



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

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



ELSEVIER



# Structure and function of SARS-CoV-2 polymerase

Hauke S Hillen<sup>1,2,3</sup>

Coronaviruses use an RNA-dependent RNA polymerase (RdRp) to replicate and express their genome. The RdRp associates with additional non-structural proteins (nsps) to form a replication–transcription complex (RTC) that carries out RNA synthesis, capping and proofreading. However, the structure of the RdRp long remained elusive, thus limiting our understanding of coronavirus genome expression and replication. Recently, the cryo-electron microscopy structure of SARS-CoV-1 RdRp was reported. Driven by the ongoing COVID-19 pandemic, structural data on the SARS-CoV-2 polymerase and associated factors has since emerged at an unprecedented pace, with more than twenty structures released to date. This review provides an overview of the currently available coronavirus RdRp structures and outlines how they have, together with functional studies, led to a molecular understanding of the viral polymerase, its interactions with accessory factors and the mechanisms by which promising antivirals may inhibit coronavirus replication.

## Addresses

<sup>1</sup> Department of Cellular Biochemistry, University Medical Center Göttingen, Humboldtallee 23, D-37073 Göttingen, Germany

<sup>2</sup> Research Group Structure and Function of Molecular Machines, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, D-37077 Göttingen, Germany

<sup>3</sup> Cluster of Excellence “Multiscale Bioimaging: from Molecular Machines to Networks of Excitable Cells” (MBExC), University of Göttingen, D-37075 Göttingen, Germany

Corresponding author:

Hillen, Hauke S ([hauke.hillen@med.uni-goettingen.de](mailto:hauke.hillen@med.uni-goettingen.de))

Current Opinion in Virology 2021, 48:82–90

This review comes from a themed issue on **Virus structure and expression**

Edited by **Adolfo H Moraes** and **Flávio Fonseca**

For complete overview about the section, refer [Virus Structure and Expression \(2021\)](#)

Available online 6th April 2021

<https://doi.org/10.1016/j.coviro.2021.03.010>

1879-6257/© 2021 The Author. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## Introduction

Coronaviruses are enveloped single-stranded RNA viruses that belong to the order of *nidovirales* [1]. In humans, four coronaviruses (229E, NL63, OC43 and HKU1) are known to circulate endemically and cause mostly mild respiratory infections [2]. In addition, three coronaviruses that can cause severe acute respiratory distress syndrome have emerged over the past two

decades [3–8]. The first two, Severe Acute Respiratory Syndrome Coronavirus 1 (SARS-CoV-1) and Middle East Respiratory Syndrome-related Coronavirus (MERS-CoV), caused outbreaks in 2002 and 2012, respectively, with high case fatality rates. The third, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), is currently causing the COVID-19 pandemic, which has led to over two million deaths worldwide [3,4,9]. All three are thought to have emerged through spillovers from animal reservoirs [4,10].

The genomes of coronaviruses are among the largest known for RNA viruses, ranging from ~26–32 kbp [11]. In addition to the structural and accessory proteins required for virion assembly, they encode for 16 non-structural proteins (nsp1-16) that drive genome expression and replication. The (+)-strand genome can serve directly as template for the expression of the non-structural proteins, which are translated as large polyproteins (ORF1a or ORF1ab) and subsequently cleaved by viral proteases [12]. In contrast, structural and accessory proteins are translated from nested subgenomic mRNAs, the biogenesis of which requires template switching of the polymerase during transcription [13]. The non-structural proteins associate to form a replication–transcription complex (RTC) that carries out RNA synthesis, capping and proofreading, a unique feature of nidoviruses [12]. At the heart of the RTC is nsp12, which harbors the RNA-dependent RNA polymerase (RdRp) active site [14]. In order to be enzymatically active, it requires the accessory factors nsp7 and nsp8, which together with nsp12 form the core RdRp complex [15]. Additional subunits are thought to bind to this complex to form the RTC, including the helicase nsp13, the proofreading exonuclease and methyltransferase nsp14, the methyltransferase nsp16, as well as nsp9 and nsp10 [12]. Like cellular mRNAs, coronaviral transcripts bear an N<sup>7</sup>-methyl guanosine cap at their 5' end which protects the RNA from degradation, facilitates translation and may mediate immune escape [12,16]. Cytosolic RNA viruses usually encode the enzymes necessary for formation of this cap, namely a triphosphatase (TPase), a guanylyltransferase (GTase) and at least one methyltransferase (MTase) [17]. While nsp13 has been suggested to act as the TPase [18] and nsp14 and nsp16 may methylate N<sup>7</sup> of the guanosine base and the 2'-OH of the ribose, respectively [19,20], no GTase activity has been identified in coronavirus proteins so far [12]. Despite a large body of functional studies, a molecular understanding of the coronavirus RTC and how it integrates different steps of viral gene

expression has been lacking, in part due to the absence of structural data.

Since the first outbreak of SARS-CoV-1 in 2002, the structures of many coronavirus proteins have been determined by X-Ray crystallography or cryo-electron microscopy (cryo-EM). However, the structure of the coronavirus RNA polymerase long remained elusive. Only recently, in 2019, the first cryo-EM structure of the SARS-CoV-1 RdRp was reported [21\*\*]. Since then, the COVID-19 pandemic has fueled scientific interest in coronavirus biology, and this has led to rapid progress in the structural characterization of coronavirus replication and gene expression. Within months after the first case reports, the cryo-EM structure of the SARS-CoV-2 RdRp was reported, and this was quickly followed by further structures of polymerase complexes in different functional states (Table 1). This review provides an overview of the currently available coronavirus RdRp structures and the molecular picture of coronavirus replication that is beginning to emerge from both structural and functional data.

### Structure of the coronavirus core RdRp complex

The first structural insights into the coronavirus polymerase were obtained only shortly before the current SARS-CoV-2 pandemic, when the cryo-EM structure of the SARS-CoV-1 core RdRp was reported [21\*\*]. This structure revealed the overall architecture of the polymerase

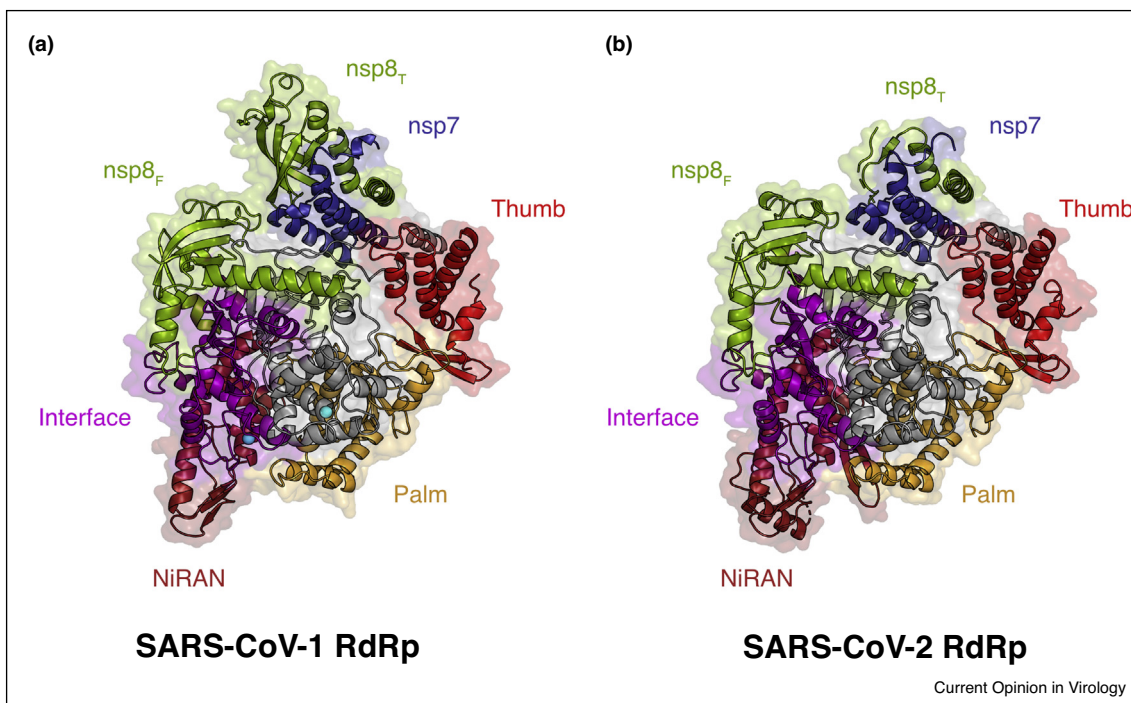
subunit nsp12 and how it interacts with the accessory factors nsp7 and nsp8 (Figure 1a). Nsp12 contains a C-terminal polymerase domain homologous to other viral RdRps and a N-terminal nidovirus-specific domain, which is essential for virus replication. This domain has been proposed to harbor a nucleotidyl transferase activity, and has hence been termed nidovirus RdRp-associated nucleotidyl transferase (NiRAN) domain [22]. Although the NiRAN domain was only partially resolved in this original structure, it was shown to resemble the nucleotide binding domains of kinases, but its molecular function remained unclear. The polymerase domain of nsp12 adopts the typical ‘right-hand’ fold of single-subunit RNA polymerases, consisting of thumb, fingers and palm subdomains [23]. These subdomains contain conserved motifs that together form the polymerase active site. Two copies of nsp8 bind to nsp12 on opposing sides of the RNA-binding cleft. These two nsp8 copies have been referred to by different designations in the literature (for example nsp8a/b [24\*\*] or nsp8-1/2 [25\*\*]). To unify the nomenclature, I here propose to refer to them according to their binding location on nsp12: Nsp8<sub>T</sub> (for ‘thumb’) binds on one side of the cleft through an interaction of its C-terminal head domain with nsp7, which in turn binds to an extension in the fingers domain of nsp12 next to the thumb. Nsp8<sub>F</sub> (for ‘fingers’) binds on the other side of cleft, and its head domain adopts a different fold and interacts with the fingers and interface domain of nsp12 directly. The N-terminal domains of both nsp8, which are highly conserved [26\*] and harbor

**Table 1**

**Available coronavirus RdRp cryo-EM structures (in order of publication)**

RdRp complex	Resolution (Å)	PDB code	Reference
SARS-CoV-1 RdRp (nsp12-nsp7-nsp8 <sub>2</sub> )	3.1	6NUR	Kirchdoerfer and Ward [21**]
SARS-CoV-1 RdRp (nsp12-nsp8)	3.5	6NUS	Kirchdoerfer and Ward [21**]
SARS-CoV-2 RdRp (nsp12-nsp7-nsp8 <sub>2</sub> )	2.9	6M71	Gao <i>et al.</i> [25**]
SARS-CoV-2 RdRp (nsp12-nsp7-nsp8 <sub>2</sub> )	3.0	7BTF	Gao <i>et al.</i> [25**]
SARS-CoV-2 RdRp (nsp12-nsp7-nsp8 <sub>2</sub> )	3.7	7BW4	Peng <i>et al.</i> [27]
SARS-CoV-2 RdRp (nsp12-nsp7-nsp8 <sub>2</sub> )	2.8	7BV1	Yin <i>et al.</i> [28**]
SARS-CoV-2 RdRp in complex with RNA and remdesivir at +1	2.5	7BV2	Yin <i>et al.</i> [28**]
SARS-CoV-2 RdRp in complex with long upstream RNA, post-translocated	2.9	6YYT	Hillen <i>et al.</i> [24**]
SARS-CoV-2 RdRp in complex with RNA, pre-translocated with remdesivir at -1	2.9	7C2K	Wang <i>et al.</i> [29**]
SARS-CoV-2 RdRp in complex RNA, post-translocated	3.3	7BZF	Wang <i>et al.</i> [29**]
SARS-CoV-2 RdRp in complex with RNA and two nsp13	3.5	6XEZ	Chen <i>et al.</i> [49**]
SARS-CoV-2 RdRp in complex with RNA and two nsp13, conformation 1	3.8	7CXN	Yan <i>et al.</i> [50**]
SARS-CoV-2 RdRp in complex with RNA and two nsp13, conformation 2	2.9	7CXM	Yan <i>et al.</i> [50**]
SARS-CoV-2 RdRp in complex with favipiravir, pre-catalytic state	3.2	7CTT	Peng <i>et al.</i> [44*]
SARS-CoV-2 RdRp in complex with favipiravir, pre-catalytic state	2.5	7AAP	Naydenova <i>et al.</i> [43*]
SARS-CoV-2 RdRp in complex with remdesivir at -3, post-translocated state	3.1	7B3B	Kokic <i>et al.</i> [38*]
SARS-CoV-2 RdRp in complex with remdesivir at -4, pre-translocated state	3.4	7B3C	Kokic <i>et al.</i> [38*]
SARS-CoV-2 RdRp in complex with AMP at -4, post-translocated state	2.8	7B3D	Kokic <i>et al.</i> [38*]
SARS-CoV-2 RdRp in complex with remdesivir at +1, -1, -2 to 3, pre-translocated state	3.9	7L1F	Bravo <i>et al.</i> [39*]
SARS-CoV-2 RdRp in complex with two nsp13 and nsp9	2.8	7CYQ	Yan <i>et al.</i> [56**]
SARS-CoV-2 RdRp in complex with suramin	2.6	7D4F	Yin <i>et al.</i> [45*]
Backtracked SARS-CoV-2 RdRp in complex with nsp13	3.4	7KRN	Malone <i>et al.</i> [54*]
Backtracked SARS-CoV-2 RdRp in complex with nsp13 <sub>2</sub>	3.6	7KRO	Malone <i>et al.</i> [54*]
Backtracked SARS-CoV-2 RdRp in complex with nsp13 (local refinement)	3.2	7KRP	Malone <i>et al.</i> [54*]

Figure 1



Structure of the core SARS-CoV RdRp.

(a) Structure of SARS-CoV-1 RdRp (PDB 6NUR) [21<sup>\*\*</sup>]. The structure is shown as cartoon representation, with surface representation overlaid transparently. Coloring: Thumb – red; Palm – orange; Fingers – grey; Interface – purple; NiRAN – firebrick; nsp8 – green; nsp7 – deep blue. Zinc atoms are shown as cyan spheres.

(b) Structure of SARS-CoV-2 RdRp (PDB 6M71) [25<sup>\*\*</sup>]. Coloring and depiction as in A.

residues important for RdRp activity [15], were not resolved in this structure, indicating conformational flexibility.

In April 2020, only months after its discovery, the first structure of the RdRp complex of the novel SARS-CoV-2 virus was determined by cryo-EM (Figure 1b) [25<sup>\*\*</sup>]. Shortly after, two further studies were published that reported similar structures [27,28<sup>\*\*</sup>]. As expected from the high degree of conservation (96.3% sequence identity in nsp12), the SARS-CoV-2 core RdRp is highly similar to that of SARS-CoV-1 and also forms a 1:1:2 complex of nsp12:nsp7:nsp8. Thus, the basic architecture of the RdRp core is conserved across these coronaviruses. The SARS-CoV-2 RdRp structures also reveal previously unresolved N-terminal parts of the NiRAN domain, which form additional interactions with the polymerase domain. However, the N-terminal domains of nsp8 are unresolved also in these structures.

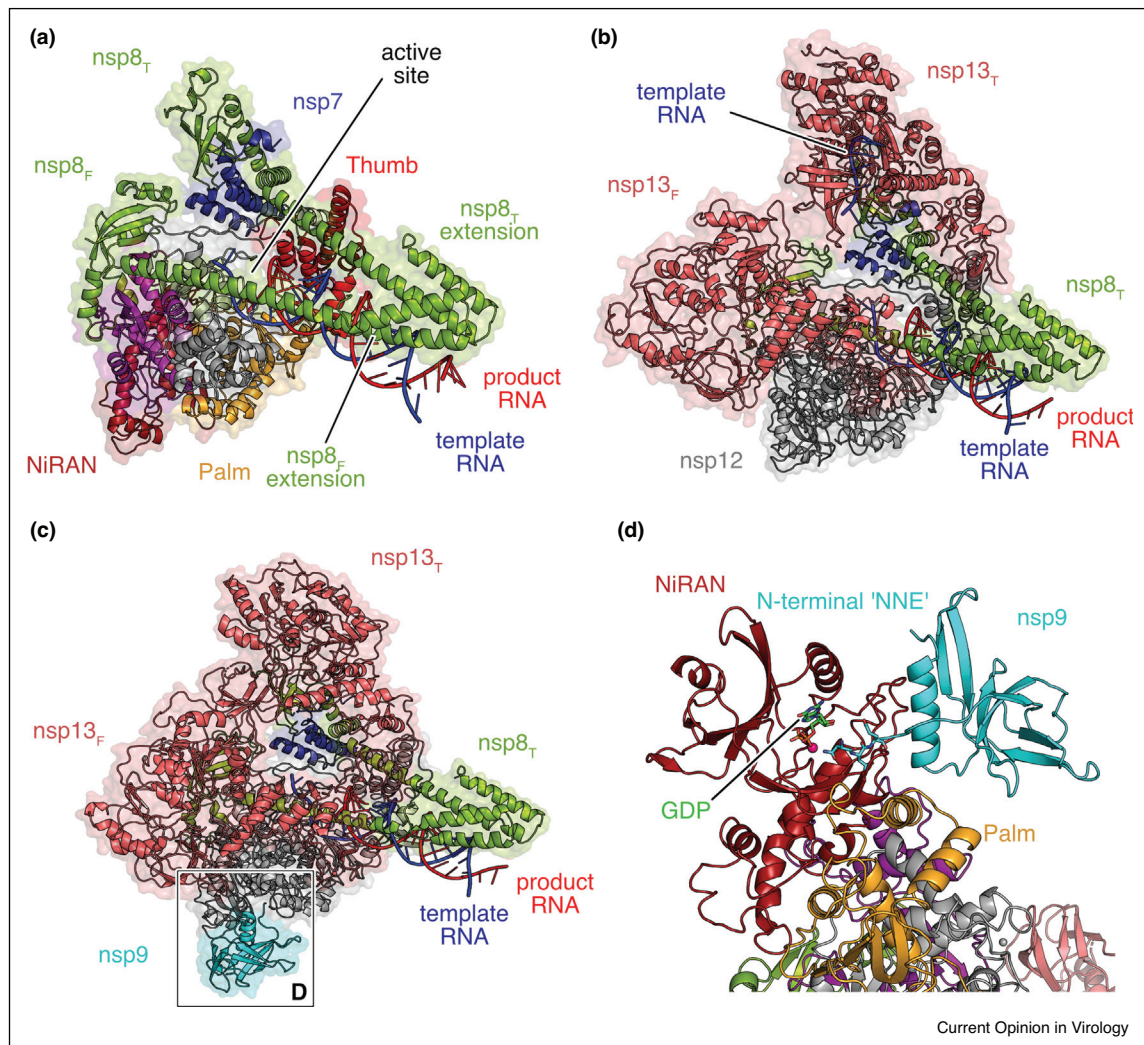
Taken together, cryo-EM studies of the RdRp of both SARS-CoV-1 and SARS-CoV-2 revealed the architecture of the polymerase complex in the absence of nucleic acids (also referred to as the ‘apo RdRp’) and provided first

insights into the interaction between nsp12 and the accessory factors nsp7 and nsp8.

### Structure of replicating coronavirus polymerase

The next important milestone came from three studies that reported the cryo-EM structures of SARS-CoV-2 RdRp in complex with different template and product RNAs, thus revealing its active conformation [24<sup>\*\*</sup>,28<sup>\*\*</sup>,29<sup>\*\*</sup>]. In all three cases, this was achieved by assembling recombinantly produced nsp12, nsp7 and nsp8 with synthetic template-product RNA duplexes, which led to stable complexes that enabled structural characterization. The structures showed that, like other viral RdRps [30,31], the enzyme does not undergo drastic rearrangements upon RNA binding and revealed how the template and product strands are positioned in the active site during RNA synthesis (Figure 2a). The active site is formed by highly conserved residues in nsp12 motifs A–E, which include the catalytic aspartates D760 and D761 that coordinate metal ions required for catalysis. Nsp12 accommodates approximately one turn of duplex RNA, and the first base of the single-stranded portion of the template strand is positioned in the +1 site to direct

Figure 2



Structures of SARS-CoV-2 RdRp functional complexes.

**(a)** Structure of replicating SARS-CoV-2 RdRp with RNA bound (PDB 6YYT) [24\*\*]. Coloring and depiction as in Figure 1. The RNA is shown without surface, with the template strand in blue and the product strand in red.

**(b)** Structure of SARS-CoV-2 RdRp in complex with nsp13 (PDB 7CXN) [50\*\*]. Depiction as in (a). Nsp13 is colored in salmon. The template strand bound to nsp13 is indicated.

**(c)** Structure of SARS-CoV-2 RdRp in complex with nsp13 and nsp9 (PDB 7CYQ) [56\*]. Depiction as in (a). Nsp9 is shown in cyan.

**(d)** Close-up view of the interaction between nsp9 and the NiRAN domain in the structure shown in (c). Proteins are shown as cartoons and colored as in Figures 1 and 2c. The N-terminal 'NNE' motif of nsp9 and bound GDP is shown as sticks.

incorporation of the next substrate nucleotide. In the absence of an incoming nucleoside triphosphate, the polymerase adopts the so-called post-translocated state in which the substrate binding site is empty, as observed previously for enteroviral and polioviral RdRps [30,31].

The structures of actively replicating SARS-CoV-2 RdRp also revealed a surprising feature of the accessory factor nsp8. In one study, the RdRp-RNA complex was assembled using an RNA that forms a long upstream duplex region, extending 28 bp from the active site [24\*\*]. In the

resulting cryo-EM structure, the previously unresolved N-terminal parts of nsp8 are visible and form long helical extensions, similar as in a previous nsp7–nsp8 crystal structure [26\*]. In the RdRp complex, these long helical extensions contact the RNA duplex that emerges during replication, and may thus act as ‘sliding poles’ on the RNA (Figure 2a). This suggests that nsp8 and nsp7 may enhance the processivity of the RdRp by stabilizing its interaction with the RNA. The nsp8 extensions are apparently mobile in the absence of a longer upstream RNA duplex, explaining why other studies that employed

shorter RNAs either could not resolve them [28\*\*] or observed conformational flexibility [29\*\*].

The structures of SARS-CoV-2 RdRp bound to RNA thus confirmed its conserved mechanism of catalysis and provided first structural insights into how the accessory factors nsp7 and nsp8 stimulate its activity.

### Cryo-EM studies unravel the mechanism of antiviral compounds

The structures of active RdRp also paved the way for studies aimed at dissecting the mechanism of antiviral compounds that target coronavirus replication. In particular, the mechanism of remdesivir, an FDA-approved drug for treating COVID-19 [32], has been studied in detail by cryo-EM. Remdesivir is a nucleoside analog that inhibits SARS-CoV-2 RdRp *in vitro* [33,34] and viral replication in cell culture [35]. Biochemical studies showed that it does not act as an immediate chain terminator, but instead leads to delayed polymerase stalling by an unknown mechanism [34,36,37\*]. Two of the studies that reported cryo-EM structures of SARS-CoV-2 RdRp-RNA complexes also investigated how remdesivir is incorporated [28\*\*,29\*\*]. In both, the triphosphate form of remdesivir (RTP) was added during sample preparation to allow for incorporation by the viral enzyme. In one case, remdesivir was incorporated at the 3' end of the product RNA but remained in the substrate binding site, thus representing a pre-translocated state [28\*\*]. In a second study, remdesivir was observed incorporated at the -1 position after the addition of one additional GTP, which remained in the pre-translocated +1 position [29\*\*]. These studies provided the structural basis for remdesivir incorporation into the product RNA and show that the polymerase can add further nucleotides to the nascent chain by translocating remdesivir. Previous data had shown that three more nucleotides can be added to the RNA chain after remdesivir incorporation before further RNA elongation is inhibited [34]. The mechanism of this delayed stalling was investigated by a study that determined cryo-EM structures of RdRp-RNA complexes assembled with RNA containing remdesivir at different positions [38\*]. While remdesivir could be accommodated by the RdRp at position -3, the assembly of the complex with an RNA containing remdesivir at -4 led to a pre-translocated state in which remdesivir was retained at position -3 and the 3' end of the RNA occupied the substrate binding site. The structures suggest that translocation of remdesivir from the -3 to the -4 position would lead to clashes between its cyano group and Ser861 on the nsp12 thumb domain, as had been predicted based on modelling [34]. This was confirmed shortly after by a second cryo-EM study, which also showed that SARS-CoV-2 RdRp stalls in the pre-translocated state after the incorporation of the fourth RTP [39\*]. Thus, the structures show that remdesivir stalls the coronavirus

polymerase by causing a translocation barrier after the addition of three more nucleotides.

The mechanism of binding of another antiviral with clinical potential against COVID-19, favipiravir, was also studied by cryo-EM. Favipiravir acts as a purine analog and inhibits SARS-CoV-2 and other RNA viruses by causing mutagenesis [40–42]. Two cryo-EM structures of SARS-CoV-2 RdRp in the presence of favipiravir revealed the nucleoside bound to the substrate binding site, but not covalently linked to the 3' end of the RNA [43\*,44\*]. These studies showed how favipiravir binds to the SARS-CoV-2 RdRp to mimic GTP, and thus suggest a mechanism for lethal mutagenesis. In addition, these snapshots of the polymerase with an incoming NTP bound provided general insights into substrate recognition by the coronavirus RdRp.

Recently, the first structure of a small molecule that is not a nucleotide analog, suramin, bound to SARS-CoV-2 RdRp was also determined by cryo-EM [45\*]. This drug has been shown to be effective against both parasitic and viral infections [46] and *in vitro* data suggest that it may also inhibit SARS-CoV-2 replication [45\*]. The structure shows that two suramin molecules can bind to SARS-CoV-2 nsp12 and this likely inhibits the polymerase by interfering with binding of the template and product RNA strands. This mechanism of inhibition differs from that previously proposed for norovirus RdRp [47], which can also bind two suramin molecules but at different locations on the polymerase than in the case of SARS-CoV-2.

These studies illustrate that the combination of structural, biochemical and cell culture data can provide a detailed understanding of the mechanism of antiviral compounds and will likely play a key role in the ongoing quest for therapeutic coronavirus RdRp inhibitors.

### Structures of SARS-CoV-2 RdRp in complex with the helicase nsp13

Structural studies have also begun to unravel how the core RdRp interacts with other non-structural proteins to form the RTC [12]. One of them, nsp13, is a dual-function protein with both helicase and triphosphatase activity [18,48]. Nsp13 is required for coronavirus replication and has been proposed to be involved in RNA synthesis and capping, but its precise role remains unclear. Two cryo-EM structures of SARS-CoV-2 RdRp in complex with RNA and nsp13 demonstrate how the helicase interacts with the core RdRp (Figure 2b) [49\*\*,50\*\*]. Two copies of nsp13 can bind to the core nsp12-nsp7-nsp8 RdRp complex on opposing sides of the RNA binding cleft. In both cases, the interaction is mediated by the nidovirus-specific zinc binding domain (ZBD) of nsp13, which interacts with the N-terminal extensions of nsp8<sub>T</sub> or nsp8<sub>F</sub>, respectively. As for nsp8, I here propose a

unifying nomenclature for these two nsp13 copies based on which copy of nsp8 they interact with: Nsp13<sub>T</sub> binds on top of nsp8<sub>T</sub> and forms contacts to the nsp8<sub>T</sub> head domain via its RecA1 domain and to the nsp12 thumb. Nsp13<sub>F</sub> binds on top of nsp8<sub>F</sub> on the other side of the cleft and appears to be less stably bound, as one study reported particles lacking nsp13<sub>F</sub> [49\*\*] and another study observed conformational flexibility [50\*\*]. In addition to the interactions with the RdRp core, the two nsp13 subunits contact each other via their helicase domains, and biochemical experiments suggest that this interaction may be important for helicase activity [50\*\*].

The structures of the RdRp-nsp13 complexes also reveal how nsp13 interacts with the RNA. The single-stranded downstream template RNA binds to the nsp13<sub>T</sub> helicase domain, suggesting that nsp13<sub>T</sub> could translocate on the template strand (Figure 2b). However, nsp13 has been shown to unwind RNA and DNA in the 5' to 3' direction [51,52], which would oppose the direction in which the RdRp translocates on the template strand during synthesis. This has led to the suggestion that nsp13 could facilitate backtracking of the coronavirus polymerase on the RNA [49\*\*,50\*\*]. Backtracking is a well-characterized feature of the evolutionarily unrelated bacterial and eukaryotic multi-subunit RNA polymerases, which facilitates rescue of stalled elongation complexes and removal of misincorporated nucleotides [53]. The hypothesis that coronavirus RdRps may backtrack is supported by a recent preprint, in which the authors report the structure of a backtracked SARS-CoV-2 RdRp-nsp13 complex assembled on a RNA scaffold with a five nucleotide mismatch at the 3' end of the product strand [54\*]. In this structure, the mismatched segment extrudes through the substrate entry channel of nsp12, demonstrating a conserved topology between backtracked viral single-subunit and cellular multi-subunit polymerases. Biochemical data in this study further suggest that nsp13 stimulates RdRp backtracking. Nsp13-mediated backtracking of the coronavirus RdRp may thus facilitate proofreading or template-switching during nested mRNA biogenesis by exposing the 3' end of the nascent RNA from the polymerase active site [49\*\*,50\*\*,54\*\*].

Together, the structures of nsp13-bound RTCs showed that nsp7 and nsp8 provide an interaction platform to which additional factors, such as nsp13, can bind. In addition, they suggest that nsp13 may enable backtracking of the RdRp complex, which may be required for co-transcriptional processes such as proofreading or template switching.

### The role of the NiRAN domain

A unique feature of the nidovirus RdRp is the NiRAN domain of nsp12. The discovery that this domain can

catalyze the transfer of nucleoside monophosphates (NMP) has led to speculations that it may be involved in RNA ligation, protein-mediated priming of RNA synthesis or transcript capping [22]. However, it is not clear whether the target substrate of nucleotidyl transfer is another protein or the RNA, and the precise role of this domain thus remains enigmatic.

The recent cryo-EM structures of SARS-CoV-2 RdRp provide structural evidence for the nucleotidyl-transferase activity of the NiRAN domain. First, the NiRAN domain resembles SelO, a conserved bacterial pseudokinase that transfers AMP to protein residues [21\*\*,49\*\*,55]. Second, structures of SARS-CoV-2 RdRp determined in the presence of ADP-BeF<sub>3</sub> [49\*\*] or GDP-BeF<sub>3</sub> [56\*\*] showed density for ADP or GTP, respectively, bound to the NiRAN domain, demonstrating that it can indeed bind nucleotides. A recent cryo-EM structure of SARS-CoV-2 RdRp in complex with nsp13 and nsp9 further reveals that the NiRAN domain interacts with nsp9 (Figure 2c) [56\*\*]. Nsp9 is a putative RNA-binding protein that forms homodimers [57–59] and is essential for viral replication [60]. It has been suggested to interact with the RTC [61], but its molecular function remains unclear. The structure shows that nsp9 binds to the NiRAN domain as a monomer, and that its N-terminus extends into the putative active site of the NiRAN domain, where GDP is bound (Figure 2d). Biochemical data in this study indicate that the NiRAN domain may catalyze transfer of GMP to the 5' end of RNA in a nsp13-dependent manner, which appears to be inhibited by nsp9. This led the authors to hypothesize that the NiRAN domain may act as the GTase required for transcript capping and that nsp9 could act as a regulatory or structural factor [56\*\*]. In contrast, a recent biochemical study suggests that the NiRAN domain catalyzes uridylation of the primary amine at the N-terminus of nsp9, lending support to a role for the NiRAN domain in protein-mediated RNA priming [62]. This activity is dependent on a 'NNE' motif at the nsp9 N-terminus, which is highly conserved among coronaviruses. Even though the authors of the nsp9-bound RdRp structure did not observe nucleotidylation of nsp9, it appears remarkably plausible based on the structure, as the first three residues of nsp9 reach into the NiRAN domain and the N-terminal aspartate is positioned within a few Ångströms of the bound nucleotide (Figure 2d) [56\*\*]. The observation that this activity is highly sensitive to the sequence at the nsp9 N-terminus [62] exemplifies the necessity for constructs with 'natural' termini for studying recombinant viral proteins *in vitro*.

Taken together, both functional and structural studies suggest that the nsp12 NiRAN domain acts as a nucleotidyl-transferase, but whether it is involved in transcript capping, protein-mediated priming or some other replication-associated process remains to be determined.

## Open questions & future directions

The unprecedented pace at which structural information on the coronavirus RdRp and associated non-structural proteins has been determined in the last months has led to a leap in our molecular understanding of coronavirus replication. However, many fundamental questions remain open. In particular, how the various non-structural proteins associate with the core RdRp to form the RTC and carry out processes such as capping and proofreading remains to be determined. Similarly, the mechanisms by which the polymerase initiates RNA synthesis and how it switches templates during nested mRNA synthesis are not known. Structural analysis will likely continue to contribute to answering these questions, as exemplified by the recent structures of SARS-CoV-2 RdRp in complex with nsp13 and nsp9. The structural and mechanistic characterization of the SARS-CoV-2 RTC may ultimately contribute to the development of new treatment options for coronavirus infections, as replication-associated processes like proofreading or capping are promising therapeutic targets. Such treatment options would not only be desperately needed during the current pandemic but also have the potential to be effective against novel emerging coronaviruses.

## Author contributions

H.S.H. prepared figures and wrote the manuscript.

## Conflict of interest statement

Nothing declared.

## Acknowledgements

I wish to thank members of the Cramer laboratory of the MPI-BPC, in particular P. Cramer, G. Kokic, L. Farnung, J. Schmitzova, C. Dienemann, D. Tegunov and F. Kabinger, for insightful discussions. H.S.H. was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) (FOR2848, SFB1190 and under Germany's Excellence Strategy — EXC 2067/1-390729940).

## References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
  - of outstanding interest
1. Hilgenfeld R, Peiris M: **From SARS to MERS: 10 years of research on highly pathogenic human coronaviruses.** *Antivir Res* 2013, **100**:286-295.
  2. Ogimi C, Kim YJ, Martin ET, Huh HJ, Chiu C-H, Englund JA: **What's new with the old coronaviruses?** *J Pediatric Infect Dis Soc* 2020, **9**:210-217.
  3. Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, Hu Y, Tao Z-W, Tian J-H, Pei Y-Y *et al.*: **A new coronavirus associated with human respiratory disease in China.** *Nature* 2020, **579**:265-269.
  4. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, Si H-R, Zhu Y, Li B, Huang C-L *et al.*: **A pneumonia outbreak associated with a new coronavirus of probable bat origin.** *Nature* 2020, **579**:270-273.
  5. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM: **Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia.** *N Engl J Med* 2012, **367**:1814-1820.
  6. Peiris J, Lai S, Poon L, Guan Y, Yam L, Lim W, Nicholls J, Yee W, Yan W, Cheung M *et al.*: **Coronavirus as a possible cause of severe acute respiratory syndrome.** *Lancet* 2003, **361**:1319-1325.
  7. Drosten C, Günther S, Preiser W, van der Werf S, Brodt H-R, Becker S, Rabenau H, Panning M, Kolesnikova L, Fouchier RAM *et al.*: **Identification of a novel coronavirus in patients with severe acute respiratory syndrome.** *N Engl J Med* 2003, **348**:1967-1976.
  8. Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, Peñaranda S, Bankamp B, Maher K, Chen M *et al.*: **Characterization of a novel coronavirus associated with severe acute respiratory syndrome.** *Science* 2003, **300**:1394-1399.
  9. Dong E, Du H, Gardner L: **An interactive web-based dashboard to track COVID-19 in real time.** *Lancet Infect Dis* 2020, **20**:533-534.
  10. Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, Wang H, Cramer G, Hu Z, Zhang H *et al.*: **Bats are natural reservoirs of SARS-like coronaviruses.** *Science* 2005, **310**:676-679.
  11. Masters PS: **The molecular biology of coronaviruses.** *Adv Virus Res* 2006, **66**:193-292.
  12. Snijder EJ, Decroly E, Ziebuhr J: **The nonstructural proteins directing coronavirus RNA synthesis and processing.** *Adv Virus Res* 2016, **96**:59-126.
  13. Sawicki SG, Sawicki DL, Siddell SG: **A contemporary view of coronavirus transcription.** *J Virol* 2006, **81**:20-29.
  14. Cheng A, Zhang W, Xie Y, Jiang W, Arnold E, Sarafianos SG, Ding J: **Expression, purification, and characterization of SARS coronavirus RNA polymerase.** *Virology* 2005, **335**:165-176.
  15. Subissi L, Posthuma CC, Collet A, Zevenhoven-Dobbe JC, Gorbalenya AE, Decroly E, Snijder EJ, Canard B, Imbert I: **One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities.** *Proc Natl Acad Sci U S A* 2014, **111**:E3900-E3909.
  16. Ferron F, Decroly E, Selisko B, Canard B: **The viral RNA capping machinery as a target for antiviral drugs.** *Antivir Res* 2012, **96**:21-31.
  17. Ghosh A, Lima CD: **Enzymology of RNA cap synthesis.** *Wiley Interdiscip Rev RNA* 2010, **1**:152-172.
  18. Ivanov KA, Ziebuhr J: **Human coronavirus 229E nonstructural protein 13: characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5'-triphosphatase activities.** *J Virol* 2004, **78**:7833-7838.
  19. Bouvet M, Debarnot C, Imbert I, Selisko B, Snijder EJ, Canard B, Decroly E: **In vitro reconstitution of SARS-coronavirus mRNA cap methylation.** *PLoS Pathog* 2010, **6**:e1000863.
  20. Chen Y, Cai H, Pan J, Xiang N, Tien P, Ahola T, Guo D: **Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase.** *Proc Natl Acad Sci U S A* 2009, **106**:3484-3489.
  21. Kirchdoerfer RN, Ward AB: **Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors.** *Nat Commun* 2019, **10**:2342.
- This paper describes the first coronavirus RdRp structure.
22. Lehmann KC, Gulyaeva A, Zevenhoven-Dobbe JC, Janssen GMC, Ruben M, Overkleef HS, van Veelen PA, Samborskiy DV, Kravchenko AA, Leontovich AM *et al.*: **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* 2015, **43**:8416-8434.
  23. Peersen OB: **A comprehensive superposition of viral polymerase structures.** *Viruses* 2019, **11**:745.
  24. Hillen HS, Kokic G, Farnung L, Dienemann C, Tegunov D, Cramer P: **Structure of replicating SARS-CoV-2 polymerase.** *Nature* 2020, **584**:154-156.



This paper describes one of the first two structures of the RNA-bound SARS-CoV-2 RdRp and for the first time revealed the nsp8 N-terminal extensions.

25. Gao Y, Yan L, Huang Y, Liu F, Zhao Y, Cao L, Wang T, Sun Q, ●● Ming Z, Zhang L *et al.*: **Structure of the RNA-dependent RNA polymerase from COVID-19 virus.** *Science* 2020, **368**:779-782.  
This paper describes the first structure of the SARS-CoV-2 apo RdRp.
26. Zhai Y, Sun F, Li X, Pang H, Xu X, Bartlam M, Rao Z: **Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer.** *Nat Struct Mol Biol* 2005, **12**:980-986.  
This paper describes the structure of a SARS-CoV-1 nsp7-nsp8 hexadecamer.
27. Peng Q, Peng R, Yuan B, Zhao J, Wang M, Wang X, Wang Q, Sun Y, Fan Z, Qi J *et al.*: **Structural and biochemical characterization of the nsp12-nsp7-nsp8 core polymerase complex from SARS-CoV-2.** *Cell Rep* 2020, **31**:107774.
28. Yin W, Mao C, Luan X, Shen D-D, Shen Q, Su H, Wang X, Zhou F, ●● Zhao W, Gao M *et al.*: **Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir.** *Science* 2020, **368**:1499-1504.  
This paper describes one of the first two structures of the RNA-bound SARS-CoV-2 RdRp and showed how remdesivir is incorporated.
29. Wang Q, Wu J, Wang H, Gao Y, Liu Q, Mu A, Ji W, Yan L, Zhu Y, ●● Zhu C *et al.*: **Structural basis for RNA replication by the SARS-CoV-2 polymerase.** *Cell* 2020, **182**:417-428.e13.  
This paper describes the structure of RNA-bound SARS-CoV-2 RdRp with remdesivir at +1.
30. Gong P, Peersen OB: **Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase.** *Proc Natl Acad Sci U S A* 2010, **107**:22505-22510.
31. Shu B, Gong P: **Structural basis of viral RNA-dependent RNA polymerase catalysis and translocation.** *Proc Natl Acad Sci U S A* 2016, **113**:E4005-E4014.
32. Ledford H: **Hopes rise for coronavirus drug remdesivir.** *Nature* 2020 <http://dx.doi.org/10.1038/d41586-020-01295-8>. ISSN 1476-4687 (online).
33. Agostini ML, Andres EL, Sims AC, Graham RL, Sheahan TP, Lu X, Smith EC, Case JB, Feng JY, Jordan R *et al.*: **Coronavirus susceptibility to the antiviral remdesivir (GS-5734) is mediated by the viral polymerase and the proofreading exoribonuclease.** *mBio* 2018, **9**:e00221-18.
34. Gordon CJ, Tchesnokov EP, Woolner E, Perry JK, Feng JY, Porter DP, Götte M: **Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency.** *J Biol Chem* 2020, **295**:6785-6797.
35. Pruijssers AJ, George AS, Schäfer A, Leist SR, Gralinski LE, Dinno KH, Yount BL, Agostini ML, Stevens LJ, Chappell JD *et al.*: **Remdesivir inhibits SARS-CoV-2 in human lung cells and chimeric SARS-CoV expressing the SARS-CoV-2 RNA polymerase in mice.** *Cell Rep* 2020, **32**:107940.
36. Tchesnokov EP, Feng JY, Porter DP, Götte M: **Mechanism of inhibition of Ebola virus RNA-dependent RNA polymerase by remdesivir.** *Viruses* 2019, **11**:326.
37. Gordon CJ, Tchesnokov EP, Feng JY, Porter DP, Götte M: **The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus.** *J Biol Chem* 2020, **295**:4773-4779 <http://dx.doi.org/10.1074/jbc.ac120.013056>.  
This paper shows biochemically that remdesivir stalls the MERS RdRp after incorporation of three additional nucleotides.
38. Kocic G, Hillen HS, Tegunov D, Dienemann C, Seitz F, ●● Schmitzova J, Farnung L, Siewert A, Höbartner C, Cramer P: **Mechanism of SARS-CoV-2 polymerase stalling by remdesivir.** *Nat Commun* 2021, **12**:279.  
This paper, together with Ref. 39, provides the structural basis of SARS-CoV-2 RdRp stalling by remdesivir.
39. Bravo JPK, Dangerfield TL, Taylor DW, Johnson KA: **Remdesivir is a delayed translocation inhibitor of SARS CoV-2 replication in vitro.** *bioRxiv* 2020 <http://dx.doi.org/10.1101/2020.12.14.422718>. 2020.12.14.422718.

This paper, together with Ref. 38, provides the structural basis of SARS-CoV-2 RdRp stalling by remdesivir.

40. Arias A, Thorne L, Goodfellow I: **Favipiravir elicits antiviral mutagenesis during virus replication in vivo.** *eLife* 2014, **3**:e03679.
41. Shannon A, Selisko B, Le N-T-T, Huchting J, Touret F, Piorkowski G, Fattorini V, Ferron F, Decroly E, Meier C *et al.*: **Rapid incorporation of favipiravir by the fast and permissive viral RNA polymerase complex results in SARS-CoV-2 lethal mutagenesis.** *Nat Commun* 2020, **11**:4682.
42. Sangawa H, Komeno T, Nishikawa H, Yoshida A, Takahashi K, Nomura N, Furuta Y: **Mechanism of action of T-705 ribosyl triphosphate against influenza virus RNA polymerase.** *Antimicrob Agents Chemother* 2013, **57**:5202-5208.
43. Naydenova K, Muir KW, Wu L-F, Zhang Z, Coscia F, Peet MJ, ●● Castro-Hartmann P, Qian P, Sader K, Dent K *et al.*: **Structural basis for the inhibition of the SARS-CoV-2 RNA-dependent RNA polymerase by favipiravir-RTP.** *bioRxiv* 2020 <http://dx.doi.org/10.1101/2020.10.21.347690>. 2020.10.21.347690.  
This paper, together with Ref. 44, shows how favipiravir binds to the SARS-CoV-2 RdRp.
44. Peng Q, Peng R, Yuan B, Wang M, Zhao J, Fu L, Qi J, Shi Y: ●● **Structural basis of SARS-CoV-2 polymerase inhibition by favipiravir.** *Innovation* 2021, **2**:100080 <http://dx.doi.org/10.1016/j.xinn.2021.100080> ISSN 2666-6758.  
This paper, together with Ref. 43, shows how favipiravir binds to the SARS-CoV-2 RdRp.
45. Yin W, Luan X, Li Z, Zhou Z, Wang Q, Gao M, Wang X, Zhou F, ●● Shi J, You E *et al.*: **Structural basis for inhibition of the SARS-CoV-2 RNA polymerase by suramin.** *Nat Struct Mol Biol* 2021, **28**:319-325.  
This paper describes the structure of suramin-bound SARS-CoV-2 apo RdRp.
46. Wiedemar N, Hauser DA, Mäser P: **100 years of suramin.** *Antimicrob Agents Chemother* 2019, **64**.
47. Mastrangelo E, Pezzullo M, Tarantino D, Petazzi R, Germani F, Kramer D, Robel I, Rohayem J, Bolognesi M, Milani M: **Structure-based inhibition of norovirus RNA-dependent RNA polymerases.** *J Mol Biol* 2012, **419**:198-210.
48. Ivanov KA, Thiel V, Dobbe JC, van der Meer Y, Snijder EJ, Ziebuhr J: **Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase.** *J Virol* 2004, **78**:5619-5632.
49. Chen J, Malone B, Llewellyn E, Grasso M, Shelton PMM, ●● Olinares PDB, Maruthi K, Eng ET, Vatandaslar H, Chait BT *et al.*: **Structural basis for helicase-polymerase coupling in the SARS-CoV-2 replication-transcription complex.** *Cell* 2020, **182**:1560-1573.e13 <http://dx.doi.org/10.1016/j.cell.2020.07.033> ISSN 0092-8674.  
This paper describes the first structure of the SARS-CoV-2 RdRp with nsp13 bound.
50. Yan L, Zhang Y, Ge J, Zheng L, Gao Y, Wang T, Jia Z, Wang H, ●● Huang Y, Li M *et al.*: **Architecture of a SARS-CoV-2 mini replication and transcription complex.** *Nat Commun* 2020, **11**:5874.  
This paper describes the structure of SARS-CoV-2 RdRp complexes with nsp13 bound; see also Ref. 49●●.
51. Seybert A, Hegyi A, Siddell SG, Ziebuhr J: **The human coronavirus 229E superfamily 1 helicase has RNA and DNA duplex-unwinding activities with 5'-to-3' polarity.** *RNA* 2000, **6**:1056-1068.
52. Tanner JA, Watt RM, Chai Y-B, Lu L-Y, Lin MC, Peiris JSM, Poon LLM, Kung H-F, Huang J-D: **The severe acute respiratory syndrome (SARS) coronavirus NTPase/helicase belongs to a distinct class of 5' to 3' viral helicases.** *J Biol Chem* 2003, **278**:39578-39582.
53. Nudler E: **RNA polymerase backtracking in gene regulation and genome instability.** *Cell* 2012, **149**:1438-1445.
54. Malone B, Chen J, Wang Q, Llewellyn E, Choi YJ, Olinares PDB, ●● Cao X, Hernandez C, Eng ET, Chait BT *et al.*: **Structural basis for backtracking by the SARS-CoV-2 replication-transcription**

**complex.** *bioRxiv* 2021 <http://dx.doi.org/10.1101/2021.03.13.435256>. 2021.03.13.435256.

This paper describes the structure of a backtracked SARS-CoV-2 RdRp-nsp13 complex.

55. Sreelatha A, Yee SS, Lopez VA, Park BC, Kinch LN, Pilch S, Servage KA, Zhang J, Jiou J, Karasiewicz-Urbańska M *et al.*: **Protein AMPylation by an evolutionarily conserved pseudokinase.** *Cell* 2018, **175**:809-821.e19.
56. Yan L, Ge J, Zheng L, Zhang Y, Gao Y, Wang T, Huang Y, Yang Y, Gao S, Li M *et al.*: **Cryo-EM structure of an extended SARS-CoV-2 replication and transcription complex reveals an intermediate state in cap synthesis.** *Cell* 2021, **184**:184-193.e10.
- This paper describes the structure of the SARS-CoV-2 RdRp with nsp13 and nsp9 bound.
57. Egloff M-P, Ferron F, Campanacci V, Longhi S, Rancurel C, Dutartre H, Snijder EJ, Gorbalenya AE, Cambillau C, Canard B: **The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world.** *Proc Natl Acad Sci U S A* 2004, **101**:3792-3796.

58. Sutton G, Fry E, Carter L, Sainsbury S, Walter T, Nettleship J, Berrow N, Owens R, Gilbert R, Davidson A *et al.*: **The nsp9 replicase protein of SARS-coronavirus, structure and functional insights.** *Structure* 2004, **12**:341-353.
59. Littler DR, Gully BS, Colson RN, Rossjohn J: **Crystal structure of the SARS-CoV-2 non-structural protein 9, Nsp9.** *Iscience* 2020, **23**:101258.
60. Frieman M, Yount B, Agnihothram S, Page C, Donaldson E, Roberts A, Vogel L, Woodruff B, Scorpio D, Subbarao K *et al.*: **Molecular determinants of severe acute respiratory syndrome coronavirus pathogenesis and virulence in young and aged mouse models of human disease.** *J Virol* 2012, **86**:884-897.
61. von Brunn A, Teepe C, Simpson JC, Pepperkok R, Friedel CC, Zimmer R, Roberts R, Baric R, Haas J: **Analysis of intraviral protein-protein interactions of the SARS coronavirus ORFome.** *PLoS One* 2007, **2**:e459.
62. Slanina H, Madhugiri R, Bylapudi G, Schultzei K, Karl N, Gulyaeva A, Gorbalenya AE, Linne U, Ziebuhr J: **Coronavirus replication-transcription complex: vital and selective NMPylation of a conserved site in nsp9 by the NiRAN-RdRp subunit.** *Proc Natl Acad Sci U S A* 2021, **118**:e2022310118.