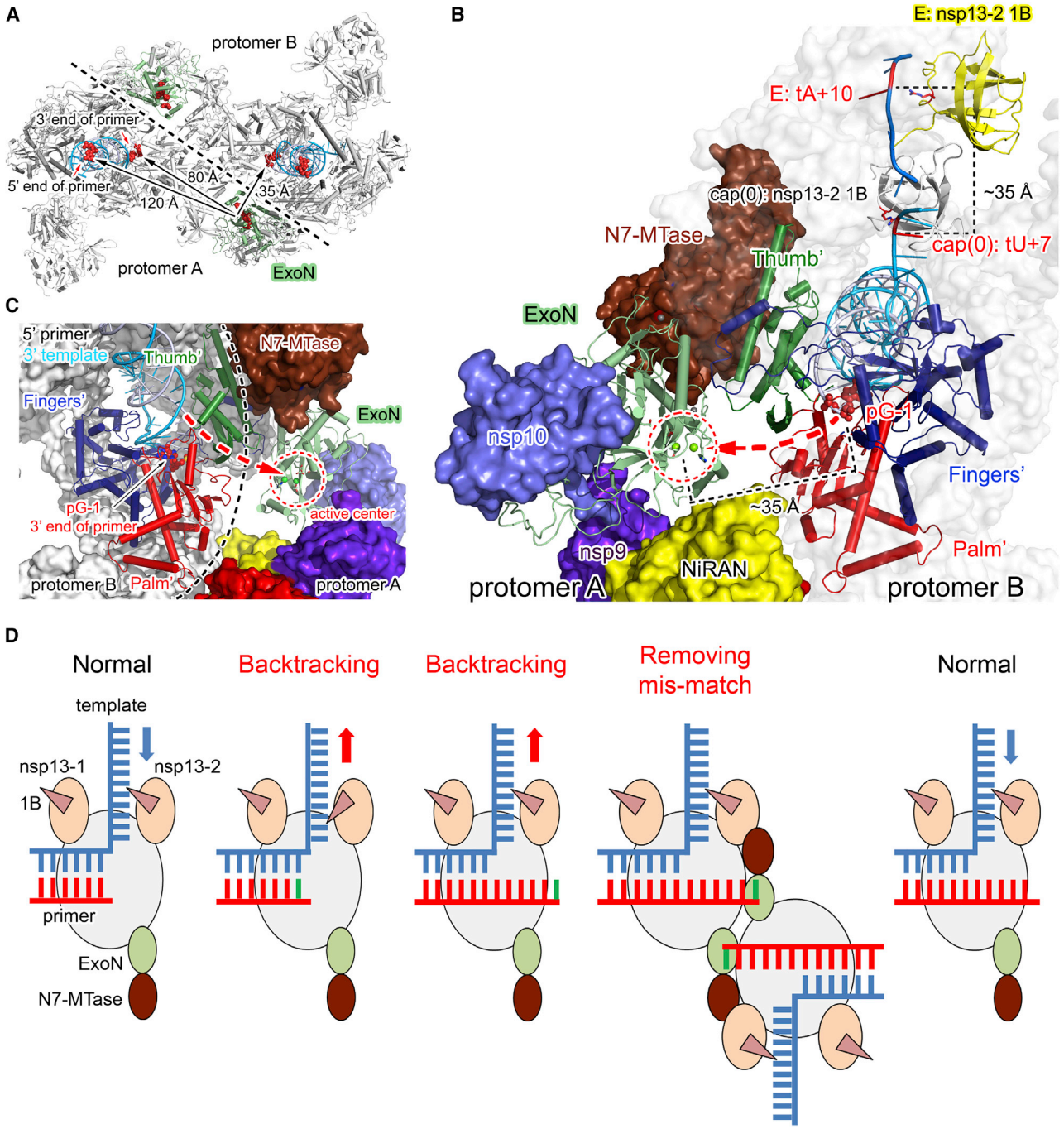


Figure 6. An *in trans* backtracking model for proofreading

(A) Distance between the catalytic center of nsp14 ExoN in one Cap(0)-RTC protomer with the 5′ end of primer RNA in the same Cap(0)-RTC protomer and the catalytic center of nsp12 polymerase in two Cap(0)-RTC protomers of the dimeric Cap(0)-RTC. The polypeptides of nsp14 ExoN are shown as a cartoon diagram in pale green, whereas other polypeptides in the dimeric Cap(0)-RTC are displayed as white cartoons. The catalytic residues in the nsp14 ExoN active center are shown as red spheres. The nucleotides at the 3′ end of the modeled primer RNA is displayed as red spheres to indicate the catalytic center of nsp12 polymerase. The nucleotide at the 5′ end of the modeled primer RNA is also displayed as red spheres. The distances are indicated by arrows with labels.

(B and C) A close-up view of the machinery for *in trans* backtracking proofreading. In protomer A, the fingers, palm, and thumb domains of nsp12 are shown as cartoons in blue, red, and green. The template and primer RNAs are represented in cartoon diagram and the nucleotide at the 3′ end of primer is shown as colored spheres. Other components in protomer A are displayed as a semi-transparent white molecular surface. In protomer B, nsp14 ExoN is shown as cartoon in pale green, while other components in protomer B are shown as a molecular surface in the same colors used in Figure 1. The catalytic residues of nsp14 ExoN in

(legend continued on next page)

1A/2A domains for RNA unwinding and the residue $_{nsp13\text{-}2}$R178 of nsp13-2 1B locates in the position of tA+10 at the 5′ end of the unpaired template RNA as presented in E-RTC. Once RTC senses the mis-incorporated nucleotide, nsp13-2 1B would undergo a large conformational changes in dCap(0)-RTC, allowing $_{nsp13\text{-}2}$R178 to lose contact with the nucleotide at the tA+10 position in E-RTC and move to the position of tU+7 as shown in dCap(0)-RTC to interact with the nucleotide located there. Subsequently, nsp13-2 1B moves back to its original position, thus dragging the template RNA toward the upstream direction. Meanwhile, the 3′ end of primer RNA with the mis-incorporated nucleotide is backtracked by 6 nucleotides toward the upstream of the 3′ end direction. It should be noted since the backtracking does not really occur in dCap(0)-RTC due to RNA and experimental conditions used to reconstitute RTCs, the binding of $_{nsp13\text{-}2}$R178 of nsp13-2 1B with tA+10 in E-RTC but with tU+7 in dCap(0)-RTC does not mean that backtracking of 4 nucleotides occurs in reality. Alternatively, the ~35 Å distance of $_{nsp13\text{-}2}$R178 of nsp13-2 1B moving from its position in Cap(0)-RTC to its position in E-RTC suggests a backtracking of 6 nucleotides would be possible. As a consequence, the 3′ end of the backtracked primer RNA with the mis-incorporated nucleotide would recruit nsp14 ExoN in another Cap(0)-RTC protomer, for further processing of the mis-incorporated nucleotide at the catalytic center of nsp14 ExoN. After the correction of the mis-incorporated nucleotide, nsp13-2 1B would return to its original position, allowing normal helicase activity to move template RNA forward and elongate primer RNA. At this stage, dCap(0)-RTC is likely to be released into the monomeric form.

A recent study showed SARS-CoV-2 possessing nsp14 with the L177F, P203L, S369F, or M501I mutations has a higher genomic mutation rate than SARS-CoV-2 with wild-type nsp14 (Takada et al., 2020). Among these, the M501I mutation occurs in nsp14 N7-MTase, which is not associated with the enzymatic activity of nsp14 ExoN, but it is located on the dimer interface and thus plays a key role in stabilizing the conformation of an helix in nsp14 N7-MTase and in the formation of dCap(0)-RTC. The impact of M501 on SARS-CoV-2 mutation rate can be explained by our proposed model. Moreover, another recent study reported a SARS-CoV-2 RdRp structure with nucleotide analogs and finds the mis-incorporated nucleotide analogs can pause the RdRp function, allowing nsp13 to induce backtracking, which forces the 3′ end of the primer RNA out the NTP-entry channel (Malone et al., 2021). The direction of the primer RNA 3′ end out the NTP-entry channel is consistent to that observed in the dCap(0)-RTC structure, this being toward the catalytic center of nsp14 ExoN.

The backtracking proofreading is widely found among DNA polymerases (Fernandez-Leiro et al., 2017; Johnson and Beese, 2004) and RNA polymerases (Cheung and Cramer, 2011; Wang et al., 2009) in eukaryotic and prokaryotic cells. Similarities could be found with the backtracking proofreading in SARS-CoV-2. For example, the distance from the catalytic center of nsp14 ExoN in one protomer to the catalytic center of polymerase of nsp12 in another protomer is 35 Å in SARS-CoV-2 dCap(0)-RTC, and the typical distance between the active centers of other polymerases and ExoN is 30~40 Å in the correction of DNA mismatch (Fernandez-Leiro et al., 2017). Unique features could be also identified in SARS-CoV-2 backtracking proofreading, including the conformational change of helicase 1B domain, as well as the specific excision of 3′ mismatched nucleotide by nsp14 ExoN but not removal of a mismatched fragment as observed in Pol II (Nudler, 2012).

Notably, in dCap(0)-RTC, the potential transfer path for the 5′ end of pre-mRNA is through the inter-protomer interface, indicating that backtracking proofreading would not concurrently occur with capping actions. This is supported by the orientation of nsp13-2 ZBD in a "suspending" state, in which nsp13-2 ZBD inserts into the paired template:primer RNA, in both mCap(0)-RTC and in Cap(−1)′-RTC as an intermediate state during capping actions (Yan et al., 2021). In contrast, nsp13-2 ZBD leaves the paired template-primer RNA in free state both in dCap(0)-RTC and in E-RTC, thus allowing the active movement of RNA in either a "backtracking" state or a "elongation" state (Figures 4B–4E). This gives an explanation for the simultaneous existences of monomeric and dimeric Cap(0)-RTC in solution, representing nsp14 as a bifunctional enzyme participating in proofreading and in capping actions.

Furthermore, the proofreading is suggested to be a factor impacting on the inhibitory efficacy of nucleotide analog inhibitors(Robson et al., 2020). For example, EIDD-1931, a cytidine analog with broad spectrum antiviral activity against coronaviruses, is the first mutagenic nucleotide analog inhibitor demonstrating the ability to evade nsp14 ExoN proofreading (Robson et al., 2020). This is the reason that EIDD-1931 cannot be recognized and excised by nsp14 ExoN, giving an explanation to its enhanced efficacy against SARS-CoV-2 containing resistance mutations to remdesivir (Sheahan et al., 2020). The structure of dCap(0)-RTC containing a primer chain with nucleotide analog inhibitor at the 3′ end is needed to provide a structural basis for the optimization of inhibitor to achieve higher efficacy against drug resistant variants.

In conclusion, the structure of SARS-CoV2 Cap(0)-RTC reported here presents a molecular basis for co-transcriptional modification of SARS-CoV-2 mRNA and reveals an *in trans* backtracking mechanism for proofreading. These results provide new insights on the understanding of the replication and transcription cycle of SARS-CoV-2 and a structural basis for the optimization of antiviral inhibitors such as EIDD-1931.

Limitations of study

Although the *in trans* backtracking model for proofreading is proposed, a structure of dCap(0)-RTC with a backtracked RNA extended from the 3′ end of primer chain at the active center of polymerase to the active center of nsp14 ExoN is needed to warrant in further study.

protomer B are shown as colored stick models and are highlighted by a red dashed circle. The boundary of the two protomers is indicated by a black dashed line. A channel for the backtracked primer RNA is indicated by a red dashed line with an arrow. A view from the opposite side is shown in (C).

(D) A schematic representation of an *in trans* backtracking model for proofreading. Triangles in brown represent nsp13 1B; blue and red bars represent the correct nucleotides in template and primer RNAs, while the green bars indicate the mis-matched nucleotides.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Protein production and purification
  - Assembly of the Cap(0)-RTC
  - Native electrophoretic mobility shift assays
  - Cryo-EM sample preparation and data collection
  - Cryo-EM image processing
  - Model building and refinement
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2021.05.033.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (U20A20135), the National Program on Key Research Project of China (2020YFA0707500 and 2017YFC0840300), Tsinghua University Spring Breeze Fund (2020Z99CFG015), and the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB37030201).

AUTHOR CONTRIBUTIONS

Z.R. and Z. Lou conceived the project and designed the experiments. L.Y., Y.Y., M.L., J.G., Y.C.H., L.Z., Y.Z., S.G., Y.Y.H., and R.Z. performed sample purification, and L.Y. prepared for RTC assembly. L.Y., Y.Y., and Y.Z. for sample native electrophoretic mobility shift assays. Y.G., Z. Liu, and T.W. collected cryo-EM data. 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.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 1, 2021
Revised: April 20, 2021
Accepted: May 19, 2021
Published: May 24, 2021

REFERENCES

Afonine, P.V., Klaholz, 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 Struct. Biol. *74*, 814–840.

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

Bouvet, M., Imbert, I., Subissi, L., Gluais, L., Canard, B., and Decroly, E. (2012). RNA 3′-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. Proc. Natl. Acad. Sci. USA *109*, 9372–9377.

Case, J.B., Ashbrook, A.W., Dermody, T.S., and Denison, M.R. (2016). Mutagenesis of S-Adenosyl-l-Methionine-Binding Residues in Coronavirus nsp14 N7-Methyltransferase Demonstrates Differing Requirements for Genome Translation and Resistance to Innate Immunity. J. Virol. *90*, 7248–7256.

Chen, Y., Cai, H., Pan, J., Xiang, N., Tien, P., Ahola, T., and Guo, D. (2009). Functional screen reveals SARS coronavirus nonstructural protein nsp14 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 nsp16/nsp10 protein complex. PLoS Pathog. *7*, e1002294.

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 *135*, 24–35.

Chen, J., Malone, B., Llewellyn, E., Grasso, M., Shelton, P.M.M., Olinares, P.D.B., Maruthi, K., Eng, E.T., Vatandaslar, H., Chait, B.T., et al. (2020). Structural Basis for Helicase-Polymerase Coupling in the SARS-CoV-2 Replication-Transcription Complex. Cell *182*, 1560–1573.

Cheung, A.C., and Cramer, P. (2011). Structural basis of RNA polymerase II backtracking, arrest and reactivation. Nature *471*, 249–253.

Conti, B.J., Kirchdoerfer, R.N., and Sussman, M.R. (2020). Mass spectrometric based detection of protein nucleotidylation in the RNA polymerase of SARS-CoV-2. bioRxiv, 2020.2010.2007.330324.

Daffis, S., Szretter, K.J., Schriewer, J., Li, J., Youn, S., Errett, J., Lin, T.Y., Schneller, S., Zust, 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.

Decroly, E., Imbert, I., Coutard, B., Bouvet, M., Selisko, 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.

Deming, D.J., Graham, R.L., Denison, M.R., and Baric, R.S. (2007). Processing of open reading frame 1a replicase proteins nsp7 to nsp10 in murine hepatitis virus strain A59 replication. J. Virol. *81*, 10280–10291.

Eckerle, L.D., Lu, X., Sperry, S.M., Choi, L., and Denison, M.R. (2007). High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. J. Virol. *81*, 12135–12144.

Eckerle, L.D., Becker, M.M., Halpin, R.A., Li, K., Venter, E., Lu, X., Scherbakova, S., Graham, R.L., Baric, R.S., Stockwell, T.B., et al. (2010). Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. PLoS Pathog. *6*, e1000896.

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

Fernandez-Leiro, R., Conrad, J., Yang, J.C., Freund, S.M., Scheres, S.H., and Lamers, M.H. (2017). Self-correcting mismatches during high-fidelity DNA replication. Nat. Struct. Mol. Biol. *24*, 140–143.

Ferron, F., Subissi, L., Silveira De Morais, A.T., Le, N.T.T., Sevajol, M., Gluais, L., Decroly, E., Vonrhein, C., Bricogne, G., Canard, B., and Imbert, I. (2018). Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. Proc. Natl. Acad. Sci. USA *115*, E162–E171.

Gao, Y., Yan, L., Huang, Y., Liu, F., Zhao, Y., Cao, L., Wang, T., Sun, Q., Ming, Z., Zhang, L., et al. (2020). Structure of the RNA-dependent RNA polymerase from COVID-19 virus. Science *368*, 779–782.

Gribble, J., Stevens, L.J., Agostini, M.L., Anderson-Daniels, J., Chappell, J.D., Lu, X., Pruijssers, A.J., Routh, A.L., and Denison, M.R. (2021). The coronavirus proofreading exoribonuclease mediates extensive viral recombination. PLoS Pathog. *17*, e1009226.

Grimm, C., Hillen, H.S., Bedenk, K., Bartuli, J., Neyer, S., Zhang, Q., Hüttenhofer, A., Erlacher, M., Dienemann, C., Schlosser, A., et al. (2019).

Structural Basis of Poxvirus Transcription: Vaccinia RNA Polymerase Complexes. Cell *179*, 1537–1550.

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.

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.

Johnson, S.J., and Beese, L.S. (2004). Structures of mismatch replication errors observed in a DNA polymerase. Cell *116*, 803–816.

Ma, Y., Wu, L., Shaw, N., Gao, Y., Wang, J., Sun, Y., Lou, Z., Yan, L., Zhang, R., and Rao, Z. (2015). Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. Proc. Natl. Acad. Sci. USA *112*, 9436–9441.

Malone, B., Chen, J., Wang, Q., Llewellyn, E., Choi, Y.J., Olinares, P.D.B., Cao, X., Hernandez, C., Eng, E.T., Chait, B.T., et al. (2021). Structural basis for backtracking by the SARS-CoV-2 replication-transcription complex. Proc. Natl. Acad. Sci. USA *118*, e2102516118.

Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. Journal of Structural Biology *152*, 36–51. https://doi.org/10.1016/j.jsb.2005.07.007.

Minskaia, E., Hertzig, T., Gorbalenya, A.E., Campanacci, V., Cambillau, C., Canard, B., and Ziebuhr, J. (2006). Discovery of an RNA virus 3′->5′ exoribonuclease that is critically involved in coronavirus RNA synthesis. Proc. Natl. Acad. Sci. USA *103*, 5108–5113.

Munster, V.J., Feldmann, F., Williamson, B.N., van Doremalen, N., Pérez-Pérez, L., Schulz, J., Meade-White, K., Okumura, A., Callison, J., Brumbaugh, B., et al. (2020). Respiratory disease in rhesus macaques inoculated with SARS-CoV-2. Nature *585*, 268–272.

Nudler, E. (2012). RNA polymerase backtracking in gene regulation and genome instability. Cell *149*, 1438–1445.

Nudler, E., Mustaev, A., Lukhtanov, E., and Goldfarb, A. (1997). The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. Cell *89*, 33–41.

Ogando, N.S., Ferron, F., Decroly, E., Canard, B., Posthuma, C.C., and Snijder, E.J. (2019). The Curious Case of the Nidovirus Exoribonuclease: Its Role in RNA Synthesis and Replication Fidelity. Front. Microbiol. *10*, 1813.

Ogando, N.S., Zevenhoven-Dobbe, J.C., van der Meer, Y., Bredenbeek, P.J., Posthuma, C.C., and Snijder, E.J. (2020). The Enzymatic Activity of the nsp14 Exoribonuclease Is Critical for Replication of MERS-CoV and SARS-CoV-2. J. Virol. *94*, e01246, 20.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

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.

Robson, F., Khan, K.S., Le, T.K., Paris, C., Demirbag, S., Barfuss, P., Rocchi, P., and Ng, W.L. (2020). Coronavirus RNA Proofreading: Molecular Basis and Therapeutic Targeting. Mol. Cell *80*, 1136–1138.

Rohou, A., and Grigorieff, N. (2015). CTFFIND4: 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.

Saberi, A., Gulyaeva, A.A., Brubacher, J.L., Newmark, P.A., and Gorbalenya, A.E. (2018). A planarian nidovirus expands the limits of RNA genome size. PLoS Pathog. *14*, e1007314.

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

Shabek, N., Ticchiarelli, F., Mao, H., Hinds, T.R., Leyser, O., and Zheng, N. (2018). Structural plasticity of D3-D14 ubiquitin ligase in strigolactone signalling. Nature *563*, 652–656.

Sheahan, T.P., Sims, A.C., Zhou, S., Graham, R.L., Pruijssers, A.J., Agostini, M.L., Leist, S.R., Schäfer, A., Dinnon, K.H., 3rd, Stevens, L.J., et al. (2020). An orally bioavailable broad-spectrum antiviral inhibits SARS-CoV-2 in human airway epithelial cell cultures and multiple coronaviruses in mice. Sci. Transl. Med. *12*, eabb5883.

Slanina, H., Madhugiri, R., Bylapudi, G., Schultheiß, K., Karl, N., Gulyaeva, A., Gorbalenya, A.E., Linne, U., and Ziebuhr, J. (2021). Coronavirus replication-transcription complex: Vital and selective NMPylation of a conserved site in nsp9 by the NiRAN-RdRp subunit. Proc. Natl. Acad. Sci. USA *118*, e2022310118.

Soon, F.F., Ng, L.M., Zhou, X.E., West, G.M., Kovach, A., Tan, M.H., Suino-Powell, K.M., He, Y., Xu, Y., Chalmers, M.J., et al. (2012). Molecular mimicry regulates ABA signaling by SnRK2 kinases and PP2C phosphatases. Science *335*, 85–88.

Sosunov, V., Sosunova, E., Mustaev, A., Bass, I., Nikiforov, V., and Goldfarb, A. (2003). Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. EMBO J. *22*, 2234–2244.

Subissi, L., Posthuma, C.C., Collet, A., Zevenhoven-Dobbe, J.C., Gorbalenya, A.E., Decroly, E., Snijder, E.J., Canard, B., and Imbert, I. (2014). One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. Proc. Natl. Acad. Sci. USA *111*, E3900–E3909.

Takada, K., Ueda, M.T., Watanabe, T., and Nakagawa, S. (2020). Genomic diversity of SARS-CoV-2 can be accelerated by a mutation in the nsp14 gene. bioRxiv, 2020.2012.2023.424231.

Wang, D., Bushnell, D.A., Huang, X., Westover, K.D., Levitt, M., and Kornberg, R.D. (2009). Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. Science *324*, 1203–1206.

Wang, W., Li, J., Sun, Q., Yu, X., Zhang, W., Jia, N., An, C., Li, Y., Dong, Y., Han, F., et al. (2017). Structural insights into the coordination of plastid division by the ARC6-PDV2 complex. Nat. Plants *3*, 17011.

Wang, Q., Wu, J., Wang, H., Gao, Y., Liu, Q., Mu, A., Ji, W., Yan, L., Zhu, Y., Zhu, C., et al. (2020). Structural Basis for RNA Replication by the SARS-CoV-2 Polymerase. Cell *182*, 417–428.

Wu, A., Peng, Y., Huang, B., Ding, X., Wang, X., Niu, P., Meng, J., Zhu, Z., Zhang, Z., Wang, J., et al. (2020). Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. Cell Host Microbe *27*, 325–328.

Yan, L., Ge, J., Zheng, L., Zhang, Y., Gao, Y., Wang, T., Huang, Y., Yang, Y., Gao, S., Li, M., et al. (2021). Cryo-EM Structure of an Extended SARS-CoV-2 Replication and Transcription Complex Reveals an Intermediate State in Cap Synthesis. Cell *184*, 184–193.

Yan, L., Zhang, Y., Ge, J., Zheng, L., Gao, Y., Wang, T., Jia, Z., Wang, H., Huang, Y., Li, M., et al. (2020). Architecture of a SARS-CoV-2 mini replication and transcription complex. Nat. Commun. *11*, 5874.

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331–332.

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
NaF	SIGMA	Cat# 201154
Be_2SO_4	ALADDIN	Cat# B106711
Deposited data		
mCap(0)-RTC EM map	This paper	EMD: 31146
mCap(0)-RTC coordinate	This paper	PDB: 7EIZ
dCap(0)-RTC EM map	This paper	EMD: 31138
dCap(0)-RTC coordinate	This paper	PDB: 7EGQ
Oligonucleotides		
Template RNA: 5′-CAUGCCAUGGCCUC UAAAAUGUCAGCUGCUCCCUAGCAU GCUACUACCGCGUAGCAUG-3′	Takara	N/A
Primer RNA: 5′-GCGGUAGUAGCAUGC UAGGGAGCAG-3′	Takara	N/A
Recombinant DNA		
pET22b-nsp12	This paper	N/A
pET22b-nsp7	This paper	N/A
pET28b-SUMO-nsp8	This paper	N/A
pET28a-nsp13	This paper	N/A
pET28a-nsp9	This paper	N/A
pGEX6p-nsp10	This paper	N/A
pGEX6p-nsp9-nsp10	This paper	N/A
pRSF-duet-nsp14	This paper	N/A
Software and algorithms		
SerialEM	Mastronarde, 2005	https://bio3d.colorado.edu/SerialEM
MotionCor2	Zheng et al., 2017	https://emcore.ucsf.edu/ucsf-software
RELION 3.0	Scheres, 2012	https://www3.mrc-lmb.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://phenix-online.org/
PyMOL	Schrodinger, LLC	Schrodinger
Other		
Superdex-200 10/300 Increase	GE Healthcare	Cat# 28990944
Hitrap-Q HP	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# UFC901096
100 kDa cutoff concentrators	Millipore	Cat# UFC910096
R0.6/1.0 200 mesh Cu holey carbon grids	Quantifoil	Cat# Q250CR-06

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 31138 and 7EGQ for dCap(0)-RTC, and 31146 and 7EIZ for mCap(0)-RTC.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

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

METHOD DETAILS

Protein production and purification

The SARS-CoV-2 nsp12 (GenBank: MN908947) gene was cloned into a modified pET-22b vector, with the C terminus possessing a 10 × His-tag. Protein was expressed in *E. coli* BL21 (DE3) as described(Yan et al., 2021). The cells were harvested and the pellets were resuspended in a buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 4 mM $MgCl_2$, 10% glycerol) and homogenized with an ultra-high-pressure cell disrupter 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 $MgCl_2$, 10% glycerol, 4 mM DTT) and buffer B (20 mM Tris-HCl, pH 8.0, 1M NaCl, 4 mM $MgCl_2$, 10% glycerol, 4 mM DTT). Next the sample was loaded onto a Superdex 200 10/300 Increase column (GE Healthcare, USA) with DEPC-treated buffer C (50 mM HEPES, pH 7.0, 100 mM NaCl, 4 mM $MgCl_2$, 2 mM GDP and 2 mM BeF_3^-). Purified nsp12 was concentrated to 4.8 mg/mL and stored at 4°C.

The SARS-CoV-2 nsp10 was cloned into pGEX-6p vector with a N-terminal GST -tag, and SARS-CoV-2 nsp14 inserted into pRSF-duet with no tag, and co-transformed into *E. coli* strain BL21 (DE3). Cells were harvested by centrifugation at 4000 rpm for 10 min, and the pellets resuspended in lysis buffer (25 mM HEPES, pH7.0, 300 mM NaCl, 4 mM $MgCl_2$, 5% glycerol). An ultra-high-pressure cell disrupter at 4°C was used for lysis,and the product was centrifuged for 14000 rpm at 4°C. Recombinant protein was purified by GST-affinity chromatography and the GST-tag was removed by PreScission protease. The complex was further purified by passage through a Hitrap SP ion-exchange column (GE Healthcare, USA) with buffer A (25 mM HEPES, pH 7.0, 4 mM $MgCl_2$, 10% glycerol, 4 mM DTT) and buffer B (25 mM HEPES, pH7.0, 1 M NaCl, 4 mM $MgCl_2$, 10% glycerol, 4 mM DTT). Then it was load onto a Superdex 200 10/300 Increase column (GE Healthcare, USA) with a buffer (50 mM HEPES, pH 7.0, 250 mM NaCl, 4 mM $MgCl_2$, 4 mM DTT). Purified nsp10/nsp14 complex was concentrated to 8 mg/mL and stored at 4°C. The nsp9-10 sample was also cloned into pGEX-6p vector, and the purification of nsp9-10/nsp14 complex was identical to the procedure for nsp10/nsp14.The preparation of nsp9 was performed as described previously(Yan et al., 2021). Nsp9 was cloned into a modified pET-28b-SUMO with the N terminus of a fusion of 6 × His-tag. The protein was expressed in *E. coli* strain BL21 (DE3). After harvesting by centrifugation, the pellets were resuspended in lysis buffer (20 mM HEPES, pH 7.0 and 150 mM NaCl) and homogenized with an ultra-high-pressure cell disrupter at 4°C. The lysate was centrifuged at 12,000 rpm for 30 min to remove cell debris. The fusion protein was purified by Ni-NTA (Novagen, USA) affinity chromatography and by application to a Superdex 200 10/300 Increase column (GE Healthcare, USA) in lysis buffer. Purified nsp9 was concentrated to 5 mg/mL and stored at 4°C.

Full-length SARS-CoV-2 nsp7 and nsp8 were co-expressed in *E. coli* BL21 (DE3) cells as a no-tagged protein and a 6 × His-SUMO fusion protein, respectively. After purification by Ni-NTA (Novagen, USA) affinity chromatography, the nsp7-nsp8 complex was eluted through on-column tag cleavage by ULP protease. The complex was further purified by using a Hitrap Q ion-exchange column (GE Healthcare, USA) and a Superdex 200 10/300 Increase column (GE Healthcare, USA) in buffer C containing 50 mM HEPES, pH 7.0, 100 mM NaCl, 4 mM $MgCl_2$, 2 mM GDP and 2 mM BeF_3^-.

Nsp13 was purified as described previously(Yan et al., 2021). The nsp13 gene was inserted into the modified pET-28a vector with a 6 × His tag attached at its N terminus, and protein was expressed in *E. coli* BL21 (DE3) cells. Cells were harvested and resuspended in lysis buffer (20 mM HEPES, pH 7.0, 150 mM NaCl, 4 mM $MgCl_2$, 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 purified by Ni-NTA (Novagen, USA) affinity chromatography and Hitrap SP ion-exchange column (GE Healthcare, USA), and finally nsp13 protein was loaded onto a Superdex 200 10/300 Increase column (GE Healthcare, USA) in buffer C. Purified nsp13 was concentrated to 8 mg/mL and stored at 4°C.

Assembly of the Cap(0)-RTC

Nsp12 was incubated with nsp7 and nsp8 at 4°C for three hours in a molar ratio of 1: 2: 2 in buffer (50 mM HEPES, pH 7.0, 100 mM NaCl and 4 mM $MgCl_2$). Next, the mixture was purified by mono Q 5/50 ion-exchange chromatography (GE Healthcare, USA), producing the nsp7-nsp8-nsp12 complex (C-RTC). C-RTC and nsp13 and RNA were mixed to form E-RTC at a 1:2:1 molar ratio as described previously(Yan et al., 2020). E-RTC was incubated with the nsp9-nsp10/nsp14 complex at a 1:1.2 molar ratio with 2 mM GDP●BeF_3^- to assemble the Cap(0)-RTC.

Native electrophoretic mobility shift assays

The binding reaction buffer contained 50 mM HEPES, pH 7.0, 100 mM NaCl, 2 mM $MgCl_2$, 2 mM GDP and 2 mM BeF_3^-. 18 μg RdRp (nsp12-nsp7-nsp8) complex protein was combined with 1.5 μg template-primer RNA, and RdRp/RNA and nsp13, or nsp9, or nsp10/14 mixed in a 1:2, or 1:1.2, or 1:1.2 molar ratio. Binding reactions were incubated for 30 min at 30°C. Reactions were run on a six lane polyacrylamide native gel (37.5:1 acrylamide:bis-acrylamide) running in 1 × TBE buffer at 150 V for 1h in 4°C. The gel was stained with ethidium bromide.

Cryo-EM sample preparation and data collection

In total, 3 μL of protein sample 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 22500 × 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 $Å^2$. All images were automatically recorded using SerialEM. A total of 12,704 images were collected with a defocus range from −2.0 μm to −1.0 μm. Statistics for data collection and refinement are in Table S1. The methods for processing are described in Figures S2 and S4.

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 ctffind4(Rohou and Grigorieff, 2015). A total of 2,039,214 particles were auto-picked using the model from SARS-CoV-2 Cap(−1)'-RTC (PDB: 7CYQ)(Yan et al., 2021) and extracted with a box size of 448 pixels in cryoSPARC(Punjani et al., 2017). The following 2D, 3D classifications and refinements were all performed in cryoSPARC. 887,588 particles were selected after two rounds of 2D classification based on complex integrity. This particle set was used to do Ab-Initio reconstruction in five classes, which were then used as 3D volume templates for heterogeneous refinement, with 135,801 particles converged into dCap(0)-RTC complex class and 80,256 particles converged into mCap(0)-RTC complex class. Next, these particles were imported into RELION 3.03(Scheres, 2012) to perform local classification to obtain one class particle with final resolution 3.35 Å and 3.78 Å, respectively. The methods are described in Figures S2 and S4.

Model building and refinement

To build the structure of SARS-CoV-2 Cap(0)-RTC complex, we started with the model of the SARS-CoV-2 nsp12 and nsp7-8 complex (PDB: 7BTF), SARS-CoV-2 nsp13 (PDB: 6ZSL), nsp9 (PDB: 6W9Q) and nsp10/nsp14 (PDB: 6C8S). These were individually placed and rigid-body fitted into the cryo-EM map using UCSF Chimera(Pettersen et al., 2004). The model was manually built in Coot(Emsley et al., 2010) with the guidance of the cryo-EM map, and with real space refinement using Phenix(Afonine et al., 2018). The data validation statistics are shown in Table S1.

QUANTIFICATION AND STATISTICAL ANALYSIS

In Figures S2 and S4, 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).

Supplemental figures

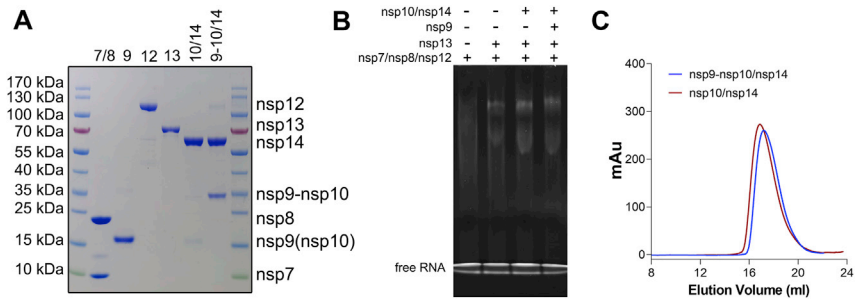


Figure S1. Biochemical analysis, related to Figure 1

(A) 10% SDS-PAGE analysis of the components used to constitute the RTCs. Lanes 1 and 8, markers; lane 2, nsp7/nsp8; lane 3, nsp9; lane 4, nsp12; lane 5, nsp13; lane 6, nsp10/nsp14 complex; lane 7, nsp9-10/nsp14 complex. The gel was stained with Coomassie blue. Because nsp14 cannot be purified in the absence of nsp10 or nsp9-nsp10, there is no sample for the individual nsp14. The molecular weights of protein markers are indicated on the left of the SDS-PAGE. (**B**) Native gel electrophoretic mobility shift assay reveals the formation of the stable complex. The 6% polyacrylamide gel was visualized with ethidium bromide to stain the RNA. (**C**) Purification of nsp10/nsp14 and nsp9-10/nsp14 complexes by size exclusion chromatography.

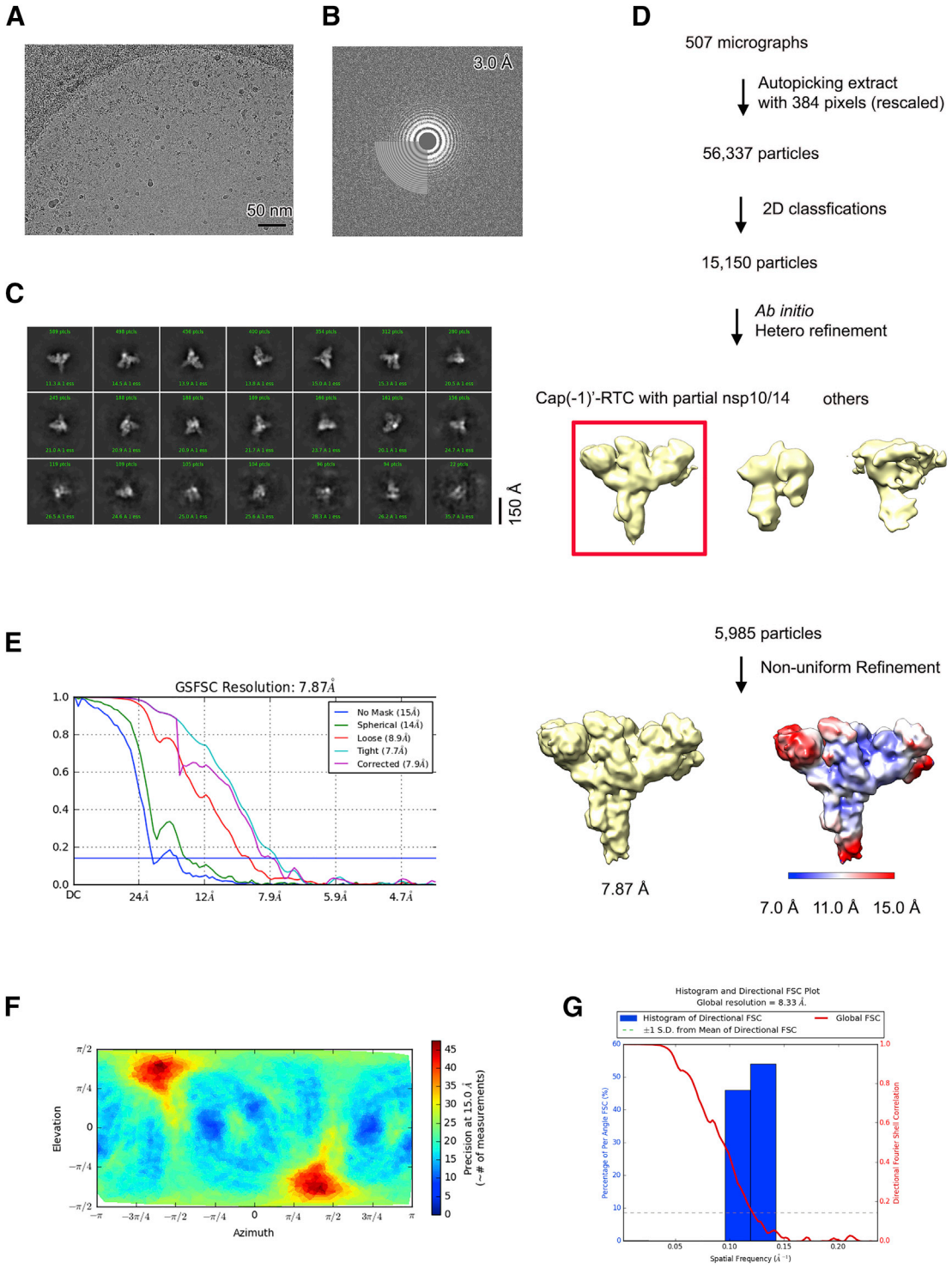


Figure S2. Cryo-EM reconstruction of SARS-CoV-2 Cap(−1)′-RTC incubated with nsp10/nsp14, related to Figure 1

(A) Raw image of the SARS-CoV-2 Cap(-)'-RTC with nsp10/14 complex particles in vitreous ice recorded at defocus values of −1.0 to −1.8 μm. Scale bar, 50 nm. (**B**) Power spectrum of the image shown in (A), with an indication of the spatial frequency corresponding to 3.0 Å resolution. (**C**) Representative class averages. The edge of each square is ~367 Å in length. (**D**) Flowchart of SARS-CoV-2 Cap(-)'-RTC with nsp10/14 reconstruction. Local resolution estimation is shown at the bottom panel. (**E**) Fourier shell correlation (FSC) of the final 3D reconstruction following gold standard refinement. FSC curves are plotted before and after masking. (**F**) Angular distribution heatmap of particles used for the refinement. (**G**) The 3DFSC sphericity analyzed with 3DFSC in cryoSPARC(Punjani et al., 2017)

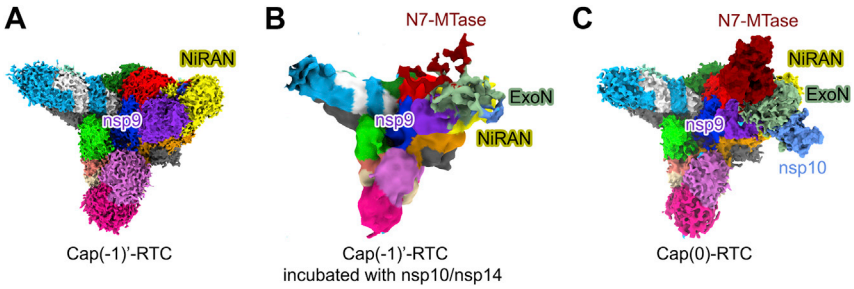
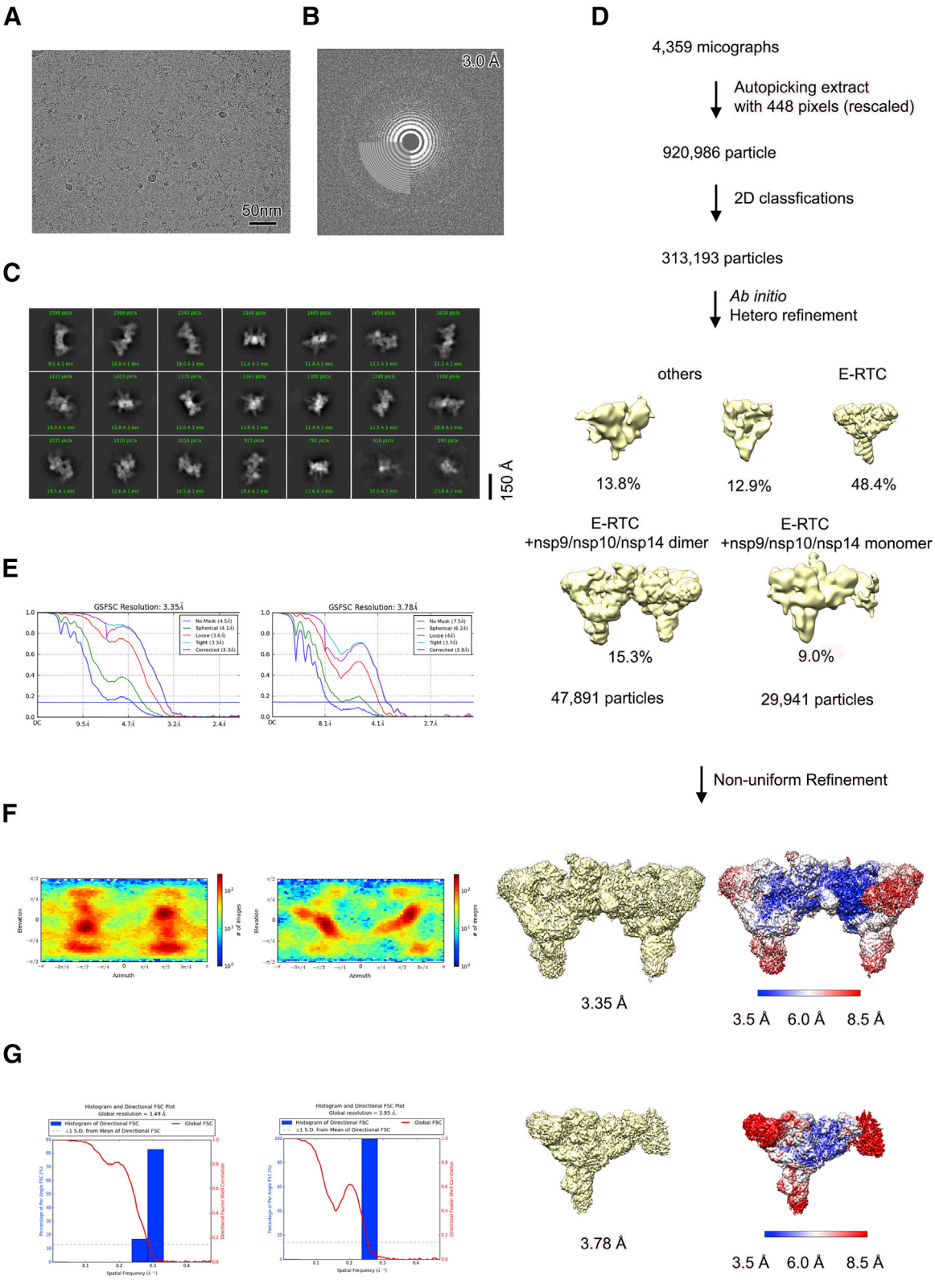


Figure S3. Comparison of cryo-EM densities, related to Figure 1

The cryo-EM densities of Cap(−1)'-RTC (**A**), initial test dataset for Cap(−1)'-RTC incubated with nsp10/nsp14 (**B**), and mCap(0)-RTC with final resolution (**C**) are shown in the same orientation and threshold. The color scheme is the same as that used in Figure 1.



(legend on next page)

Figure S4. Cryo-EM reconstruction of SARS-CoV-2 Cap(0)-RTC, related to Figure 1

(A) Raw image of the dCap(0)-RTC and mCap(0)-RTC particles in vitreous ice recorded at defocus values of −1.0 to −1.8 μm. Scale bar, 50 nm. (**B**) Power spectrum of the image shown in (A), with an indication of the spatial frequency corresponding to 3.0 Å resolution. (**C**) Representative class averages. The edge of each square is ~367 Å. (**D**) Flowchart of SARS-CoV-2 dCap(0)-RTC and mCap(0)-RTC reconstruction. Local resolution estimation is shown at the bottom panel.(**E**) Fourier shell correlation (FSC) of the final 3D reconstruction following gold standard refinement. FSC curves are plotted before and after masking. (**F**) Angular distribution heatmap of particles used for the refinement. (**G**) The 3DFSC sphericity analyzed with 3DFSC in cryoSPARC(Punjani et al., 2017).

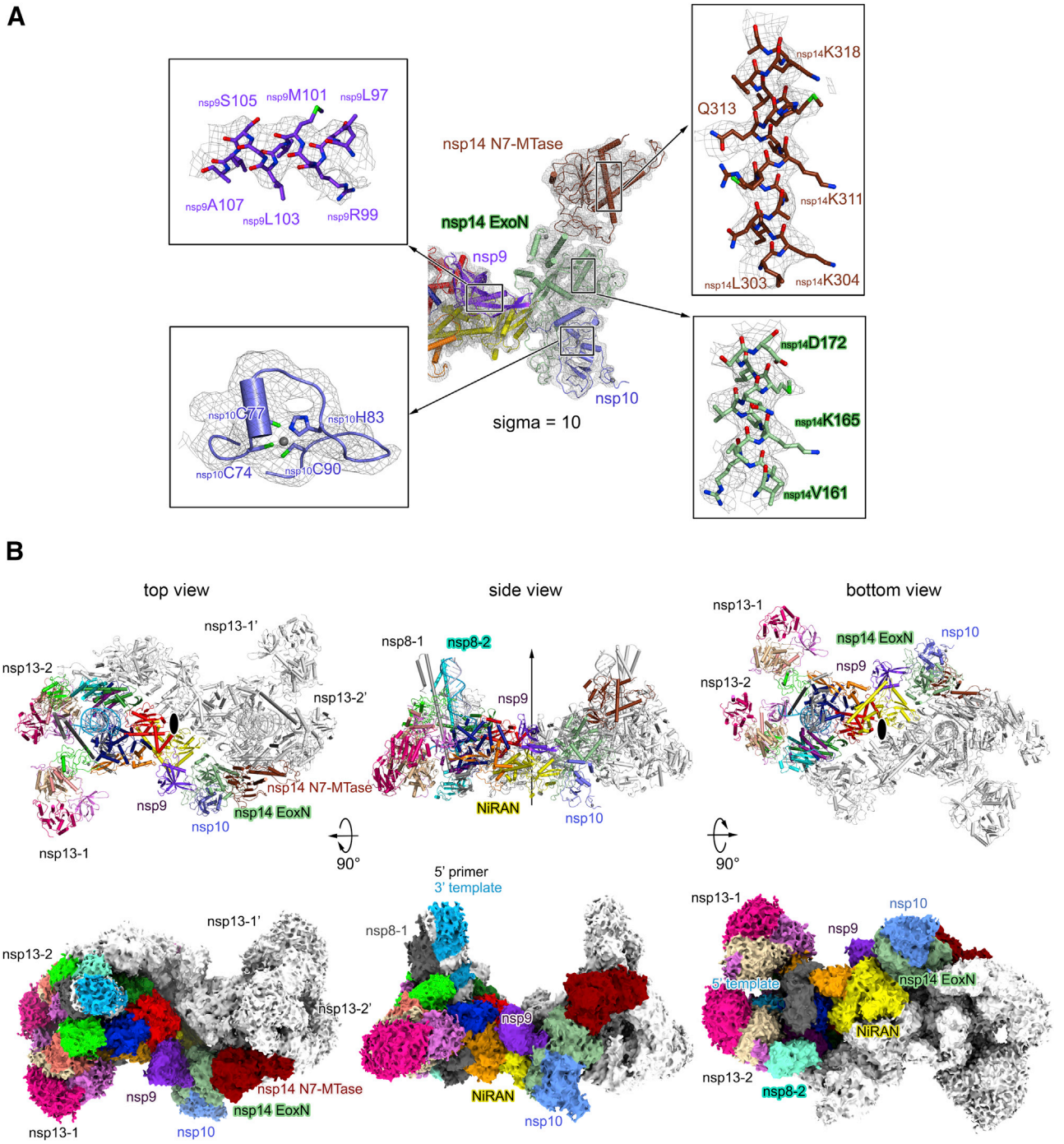
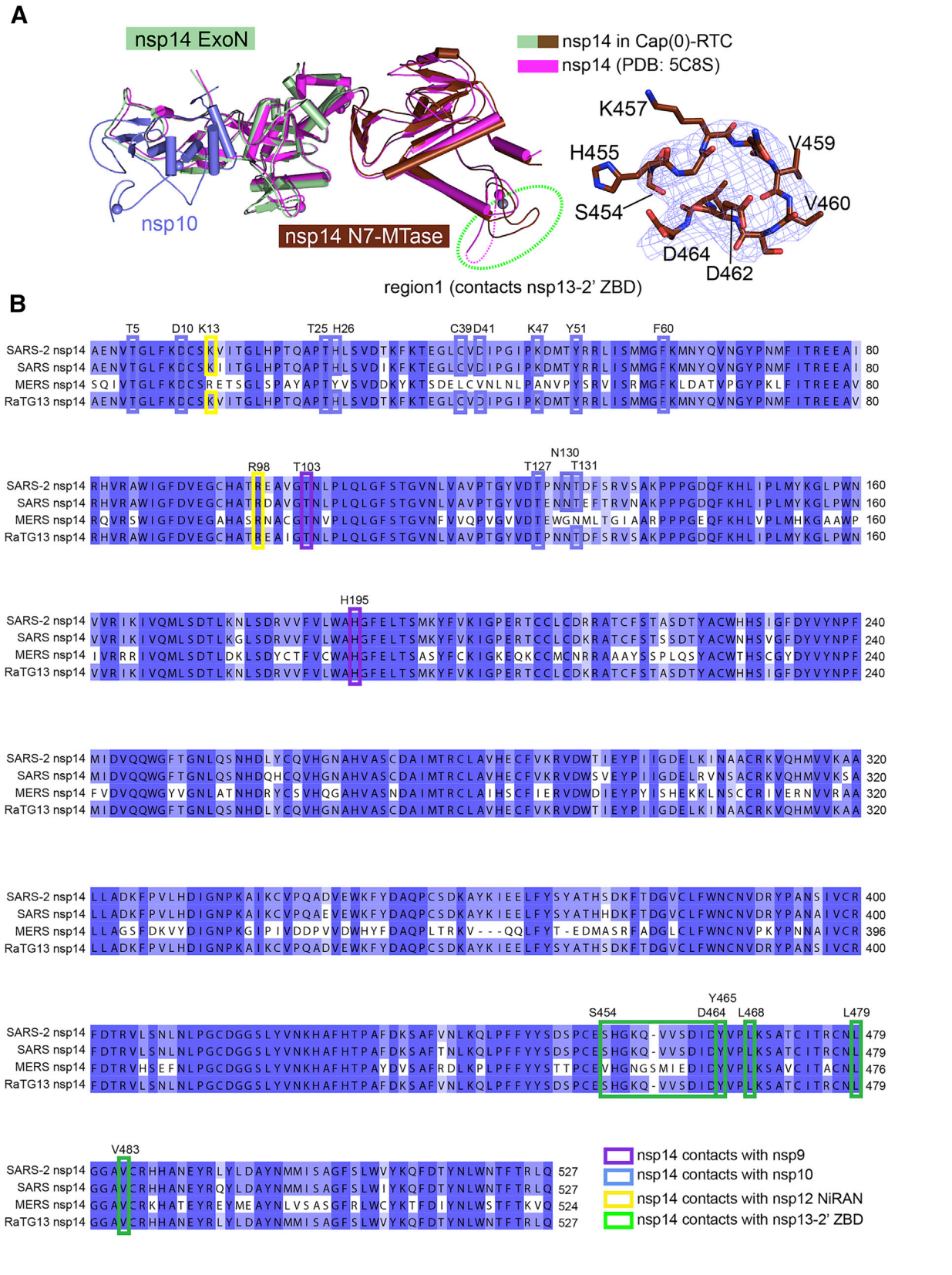


Figure S5. Density and structure of Cap(0)-RTC, related to Figure 1

(A) Density of Cap(0)-RTC (Related to Figure 1). Structures of Nsp12 NiRAN, nsp9, nsp10 and nsp14 are overlaid with the cryo-EM map of dCap(0)-RTC protomer in the middle panel. Four parts are shown in enlarged panels. The polypeptides of Cap(0)-RTC are displayed as colored cartoons; residues in the enlarged panels are exhibited as colored sticks; the cryo-EM densities are shown as gray mesh with the sigma value of 10. The color scheme is the same as used in Figure 1. (**B**) Structure of dCap(0)-RTC (Related to Figure 1C). Structure of dCap(0)-RTC in cartoon diagrams (upper panels) and the cryo-EM densities (bottom panels) are shown in three perpendicular views. The two-fold axis is indicated as ellipses (left and right panel) or an arrow (middle panel). The dashed lines roughly indicate the boundary of two Cap(0)-RTC protomers in dCap(0)-RTC (C). Compared to Figure 1C, one Cap(0)-RTC protomer is shown in colored scheme, while the other protomer is colored in white.



(legend on next page)

Figure S6. Structure of nsp10/nsp14 and sequence comparison, related to Figure 2

(A) The structures of nsp10/nsp14 complex in Cap(0)-RTC and in crystallo(Ma et al., 2015) (PDB code: 5C8S) are aligned. Nsp14 as it appears in the crystal structure is colored in magenta; nsp14 in Cap(0)-RTC and nsp10 are colored as the same scheme in Figure 1. The cryo-EM density (at a sigma value of 8.5) of the loop region spanning residues $_{nsp14}$S454-$_{nsp14}$D464 is shown in the right panel. (**B**) The alignment of nsp14 encoding by SARS-2, SARS, MERS, RaTG13. The residues in blue, light blue, or white shading indicate the identical, conserved or non-conserved residues, respectively. The interacting residues of nsp14 are highlighted by colored frames.