ANNUAL REVIEWS

Annual Review of Biochemistry

The Life of SARS-CoV-2 Inside Cells: Replication–Transcription Complex Assembly and Function

Zhiyong Lou[1] and Zihe Rao[1,2,3,4,5,6]

[1]Ministry of Education Key Laboratory of Protein Science, School of Medicine, Tsinghua University, Beijing, China; email: louzy@mail.tsinghua.edu.cn, raozh@tsinghua.edu.cn

[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]State Key Laboratory of Medicinal Chemical Biology, College of Life Sciences and College of Pharmacy, Nankai University, Tianjin, China

[5]National Laboratory of Biomacromolecules, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China

[6]Guangzhou Laboratory, Guangzhou, China

Annu. Rev. Biochem. 2022. 91:381–401

The *Annual Review of Biochemistry* is online at biochem.annualreviews.org

https://doi.org/10.1146/annurev-biochem-052521-115653

Keywords

SARS-CoV-2, virus life cycle, replication, transcription, nonstructural protein, replication–transcription complex, assembly, structure

Abstract

The persistence of the coronavirus disease 2019 (COVID-19) pandemic has resulted in increasingly disruptive impacts, and it has become the most devastating challenge to global health in a century. The rapid emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants challenges the currently available therapeutics for clinical application. Nonstructural proteins (also known as replicase proteins) with versatile biological functions play central roles in viral replication and transcription inside the host cells, and they are the most conserved target proteins among the SARS-CoV-2 variants. Specifically, they constitute the replication–transcription complexes (RTCs) dominating the synthesis of viral RNA. Knowledge of the

molecular mechanisms of nonstructural proteins and their assembly into RTCs will benefit the development of antivirals targeting them against existing or potentially emerging variants. In this review, we summarize current knowledge of the structures and functions of coronavirus nonstructural proteins as well as the assembly and functions of RTCs in the life cycle of the virus.

Contents

INTRODUCTION

As of February 2022, the coronavirus disease 2019 (COVID-19) pandemic has caused over 396 million infections and over 5.7 million deaths worldwide (1). It has become one of the most disastrous infectious diseases in human history.

The causative agent of COVID-19, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is a positive-sense, single-stranded RNA (+ssRNA) virus in the *Betacoronavirus* genus of the *Orthocoronavirinae* subfamily, *Coronaviridae* family, Nidovirales order (2). Coronaviruses have a long history of circulating in the human population. HCoV-229E and HCoV-OC43 were first identified in 1966–1967 and shown to cause mild respiratory tract infections and symptoms similar to the common cold (3, 4). Since 2003, severe acute respiratory syndrome virus (SARS-CoV), Middle East respiratory syndrome virus (MERS-CoV), and SARS-CoV-2,

all with severe pathologies, have emerged, highlighting the need to understand and defeat these viruses.

CoV has a remarkably large +ssRNA genome (approximately 30 kb); thus, it employs a set of unique mechanisms to ensure efficient synthesis of viral RNA (5). After the virus enters the host cell, the successive replication and transcription of CoV involve negative-sense RNA synthesis to create a full-length complementary RNA (cRNA) or subgenomic RNA (sgRNA) template and the subsequent copying of cRNA or sgRNA into positive-sense progeny genomic RNA (gRNA) or subgenomic mRNA (6). Both CoV gRNA and mRNA are capped at their 5′ end with cap(1) ($^{7Me}GpppA_{2'OMe}$) structure and are polyadenylated at their 3′ end, but polyuridylic acid (instead of the cap structure) is decorated at the 5′ end of cRNA (7–9). The 5′ cap(1) structure on CoV gRNA/mRNA, in which an N7-methyl-guanine moiety is linked to the first 5′ nucleotide of the nascent pre-gRNA/mRNA and the ribose 2′-*O* position of the first transcribed nucleotide is further methylated (8, 10, 11), is important for promoting initiation of translation, protecting mRNAs, and helping the virus escape host immune recognition (12). The cotranscriptional capping of gRNA/mRNA is exerted by four processive actions (13). First, an RNA 5′-triphosphatase (RTPase) removes the γ-phosphate from the 5′-triphosphate end (pppA) of the nascent pre-gRNA/mRNA to generate the 5′-diphosphate end (ppA) (14, 15). Second, a guanylyltransferase (GTase) transfers a guanosine monophosphate to the remaining 5′-diphosphate end to yield the cap core (GpppA) (13). Third, an N7-methyltransferase (N7-MTase) methylates the first guanine at the N7 position to produce the cap(0) ($^{7Me}GpppA$) (16). In the final step, a 2′-*O*-methyltransferase (2′-*O*-MTase) converts cap(0) into cap(1) ($^{7Me}GpppA_{2'OMe}$) by methylating the ribose 2′-*O* position of the first transcribed nucleotide to complete the capping process (10, 11, 17).

The genome of SARS-CoV-2 encodes 4 structural proteins, 16 nonstructural proteins (nsp1–nsp16), and several accessory proteins (18). Nonstructural proteins, also known as replicases, play essential roles in every aspect of the life cycle of the virus inside the host cells (**Figure 1**). After entry into the host cells, nonstructural proteins are translated and subsequently assemble the replication–transcription complex (RTC). This complex is responsible for a variety of processes, including all key steps in replicating and transcribing the viral genome, forming a molecular pore on double-membrane vesicles (DMVs), and regulating protein translation in host ribosomes.

Since the outbreak of severe acute respiratory syndrome (SARS) in 2003, our understanding of the structures and functions of CoV-encoded nonstructural proteins has grown rapidly. In particular, benefiting from great advances in single-particle cryogenic electron microscopy (cryo-EM) technology, investigations of the assembly of the CoV RTC have made significant progress, allowing for visualization of the dynamics of CoV living in host cells. In this review, we discuss the key structural knowledge of CoV nonstructural proteins gained from studies of SARS-CoV-2 and other CoVs and their assembly into complicated RTCs. Structural proteins and accessory proteins also play essential roles in CoV proliferation; however, due to space limitations, we focus only on nonstructural proteins and RTCs.

NONSTRUCTURAL PROTEINS

nsp1

The protein nsp1 is the N-terminal cleavage product of the replicase polyprotein encoded by the CoV genome and is a virulence factor in SARS-CoV-2 infection. It folds into an N-terminal domain with a β-barrel structure (19–21) and a C-terminal domain adopting a helix-turn-helix

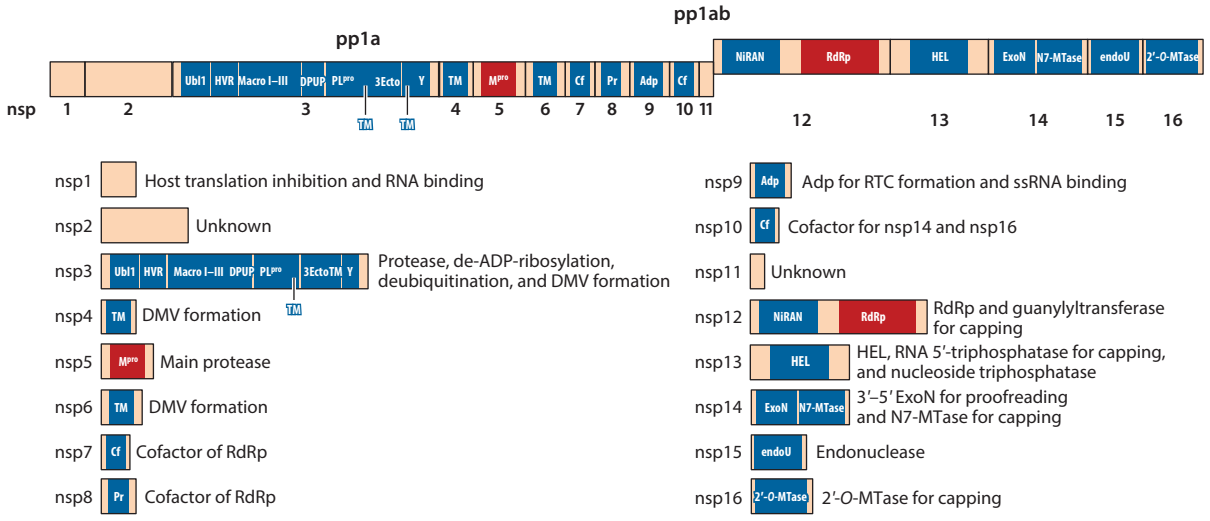


Figure 1

Nonstructural proteins encoded by SARS-CoV-2. Domains in nsp1–nsp16 with known functions are indicated. Abbreviations: 2′-*O*-MTase, 2′-*O*-methyltransferase; 3Ecto, ectodomain; Adp, adaptor; Cf, cofactor; DMV, double-membrane vesicle; DPUP, domain preceding Ubl2 and PL^{pro}; endoU, endoribonuclease; ExoN, exoribonuclease; HEL, helicase; HVR, hypervariable region; Macro, macrodomain; M^{pro}, main protease; N7-MTase, N7-methyltransferase; NiRAN, nidovirus RdRp-associated nucleotidyltransferase; nsp, nonstructural protein; PL^{pro}, papain-like protease domain; Pr, primase; RdRp, RNA-dependent RNA polymerase; RTC, replication–transcription complex; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; ssRNA, single-stranded RNA; TM, transmembrane region; Ubl, ubiquitin-like domain; Y, Y1 and CoV-Y.

(HTH) fold (22–24) (**Figure 2*a***). The C-terminal domain of SARS-CoV-2 or SARS-CoV nsp1 occludes the entrance of the mRNA channel in the host ribosome, thus inhibiting the translation of targeting antiviral signaling pathways to suppress host innate immunity (22–26). Meanwhile, its N-terminal domain binds to the 5′ untranslated region (5′-UTR) of the viral genome with high affinity (0.31 μM) (27). Also, nsp1 promotes the degradation of host mRNAs by cleavage of their 5′-UTR, which in turn leads to accelerated cellular 5′-3′ exoribonuclease 1–mediated mRNA decay (28, 29). The binding of nsp1 to the viral genome 5′-UTR with high affinity may play a role in helping the virus escape mRNA decay mediated by nsp1 (30). A recent study showed that the SARS-CoV-2 nsp1 N-terminal domain forms a stable complex with DNA polymerase α in the human primosome (31). In that work, although the full-length nsp1 was used for structural study, only the N-terminal domain could be traced in the experimental cryo-EM density, indicating the N- and C-terminal domains may simultaneously exert their different functions.

nsp2

One of the most variable nonstructural proteins among CoVs, in terms of both its primary sequence and biological function, is nsp2. It is related to viral replication, protein expression, localization, and antagonism of the regular activity of host cells (32, 33). However, the precise function of nsp2 in the CoV life cycle is not fully understood. We have determined the crystal structures of the N- and C-terminal fragments of SARS-CoV nsp2 [Protein Data Bank identifier (PDB ID): 7FA1 and 7FAC] and infectious bronchitis virus (IBV) nsp2 (PDB ID: 7F52) (Z. Lou & Z. Rao, unpublished results). The protein nsp2 adopts a multidomain architecture containing a zinc finger–rich domain (ZRD) at the N terminus composed of three zinc fingers, a helix-rich domain

in the middle resembling an armadillo repeat–like fold, and a C-terminal domain consisting mostly of loops and coils with high structural flexibility (**Figure 2*b***). Comparison of nsp2 structure with entries in the PDB generates no obvious homologues. The structure-function relationship of nsp2 needs further investigation.

nsp3

CoV nsp3 is the largest protein and one of three membrane-associated proteins encoded by the CoV genome. It generally adopts eight featured domains: the ubiquitin-like domain 1 (Ubl1), the Glu-rich acidic domain [also known as the hypervariable region (HVR)], macrodomains I–III (Macro I–III), a domain preceding Ubl2 and the papain-like protease domain containing Ubl2 (PL^{pro}) (DPUP), PL^{pro}, the ectodomain (3Ecto) [also known as the zinc-finger (ZF) domain] located between two transmembrane regions (TM1 and TM2), and two Y domains (Y1 and CoV-Y). This protein plays versatile roles in CoV proliferation, including self-release from pp1a/pp1ab by its PL^{pro}, interacting with virus or host proteins, and, in particular, participating in forming RTCs and DMVs (34–39). The structures of many fragments of CoV nsp3 have been determined separately (reviewed in 40), but the structure of the full-length nsp3 remains unresolved.

nsp4 and nsp6

Alongside nsp3, CoV encodes another two membrane-spanning proteins: nsp4 and nsp6 (41, 42). The 56-kDa nsp4 is released from polypeptide by the combined activity of the nsp3 and nsp5 proteases. The protein nsp4 is predicted to have four transmembrane domains: an N-terminal transmembrane domain, a luminal domain, three closely spaced additional transmembrane regions, and a small C-terminal domain (approximately 100 residues) that is exposed to the cytosol (41, 42). The structure of nsp4 is largely unknown, although the structure of its C-terminal domain has been determined by crystallography (43, 44). The other membrane-spanning protein, nsp6, is an integral membrane protein (42) and activates autophagy (45). Structural details about nsp6 are unavailable. Although the molecular details for nsp4 and nsp6 are largely unknown, they, together with nsp3, are able to induce DMVs as CoV replication organelles (38) and form a hexameric crown-shaped pore structure that spans DMVs as the opening to the cytosol (46). This pore structure is proposed to interact with RTCs inside DMVs and export RNA to the cytosol (46).

nsp5

The next nonstructural protein, nsp5, is a 3C-like protease ($3CL^{pro}$) bearing a typical chymotrypsin-like fold (47–51). It proteolytically releases nsp4/nsp16 from pp1a/pp1ab as individual proteins (34, 52–54), and therefore, it is known as the main proteinase (M^{pro}). The catalytic site of nsp5 can be summarized with the P4-P2′ consensus motif (small)-X-(L/I/V/F/M)-Q↓(S/A/G), where X is any amino acid and ↓ represents the cleavage site (54). Because of its crucial function and the absence of homologues in humans, nsp5 is one of the most attractive targets for antiviral development (e.g., PF-07321332) against SARS-CoV-2 and other CoVs (47, 55–62).

The three domains of nsp5 are shown in **Figure 2*c*** (47, 48, 55). Domains I and II have antiparallel β-barrel structures, clamping the substrate-binding site. Domain III contains five α-helices arranged into a largely antiparallel globular cluster. Moreover, nsp5 forms a homodimer in the crystal, in which the N terminus (also called the N finger) of one protomer helps shape the S1 pocket and the oxyanion hole of the other protomer; hence, dimerization is a prerequisite for its catalytic activity (47, 48).

nsp7 and nsp8

Together, nsp7 and nsp8 form a stable complex (**Figure 2*d*, subpanel *i***) and are the cofactors of polymerase nsp12 to form RTCs. Biochemical studies identified that nsp8 is a second polymerase unique in CoVs and that the nsp7/nsp8 complex is capable of de novo RNA synthesis and extension with low fidelity on ssRNA templates, thus leading to the proposal that nsp8 serves as an RNA primase to synthesize short oligonucleotide primers for subsequent extension in the nsp12 polymerase (63, 64).

The protein nsp7 folds into a compact helical bundle, whereas the structure of nsp8 has two portions, among which the C-terminal part has a central β-sheet flanked by helices and the N-terminal part has a remarkably long helical architecture. The structure of the SARS-CoV

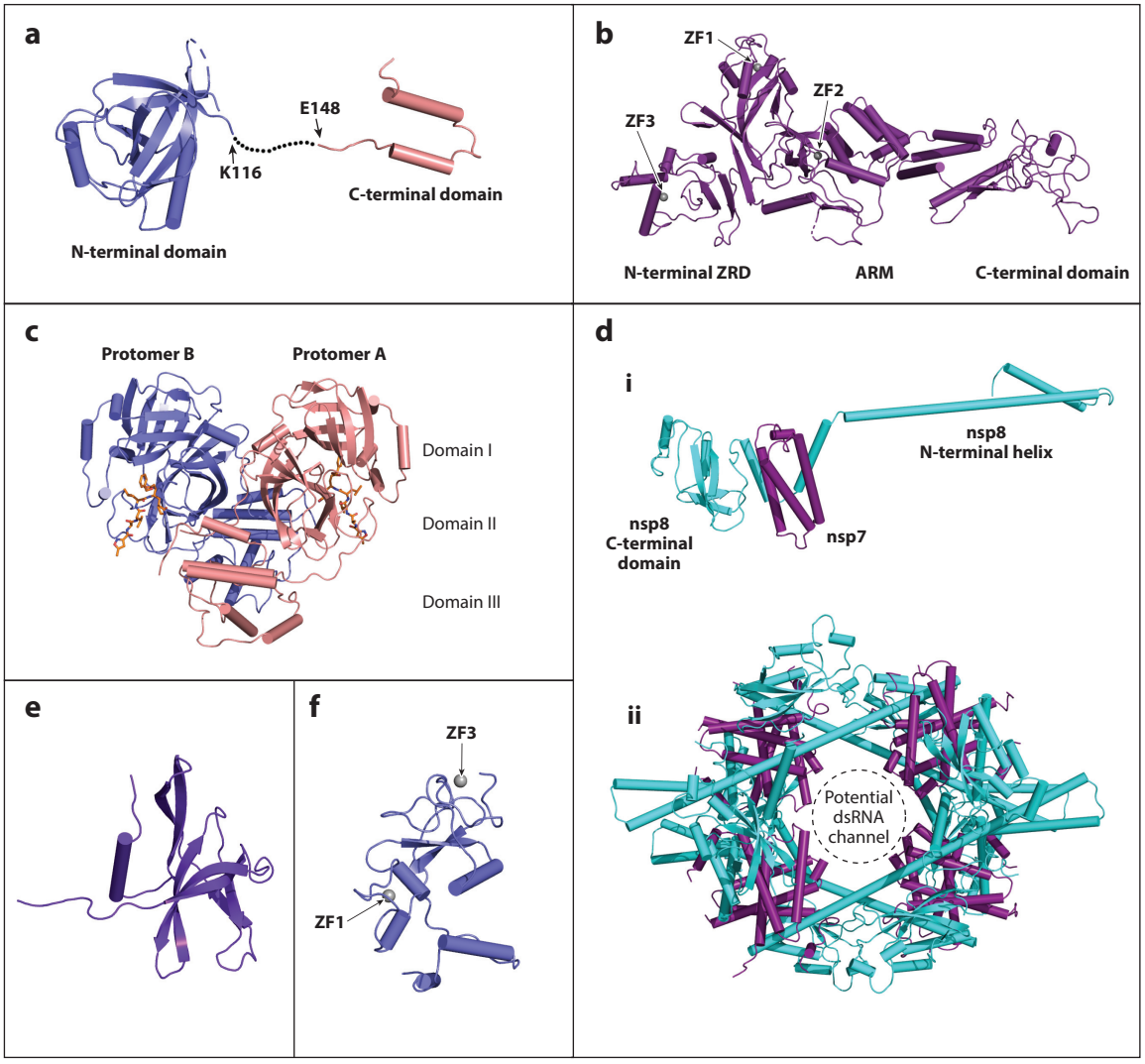


Figure 2

(*Continued*)

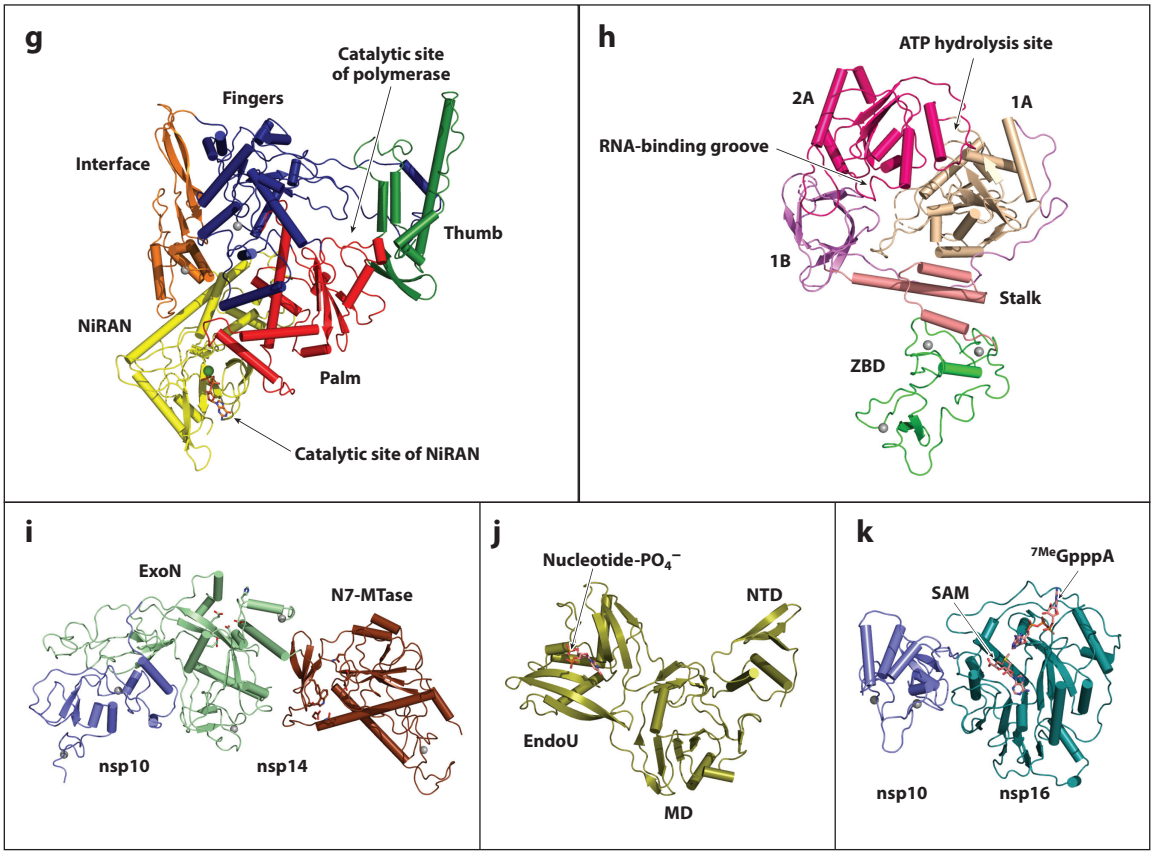


Figure 2

SARS-CoV-2 nonstructural proteins. (*a*) Structure of nsp1. The crystal structure of the N-terminal domain, which ends at residue K166 (PDB ID: 7K3N) (20), and the structure of the C-terminal helix-turn-helix fold (starting from residue E148) in the nsp1/40S ribosome complex (PDB ID: 6ZOK) (22) are placed together; the dashed line indicates the linker region between them. (*b*) Structure of nsp2 assembled by its N- and C-terminal fragments (PDB ID: 7FA1 and 7FAC). (*c*) Structure of nsp5 (PDB ID: 6LU7) (55). (*d*, *i*) Structure of the nsp7/nsp8 complex in the RTC (PDB ID: 7EGQ) (76). (*d*, *ii*) The hexadecameric crystal structure of SARS-CoV nsp7/nsp8 (PDB ID: 2AHM) (65). (*e*) Structure of nsp9 in the RTC (PDB ID: 7EGQ) (76). (*f*) Structure of nsp10 in the RTC (PDB ID: 7EGQ) (76). (*g*) Structure of nsp12 in the RTC (PDB ID: 7CYQ) (13). (*h*) Structure of nsp13 in the RTC (PDB ID: 7CXM) (89). (*i*) Structure of nsp10/nsp14 in the RTC (PDB ID: 7EGQ) (76). The residues in the catalytic centers of ExoN and N7-MTase are shown in stick mode. (*j*) Structure of the nsp15 monomer in complex with a nucleotide (PDB ID: 7K0R) (101). (*k*) Structure of nsp10/nsp16 in complex with SAM and ^{7Me}GpppA (PDB ID: 6WKS) (109). Abbreviations: ARM, armadillo repeat–like fold domain; dsRNA, double-stranded RNA; endoU, endoribonuclease; ExoN, exonuclease; MD, middle domain; N7-MTase, N7-methyltransferase; NiRAN, nidovirus RdRp-associated nucleotidyltransferase; nsp, nonstructural protein; NTD, N-terminal domain; PDB ID, Protein Data Bank identifier; RTC, replication–transcription complex; SAM, *S*-adenosylmethionine; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; ZBD, zinc-binding domain; ZF, zinc finger; ZRD, ZF-rich domain.

nsp7/nsp8 complex was first determined to be a hexadecameric ring-shaped structure with a postulated double-stranded RNA (dsRNA)-binding channel (65) (**Figure 2*d*, subpanel *ii***). The N-terminal helix of nsp8 has two distinct conformations in this hexadecameric complex (65). Similarly, the nsp8 N-terminal helix was found to have multiple conformations, either in complex with nsp7 or in RTCs (66, 67), and was shown to stabilize the template–product duplex synthesized by nsp12 polymerase in RTCs (68, 69).

nsp9

The protein nsp9 is a single-stranded nucleic acid–binding protein (70, 71) (**Figure 2*e***) and plays a role in CoV replication (72). It exhibits topological similarities with oligonucleotide- and oligosaccharide-binding (OB-fold) proteins and binds long oligonucleotides with modest affinity (70, 71). Crystallography studies find that nsp9 is presumed to form a biological dimer utilizing the interaction between its C-terminal α-helices with a GXXXG motif (70, 71, 73–75). Mutations at the nsp9 dimerization interface significantly attenuate virus proliferation (75). In recent studies, monomeric nsp9 was found to tightly bind with nsp12 nidovirus RNA-dependent RNA polymerase (RdRp)-associated nucleotidyltransferase (NiRAN), leading to attenuated NiRAN GTase activity (13), and it subsequently plays a role in recruiting nsp10/nsp14 to form Cap(0)-RTC (see the section titled Replication–Transcription Complexes) (76). Interestingly, the interfaces for nsp9 dimerization and for nsp9 interaction with other components in Cap(0)-RTC largely overlap, suggesting the homodimeric form of nsp9 may structurally transition before its assembly into RTCs (13).

nsp10

The ZF protein nsp10 has a molecular weight of approximately 15 kDa. It contains two ZFs with the sequence motifs C-(X)2-C-(X)5-H-(X)6-C and C-(X)2-C-(X)7-C-(X)-C (77) (**Figure 2*f***). It was first reported to have a spherical dodecameric architecture in a crystallography study (77) but was later shown to function as a monomer to stimulate the enzymatic activities of nsp14 and nsp16 (11, 78, 79).

nsp12

As the most critical nonstructural protein for virus replication and transcription, nsp12 is the RdRp that catalyzes the synthesis of viral RNA by using nsp7 and nsp8 as cofactors (80, 81). The structure of SARS-CoV-2 nsp12 has three domains: a NiRAN domain, an interface domain in the middle, and a C-terminal right-hand polymerase (Pol) domain (**Figure 2*g***). This Pol domain adopts the conserved architecture of the viral polymerase family (82), being composed of fingers, palm, and thumb subdomains. In the central channel of nsp12 Pol, motifs A–G constitute the catalytic center for RNA synthesis. The N-terminal NiRAN domain is a unique feature in nidovirus RdRp and was first found to have nucleotidyltransferase activity with G/U specificity (83). SARS-CoV-2 nsp12 NiRAN was recently demonstrated to function as a GTase, catalyzing the second capping action to form the cap core structure GpppA from the ppA precursor (13, 84). Moreover, one study reported that SARS-CoV-2 nsp12 NiRAN can NMPylate nsp9 in biochemical assays (85), but another study suggested nsp7 is the substrate (86). The posttranslational modification of nsp9 by nsp12 NiRAN is hypothesized to play a role in RTC formation (76).

nsp13

The protein nsp13 is a superfamily 1 (SF1) helicase with multiple enzymatic functions, including duplex unwinding, nucleoside triphosphatase, and RTPase activity (15, 87). To guide the synthesis of RNA, nsp13 unwinds the template–product duplexes and preprocesses the highly ordered structure in the ssRNA template (88, 89). Unlike most SF1 helicases and +ssRNA virus SF2 helicases with a 3′–5′ polarity, the recombinant apo CoV helicases unwind duplexes in the 5′–3′ direction (88–91). Moreover, CoV nsp13 also adopts RTPase activity to facilitate the first capping action (10, 11, 13–17, 90).

SARS-CoV-2 nsp13 has a fold typical of an SF1 helicase, containing an N-terminal zinc-binding domain (ZBD), two core Rec-like domains (1A and 2A), and an inserted domain 1B

connected to a ZBD via a stalk region (88, 89) (**Figure 2*b***). The 1A and 2A domains clamp the catalytic site for ATP hydrolysis, while the 1A, 2A, and 1B domains constitute an RNA-binding groove to adopt the ssRNA. In recent works, the 1B domain has been found to show a remarkable conformational change, leaving its original position adjacent to the 1A and 2A domains, indicating a possible role to provide a driven force for template backtracking (76, 92). The ssRNA binds in the RNA-binding groove in the 3′–5′ direction (88, 89), which is opposite to the 5′–3′ polarity identified in biochemical assays (88–91). This is proposed to be compatible with a backtracking unwinding (88, 89), but the mechanism for unwinding the nascent template–product RNA duplex needs further investigation.

nsp14

A bifunctional enzyme, nsp14 bears an N-terminal exonuclease (ExoN) domain implicated in proofreading to maintain replication fidelity (93–95) and a C-terminal N7-MTase domain catalyzing the third capping action to yield the cap(0) (16) (**Figure 2*i***). The C-terminal N7-MTase adopts an atypical *M*Tase fold with an *S*-adenosylmethionine (SAM) binding motif (79, 96, 97), which accommodates SAM in the space adjacent to the GpppA-binding site (79). The N-terminal ExoN possesses 3′–5′ ExoN activity proposed to correct errors in replication and transcription by removing misincorporated nucleotides from the 3′ end of nascent product RNA (96, 98, 99), featuring an in *trans* backtracking proofreading mechanism (76). The full function of nsp14 requires its association with nsp10 (79); however, nsp10 exclusively binds with and enhances the activity of nsp14 ExoN (10, 100) and does not impact the nsp14 N7-MTase (10, 79).

nsp15

Coronavirus nsp15 is a uridine-specific endoribonuclease (endoU) conserved across coronaviruses. The folding of nsp15 includes an N-terminal domain and a middle domain with unclear function, as well as a C-terminal endoU domain (101–103) (**Figure 2*j***). Although the role of its oligomerization is not clear, nsp15 is active only in a hexameric form (102). This protein preferentially cleaves RNA substrates' 3′ end with uridine specificity (104, 105). Lines of evidence have shown that nsp15 is important in regulating the accumulation of viral RNA and preventing activation of RNA-activated antiviral responses (106–108); however, the specific role and working mechanism of nsp15 in the virus life cycle are not fully understood.

nsp16

The last nonstructural protein encoded by the CoV genome, nsp16 functions as a 2′-*O*-MTase to catalyze the last capping action (109, 110). Loop regions originating from multiple strands in the central β-sheet form pockets to accommodate cap(0) as the substrate and SAM as the methyl donor to yield cap(1) (109, 110) (**Figure 2*k***). Compared to SARS-CoV nsp16, SARS-CoV-2 nsp16 was reported to have a unique adenosine-binding pocket in a position approximately 25 Å from the back of the catalytic pocket and may play a role in the binding of pre-mRNA during capping (10, 11, 78, 109). The 2′-*O*-MTase activity of CoV nsp16 can be enhanced when nsp16 forms a complex with cofactor nsp10 (10, 11, 77, 78). However, the binding of nsp10 does not induce any remarkable conformational change in either nsp10 or nsp16 (10, 11, 78).

REPLICATION–TRANSCRIPTION COMPLEXES

The individual nonstructural proteins organize to form a set of protein machineries, or RTCs, for successive replication and transcription inside the host cells. The first attempt to understand the highly ordered organization of CoV nonstructural proteins was achieved by a crystallography

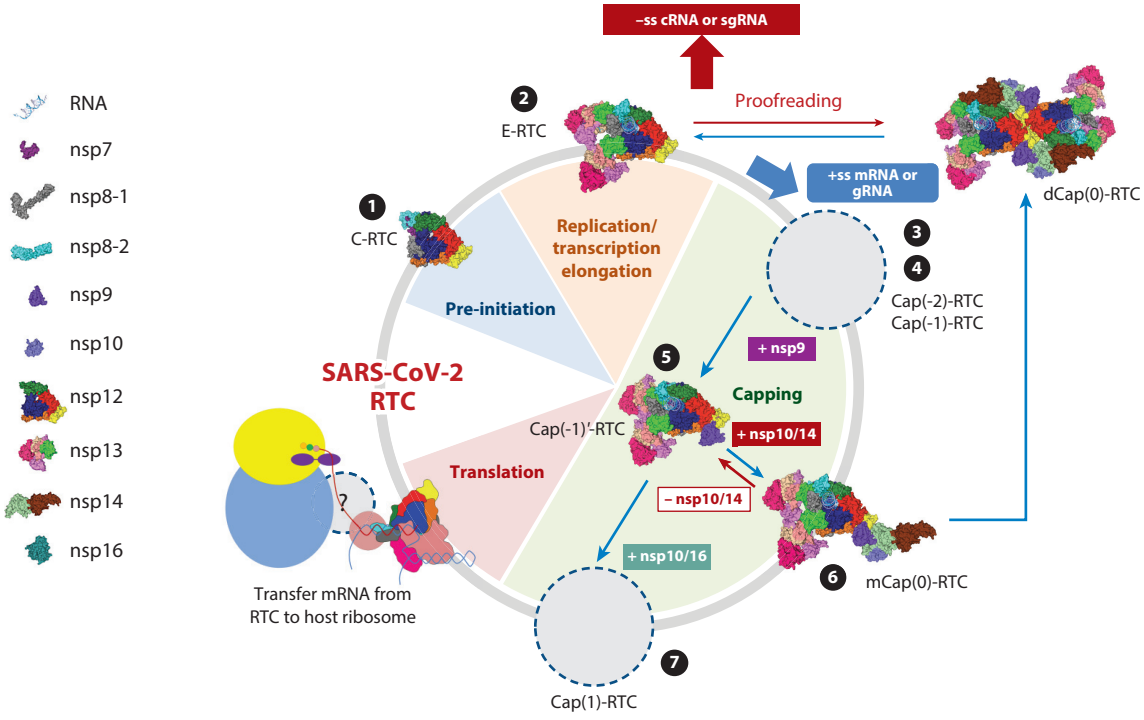


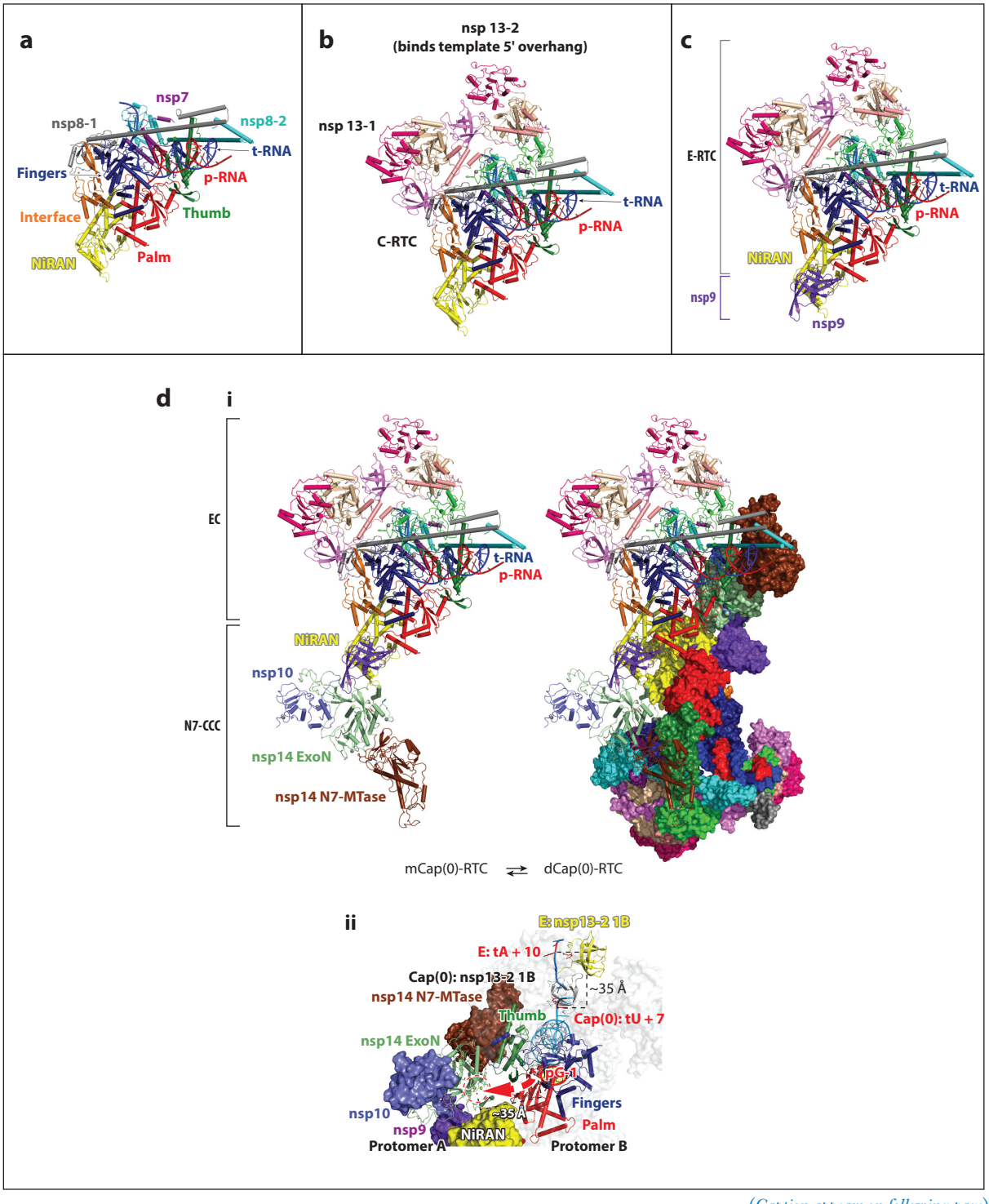
Figure 3

Summary of RTCs in the SARS-CoV-2 life cycle. The unavailable structures of Cap(-2)-RTC, Cap(-1)-RTC, and RTC-to-ribosome are indicated by a dashed circle or model. Abbreviations: cRNA, complementary RNA; C-RTC, central RTC; E-RTC, elongation RTC; gRNA, genomic RNA; nsp, nonstructural protein; RTC, replication–transcription complex; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; sgRNA, subgenomic RNA; ss, single-stranded.

study of the SARS-CoV nsp7/nsp8 complex (65). Due to the large molecular size of the complex and the dynamic process of RTC formation, the study of RTC assembly and function stalled until the structure of the SARS-CoV nsp12/nsp7/nsp8 complex was determined by single-particle cryo-EM technology in 2019 (80). Benefiting from the powerful support provided by single-particle cryo-EM, our understanding of CoV RTC assembly and function has significantly improved, mainly based on the studies of SARS-CoV-2 RTCs. Based on our current knowledge, we define the CoV replication and transcription cycle in host cells as including seven key stages (**Figure 3**). We summarize these below.

C-RTC

Prior to RNA synthesis, nsp12 interacts with cofactors nsp7 and nsp8 to form the central RTC (C-RTC). The C-RTC is the central component constituting all states of the RTC. SARS-CoV and SARS-CoV-2 C-RTCs are both composed of one nsp12, one nsp7, and two nsp8 molecules (67, 80). One individual nsp8 (nsp8-1) mainly contacts the nsp12 interface and fingers subdomains, whereas another nsp8 (nsp8-2), together with nsp7, interacts with the nsp12 fingers and thumb subdomains. In the apo form of C-RTC, the N-terminal helices of both nsp8s are not visible (67). But in the presence of the template–product duplex, nsp8 N-terminal helices are displayed as long helices and lift the duplex extending from the catalytic center of nsp12 Pol like a crane (68, 69) (**Figure 4*a***).



(Caption appears on following page)

Figure 4 (*Figure appears on preceding page*)

(*a*) Structure of the C-RTC. So that the complete structure of the nsp8 N-terminal helix can be viewed, the structure of the C-RTC is shown as it is in the E-RTC with the template–product duplex (PDB ID: 7CXM) (89). (*b*) Structure of the E-RTC (PDB ID: 7CXM) (89). (*c*) Structure of Cap(-1)′-RTC (PDB ID: 7CYQ) (13). (*d*, *i*) Structures of mCap(0)-RTC (PDB ID: 7EIZ) and dCap(0)-RTC (PDB ID: 7EGQ) (76). (*d*, *ii*) A proposed model for in *trans* backtracking proofreading. In the Cap(0)-RTC protomer A, the fingers, palm, and thumb subdomains of nsp12 are shown as cartoons in blue, red, and green. The t-RNA and p-RNA are represented by blue and red helices, respectively, and the nucleotides at the 3′ end of the primer are shown as colored spheres. Other parts in protomer A are covered by a white surface. In the Cap(0)-RTC protomer B, the nsp14 ExoN is shown as a cartoon in pale green, while other components in protomer B are shown as a molecular surface. The catalytic center of the nsp14 ExoN in protomer B is highlighted by a red dashed circle. A pathway for the backtracked product RNA is indicated by a red dashed arrow. Abbreviations: CCC, cotranscription capping complex; C-RTC, central RTC; EC, elongation complex; E-RTC, elongation RTC; ExoN, exonuclease; N7-MTase, N7-methyltransferase; NiRAN, nidovirus RNA-dependent RNA polymerase–associated nucleotidyltransferase; nsp, nonstructural protein; PDB ID, Protein Data Bank identifier; p-RNA, product RNA; RTC, replication–transcription complex; tA + 10, adenosine monophosphate at the 10 position in the t-RNA chain; t-RNA, template RNA; tU + 7, urydine monophosphate at the 7 position in the t-RNA chain.

E-RTC

Because there are many highly ordered structures in the RNA genome (111), the C-RTC alone is not likely to initiate the synthesis of RNA. To achieve this, nsp13 must be recruited to the C-RTC to preprocess these highly ordered structures, thus generating the template to guide RNA synthesis in the C-RTC. The complex formed by nsp13 and the C-RTC is named the elongation RTC (E-RTC) (89) (**Figure 4*b***). Because of the poor binding affinity of nsp13 with the C-RTC, a template–product duplex with overhang at the template 5′ end mimicking the single-stranded template extending from nsp13 to the nsp12 Pol catalytic center was designed to help capture the structure of the E-RTC (88, 89). An unexpected structural feature is that two nsp13 molecules bind to the C-RTC; one binds with the template 5′ overhang, but the other does not bind to RNA (88, 89). The interactions between two nsp13 molecules play a role in enhancing the helicase activity of the E-RTC (89). Another key finding is that the template 5′ overhang binds to the RNA-binding groove of one nsp13 in the 3′–5′ direction (88, 89), which is contrary to the 5′–3′ polarity identified by biochemical analysis (88–91). A backtracking model has thus been proposed to concert these structural and biochemical results (88, 89). However, this model raises the challenging question of how to unwind the template–product duplex synthesized by the polymerase. Although the structure of an RTC with an additional nsp13 further recruited onto the E-RTC to unwind the upstream end of the template–product duplex is still unavailable, nsp13 is likely to be the factor to unwind this duplex. The structural determination of such an RTC state warrants further investigation.

Cap(-2)-RTC and Cap(-1)-RTC

Both CoV gRNA and mRNA in positive sense are capped at their 5′ end, but cRNA or sgRNA in negative sense is likely not capped (7–9). To yield cRNA or sgRNA with the guidance of gRNA, the product cRNA or sgRNA unwound by helicase is likely to be directly released to the environment, since it does not need further modification at its 5′ end. But to synthesize gRNA or mRNA by using cRNA or sgRNA as the template, a four-step processive capping action is required to decorate cap(1) at the 5′ end. The first and second capping actions are facilitated by nsp13 RTPase activity (10, 11, 13–17, 90) and nsp12 NiRAN GTase activity (13), respectively. We name RTCs that generate the ppA end and the GpppA cap core at the 5′ end of product RNA as Cap(-2)-RTC and Cap(-1)-RTC, respectively; the use of Cap(-2) and Cap(-1) to refer to ppA and GpppA structures is not a formal scientific expression, and we used these terms here for ease of representation.

To date, no structures for Cap(-2)-RTC and Cap(-1)-RTC are available. However, the extremely large distance (over 120 Å in linear distance) between the 5′ end of product RNA of the duplex in the E-RTC and the catalytic center of nsp12 NiRAN strongly suggests that either (*a*) some other component or components would be recruited in the RTC to constitute a pathway that delivers the product 5′ end to nsp12 NiRAN for an in *cis* capping or (*b*) a special oligomerization state of the RTC would be formed to allow the close positioning of nsp12 NiRAN in one RTC with the product 5′ end in another RTC, thus reasoning an in *trans* mechanism for the first two capping actions.

Cap(-1)′-RTC

Prior to the third and fourth capping actions exerted in nsp14 and nsp16, a key intermediate state of the RTC is required, which is formed by supplying nsp9 to the E-RTC and is named Cap(-1)′-RTC (13) (**Figure 4*c***). Although revealed to be a dimer in crystallographic structures (70, 71, 73–75), nsp9 is recruited into Cap(-1)′-RTC as a monomer and functions as an adaptor for the further recruitment of nsp10/nsp14 or nsp10/nsp16 to form the cotranscription capping complex (CCC) for capping. Moreover, the binding of nsp9 is likely to suspend the RNA elongation and capping by two proposed mechanisms. First, the N terminus of nsp9 inserts deeply into the catalytic center of nsp12 NiRAN, forming stable interactions with a cation and a GDP bound there (13). The existence of nsp9 clearly attenuates nsp12 NiRAN GTase activity and GpppA formation (13). Second, the binding of nsp9 leads to a significant change in the orientation of nsp13-2, which binds the template 5′ overhang in the E-RTC (13). Upon nsp9 binding, nsp13-2 shifts as a rigid body and inserts its ZBD into the second minor groove of the template–product duplex like a bolt to obstruct the duplex's upstream movement. In this context, nsp9 not only provides a platform for the formation of CCC but is also likely to functionally suspend replication and transcription after the second capping action. As discussed in the section titled nsp9, this protein is a single-stranded nucleic acid–binding protein and shares topological homology with OB-fold proteins to bind long oligonucleotides (70, 71). Further structural and biological investigation is needed to find out whether nsp9 in RTCs can participate in RNA binding during replication and transcription or capping.

Cap(0)-RTC

The recruitment of the nsp10/nsp14 complex by nsp9 to Cap(-1)′-RTC results in the formation of Cap(0)-RTC (**Figure 4*d***, **subpanels *i,ii***). Cap(0)-RTC exists as an equilibrium of a monomeric form and a dimeric form, either in solution or in the cryo-EM captured state (76). In the monomeric Cap(0)-RTC [mCap(0)-RTC], nsp9 and nsp12 NiRAN exclusively interact with nsp14 to form N7-CCC together with nsp10. In N7-CCC, several highly positively charged regions were identified; these are presumably the pathway by which pre-gRNA and mRNA transfer from the catalytic center of nsp12 NiRAN to nsp14 N7-MTase. Alongside N7-CCC, the other components in mCap(0)-RTC (i.e., nsp12 Pol and nsp13) are named the elongation complex (EC), with EC and N7-CCC constituting a dumbbell-shaped architecture of mCap(0)-RTC.

Most importantly, the dimeric Cap(0)-RTC [dCap(0)-RTC] suggests the existence of an in *trans* backtracking mechanism of proofreading (76). In dCap(0)-RTC, the catalytic center of the nsp14 ExoN of one Cap(0)-RTC protomer faces the catalytic center of nsp12 Pol and is located approximately 35 Å from the 3′ end of product RNA. In a proposed model, once RTC senses a nucleotide is mistakenly incorporated into the product RNA, nsp13-2 1B undergoes large conformational changes to drag the template RNA in the reverse direction of RNA elongation, thus leading the 3′ end of product RNA with the misincorporated nucleotide to backtrack to

nsp14 ExoN for correction (76). After the correction of the mismatched nucleotide, nsp13-2 1B returns to its normal position for helicase function. The significant conformational change of nsp13-2 1B is not observed in any other SF1 or SF2 helicases. Whether this is a unique feature in the SARS-CoV-2 helicase or a similar mechanism could also function in other SF1 or SF2 helicases needs further investigation.

Cap(1)-RTC

In the last step of cap synthesis, an nsp10/nsp16 complex with 2′-*O*-MTase activity is recruited to finalize the capping process. We name this complex Cap(1)-RTC. Although the architecture of Cap(1)-RTC is currently unpublished, we have perspectives on Cap(1)-RTC based on the knowledge acquired from Cap(0)-RTC. One key question is whether nsp10/nsp14 and nsp10/nsp16 would together assemble a complete RTC with both N7-CCC and 2′-*O*-CCC. Some evidence suggests this is unlikely. First, the footprints of nsp14 and nsp16 on nsp10 overlap (100), suggesting the recruitment of nsp14 and nsp16 by one nsp10 into RTCs is less possible. Moreover, if we superpose the structures of Cap(0)-RTC (76) with nsp10/nsp16 (109, 110), the position of nsp16 significantly conflicts with the interprotomer interface in dCap(0)-RTC, which may negatively impact proofreading in dCap(0)-RTC. Therefore, a model for the sequential formation of N7-CCC and 2′-*O*-CCC in two subsequent capping actions is the most feasible.

OTHER NONSTRUCTURAL PROTEINS RELATED TO THE RTC

Transfer mRNA from the RTC to the Host Ribosome

CoVs usually replicate and transcribe their genomes in a position associated with modified host membranes, such as DMVs, which are the most abundant component of the viral replication organelle derived from transformed endoplasmic reticulum (ER) membranes (112–114). DMVs package the dsRNA replication intermediate and may offer a microenvironment in which to shield viral RNA from host innate immunity (112–114). A recent work has shown that nsp3, nsp4, and nsp6 form a hexameric ring-shaped pore on DMVs, thus providing an opening in the channel to export mRNA or gRNA to the cytosol for translation or wrapping by nucleocapsid proteins (46). Would gRNA or mRNA be generated in the DMV's interior and subsequently be exported to the cytosol in a free form? A subtomogram analysis of CoV-induced DMVs suggests the RTC is associated with the luminal side of the nsp3/nsp4/nsp6 pore (46), indicating mRNA could be exported to the cytosol immediately after it is synthesized in the RTC. Moreover, the N-terminal domain of nsp1 binds the 5′-UTR with high affinity (27). It is conceivable that nsp1 in the cytosol can bind to the mRNA 5′-UTR after it is exported from DMVs. This possibility raises another concern that nsp1 could mediate a way to transfer mRNA to the host ribosome, since the nsp1 C-terminal HTH fold tightly binds to host ribosome. We hypothesize that this might be homologous to dengue virus delivering mRNA from the vesicle's interior to an ER membrane–associated ribosome adjacent to the opening of the vesicle (115). In this proposal, nsp1 might function like a shuttle to deliver the mRNA 5′ end to the host ribosome. However, this assumption needs further study.

nsp2 and nsp15 with RTC

Whether nsp2 and nsp15 could participate in the formation of RTCs and what their function is in the CoV life cycle are currently unclear.

The protein nsp2 has an N-terminal ZRD domain bearing three ZFs, and zinc-coordinating proteins often play a role in nucleic acid or protein binding, similar to eukaryotic transcription

factors (116). Previous studies have shown the colocalization of nsp2 and nsp8, as well as RNA products, with RTCs (32). It is plausible that nsp2 could participate in RTC formation to stabilize the nucleic acid together with nsp8 or other nonstructural proteins for optimal function.

The protein nsp15 is encoded by the genome near the 5′ end. Although its function in the virus life cycle is not clear, the high specificity of nsp15 endoU with uridine (104, 105) suggests specific RNAs could be designed to test whether it could be involved in RTC formation and function.

SUMMARY POINTS

1. The structures of most nonstructural proteins have been well elucidated, but some [e.g., nonstructural protein 3 (nsp3)] need further study.
2. A set of replication–transcription complexes (RTCs) are formed by nonstructural proteins to facilitate coronavirus (CoV) replication and transcription.
3. The elongation complex and cotranscription capping complex constitute the basic architecture of a complete RTC.

FUTURE ISSUES

1. How is RNA synthesis initiated in nsp12 polymerase? Could the structure of nsp8 in a state for primer synthesis be compatible with the central RTC or elongation RTC?
2. Is nsp13 the exact protein to unwind the template–product duplex synthesized by polymerase? How are nsp13 RTPase and duplex unwinding activities compatible in the RTC?
3. How do RTCs distinguish and regulate continuous replication and discontinuous transcription?
4. How does nsp12 polymerase sense a mismatched nucleotide? And how does Cap(0)-RTC sense the correction (removing) of the mismatched nucleotide?
5. How do the four capping actions subsequently occur in the RTC?
6. How does nsp10/nsp16 form the RTC with other components to finalize the last capping actions?
7. How does mRNA transfer from the RTC to the host ribosome? Can translation with transcription occur in CoVs like it does in prokaryotes or other viruses (e.g., dengue virus)?

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We apologize to all the investigators who may not be appropriately cited owing to space limitations. We thank Prof. Luke W. Guddat for reviewing the manuscript. The research is funded 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 (2020Z99CFG015).

LITERATURE CITED

1. WHO (World Health Organ.). 2022. WHO coronavirus disease (COVID-19) dashboard. *World Health Organ.* **https://covid19.who.int/**
2. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, et al. 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579:270–73
3. Hamre D, Procknow JJ. 1966. A new virus isolated from the human respiratory tract. *Proc. Soc. Exp. Biol. Med.* 121:190–93
4. McIntosh K, Dees JH, Becker WB, Kapikian AZ, Chanock RM. 1967. Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease. *PNAS* 57:933–40
5. Sola I, Almazán F, Zúñiga S, Enjuanes L. 2015. Continuous and discontinuous RNA synthesis in coronaviruses. *Annu. Rev. Virol.* 2:265–88
6. Robson F, Khan KS, Le TK, Paris C, Demirbag S, et al. 2020. Coronavirus RNA proofreading: molecular basis and therapeutic targeting. *Mol. Cell* 79:710–27
7. Lai MM, Stohlman SA. 1981. Comparative analysis of RNA genomes of mouse hepatitis viruses. *J. Virol.* 38:661–70
8. van Vliet ALW, Smits SL, Rottier PJM, de Groot RJ. 2002. Discontinuous and non-discontinuous subgenomic RNA transcription in a nidovirus. *EMBO J.* 21:6571–80
9. Peng YH, Lin CH, Lin CN, Lo CY, Tsai TL, Wu HY. 2016. Characterization of the role of hexamer AGUAAA and poly(A) tail in coronavirus polyadenylation. *PLOS ONE* 11:e0165077
10. Bouvet M, Debarnot C, Imbert I, Selisko B, Snijder EJ, et al. 2010. In vitro reconstitution of SARS-coronavirus mRNA cap methylation. *PLOS Pathog.* 6:e1000863
11. Chen Y, Su C, Ke M, Jin X, Xu L, 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
12. Daffis S, Szretter KJ, Schriewer J, Li J, Youn S, et al. 2010. 2′-*O* methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature* 468:452–56
13. Yan L, Ge J, Zheng L, Zhang Y, Gao Y, et al. 2021. Cryo-EM structure of an extended SARS-CoV-2 replication and transcription complex reveals an intermediate state in cap synthesis. *Cell* 184:184–93.e10
14. Ivanov KA, Ziebuhr J. 2004. Human coronavirus 229E nonstructural protein 13: characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5′-triphosphatase activities. *J. Virol.* 78:7833–38
15. Ivanov KA, Thiel V, Dobbe JC, van der Meer Y, Snijder EJ, Ziebuhr J. 2004. Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase. *J. Virol.* 78:5619–32
16. Chen Y, Cai H, Pan J, Xiang N, Tien P, et al. 2009. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *PNAS* 106:3484–89
17. Decroly E, Imbert I, Coutard B, Bouvet M, Selisko B, et al. 2008. Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing (nucleoside-2′*O*)-methyltransferase activity. *J. Virol.* 82:8071–84
18. Wu A, Peng Y, Huang B, Ding X, Wang X, et al. 2020. Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell Host Microbe* 27:325–28
19. Almeida MS, Johnson MA, Herrmann T, Geralt M, Wüthrich K. 2007. Novel β-barrel fold in the nuclear magnetic resonance structure of the replicase nonstructural protein 1 from the severe acute respiratory syndrome coronavirus. *J. Virol.* 81:3151–61
20. Semper C, Watanabe N, Savchenko A. 2021. Structural characterization of nonstructural protein 1 from SARS-CoV-2. *iScience* 24:101903
21. Clark LK, Green TJ, Petit CM. 2021. Structure of nonstructural protein 1 from SARS-CoV-2. *J. Virol.* 95:e02019-20
22. Schubert K, Karousis ED, Jomaa A, Scaiola A, Echeverria B, et al. 2020. SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. *Nat. Struct. Mol. Biol.* 27:959–66

23. Thoms M, Buschauer R, Ameismeier M, Koepke L, Denk T, et al. 2020. Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. *Science* 369:1249–55
24. Yuan S, Peng L, Park JJ, Hu Y, Devarkar SC, et al. 2020. Nonstructural protein 1 of SARS-CoV-2 is a potent pathogenicity factor redirecting host protein synthesis machinery toward viral RNA. *Mol. Cell* 80:1055–66.e6
25. Narayanan K, Huang C, Lokugamage K, Kamitani W, Ikegami T, et al. 2008. Severe acute respiratory syndrome coronavirus nsp1 suppresses host gene expression, including that of type I interferon, in infected cells. *J. Virol.* 82:4471–79
26. Kumar A, Ishida R, Strilets T, Cole J, Lopez-Orozco J, et al. 2021. SARS-CoV-2 nonstructural protein 1 inhibits the interferon response by causing depletion of key host signaling factors. *J. Virol.* 95:e0026621
27. Vankadari N, Jeyasankar NN, Lopes WJ. 2020. Structure of the SARS-CoV-2 Nsp1/5′-untranslated region complex and implications for potential therapeutic targets, a vaccine, and virulence. *J. Phys. Chem. Lett.* 11:9659–68
28. Kamitani W, Narayanan K, Huang C, Lokugamage K, Ikegami T, et al. 2006. Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. *PNAS* 103:12885–90
29. Kamitani W, Huang C, Narayanan K, Lokugamage KG, Makino S. 2009. A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. *Nat. Struct. Mol. Biol.* 16:1134–40
30. Tanaka T, Kamitani W, DeDiego ML, Enjuanes L, Matsuura Y. 2012. Severe acute respiratory syndrome coronavirus nsp1 facilitates efficient propagation in cells through a specific translational shutoff of host mRNA. *J. Virol.* 86:11128–37
31. Kilkenny ML, Veale CE, Guppy A, Hardwick SW, Chirgadze DY, et al. 2021. Structural basis for the interaction of SARS-CoV-2 virulence factor nsp1 with DNA polymerase α–primase. *Protein Sci*. 31:333–44
32. Gadlage MJ, Graham RL, Denison MR. 2008. Murine coronaviruses encoding nsp2 at different genomic loci have altered replication, protein expression, and localization. *J. Virol.* 82:11964–69
33. Cornillez-Ty CT, Liao L, Yates JR III, Kuhn P, Buchmeier MJ. 2009. Severe acute respiratory syndrome coronavirus nonstructural protein 2 interacts with a host protein complex involved in mitochondrial biogenesis and intracellular signaling. *J. Virol.* 83:10314–18
34. Ziebuhr J, Snijder EJ, Gorbalenya AE. 2000. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J. Gen. Virol.* 81:853–79
35. Imbert I, Snijder EJ, Dimitrova M, Guillemot JC, Lécine P, Canard B. 2008. The SARS-Coronavirus PLnc domain of nsp3 as a replication/transcription scaffolding protein. *Virus Res.* 133:136–48
36. Pfefferle S, Schöpf J, Kögl M, Friedel CC, Müller MA, et al. 2011. The SARS-coronavirus-host interactome: identification of cyclophilins as target for pan-coronavirus inhibitors. *PLOS Pathog.* 7:e1002331
37. van Hemert MJ, van den Worm SHE, Knoops K, Mommaas AM, Gorbalenya AE, Snijder EJ. 2008. SARS-coronavirus replication/transcription complexes are membrane-protected and need a host factor for activity in vitro. *PLOS Pathog.* 4:e1000054
38. Angelini MM, Akhlaghpour M, Neuman BW, Buchmeier MJ. 2013. Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles. *mBio* 4:e00524-13
39. Snijder EJ, van der Meer Y, Zevenhoven-Dobbe J, Onderwater JJ, van der Meulen J, et al. 2006. Ultrastructure and origin of membrane vesicles associated with the severe acute respiratory syndrome coronavirus replication complex. *J. Virol.* 80:5927–40
40. Lei J, Kusov Y, Hilgenfeld R. 2018. Nsp3 of coronaviruses: structures and functions of a large multi-domain protein. *Antivir. Res.* 149:58–74
41. Clementz MA, Kanjanahaluethai A, O'Brien TE, Baker SC. 2008. Mutation in murine coronavirus replication protein nsp4 alters assembly of double membrane vesicles. *Virology* 375:118–29
42. Oostra M, te Lintelo EG, Deijs M, Verheije MH, Rottier PJM, de Haan CAM. 2007. Localization and membrane topology of coronavirus nonstructural protein 4: involvement of the early secretory pathway in replication. *J. Virol.* 81:12323–36
43. Xu X, Lou Z, Ma Y, Chen X, Yang Z, et al. 2009. Crystal structure of the C-terminal cytoplasmic domain of non-structural protein 4 from mouse hepatitis virus A59. *PLOS ONE* 4:e6217

44. Manolaridis I, Wojdyla JA, Panjikar S, Snijder EJ, Gorbalenya AE, et al. 2009. Structure of the C-terminal domain of nsp4 from feline coronavirus. *Acta Crystallogr. D Biol. Crystallogr.* 65:839–46
45. Cottam EM, Maier HJ, Manifava M, Vaux LC, Chandra-Schoenfelder P, et al. 2011. Coronavirus nsp6 proteins generate autophagosomes from the endoplasmic reticulum via an omegasome intermediate. *Autophagy* 7:1335–47
46. Wolff G, Limpens RWAL, Zevenhoven-Dobbe JC, Laugks U, Zheng S, et al. 2020. A molecular pore spans the double membrane of the coronavirus replication organelle. *Science* 369:1395–98
47. Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R. 2003. Coronavirus main proteinase ($3CL^{pro}$) structure: basis for design of anti-SARS drugs. *Science* 300:1763–67
48. Yang H, Yang M, Ding Y, Liu Y, Lou Z, et al. 2003. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *PNAS* 100:13190–95
49. Xue X, Yu H, Yang H, Xue F, Wu Z, et al. 2008. Structures of two coronavirus main proteases: implications for substrate binding and antiviral drug design. *J. Virol.* 82:2515–27
50. Zhao Q, Li S, Xue F, Zou Y, Chen C, et al. 2008. Structure of the main protease from a global infectious human coronavirus, HCoV-HKU1. *J. Virol.* 82:8647–55
51. Ren Z, Yan L, Zhang N, Guo Y, Yang C, et al. 2013. The newly emerged SARS-like coronavirus HCoV-EMC also has an "Achilles' heel": current effective inhibitor targeting a 3C-like protease. *Protein Cell* 4:248–50
52. Lai MM, Cavanagh D. 1997. The molecular biology of coronaviruses. *Adv. Virus Res.* 48:1–100
53. Lu Y, Lu X, Denison MR. 1995. Identification and characterization of a serine-like proteinase of the murine coronavirus MHV-A59. *J. Virol.* 69:3554–59
54. Snijder EJ, Decroly E, Ziebuhr J. 2016. The nonstructural proteins directing coronavirus RNA synthesis and processing. *Adv. Virus Res.* 96:59–126
55. Jin Z, Du X, Xu Y, Deng Y, Liu M, et al. 2020. Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors. *Nature* 582:289–93
56. Jin Z, Zhao Y, Sun Y, Zhang B, Wang H, et al. 2020. Structural basis for the inhibition of SARS-CoV-2 main protease by antineoplastic drug carmofur. *Nat. Struct. Mol. Biol.* 27:529–32
57. Yang H, Xie W, Xue X, Yang K, Ma J, et al. 2005. Design of wide-spectrum inhibitors targeting coronavirus main proteases. *PLOS Biol.* 3:e324
58. Ghosh AK, Gong G, Grum-Tokars V, Mulhearn DC, Baker SC, et al. 2008. Design, synthesis and antiviral efficacy of a series of potent chloropyridyl ester-derived SARS-CoV 3CLpro inhibitors. *Bioorg. Med. Chem. Lett.* 18:5684–88
59. Yang H, Bartlam M, Rao Z. 2006. Drug design targeting the main protease, the Achilles' heel of coronaviruses. *Curr. Pharm. Des.* 12:4573–90
60. Bartlam M, Yang HT, Rao ZH. 2005. Structural insights into SARS coronavirus proteins. *Curr. Opin. Struct. Biol.* 15:664–72
61. Lou Z, Sun Y, Rao Z. 2014. Current progress in antiviral strategies. *Trends Pharmacol. Sci.* 35:86–102
62. Owen DR, Allerton CMN, Anderson AS, Aschenbrenner L, Avery M, et al. 2021. An oral SARS-CoV-2 M^{pro} inhibitor clinical candidate for the treatment of COVID-19. *Science* 374:1586–93
63. Imbert I, Guillemot JC, Bourhis JM, Bussetta C, Coutard B, et al. 2006. A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J.* 25:4933–42
64. te Velthuis AJW, van den Worm SHE, Snijder EJ. 2012. The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. *Nucleic Acids Res.* 40:1737–47
65. Zhai Y, Sun F, Li X, Pang H, Xu X, et al. 2005. Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer. *Nat. Struct. Mol. Biol.* 12:980–86
66. Li S, Zhao Q, Zhang Y, Zhang Y, Bartlam M, et al. 2010. New nsp8 isoform suggests mechanism for tuning viral RNA synthesis. *Protein Cell* 1:198–204
67. Gao Y, Yan L, Huang Y, Liu F, Zhao Y, et al. 2020. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* 368:779–82
68. Hillen HS, Kokic G, Farnung L, Dienemann C, Tegunov D, Cramer P. 2020. Structure of replicating SARS-CoV-2 polymerase. *Nature* 584:154–56

69. Wang Q, Wu J, Wang H, Gao Y, Liu Q, et al. 2020. Structural basis for RNA replication by the SARS-CoV-2 polymerase. *Cell* 182:417–28.e13
70. Egloff MP, Ferron F, Campanacci V, Longhi S, Rancurel C, et al. 2004. The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. *PNAS* 101:3792–96
71. Sutton G, Fry E, Carter L, Sainsbury S, Walter T, et al. 2004. The nsp9 replicase protein of SARS-coronavirus, structure and functional insights. *Structure* 12:341–53
72. Frieman M, Yount B, Agnihothram S, Page C, Donaldson E, et al. 2012. Molecular determinants of severe acute respiratory syndrome coronavirus pathogenesis and virulence in young and aged mouse models of human disease. *J. Virol.* 86:884–97
73. Littler DR, Gully BS, Colson RN, Rossjohn J. 2020. Crystal structure of the SARS-CoV-2 non-structural protein 9, Nsp9. *iScience* 23:101258
74. Kleiger G, Grothe R, Mallick P, Eisenberg D. 2002. GXXXG and AXXXA: common α-helical interaction motifs in proteins, particularly in extremophiles. *Biochemistry* 41:5990–97
75. Miknis ZJ, Donaldson EF, Umland TC, Rimmer RA, Baric RS, Schultz LW. 2009. Severe acute respiratory syndrome coronavirus nsp9 dimerization is essential for efficient viral growth. *J. Virol.* 83:3007–18
76. Yan L, Yang Y, Li M, Zhang Y, Zheng L, et al. 2021. Coupling of N7-methyltransferase and 3′-5′ exoribonuclease with SARS-CoV-2 polymerase reveals mechanisms for capping and proofreading. *Cell* 184:3474–85.e11
77. Su D, Lou Z, Sun F, Zhai Y, Yang H, et al. 2006. Dodecamer structure of severe acute respiratory syndrome coronavirus nonstructural protein nsp10. *J. Virol.* 80:7902–8
78. Decroly E, Debarnot C, Ferron F, Bouvet M, Coutard B, et al. 2011. Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2′-O-methyltransferase nsp10/nsp16 complex. *PLOS Pathog.* 7:e1002059
79. Ma Y, Wu L, Shaw N, Gao Y, Wang J, et al. 2015. Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. *PNAS* 112:9436–41
80. Kirchdoerfer RN, Ward AB. 2019. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nat. Commun.* 10:2342
81. Ziebuhr J. 2005. The coronavirus replicase. *Curr. Top. Microbiol. Immunol.* 287:57–94
82. McDonald SM. 2013. RNA synthetic mechanisms employed by diverse families of RNA viruses. *Wiley Interdiscip. Rev. RNA* 4:351–67
83. Lehmann KC, Gulyaeva A, Zevenhoven-Dobbe JC, Janssen GM, Ruben 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.* 43:8416–34
84. Walker AP, Fan H, Keown JR, Grimes JM, Fodor E. 2021. Identification of guanylyltransferase activity in the SARS-CoV-2 RNA polymerase. bioRxiv 2021.03.17.435913. **https://doi.org/10.1101/2021.03.17.435913**
85. Slanina H, Madhugiri R, Bylapudi G, Schultheiß K, Karl N, et al. 2021. Coronavirus replication–transcription complex: vital and selective NMPylation of a conserved site in nsp9 by the NiRAN-RdRp subunit. *PNAS* 118:e2022310118
86. Conti BJ, Kirchdoerfer RN, Sussman MR. 2020. Mass spectrometric based detection of protein nucleotidylation in the RNA polymerase of SARS-CoV-2. bioRxiv 2020.10.07.330324. **https://doi.org/10.1101/2020.10.07.330324**
87. Jia Z, Yan L, Ren Z, Wu L, Wang J, et al. 2019. Delicate structural coordination of the Severe Acute Respiratory Syndrome coronavirus Nsp13 upon ATP hydrolysis. *Nucleic Acids Res.* 47:6538–50
88. Chen J, Malone B, Llewellyn E, Grasso M, Shelton PMM, et al. 2020. Structural basis for helicase-polymerase coupling in the SARS-CoV-2 replication-transcription complex. *Cell* 182:1560–73.e13
89. Yan L, Zhang Y, Ge J, Zheng L, Gao Y, et al. 2020. Architecture of a SARS-CoV-2 mini replication and transcription complex. *Nat. Commun.* 11:5874
90. Seybert A, van Dinten LC, Snijder EJ, Ziebuhr J. 2000. Biochemical characterization of the equine arteritis virus helicase suggests a close functional relationship between arterivirus and coronavirus helicases. *J. Virol.* 74:9586–93

91. Seybert A, Hegyi A, Siddell SG, Ziebuhr J. 2000. The human coronavirus 229E superfamily 1 helicase has RNA and DNA duplex-unwinding activities with 5′-to-3′ polarity. *RNA* 6:1056–68
92. Chen J, Wang Q, Malone B, Llewellyn E, Pechersky Y, et al. 2021. Ensemble cryo-electron microscopy reveals conformational states of the nsp13 helicase in the SARS-CoV-2 helicase replication-transcription complex. bioRxiv 2021.11.10.468168. **https://doi.org/10.1101/2021.11.10.468168**
93. Minskaia E, Hertzig T, Gorbalenya AE, Campanacci V, Cambillau C, et al. 2006. Discovery of an RNA virus 3′→5′ exoribonuclease that is critically involved in coronavirus RNA synthesis. *PNAS* 103:5108–13
94. Eckerle LD, Becker MM, Halpin RA, Li K, Venter E, et al. 2010. Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. *PLOS Pathog.* 6:e1000896
95. Eckerle LD, Lu X, Sperry SM, Choi L, Denison MR. 2007. High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *J. Virol.* 81:12135–44
96. Ferron F, Subissi L, Silveira De Morais AT, Le NTT, Sevajol M, et al. 2018. Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. *PNAS* 115:E162–71
97. Case JB, Ashbrook AW, Dermody TS, Denison MR. 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.* 90:7248–56
98. Munster VJ, Feldmann F, Williamson BN, van Doremalen N, Pérez-Pérez L, et al. 2020. Respiratory disease in rhesus macaques inoculated with SARS-CoV-2. *Nature* 585:268–72
99. Ogando NS, Ferron F, Decroly E, Canard B, Posthuma CC, Snijder EJ. 2019. The curious case of the nidovirus exoribonuclease: its role in RNA synthesis and replication fidelity. *Front. Microbiol.* 10:1813
100. Bouvet M, Imbert I, Subissi L, Gluais L, Canard B, Decroly E. 2012. RNA 3′-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. *PNAS* 109:9372–77
101. Pillon MC, Frazier MN, Dillard LB, Williams JG, Kocaman S, et al. 2021. Cryo-EM structures of the SARS-CoV-2 endoribonuclease Nsp15 reveal insight into nuclease specificity and dynamics. *Nat. Commun.* 12:636
102. Xu X, Zhai Y, Sun F, Lou Z, Su D, et al. 2006. New antiviral target revealed by the hexameric structure of mouse hepatitis virus nonstructural protein nsp15. *J. Virol.* 80:7909–17
103. Zhang L, Li L, Yan L, Ming Z, Jia Z, et al. 2018. Structural and biochemical characterization of endoribonuclease Nsp15 encoded by Middle East respiratory syndrome coronavirus. *J. Virol.* 92:e00893-18
104. Kang H, Bhardwaj K, Li Y, Palaninathan S, Sacchettini J, et al. 2007. Biochemical and genetic analyses of murine hepatitis virus Nsp15 endoribonuclease. *J. Virol.* 81:13587–97
105. Bhardwaj K, Guarino L, Kao CC. 2004. The severe acute respiratory syndrome coronavirus Nsp15 protein is an endoribonuclease that prefers manganese as a cofactor. *J. Virol.* 78:12218–24
106. Hackbart M, Deng X, Baker SC. 2020. Coronavirus endoribonuclease targets viral polyuridine sequences to evade activating host sensors. *PNAS* 117:8094–103
107. Deng X, van Geelen A, Buckley AC, O'Brien A, Pillatzki A, et al. 2019. Coronavirus endoribonuclease activity in porcine epidemic diarrhea virus suppresses type I and type III interferon responses. *J. Virol.* 93:e02000-18
108. Kindler E, Gil-Cruz C, Spanier J, Li Y, Wilhelm J, et al. 2017. Early endonuclease-mediated evasion of RNA sensing ensures efficient coronavirus replication. *PLOS Pathog.* 13:e1006195
109. Viswanathan T, Arya S, Chan SH, Qi S, Dai N, et al. 2020. Structural basis of RNA cap modification by SARS-CoV-2. *Nat. Commun.* 11:3718
110. Krafcikova P, Silhan J, Nencka R, Boura E. 2020. Structural analysis of the SARS-CoV-2 methyltransferase complex involved in RNA cap creation bound to sinefungin. *Nat. Commun.* 11:3717
111. Sun L, Li P, Ju X, Rao J, Huang W, et al. 2021. In vivo structural characterization of the SARS-CoV-2 RNA genome identifies host proteins vulnerable to repurposed drugs. *Cell* 184:1865–83.e20
112. Knoops K, Kikkert M, van den Worm SHE, Zevenhoven-Dobbe JC, van der Meer Y, et al. 2008. SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLOS Biol.* 6:e226

113. Snijder EJ, Limpens RWAL, de Wilde AH, de Jong AWM, Zevenhoven-Dobbe JC, et al. 2020. A unifying structural and functional model of the coronavirus replication organelle: tracking down RNA synthesis. *PLOS Biol*. 18:e3000715
114. Ulasli M, Verheije MH, de Haan CAM, Reggiori F. 2010. Qualitative and quantitative ultrastructural analysis of the membrane rearrangements induced by coronavirus. *Cell Microbiol*. 12:844–61
115. Welsch S, Miller S, Romero-Brey I, Merz A, Bleck CK, et al. 2009. Composition and three-dimensional architecture of the dengue virus replication and assembly sites. *Cell Host Microbe* 5:365–75
116. Luscombe NM, Austin SE, Berman HM, Thornton JM. 2000. An overview of the structures of protein-DNA complexes. *Genome Biol*. 1:reviews001.1