Nucleic Acids Research, 2025, **53**, gkaf450
https://doi.org/10.1093/nar/gkaf450
Nucleic Acid Enzymes

A post-assembly conformational change makes the SARS-CoV-2 polymerase elongation-competent

Misha Klein[1,†], Arnab Das[1,†], Subhas C. Bera[2,†], Thomas K. Anderson[3], Dana Kocincova[4], Hery W. Lee[4], Bing Wang[5], Flavia S. Papini[2], John C. Marecki[6], Jamie J. Arnold[7], Craig E. Cameron[7], Kevin D. Raney[6], Irina Artsimovitch[5], Mathias Götte[4], Robert N. Kirchdoerfer[3], Martin Depken[8,*], David Dulin[1,2,*]

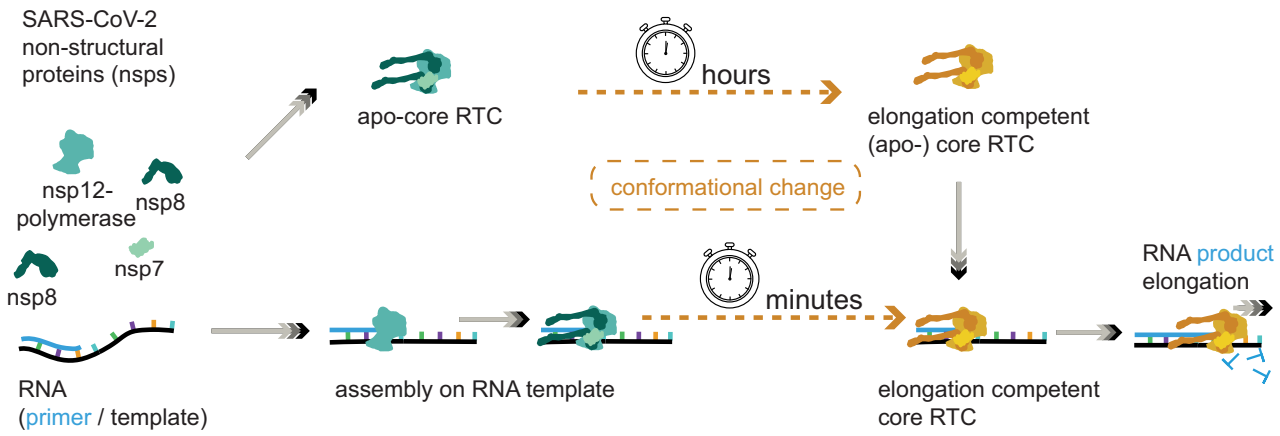
[1]Department of Physics and Astronomy, and LaserLaB Amsterdam, Vrije Universiteit Amsterdam, De Boelelaan 1100, 1081 HZ Amsterdam, The Netherlands
[2]Junior Research Group 2, Interdisciplinary Center for Clinical Research, Friedrich-Alexander-University Erlangen-Nürnberg (FAU), Cauerstr. 3, 91058 Erlangen, Germany
[3]Department of Biochemistry and Institute for Molecular Virology, University of Wisconsin-Madison, Madison, WI 53706, United States
[4]Department of Medical Microbiology and Immunology, University of Alberta, Edmonton, Alberta T6G 2R3, Canada
[5]Department of Microbiology and The Center for RNA Biology, The Ohio State University, Columbus, OH 43210, United States
[6]Department of Biochemistry and Molecular Biology, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, AR 72205, United States
[7]Department of Microbiology and Immunology, University of North Carolina School of Medicine, Chapel Hill, NC 27599, United States
[8]Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Van der Maasweg 9, 2629 HZ Delft, The Netherlands

[*]To whom the correspondence may be addressed. Email. d.dulin@vu.nl
Correspondence may also be addressed to Martin Depken. Email: s.m.depken@tudelft.nl
[†]These authors contributed equally to the present work.

Abstract

Coronaviruses (CoVs) encode 16 nonstructural proteins (nsps), most of which form the replication–transcription complex (RTC). The RTC contains a core composed of one nsp12 RNA-dependent RNA polymerase (RdRp), two nsp8s, and one nsp7. The core RTC recruits other nsps to synthesize all viral RNAs within the infected cell. While essential for viral replication, the mechanism by which the core RTC assembles into a processive polymerase remains poorly understood. We show that the core RTC preferentially assembles by first having nsp12-polymerase bind to the RNA template, followed by the subsequent association of nsp7 and nsp8. Once assembled on the RNA template, the core RTC requires hundreds of seconds to undergo a conformational change that enables processive elongation. In the absence of RNA, the (apo-)RTC requires several hours to adopt its elongation-competent conformation. We propose that this obligatory activation step facilitates the recruitment of additional nsps essential for efficient viral RNA synthesis and may represent a promising target for therapeutic interventions.

Graphical abstract



Received: October 1, 2024. **Revised:** April 14, 2025. **Editorial Decision:** May 6, 2025. **Accepted:** May 16, 2025

Introduction

The COVID-19 pandemic is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) with devastating impacts on public health that are still ongoing (WHO: https://data.who.int/dashboards/covid19/cases?n=c). Novel CoVs may also emerge in the future, as the present pandemic comes close on the heels of the middle east respiratory syndrome coronavirus (MERS-CoV) in 2012 and SARS-CoV-1 in 2002. Viral genome replication and transcription are important targets for therapeutic interventions due to their conserved nature across CoVs, and treatments based on nucleotide analogs, such as remdesivir [1] and molnupiravir [2], have proven to be effective. However, CoVs have a great ability to evolve and adapt, either via mutation or recombination [3–5]. Having a precise understanding of the molecular determinants in CoV replication and transcription will help design future antiviral drugs to counter such adaptation [6–12].

CoVs are positive (+)RNA viruses with a ∼30 kb long single-stranded (ss) genome. The 5′-proximal two-thirds encode for the sixteen nonstructural proteins (nsp(s)) [13]. These nsps are encoded on two open reading frames separated by a ribosomal frameshifting sequence within the nsp12-polymerase gene. Translation produces two polyproteins: one spanning from nsp1 to nsp11, and the other from nsp1 to nsp16. These polyproteins are subsequently processed into individual nsps by two proteases—nsp3 (papain-like protease) and nsp5 (3C-like protease). Notably, nsp5 processes all nsps from nsp4 to nsp16 [14]. Most of the nsps associate to form the replication–transcription complex (RTC), which synthesizes all viral RNAs in the infected cell [15, 16]. The minimal complex capable of RNA synthesis consists of the nsp12-polymerase associated with nsp7 and nsp8 in a 1:1:2 stoichiometry (hereafter "core RTC") [10, 17–20]. Nsp7 and nsp8 are required for *in vitro* RNA primer-extension by the core RTC [21–24].

Previous studies have shown that it takes several minutes for the CoV core RTC to perform detectable primer-extension in *in vitro* bulk assays [23, 24]. This is particularly striking, as the SARS-CoV-2 core RTC is the fastest RNA polymerase reported to date, with a nucleotide addition rate up to ∼180 nt/s at 37°C [11]. The apparent lag time in primer-extension assays, therefore, seems to stem from a step prior to RNA synthesis and has been proposed to arise from a stepwise assembly of the individual nsps into the core RTC [22, 23, 25]. However, no quantitative assessment of different contributors to the lag times has been reported to date. Characterizing the mechanism through which nsp7 and nsp8 regulate primer-extension by the core RTC requires a more in-depth comparison of their influence, both before (hereafter "activation phase") and during (hereafter "elongation phase") RNA synthesis (Fig. 1A).

We set out to understand the path the CoV core RTC proteins take towards polymerase activity by using a high-throughput single-molecule magnetic tweezers assay [26] in conjunction with kinetic modelling. Previously, we used the same assay to study the elongation phase of the SARS-CoV-2 core RTC in-depth [26]. Here, we used our ability to detect the onset of elongation from individual RNA primer-templates to differentiate between the time spent by single core RTCs on activation and elongation. Our data and kinetic model reveal that a core RTC is not in an elongation-competent conformer upon assembly. As such, the core RTC must undergo a conformational change that allows it to enter the elongation phase. We start by confirming that a core RTC (reconstituted from individually expressed proteins) exhibits a considerable lag-time prior to elongation [23, 24]. The lag time does not stem from the time required to assemble the complex since activation times remain substantial at saturating protein concentrations. Instead, nsp7, nsp8, and the template RNA appear to greatly limit the conformational space to be explored by the core RTC, allowing it to adopt its elongation-competent conformer within minutes. The conformational change occurs prior to the elongation phase as factors influencing the kinetics of the activation phase have no effect on the elongation phase. Finally, even in the presence of nsp13-helicase an average activation time of several hundreds of seconds remains. Our results reveal a mandatory step towards assembling a complete and functional CoV-RTC, opening new avenues for antiviral drug development.

Materials and methods

Purification and recombinant protein expression of nsp7 and nsp8 from SARS-CoV-2

This protocol was described in detail in [6, 26]. The SARS-CoV-2 nsp7 and nsp8 genes were cloned into pET46 (Novagen) with an N-terminal 6 histidine tag and a TEV protease site for purification. The proteins were recombinantly expressed in Rosetta2 pLys *Escherichia coli* (Novagen) and purified using Ni-NTA agarose beads. Nsp7 and nsp8 proteins were further purified by size-exclusion chromatography. Purified proteins were dialyzed into storage buffer (10 mM HEPES, pH 7.4, 300 mM NaCl, and 2 mM dithiothreitol (DTT)), aliquoted, and stored at −80°C.

Purification and recombinant protein expression of nsp12-polymerase in *Sf9* cells

This protocol was described in detail in [6, 26]. The SARS-CoV-2 nsp12-polymerase gene was cloned into pFastBac with a TEV site and strep tag (Genscript) for efficient purification. The bacmid was created in DH10Bac *E. coli* (Life Technologies) and amplified in *Sf9* cells (Expression Systems) with Cellfectin II (Life Technologies) to generate recombinant baculovirus. Sf21 cells were infected, and the cell lysate cleared by centrifugation and filtration. The protein was purified via strep-tactin agarose and further polished in size-exclusion chromatography. Purified protein was dialyzed into storage buffer (10 mM HEPES, pH 7.4, 300 mM NaCl, 0.1 mM $MgCl_2$, and 2 mM Tris(2-carboxyethyl)phosphine (TCEP)), aliquoted, and stored at −80°C.

Purification and recombinant protein expression of nsp12-polymerase in *E. coli*

The protocol used to obtain nsp12-RdRp used for these experiments was described in detail in [27]. Nsp12-polymerase was overexpressed in *E. coli* BL21(DE3) cells (Novagen). The cleared lysate was applied to Ni21-NTA resin (Cytiva) and the eluted protein was further purified by anion exchange chromatography and size-exclusion chromatography. Purified protein was dialyzed into storage buffer (20 mM HEPES, pH 7.5, 150 mM KCl, 45% glycerol, 1 mM $MgCl_2$, and 1 mM DTT), aliquoted, and stored at −80°C.

Purification of co-translated SARS-CoV-2 core RTC

The protocol was described in detail in [12]. The pFastBac-1 (Invitrogen, Burlington, ON, Canada) plasmid with the codon-optimized synthetic DNA sequences (GenScript, Piscataway, NJ, USA) coding for SARS-CoV-2 (NCBI: QHD43415.1) nsp5, nsp7, nsp8, and nsp12 were used as a starting material for protein expression in insect cells (*Sf9*, Invitrogen, Waltham, MA, USA) [8, 9, 12, 28]. We employed the MultiBac (Geneva Biotech, Indianapolis, IN, USA) system for protein expression in insect cells (Sf9, Invitrogen) according to published protocols [29, 30]. SARS-CoV-2 protein complexes were purified using nickel-nitrilotriacetic acid affinity chromatography of the nsp8 N-terminal 8-histidine tag according to the manufacturer's specifications (Thermo Fisher Scientific, Waltham, MA, USA). Purified protein was dialyzed into storage buffer [100 Tris–HCl (pH 8), 1000 mM NaCl, 5 mM TCEP, 0.01% Tween-20, 200 mM imidazole, and 40% glycerol], aliquoted, and stored at −20°C.

Purification and recombinant protein expression of nsp13-helicase

The coding sequence for nsp13 from the SARS CoV-2 Washington isolate (Genbank MN985325) was synthesized as an *E. coli* codon-optimized fragment (GenScript, Piscataway NJ) and cloned into the *Bsa*I site of the pSUMO plasmid (LifeSensors, Malvern, PA) to produce an *N*-terminal six histidine-tagged SUMO-nsp13 fusion cassette (6XHis-SUMO-nsp13). Final plasmids were sequence-verified through the UAMS Sequencing Core Facility using a 3130XL Genetic Analyzer (Applied Biosystems, Foster City, CA). The SUMO-nsp13 construct was transformed into Rosetta2 cells, and colonies were grown overnight at 37°C in NZCYM (Research Products International, Mount Prospect, IL) supplemented with kanamycin (50 μg/ml) and chloramphenicol (25 μg/ml). The cultures were diluted 1:100 into fresh antibiotic-containing NZCYM media and grown to an $OD_{600\,nm}$ of 0.8–1. The bacterial media was supplemented with 0.1 mM $ZnSO_4$ and 0.2% dextrose and cooled on ice for 10 min. Protein expression was induced with 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 18°C for 12–16 h. The cells were harvested by centrifugation at 4000 × *g* for 15 min at 4°C, and pellets stored at −80°C. All purifications steps were carried out on ice or at 4°C. Pellets were resuspended in lysis buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 1 mM β-mercaptoethanol, 10% glycerol, and 20 mM imidazole) supplemented with 2 mM phenylmethylsulfonyl fluoride (PMSF) and 1× EDTA-free protease inhibitor cocktail (Pierce). Bacteria were lysed by microfluidization and the lysate clarified by centrifugation at 17 000 × *g* for 1 h at 4°C. The His-tagged SUMO-nsp13 was passed through a HisTrap FF column (Cytiva) equilibrated in lysis buffer at 1 ml/min using a Cytiva Akta FPLC. The affinity resin was washed with 20 column volumes of lysis buffer, and the protein eluted with 10 column volumes of lysis buffer containing 200 mM imidazole. The pooled SUMO-nsp13-containing fractions were dialyzed overnight into two changes of 20 mM imidazole-containing lysis buffer, and the SUMO tag cleaved with ULP-1 for 4 h at 4°C. Digestion was confirmed by SDS–PAGE analysis. The His6-ULP-1 and His6-SUMO proteins were separated from the native nsp13 with a second round of Ni^{2+}-affinity chromatography as before. Nsp13-containing flow-through fractions were pooled, dialyzed overnight against two changes of low salt buffer (50 mM sodium phosphate, pH 6.8, 150 mM NaCl, 4 mM β-mercaptoethanol, 0.5 mM EDTA, and 10% glycerol) and passed through a HighTrap SP (Cytiva) cation exchange column. Under these conditions, nsp13 did not adhere to the SP column and the flow-through was collected in multiple fractions. The nsp13 was concentrated with an Amicon Ultra-15 centrifugation filter unit to a volume of ∼1.5 ml and loaded on to a Sephacyl S200-HR HiPrep 26/60 column (Cytiva) equilibrated with nsp13 Storage Buffer (25 mM HEPES, pH 7.5, 150 mM NaCl, 0.5 mM TCEP, and 20% glycerol). The final nsp13 was quantified by UV spectrophotometry at 280 nm using the expected extinction coefficient of 68 785 M^{-1} cm^{-1} and confirmed using the BCA Protein Assay (Pierce). Protein samples were aliquoted, flash frozen, and stored at −80°C.

RNA hairpin fabrication

The fabrication of the RNA hairpin has been described in detail in [31]. The RNA hairpin is made of a 499 bp double stranded RNA stem terminated by a 20 nt loop that is assembled from three ssRNA annealed together (Supplementary Fig. S1A). The RNA stem is flanked by two spacers, ∼800 bp each, containing a biotin- and a digoxigenin-handle, respectively. A gap of 25 nt between the biotin-handle and the hairpin stem serves as the loading site for the polymerase (Supplementary Fig. S1A). At forces >22 pN, the hairpin opens and frees up a 1043 nt ssRNA template for the SARS-CoV-2 polymerase (Supplementary Fig. S1B). The RNA construct was synthesized by amplifying DNA fragments in PCR and *in vitro* transcribing them (NEB HiScribe T7 High Yield RNA Synthesis Kit) after purification (Monarch PCR and DNA Cleanup Kit). ssRNA fragments containing biotin- or digoxigenin-labels were synthesized with biotin-UTP or digoxigenin-UTP (Jena Biosciences) in the reaction. Transcripts were mono-phosphorylated (Antarctic Phosphatase and T4 Polynucleotide Kinase), annealed, and ligated. The RNA template sequence is provided in [26].

Flow cell assembly and surface functionalization

The fabrication procedure for flow cells has been described in detail in [32]. To summarize, we sandwiched a double layer of Parafilm by two #1 coverslips, the top one having one hole at each end serving as inlet and outlet, the bottom one being coated with a 0.1% m/V nitrocellulose dissolved in amyl acetate solution. The flow cell is mounted into a custom-built holder and rinsed with ∼1 ml of 1× phosphate-buffered saline (PBS) solution. 3-μm diameter polystyrene reference beads are attached to the bottom coverslip surface by incubating 100 μl of a 1:1000 dilution in PBS (LB30, Sigma–Aldrich, stock conc.: 1.828 × 10^{11} particles per milliliter) for ∼3 min. The tethering of the magnetic beads by the RNA hairpin construct relies on a digoxygenin/anti-digoxygenin and biotin-streptavidin attachment at the coverslip surface and the magnetic bead, respectively. Therefore, following a thorough rinsing of the flow cell with PBS, 50 μl of anti-digoxigenin (50 mg/ml in PBS) is incubated for 30 min. The flow cell was rinsed with 1 ml of 10 mM Tris, 1 mM EDTA pH 8.0, 750 mM NaCl, and 2 mM sodium azide buffer to remove excess of anti-digoxigenin followed by rinsing with another 0.5 ml of 1× TE buffer (10 mM Tris, 1 mM EDTA pH 8.0 supplemented with 150 mM NaCl, and 2 mM sodium azide). The surface is then passivated by incubating bovine serum albu-

min (BSA, New England Biolabs, 10 mg/ml in PBS, and 50% glycerol) for 30 min and rinsed with 1× TE buffer.

Single-molecule SARS-CoV-2 primer-extension experiments

Twenty microliters of streptavidin coated Dynabeads M-270 magnetic beads (ThermoFisher) were mixed with ∼0.1 ng of RNA hairpin (total volume 40 μl) and incubated for ∼5 min before rinsing with ∼2 ml of 1× TE buffer to remove any unbound RNA and the excess of magnetic beads. RNA tethers were sorted for functional hairpins by looking for the characteristic jump in extension of the correct length (∼ 0.6 μm at 22 pN) due to the sudden opening of the hairpin during a force ramp experiment (Supplementary Fig. S1C) ([26, 31]). The flow cell was subsequently rinsed with 0.5 ml reaction buffer (50 mM HEPES pH 7.9, 10 mM DTT, 2 mM EDTA, and 5 mM $MgCl_2$). After starting the data acquisition, the hairpin tether quality was tested by ramping the force up to monitor the typical cooperative opening signature of a hairpin, i.e. a vertical jump of the magnetic bead by ∼0.6 μm when reaching a critical opening force of ∼22 pN (Supplementary Fig. S1D). The force was subsequently decreased to and maintained at 25 pN, unless specified otherwise. Hundred microliters of reaction buffer containing the proteins and 500 μM of {A, U, C, G}TP (Jena Biosciences) were flushed into the flow chamber (Fig. 1A). SARS-CoV-2 core RTC activity traces were spotted as a downward movement of the bead, indicating the conversion of the ssRNA template into double-stranded (ds) RNA, and therefore a shortening of the tether (Fig. 1A). The recordings lasted 30 min. A temperature of 25°C was maintained during all experiments. A custom written LabView routine controlled the data acquisition and the (*x*-, *y*-, *z*-) positions analysis/tracking of both the magnetic and reference beads (Sigma) in real-time [33]. Mechanical drift correction was performed by subtracting the position of the reference bead from that of the magnetic beads, and further stabilized by an automated focusing routine that adaptively moves the objective to keep the reference bead at the same focal plane [26]. The camera frame rate was fixed at 58 Hz.

Single-molecule experiments after pre-incubating proteins

For the experiments performed in Fig. 3A and Supplementary Fig. S3, 20 μM of nsp12 (of either expression system) was added to 180 μM of both nsp8 and nsp7. Next, the mixture was placed on a heating block at 25°C for 4.5 h (nsp12-polymerase expressed in *E. coli*) or 6.5 h (nsp12-polymerase expressed in *Sf9*). After the specified waiting time, the proteins were diluted in reaction buffer to a final concentration of 0.2 μM of nsp12-polymerase and 1.8 μM of both nsp8 and nsp7. Finally, we proceeded by performing the primer-extension experiment as described above.

Single-molecule experiments after pre-incubating proteins with RNA

After forming the RNA tethers as described above, proteins were added at specified concentration to the flow chamber. The recording was started. After 10 min, we removed any access of proteins in solution by rinsing the flow chamber with 500 μl of reaction buffer. Finally, primer-extension experiments were performed after the system was complemented with the missing reagents (Fig. 3B and C, Supplementary Figs S4 and S5). In Fig. 3B, and Supplementary Figs S4 and S5, the proteins were incubated together with 500 μM of NTP to further confirm no activity can occur unless all nsps are present.

Data processing of RNA synthesis traces (elongation phase).

The change in extension in micron, resulting from the ssRNA to dsRNA conversion, was subsequently converted into replicated nucleotides, low-pass filtered at 2 Hz, and the dwell times were extracted using a window of 10 nt as described in [11, 26, 34, 35]. Elongation times and product lengths were inferred from the final datapoint before the trajectory ended of the filtered traces (Fig. 4A). Product lengths shorter than the full length of the template were only considered if the bead stopped moving downwards before the end of the recording. If the tether broke before the full product was synthesized (seen as the loss of the bead), that event was excluded from both elongation time and product length measurements.

Data processing of activation times

When RNA extension could be observed in the trace, the activation time was taken as the time between ending the addition of reagents into the flow chamber and the first datapoint when the bead starts moving downwards (Fig. 1C). Only tethers with an open hairpin before addition of the reagents were considered (Supplementary Fig. S1D). Polymerase activity when the elongation phase started while still adding reagents was clearly recognized (Fig. 1E), but the start time/position of these events could not be determined with high accuracy, and these were assigned an activation time of zero seconds. Tethers that remained until the end of the recording were considered to determine the fraction of tethers with activity (Supplementary Figs S2, S7, and S8). This showed that a considerable portion of traces showed no sign of RNA extension during the 30 min recordings. Assuming that activation on different tethers are independent, the activation efficiency results from a Bernoulli trial. The experimental estimate (η_A^{exp}) of the primer-extension efficiency (η_A) is the total number of tethers on which we record activity (N_{rec}) over to total number of quality RNA hairpin tethers (N_{HP}):

$$\eta_A^{exp} = \frac{N_{rec}}{N_{HP}} \text{and} \delta\eta_A^{exp} = \sqrt{\frac{\eta_A^{exp}\left(1-\eta_A^{exp}\right)}{N_{HP}}} \quad (1)$$

We used that η_A^{exp} is an unbiased estimator of η_A, and plot the 95% confidence intervals ($\eta_A^{exp} \pm 2\delta\eta_A^{exp}$) in Supplementary Fig. S2.

To extract the maximum likelihood estimator of the single-exponential time constant that best describes the distribution of activation times we included events missed due to the limited experimental time. This method has been described in detail in [36]. Briefly, the maximum-likelihood estimator (MLE) for the time constant in a single-exponential distribution would equal the average of all recorded times (Supplementary Fig. S2),

$$\Delta t_A^{rec} = \frac{1}{N_{rec}} \sum_n^{N_{rec}} \Delta t_A^{(n)} \quad (2)$$

when there is no limit on the experimental observation time. In the above, $\Delta t_A^{(n)}$ corresponds to the activation time recorded on the n^{th} out of the N_{rec} traces that initiate. With a finite ex-

perimental time, N_{cut} traces will not initiate within the experimental time T_{cut}, and the MLE that accounts for this is given by:

$$\Delta t_A^{MLE} = \Delta t_A^{rec}\left(1 + \frac{N_{cut}T_{cut}}{N_{rec}\Delta t_A^{rec}}\right) \quad (3)$$

To obtain an error estimate for Δt_A^{MLE}, 1000 bootstrap samples were drawn from the collection of $N_{HP} = N_{rec} + N_{cut}$ traces. In this manner, we get a distribution of estimated activation times (Equation 3). The resulting 95% confidence intervals were used as the error estimates shown in Fig. 2. In Supplementary Figs S7 and S8, we overlay the distributions of recorded activation times with the exponential distribution with time constant Equation 3 and total probability η_A^{exp}.

The MLE of Equation 3 was only applied to the experiments involving the individually purified nsps. At saturating concentrations of the purified complex, activation typically started before we stopped flushing reagents into the flow chamber (Figs 1E and 2E). Yet not all RNA tethers got converted into dsRNA during our recording (Equation 1 and Supplementary Fig. S2J). The measured activation efficiencies (Equation 1 and Supplementary Fig. S2J) showed little to no dependence on the concentration of the complex in this case, indicating that they do not originate in events missed due to our limited observational time; consequently, we exclude these from our considerations. Figure 2E and Supplementary Figure S2I therefore show the MLE of Equation 2. Error estimates are 95% confidence intervals from 1000 bootstrap samples drawn from the N_{rec} recorded times.

Modeling the activation time of the reconstituted core RTC

The reaction schema underlying our kinetic model is shown in Supplementary Fig. S6. We split the activation time into two parts: (i) a "slow" process that includes the nsp7- and nsp8-dependent "activation step" (Figs. 5 and Supplementary Fig. S6), and (ii) the time spent by the core RTC after being stabilized into an elongation-competent conformation, but before incorporation of the first nucleotide (Supplementary Fig. S6):

$$\Delta t_A^{mod} = \frac{1}{k_{slow}} + \frac{1}{k_{fast}^{recon.}} \quad (4)$$

We will start by describing the rate-limiting transition (k_{slow}). Our results showed activation is orders of magnitudes faster in the presence of the RNA template (Fig. 3A and B, and Supplementary Fig. S3); hence, we limit the activating conformational change to the RNA bound state. With this simplification the rate of the "slow" step in Equation 4 equals the equilibrated fraction P_{bnd} of tethers with an assembled core RTC, multiplied by the intrinsic rate of undertaking the conformational change (k_{conf}):

$$k_{slow} = k_{conf} \times P_{bnd} \quad (5)$$

Let c_7, c_8, and c_{12} denote the concentrations of nsp7, nsp8, and nsp12-polymerase respectively. Given core RTC assembly and RNA binding happen before the rate limiting step, their respective probabilities of occurance should be equilibrated, i.e. satisfy the law of mass action and detailed balance:

$$c_{RTC} = K_{EQ}^{RTC} c_7 c_8^2 c_{12} \quad (6)$$

$$P_{bnd} = K_{EQ}^{bnd}(1 - P_{bnd})\, c_{RTC} \quad (7)$$

where P_{bnd} is the fraction of the RNA tethers with an assembled core RTC bound to it. We also introduced the equilibrium constants for assembling the core RTC (K_{EQ}^{RTC}) and binding the core RTC to the RNA (K_{EQ}^{bnd}). The latter having a force-dependence as shown below. Given the large disparity in apparent saturating concentrations using pulled-down versus reconstituted core RTCs (Fig. 2), we assumed the total number of free nsps that complex remains small enough to not affect their concentration in bulk, see below for the precise condition on K_{EQ}^{RTC} and K_{EQ}^{bnd}. While losing some generality, we avoid allot of complexity in the model. This allows us to use the model purely to focus on the effect of having a rate-limiting step after equillibrated binding and assembly. Combining Equations (5–7) yields:

$$k_{slow} = k_A \frac{c_7 c_8^2 c_{12} K_{EQ}^{bnd} K_{EQ}^{RTC}}{1 + c_7 c_8^2 c_{12} K_{EQ}^{bnd} K_{EQ}^{RTC}} \quad (8)$$

After rearranging the core RTC into an elongation-competent conformation, the protein complex can still potentially unbind from the RNA before incorporating the first NTP (Supplementary Fig. S6). The above used that this process is much faster than the conformational change ($k_{ub} \gg k_{conf}$). However, this assumption no longer holds true when competing with NTP incorporation (which occurs at $k_{NTP} \sim 70$ nt/s) [26]. Instead of assuming equilibrium, we solve for the fraction of extended RNA molecules (P_{NTP}) from the flux-balances of RNA species after the core RTC changed conformation (Supplementary Fig. S6):

$$\partial_t P_{ub} = -k_{slow} P_{ub}(t) + k_{ub} P_{bnd}(t) \quad (9)$$

$$\partial_t P_{bnd} = +k_{slow} P_{ub}(t) - (k_{ub} + k_{NTP}) P_{bnd}(t) \quad (10)$$

$$\partial_t P_{NTP} = k_{NTP} P_{bnd}(t) \quad (11)$$

Here, $P_{ub}(t)$ and $P_{bnd}(t)$ represent the time-dependent fractions of RNA molecules (not) bound by a core RTC stabilized in its active conformation. The concentration of RNA molecules on the flow cell's surface (picomolar range) is negligible when compared to the amount of protein available in solution. A core RTC that unbinds from the RNA is therefore replaced by a different complex that must still undergo the conformational change. Hence, the bare RNA enters the state of being occupied by an activated core RTC at "an effective binding rate" of k_{slow} (Supplementary Fig. S6, and Equations 10 and 11). Incorporation of the first nucleotide is treated as irreversible (Fig. 5 and Supplementary Fig. S6). Therefore, any additional number of RNA molecules that are extended between times t and dt also equals the probability of having started NTP incorporation within that time frame; the first-passage time distribution.

$$\Psi_{NTP}^{recon./pull-down}(t) = \partial_t P_{NTP}^{recon./pull-down} \quad (12)$$

The superscripts "recon."/"pull-down" are used to distinguish between the solutions to Equations 9 and 11 under the initial conditions appropriate to the reconstituted/pulled-down core RTC systems respectively. The reconstituted core RTC predominantly bound the RNA before changing conformation (Fig. 5 and Equation 5, $P_{bnd}^{recon.}(t=0) = 1$, $P_{ub}^{recon.}(t=0) = P_{NTP}^{recon.}(t=0) = 0$). The inverse transition rate

into the elongation phase equals the average first-passage time:

$$\frac{1}{k_{\text{fast}}^{\text{recon.}}} = \int_0^\infty t\Psi_{\text{NTP}}^{\text{recon.}}(t)\,dt = \frac{1}{k_{\text{NTP}}}\left(1+\frac{k_{\text{ub}}}{k_{\text{slow}}}\right) \quad (13)$$

Modeling the activation time of the purified core RTC

Given a pulled-down core RTC has already adopted the proper elongation-competent conformation, its activation time only includes the fast component:

$$\Delta t_A^{\text{mod}} = \frac{1}{k_{\text{fast}}^{\text{pull-down}}} \quad (14)$$

Following the same strategy as done for $k_{\text{fast}}^{\text{recon.}}$, we once again determine the mean first passage time to incorporate the first nucleotide. Different from the case presented above, co-translated complexes in solution already are activated and thus rebind with an intrinsic binding rate (k_{bnd}). Assigning the same affinity of the purified and reconstituted RTC to the RNA (at a reference concentration of 1 μM), we invoke the law of mass action again:

$$k_{\text{bnd}} = K_{\text{EQ}}^{\text{bnd}} k_{\text{ub}} c_{\text{RTC}} \quad (15)$$

where we used $K_{\text{EQ}}^{\text{bnd}} \equiv k_{\text{bnd}}^{1\ \mu\text{M}}/k_{\text{ub}}$; Supplementary Fig. S6. Replacing k_{slow} by k_{bnd} in Equations 9 and 10, results in the inverse reaction rate under the initial condition that the complex must first bind the RNA ($P_{\text{ub}}^{\text{pull-down.}}(t=0)=1$, $P_{\text{bnd}}^{\text{pull-down.}}(t=0) = P_{\text{NTP}}^{\text{pull-down.}}(t=0)=0$)):

$$\frac{1}{k_{\text{fast}}^{\text{purif.}}} = \int_0^\infty t\Psi_{\text{NTP}}^{\text{pull-down.}}(t)\,dt = \frac{1}{k_{\text{bnd}}}\left(1+\frac{k_{\text{ub}}}{k_{\text{NTP}}}\right)+\frac{1}{k_{\text{NTP}}} \quad (16)$$

Incorporating force-dependent binding

At equilibrium, the fraction of bound complex follows the Boltzmann distribution:

$$\frac{P_{\text{bnd}}}{1-P_{\text{bnd}}} = e^{+\frac{\Delta G_{\text{bnd}}}{k_B T}} \quad (17)$$

Here, we assigned a free-energy difference for binding the (assembled) RTC on the RNA (ΔG_{bnd}). The magnetic tweezers apply a constant force (F) to the template strand along the vertical direction (z-direction), thereby putting work into the system. Interactions with the template strand favor binding at lower force (Fig. 2F), i.e. $\Delta G_{\text{bnd}}(F) \equiv \Delta G_{\text{bnd}}^{0\ \text{pN}} + F\delta z$, with δz the length change induced to the RNA by binding of the RTC. Defining a characteristic force of $F_0 = k_B T/\delta z$, Equations 7 and 17 imply a force-dependent equilibrium constant,

$$K_{\text{EQ}}^{\text{bnd}}(F) \propto e^{-\frac{F}{F_0}} \quad (18)$$

Equation 18 was applied in Equations 13 and 16 to determine force-dependent activation times. The model parameter $K_{\text{EQ}}^{\text{bnd}} = k_{\text{bnd}}^{0\ \text{pN and}\ 1\ \mu\text{M}}/k_{\text{ub}}$ is the constant of proportionality in Equation 18, denoting the equilibrium constant in the absence of an applied force. For notational convenience, the equations above still contain $K_{\text{EQ}}^{\text{bnd}}$, where its force-dependence (Equation 18) is implied.

Comparing (half-)saturating for RTC complex formation and RNA binding

The presented model requires the number of free nsps to neglibly change as new complexes get formed. We justified this assumption by comparing the half-saturating concentrations in Fig. 2D ($[\text{nsp12}]_{50}$) and Fig. 2E ($[\text{Pulled}-\text{down core RTC}]_{50}$). Here, we show the condition on the equilibrium constants in our model this implies.

The concentration of pulled-down core RTC at which the activation rate drops to half-saturation can be directly obtained through setting the bound fraction equal to $\frac{1}{2}$ in Equation 7 (which is also valid for the pulled-down complex).

$$[\text{Pulled}-\text{down core RTC}]_{50} = \frac{1}{K_{\text{EQ}}^{\text{bnd}}} \quad (19)$$

Given one core RTC contains one copy of nsp12, the above concentration can be directly compared to the half-saturating concentration of nsp12 needed when reconstituting the complex. The data shown in Fig. 2D are taken while mainting a constant stoichiometry of [nsp8] = [nsp7] = 9[nsp12]. Setting the bound fraction in Equation 8 equal to $\frac{1}{2}$ relates $[\text{nsp12}]_{50}$ to both equilibrium constants:

$$[\text{nsp12}]_{50} = \sqrt[4]{\frac{1}{729\ K_{\text{EQ}}^{\text{bnd}} K_{EQ}^{\text{RTC}}}} \quad (20)$$

Figure 2D and E showed that $[\text{nsp12}]_{50} \gg [\text{Pulled}-\text{down core RTC}]_{50}$, which according to Equations 18–20 implies:

$$729\ K_{EQ}^{\text{RTC}} \ll \left(K_{\text{EQ}}^{\text{bnd}} \times e^{-\frac{25}{F_0}}\right)^3 \quad (20)$$

After fitting our model to the data, we obtained K_{EQ}^{RTC}, K_{EQ}^{bnd}, and F_0 values for which Equation 20 is satisfied.

Fitting procedure using simulated annealing

The Simulated Annealing algorithm [37] is a commonly used algorithm for high-dimensional optimization problems. We used a custom-built Python code that has been more extensively described in [38]. We optimized the loss-function (χ^2) with respect to our model parameters $\log_{10}(k_{\text{conf}}/1\ s)$, $\log_{10}(K_{\text{EQ}}^{\text{RTC}}/1\ \mu\text{M})$, $\log_{10}(K_{\text{EQ}}^{\text{bnd}}/(1\ \mu\text{M})^3)$, $\log_{10}(k_{\text{ub}}/1\ s)$, and $\log_{10}(F_0/1\ \text{pN})$. Model parameters have been transformed to make all fitted parameters lie within the same range.

$$\chi^2 = \sum_{i\in\begin{pmatrix}\text{all experiments}\\ \text{used for training}\end{pmatrix}} \frac{\left(\Delta t_{\text{A}}^{\text{mod, }(i)} - \Delta t_{\text{A}}^{\text{exp, }(i)}\right)^2}{\left(\delta t_A^{\text{exp, (i)}}\right)^2} \quad (21)$$

That is, for every primer-extension experiment represented in Fig. 2, we minimize the sum of differences of our model's predictions to the corresponding datapoint, weighted by the experimental error ($\delta t_A^{\text{exp, (i)}}$) calculated as described above. Trial moves were generated by adding uniform noise of magnitude δ to the present value of each model parameter. The process was initiated with a noise strength of $\delta = 1.0$. In the initiation cycle, the temperature was adjusted until we had an acceptance fraction of 40%–60% over 1000 trial moves, based on the Metropolis condition. After this intitial cycle, the temperatures followed an exponential cooling scheme with a 1% cooling rate ($T_{k+1} = 0.99T_k$). At every temperature, we

adjusted the noise strength δ until an acceptance fraction of 40%–60% was reached over 1000 trial moves. Once the desired acceptance fraction was reached, an additional 1000 trial moves were performed to allow the system to equilibrate before the next cooling step. Once the temperature had dropped to a factor 10^{-4} of its initial value, we applied the stop condition:

$$\left|\tilde{\chi}_k^2 - \tilde{\chi}_{k-1}^2\right| \leq 10^{-5}\tilde{\chi}_{k-1}^2 \tag{22}$$

In the above, $\tilde{\chi}_k^2$ denotes our cost function averaged over the last 1000 trial moves performed at temperature T_k. The optimization procedure was repeated 40 times (Supplementary Fig. S9). The parameter values with the lowest χ^2 were used to generate the curves shown in Fig. 2.

Results

Slow activation by reconstituted SARS-CoV-2 core RTCs

To determine the kinetics of the CoV polymerase complex during activation and elongation phases, we built upon our previously developed single-molecule high-throughput magnetic tweezers assay that monitors primer-extension of dozens of individual polymerases simultaneously [26, 34, 35, 39]. Magnetic beads were tethered to the glass surface of a flow chamber by a 1043 nt ssRNA template that included a ∼800 bp primer on the bead proximal side [31] (Fig. 1A, and Supplementary Fig. S1A and B, "Materials and methods" section). A pair of permanent magnets located above the flow chamber at a fixed height applied a constant attractive force (25 pN, unless specified otherwise) to the magnetic beads that stretched the RNA tethers [32, 40, 41]. The three-dimensional position of each bead was tracked in real-time, providing the extension of their respective tether [33]. The RNA construct was designed to form a ∼500 bp hairpin when relaxed. The clear signature of the hairpin opening when rapidly increasing the force was used to select for tethers with a properly presented primer (Supplementary Fig. S1C and D). During the elongation phase, the core RTC converted the ssRNA template into dsRNA, which decreased the extension of the tether [26].

Initial primer-extension assays were performed using a core RTC reconstituted from nsp7, nsp8, and nsp12-polymerase individually expressed and purified in *E. coli* (Fig. 1B and "Materials and methods" section). The proteins were added together, and to the RNA, just before the start of the recording ("Materials and methods" section). We can define the activation time Δt_A as the time from injecting the core RTC components and NTPs until the start of the elongation phase when the tether extension starts decreasing (Fig. 1A and C, and "Materials and methods" section). These single-molecule experiments showed that the median activation time for the reconstituted core RTC (∼5–10 min) greatly exceeded the elongation time (∼20–40 s) (Fig. 1C and D). To uncover the process(es) giving rise to the observed activation times, we utilized a combination of different strategies to express and/or reconstitute the core RTC.

To perform primer-extension experiments with a core RTC expressed, processed and assembled in the cell, we expressed the complex in *Sf9* cells from a bacmid encoding for nsp12-polymerase, nsp8, nsp7, as well as nsp5-protease, and pulled down and purified the fully assembled complex (hereafter, "pulled-down core RTC") (Fig. 1B and "Materials and methods" section). Strikingly, the typical activation time decreased to mere seconds when repeating the experiment using a comparable concentration of pulled-down core RTC (Fig. 1D and E). We note that all experiments in Fig. 1 were performed at saturating concentrations of core RTC proteins (see below for discussion on Fig. 2).

Given that individual nsps were expressed in *E. coli* and the pulled-down core RTC was expressed in *Sf9* cells (Fig. 1B), we first evaluated whether the different expression system could promote the purified complex' activation. To this end, we repeated our single-molecule experiments while reconstituting the core RTC with a nsp12-polymerase expressed in *Sf9* cells (Fig. 1B and F, and "Materials and methods" section). The resulting reconstituted core RTC, with nsp12-polymerase expressed in insect cells, did not activate rapidly (Fig. 1D and F). We concluded that the difference between bacteria- and insect-derived polymerases cannot account for most of the observed differences in activation times when comparing either reconstituted core RTC to their pulled-down counterpart.

Assembly and RNA binding of the core RTC cannot account for the observed activation times

As opposed to the pulled-down core RTC, the individual proteins of the reconstituted core RTC must come together to assemble into a complex. We questioned whether the time required to assemble the reconstituted core RTC can explain its slower activation as has been suggested in literature ([23, 24]). To this end, we performed primer-extensions under varying concentrations of one of the three nsps, while keeping the other two unchanged. Here, we used the nsp12-polymerase expressed in *E. coli* to reconstitute the core RTC. While the average activation times measured responded to the varying protein concentrations, we noticed a stronger effect on the fraction of RNA primer-templates that got extended during our 30 min recording (Supplementary Fig. S2). This indicated a significant and varying proportion of events (up to ∼75% at the lowest nsp concentrations) were missed as they took longer than the duration of our experiment. The estimated (effective) activation times shown in Fig. 2A–D, and F are maximum likelihood estimates accounting for the fraction of events that exceed the observation time ("Materials and methods" section).

Elongation started sooner on average at higher concentrations of nsp7, with effective activation times dropping by ∼4 fold over the tested concentration range (Fig. 2A). However, at saturating nsp7 concentrations (>1.8 μM), a substantial effective activation time of ∼1000 s remained (Fig. 2A), i.e. still significantly slower than the pulled-down core RTC (Fig. 1E). A similar trend was seen when increasing either the nsp8 or nsp12-polymerase concentration (Fig. 2B and C), i.e. the activation time decreased until an average time of ∼1000 s remained at saturation, meaning for concentrations of nsp8 ≥ 1.8 μM and nsp12-polymerase ≥ 0.2 μM. The remaining time can therefore not be explained by the need to bring the complex together.

To evaluate whether the time required to find and bind the RNA accounts for the remaining activation time, we fixed the stoichiometry (nsp12:nsp7:nsp8 of 1:9:9) and varied the overall concentration both with nsp12-polymerase expressed in bacterial and insect cells (Fig. 2D). In both cases, an effective activation time of ∼1000 s remained, even above saturating

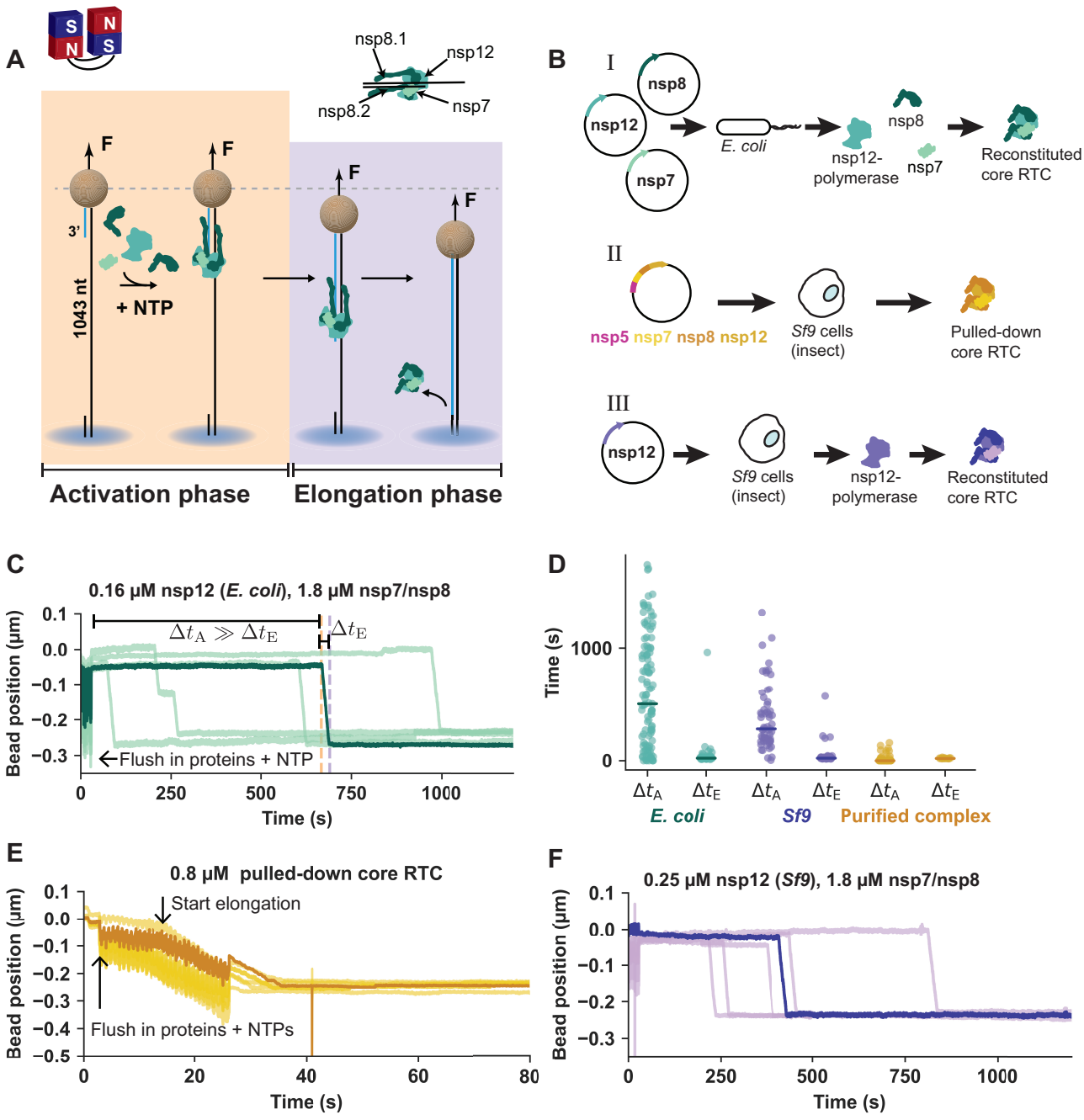


Figure 1. Slow activation by reconstituted SARS-CoV-2 core RTCs. (**A**) Schematic of the magnetic tweezers assay to monitor activation and elongation by the SARS-CoV-2 core RTC ("Materials and methods" section). (**B**) Schematic representation different methods of obtaining the core RTC used throughout this work. (I) nsp7, nsp8, and nsp12-polymerase were separately and recombinantly expressed in *Escherichia coli*, (II) pull-down of a core RTC complex after expressing a bacmid containing nsp5-Mpro (main protease), nsp8, nsp7, and nsp12-polymerase in *Sf9* cells, (III) nsp7, nsp8, and nsp12-polymerase were separately and recombinantly expressed in *E. coli*, and nsp12-polymerase was recombinantly expressed in *Sf9* ("Materials and methods" section). Color coding for the different core RTCs is kept in subsequent figures to represent to related data. (**C**) Example time traces for the reconstituted SARS-CoV-2 core RTC using 0.16 μM nsp12-polymerase [see (I) in panel B] and 1.8 μM of nsp7 and nsp8. The orange (left-most) and purple (right-most) dashed lines indicate the end of the activation and elongation phases. Δt_A and Δt_E are their respective durations. (**D**) Δt_A and Δt_E for all recorded traces in experiments of which representative subsets are shown in panels (C), (E), and (F). Horizontal markers indicate the group medians. (**E**) Example time traces using 0.8 μM of the pulled-down SARS-CoV-2 core RTC [see (II) in panel B]. (**F**) Example time traces for the reconstituted SARS-CoV-2 core RTC using 0.25 μM nsp12-polymerase expressed in *Sf9* [see (III) in panel B] and 1.8 μM of nsp7 and nsp8.

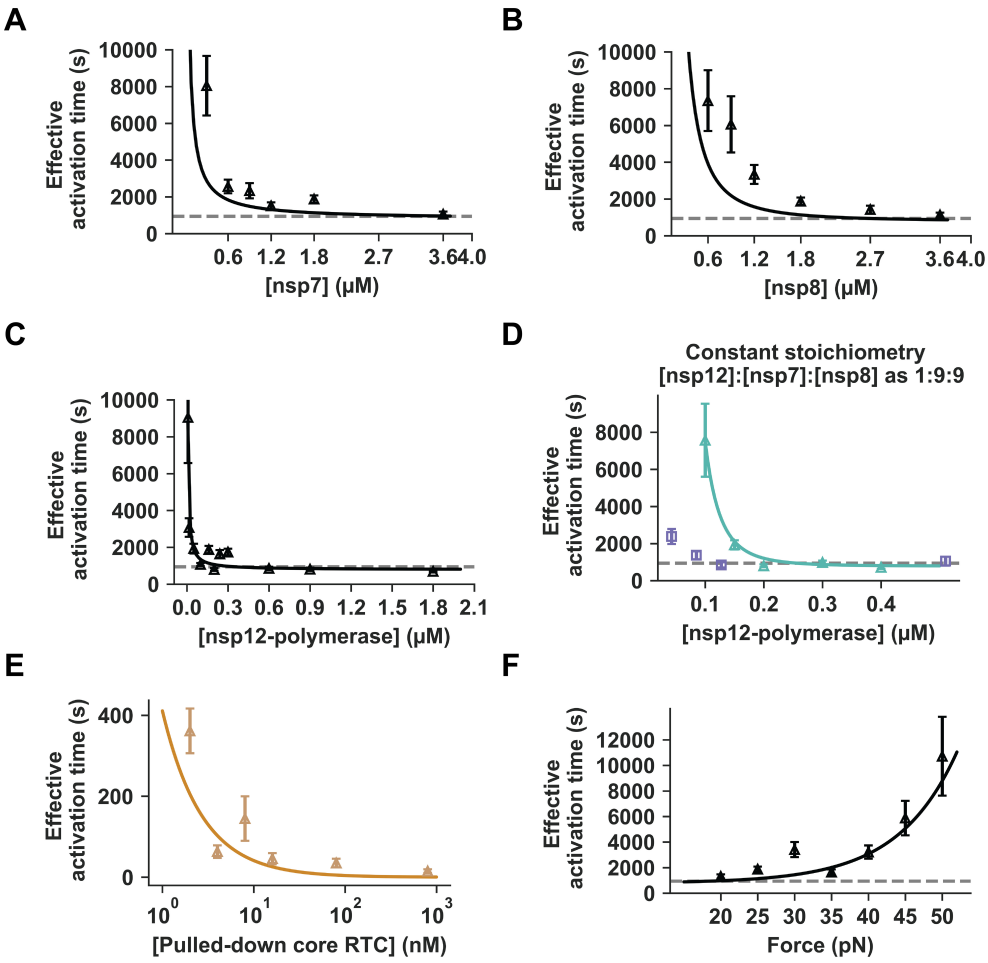


Figure 2. Assembly and RNA binding of the core RTC cannot fully account for the observed activation times. When not being varied, experiments were performed using 0.16 μM of nsp12-polymerase, 1.8 μM nsp7, 1.8 μM nsp8, and 25 pN. Effective activation times across (**A**) nsp7 concentration, (**B**) nsp8 concentration, and (**C**) nsp12-polymerase concentration (expressed in *E. coli*). (**D**) Activation times versus nsp12-polymerase concentration, while maintaining a constant stoichiometry of [nsp12-polymerase]:[nsp7]:[nsp8] as 1:9:9 in solution. Core RTC reconstituted using nsp12-polymerase expressed in *E. coli* (turquoise triangles) or *Sf9* (purple squares). (**E**) Activation times versus concentration of pulled-down core RTC. (**F**) Activation times versus force (reconstituted core RTC, all proteins expressed in *E. coli*). Error bars represent 95% confidence intervals estimated through bootstrapping ("Materials and methods" section). Effective activation times shown incorporate the fraction of events lasting longer than the recording ("Materials and methods" section). Solid curves are the best fit of our mechanochemical model ("Materials and methods" section, Fig. 5 and Supplementary Fig. S6). Dashed horizontal lines are shown to guide-the-eye.

proteins concentration, showing that this time is both independent of host-specific post-translational modifications on the nsp12-polymerase and core RTC assembly but must be intrinsic to the core RTC components. On the other hand, when we varied the concentration of the pulled-down core RTC, an activation time of mere seconds was reached at saturation, i.e. activity was detected while flushing in reaction buffer with all proteins into the flow chamber (Figs 1E and 2E). The saturating concentration of the pulled-down core RTC is in the tens of nanomolar range (Fig. 2E), while hundreds of nanomolar of the reconstituted core RTC were required to saturate activity (Fig. 2D). As the pulled-down core RTC only needs to find the primer, a lower saturation concentration suggests that the affinity of the complex to the RNA is much greater than that of the proteins to themselves (at the stoichiometry used).

The force applied by the magnetic tweezers provides an additional gauge of the affinity of the core RTC for the primer-template. Namely, elevated tension in the template strand (above ~35 pN) destabilizes the primer-template junction, hindering a proper placement of the terminal base in the polymerase's active site upon binding [26]. Furthermore, applying tension probes possible conformational changes of core RTC along the direction of the force (e.g. the upstream RNA contacts made by the nsp8 tails [18, 20]), and any conformational change orthogonal to the direction of the force remains unaffected. While we observe a steep increase in effective activation time at forces beyond 40 pN, the effective activation time remains ~1000 s at lower forces (Fig. 2F). Taken together, neither the time needed to assemble the core RTC nor the time spent on binding to the RNA can fully account for the observed activation times. We thus conclude that another process must take place to enable the reconstituted core RTC to enter the elongation phase.

A slow conformational change renders the core RTC elongation competent

As the duration of the activation phase still greatly exceeded that of the elongation phase at lower forces (Fig. 2F), we

reasoned that the remaining time should be accounted for by processes other than direct interactions along the template RNA. If a slow conformational change is to occur after the core RTC assembled, providing sufficient time for the nsps to interact should induce rapid activation upon introduction of the nucleotides. We pre-incubated nsp12-polymerase with nsp7 and nsp8 for over 4 h prior to conducting the single-molecule primer extension assay (Fig. 3A). To counteract any loss of active proteins during incubation, proteins were added together at 100-fold the concentration used in the flow chamber during the experiment ("Materials and methods" section). Rapid activation was indeed observed with pre-incubated nsp7, nsp8, and nsp12-polymerase individually expressed in *E. coli* (Fig. 3A). Given sufficient time, reconstituted core RTCs can show the same phenotypical short activation times as the purified complex (86% of events started directly upon introduction of NTP). However, in experiments performed without any pre-incubation up to ∼80% of core RTCs can activate within 600 s (Supplementary Fig. S2). Incubating the nsps for 10 min (nsp12-polymerase expressed in *E. coli*) did not result in any rapidly activating core RTCs (29% activated rapidly) (Supplementary Fig. S3). The core RTC reconstituted with nsp12-polymerase expressed in *Sf9* did not show rapid activation even after incubation with the co-factors for over 6.5 h (Fig. 3A and Supplementary Fig. S3B). If the pulled-down core RTC only activated rapidly because the proteins resided together in the same cell for hours, what then allowed the reconstituted core RTCs to activate within 10 min during our experiments? We surmised that access to the template RNA may have aided the reconstituted system. Introducing the nsps to the RNA in the flow chamber for 10 min indeed resulted in rapid activation upon addition of nucleotides (82% of events using nsp12-polymerase expressed in *E. coli*; 95% of events using nsp12-polymerase expressed in *Sf9*) (Fig. 3B and Supplementary Fig. S3B; "Materials and methods" section). Taken together, independent of the expression system, pre-incubating the nsps with the RNA reproduced the rapid activation observed with the purified complex (Fig. 3B and Supplementary Fig. S3B, $P < 10^{-5}$ comparing incubation of 10 minutes with/without RNA using the Mann–Whitney *U*-test). The activation time of ∼1000 s that remained at saturating protein concentrations (Fig. 2A–D) can thus be accounted for by a process that took place during the incubation with the RNA (Fig. 3A and B). Given the majority of this time is neither spent on assembling the core RTC nor binding the RNA, we concluded that the core RTC must undergo a conformational change that renders it elongation competent. While the protein complex, at least in part, rearranges itself orthogonally to the RNA template (Fig. 2F), interactions with the template RNA greatly stabilize the elongation-competent conformation (Fig. 3B).

The conformational change enabling elongation requires both nsp7 and nsp8

Having established the existence and necessity of a conformational change for RNA synthesis activity, we next questioned whether nsp7 and nsp8 facilitate it. If the co-factors are a necessary condition, allowing only nsp12-polymerase to access the RNA for 10 min should not result in the rapid activation as seen in the experiments represented in Fig. 3B. As expected, adding nsp12-polymerase alone with NTP to the RNA did not result in any activity (Fig. 3C). We subsequently removed all free-floating proteins and thereafter injected the reaction buffer containing nsp7, nsp8, and NTP ("Materials and methods" section). The newly added co-factors can only form a core RTC if a nsp12-polymerase is still bound to an RNA template after rinsing the flow chamber. We successfully recovered activity, but only after ∼10 min (Fig. 3C). The nsp12-polymerase preincubated with the RNA template was not yet elongation-competent but required co-factors to enact the needed conformational change.

To establish whether both nsp7 and nsp8 are needed to induce the conformational change we repeated the experiment described in Fig. 3C, though this time we added one of the two co-factors together with nsp12-polymerase during pre-incubation with the RNA in the flow chamber (Supplementary Fig. S4). Not only are both co-factors required to activate the core RTC, as no activity was recorded after incubation despite the presence of NTP (Supplementary Fig. S4), initially omitting one of them severely lowered or even abolished the primer-extension activity. Namely, nsp12-polymerase incubated with nsp7 and the RNA template could only rarely be activated through addition of new nsp7 and nsp8 (∼5% activity versus ∼50% activity when incubating the three proteins with the RNA). Moreover, in the rare cases of activation observed (4 events in total, ∼80 tethers surviving the full duration of the measurement) it was only after an activation time of more than an hour (Supplementary Fig. S4A). Allowing nsp12-polymerase and nsp8 to preincubate with the RNA template completely blocked the RTC from activation (Supplementary Fig. S4B). No activity was observed within an hour from the point at which all nsps were present (Supplementary Fig. S4B). These results suggest nsp12-polymerase explores a conformation space, only to be locked into a stable conformer by the co-factors nsp7 and nsp8. When only one of the two co-factors is available, the polymerase appears to adopt a conformation that prevents RNA synthesis. Escaping this state requires significantly more time (Supplementary Fig. S4).

The elongation-competent conformation of the core RTC is stable for several hours but is reversed upon loss of nsp7 and nsp8

To establish if a core RTC stably resides in its active conformer, we pre-incubated the pulled-down core RTC for over eight hours in reaction buffer and at room temperature before performing primer-extensions ("Materials and methods" section). We still observed rapid activation on ∼25% of all tethers upon NTP addition (Supplementary Fig. S5), indicating that a substantial fraction of the complexes was still elongation-competent. To test if some of the tethers showed no activity due to disassembly of a bound core RTC, we waited for an additional 10 min, rinsed the flow chamber, and introduced new nsp7, nsp8, and NTP in an attempt to rescue any nsp12-polymerase bound without co-factor (Supplementary Fig. S5). Re-introduction of new nsp7 and nsp8 led to new activity (Supplementary Fig. S5). Together with the reduction in saturating concentration for the pulled-down core RTC (i.e. pre-assembled) discussed above, Fig. 2D and E), these results indicate that nsp12-polymerase binds the RNA template with a higher affinity than nsp7 and nsp8 bind to nsp12-polymerase. A long-lasting activation phase was observed again, indicating that the core RTC reverts to its inactive form upon loosing co-factors (Supplementary Fig. S5).

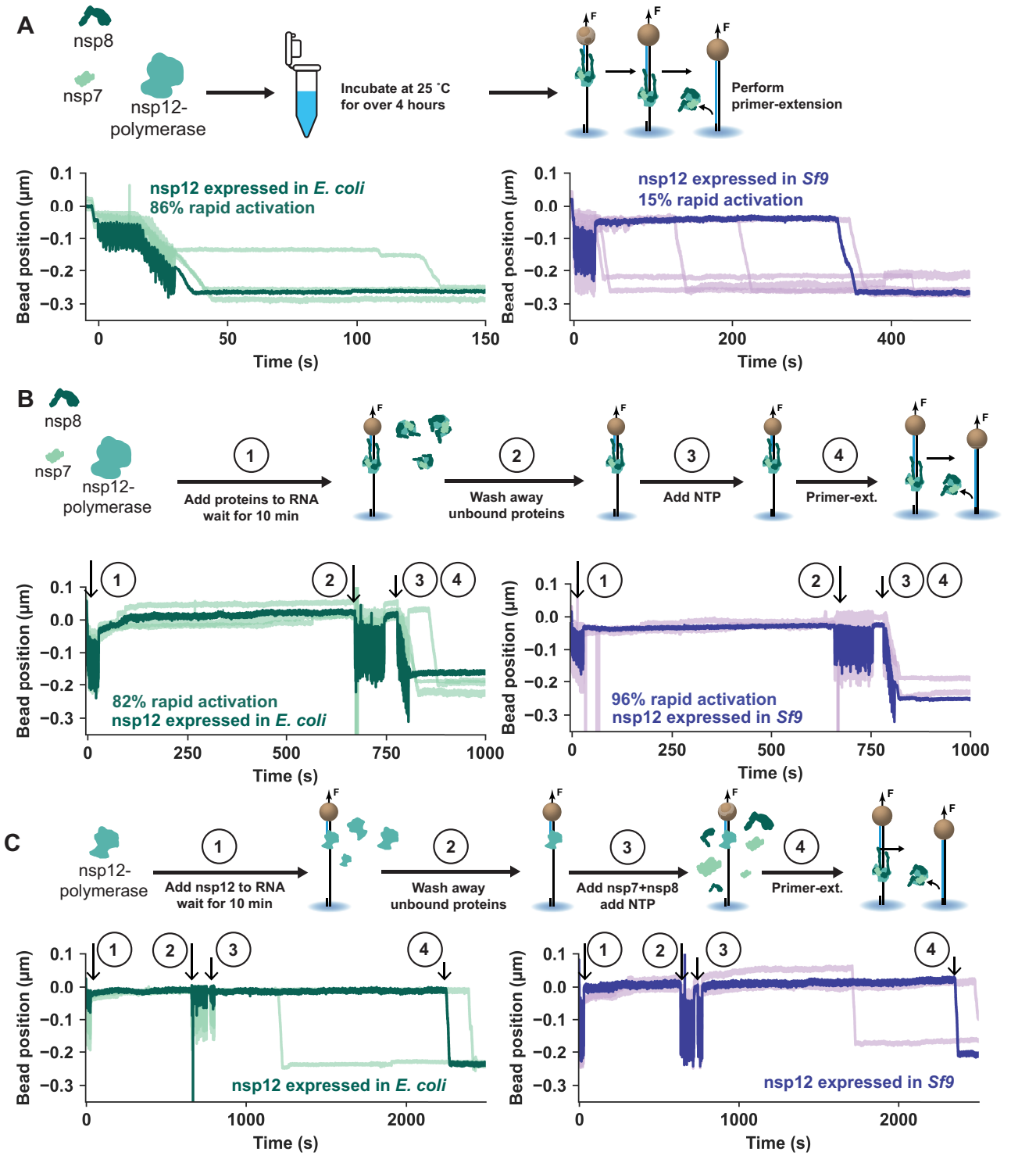


Figure 3. Presence of the template RNA reduces activation times. Example time traces are shown for a reconstituted core RTC. (**A**) Nsps were added together 4.5 hours (nsp12-polymerase expressed in *E. coli*) or 6.5 h (nsp12-polymerase expressed in *Sf9*) prior to diluting them to 0.2 µM nsp12, 1.8 µM nsp7, and nsp8 and performing primer-extensions. In panels (A) and (B), percentages of rapidly activating polymerases represent the fraction of traces in which activity started during phase 1 (flushing the reagents into the flow chamber). (**B**) (1) The proteins were added to the RNA tethering the magnetic beads in the flow chamber. (2) Ten minutes later 500 µl of reaction buffer was used to rinse the flow chamber from free-floating proteins. (3) Finally, ribonucleotides were added and (4) activation typically started during step (3). (**C**) (1) After tethering the magnetic beads to the RNA, 0.2 µM of nsp12-polymerase, and 500 µM of NTP were added. (2) Ten minutes later 500 µl of reaction buffer was used to remove free-floating proteins. (3) The nucleotides were reintroduced together with 1.8 µM of nsp7 and nsp8. (4) RTC elongation activity.

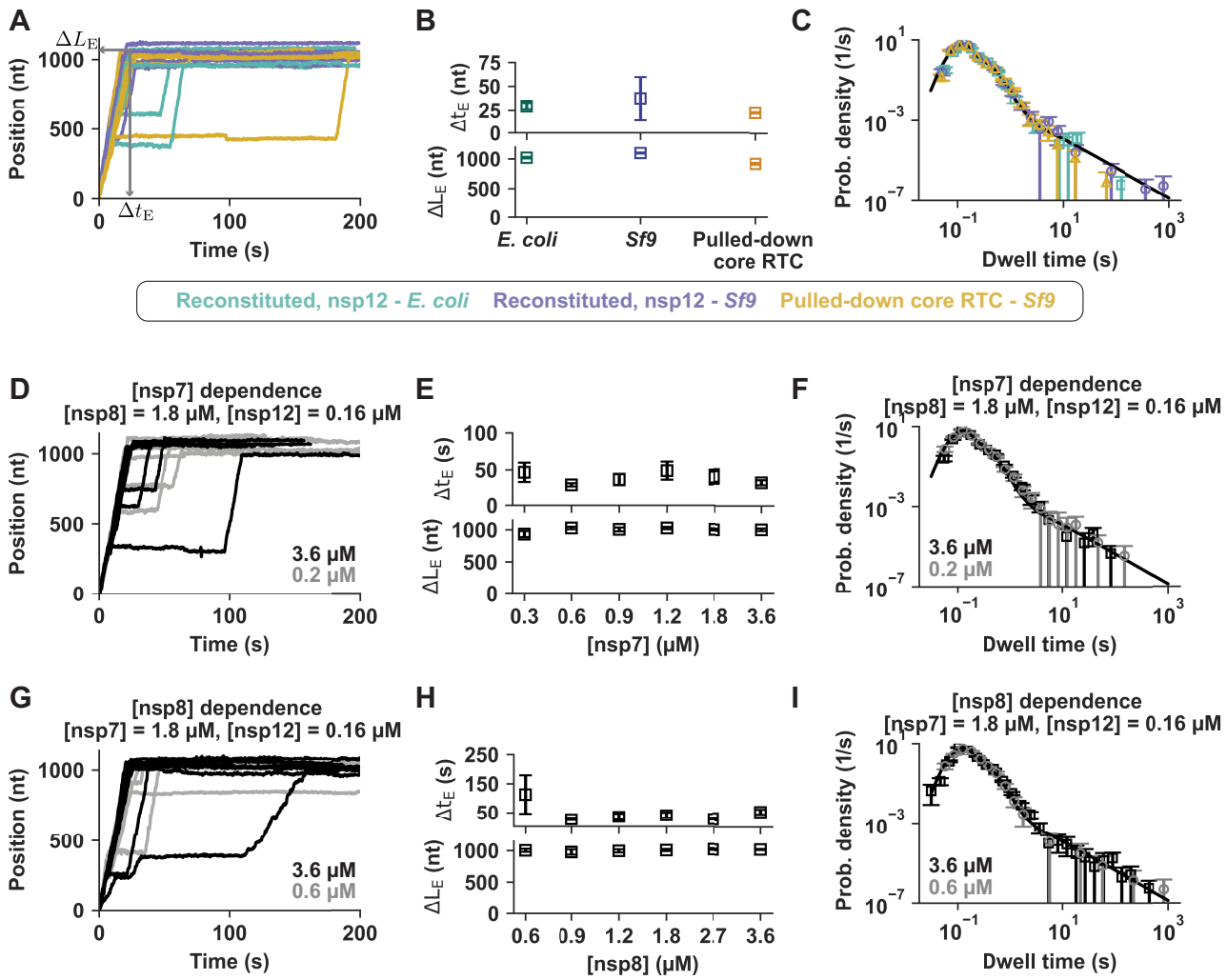


Figure 4. The activation phase and the mode of core RTC proteins expression do not influence the kinetics of the elongating core RTC. (**A**) Example time traces during elongation phase of either reconstituted core RTC with 0.16 μM nsp12-polymerase expressed in *E. coli*, and 1.8 μM nsp7 and nsp8 (turquoise), reconstituted core RTC with 0.2 μM nsp12-polymerase expressed in *Sf9* and 1.8 μM nsp7 and nsp8 (purple), or 2 nM of pulled-down core RTC (yellow). (**B**) Mean elongation times (Δt_E) and product lengths (ΔL_E), and (**C**) dwell time distributions for the entire set of activity traces recorded. The solid line in panel (C) represents the distribution we previously reported in [26]. (**D**) Example time traces during elongation phase of a reconstituted core RTC (nsp12-polymerase expressed in *E. coli*) at 1.8 μM (black) and 0.2 μM (gray) of nsp7. (**E**) Mean elongation times and product lengths, and (**F**) dwell time distributions across nsp7 concentrations. (**G**) Example time traces during elongation of reconstituted core RTC (nsp12-polymerase expressed in *E. coli*) at 1.8 μM (black) and 0.6 μM (gray) of nsp8. (**H**) Mean elongation times and product lengths, and (**I**) dwell time distribution across nsp8 concentrations. All error bars represent 95% confidence intervals determined as described in "Materials and methods" section.

Taken together, the core RTC must undergo a conformational change to initiate RNA synthesis (Fig. 3A and B). This conformational change requires both nsp7 and nsp8 to be present during the assembly process to prevent the complex from being stabilized in an inactive form (Fig. 3C and Supplementary Fig. S4). After successfully having undergone the correct transition, the core RTC is stably locked into an active conformation for several hours (Supplementary Fig. S5).

Dynamics during the elongation phase are independent of those during the activation phase

Next, we set out to determine whether the long lag-times prior to detectible primer-extension in bulk [23, 24] are completely explained by the time required to escape the activation phase. To this end, we examined the dynamics during the ensuing elongation phase. Our single-molecule experiments allow us to directly tell if an increased yield of product RNA is due to more core RTCs successfully entering the elongation phase, or due to an enhanced rate of RNA synthesis. During the elongation phase, the tethered magnetic bead moved downwards (Fig. 1A and C). Converting the height drop to the fraction of ssRNA that has been converted into dsRNA gives rise to the position of the core RTC along the template [34] (Fig. 4A). Comparing the pulled-down and reconstituted core RTCs, we observed no discernible difference amongst their time trajectories (Fig. 4A). Given the rich, stochastic, variation amongst traces within an experiment we quantified the duration of the elongation phase (Δt_E) together with the resulting length of the product (ΔL_E) (Fig. 4A and B). No significant difference in either quantity was detected amongst the three expression systems tested (Fig. 4B). To completely characterize the dynamics of the obtained time traces we further extracted the distribution of times required to add ten consecutive nucleotides [11, 26, 35, 39, 42]. None of the three dwell time distributions shown in Fig. 4C were significantly different from the

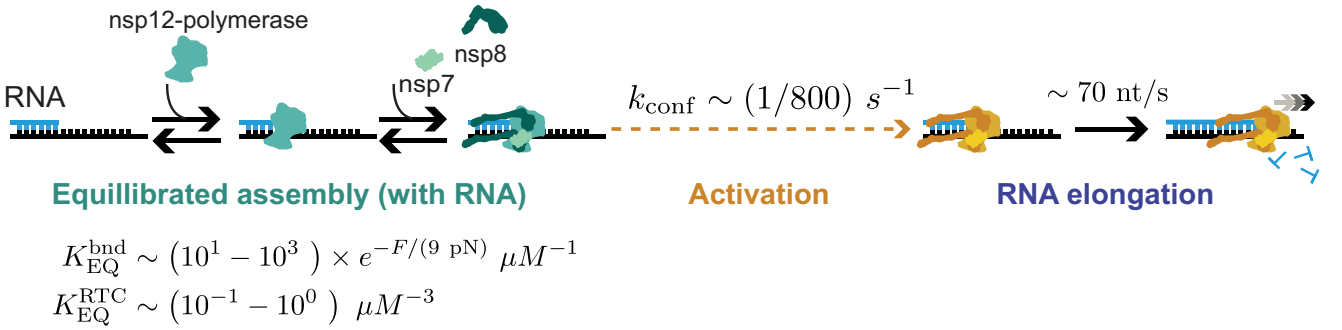


Figure 5. The activation of the core RTC into a processive polymerase is rate-limited by a slow conformational change following assembly. Nsp12-polymerase rapidly binds to the RNA and thereafter recruits the co-factors nsp7 and nsp8, followed by a slow and irreversible conformational change that activates the core RTC into processive elongation. Through fitting our kinetic model (Supplementary Fig. S6 and "Materials and methods" section) to the data shown in Fig. 2 we infer this conformational transition has an intrinsic rate of ~1 in 800 s (k_{conf}). Furthermore, our model shows that binding of the nsp12-polymerase to the RNA is orders of magnitudes tighter (in absence of an applied force) than the binding of the nsps within the core RTC ($(K_{\mathrm{EQ}}^{\mathrm{bnd}})^3 \gg K_{\mathrm{EQ}}^{\mathrm{RTC}}$).

distribution we previously reported to hold true for a (reconstituted) core RTC [26]. The pause-free elongation rate can be estimated by the location of the main peak seen in the histograms [26], and the three of them coincide in Fig. 4C. We found no evidence for any influence of the choice of expression system on the core RTC's dynamics during the elongation phase. As the activation times increased with decreasing co-factor concentration(s) (Fig. 2A and B), we examined the elongation phases under varying nsp7 (Fig. 4D–F) or nsp8 (Fig. 4G–I) concentrations. The concentration of co-factors in solution had no influence on elongation times, product lengths, or the dwell time distributions. All measured distributions coincide with our previously reported one [11, 26].

We conclude that any lag-time before RNA production seen in bulk biochemistry experiments [23, 24] stems entirely from altered dynamics during the activation phase. Our previous model of the elongation phase in the reconstituted system [26] should therefore hold true irrespective of the dynamics of the activation phase.

Activation of the core RTC follows its equilibrated assembly

To further test if the existence of a conformational change after core RTC assembly can quantitatively describe our data we built a kinetic model based on our above findings. Figure 5 shows the predominant assembly and activation pathway of the core RTC towards processive RNA synthesis. The complete reaction pathway underlying the kinetic model that was fitted to the data is shown in Supplementary Fig. S6 ("Materials and methods" section). The core RTC can sequentially assemble on the RNA, starting with nsp12-polymerase, followed by the addition of nsp7 and nsp8 (Fig. 3C and Supplementary Fig. S5). Both co-factors (nsp7 and nsp8) must be present to obtain active core RTCs to any appreciable degree (Supplementary Fig. S4). Furthermore, the core RTC activates within minutes in the presence of RNA, whereas the apo form, i.e. without RNA, requires several hours of incubating the nsps to see rapid activation (Figs 3A and B, and 5).

To find the least complex model that explains our data, we attempted to summarize the activation phase as having a single step that sets its duration. From a modeling perspective, having such a single characteristic timescale implies activation times should be (single-)exponentially distributed. We found such a single-exponential distribution ("Materials and methods" section), with its characteristic timescale being dependent on protein concentrations (Supplementary Fig. S7) and force (Supplementary Fig. S8), to be consistent with the distributions of activation times for individual primer-extension events. Any distribution other than a single-exponential would be the result of a more complex kinetic model, for which our measured distributions provide insufficient evidence (Supplementary Figs S7 and S8). While the path towards a processive core RTC involves multiple intermediates of binding the RNA, assembly of a (intermediate) complex [25], and the conformational change we discover here, our data thus imply a single step sets the activation time (Supplementary Figs S7 and S8). We observed long activation times of several minutes, even at saturating protein concentrations (Fig. 2A–D), indicating that the time determining step occurs after assembly of the core RTC onto the RNA. Hence, we treated both assembly and RNA binding to equilibrate (with equilibrium constants $K_{\mathrm{EQ}}^{\mathrm{bnd}}$ and $K_{\mathrm{EQ}}^{\mathrm{RTC}}$, respectively) on a time scale much shorter than the (effective) activation time, requiring a conformational change occurring at a rate k_{conf} (Fig. 5 and Supplementary Fig. S6, "Materials and methods" section).

It took micromolar of (co-)factors (Fig. 2A–C and Supplementary Fig. S2C–F), while just tens of nanomolar of the pulled-down core RTC was needed to saturate activity (Fig. 2E, and Supplementary Fig. S2I and J). Furthermore, activity can be restored on RNA initially bound by nsp12-polymerase alone by adding new nsp7 and nsp8 (Fig. 3C and Supplementary Fig. S5). Hence, we assume that (i) $K_{\mathrm{EQ}}^{\mathrm{bnd}}$ is independent of the core RTC being activated, and (ii) the core RTC binds much tighter to the RNA (via its nsp12-polymerase) than the complex is held together, i.e. $[\mathrm{nsp12}]_{50} \gg [\mathrm{Pulled-down\ RTC}]_{50}$ at 25 pN (Supplementary Fig. S8, "Materials and methods" section).

The pulled-down core RTC showed rapid elongation upon addition into the flow chamber (Fig. 1E), suggesting this complex had already gone through activation, likely during protein expression. Furthermore, the activated conformation of the core RTC proved to be reasonably stable for over 8 h of pre-incubation (Supplementary Fig. S5). We therefore modelled the activation step as irreversible (Fig. 5 and Supplementary Fig. S6, "Materials and methods" section). Having adopted the elongation-competent conformation, the activation time is only limited by the rate of nu-

cleotide incorporation ($k_{NTP} \sim 70\frac{nt}{s}$ at 25°C, Fig. 4) [26]. The kinetic model that we fitted to the data globally (Fig. 2 and Supplementary Fig. S6, "Materials and methods" section) quantitatively captures all experiments in a unified manner (Supplementary Fig. S9 shows the distribution of fit parameters with fit results within 1% of the best fit, "Materials and methods" section). Although the model allows for multiple cycles of core RTCs binding and activating prior to starting RNA synthesis (Supplementary Fig. S6, "Materials and methods" section), we find a good agreement with our data when the unbinding rate (k_{ub}^{-1} between 1 and 20 s, Supplementary Fig. S9) is much lower than the rate of nucleotide addition (70 s^{-1}, $k_{NTP}^{-1} << k_{ub}^{-1}$). Taken together, following equilibrated assembly (K_{EQ}^{RTC} between 0.06 and 0.99 μM^{-3}, Supplementary Fig. S9) and RNA binding (K_{EQ}^{bnd} between 60 and 980 μM^{-1} in the absence of force, with a characteristic force F_0 between 7.6 and 8.2 pN, Supplementary Fig. S9, "Materials and methods" section), the rate of starting RNA synthesis is truly limited by the time required for the first core RTC to undergo the conformational change (k_{conf}^{-1} between 773 and 800 s, Supplementary Fig. S9).

SARS-CoV-2 core RTC activation remains slow in the presence of nsp13-helicase

During a viral infection, the core RTC associates with additional nsps to form an extended RTC, such as nsp13-helicase [15, 43]. We addressed whether the associated with nsp13-helicase to the core RTC results in a similarly large barrier towards elongation-competence. We performed primer-extension experiments with the three core RTC proteins at saturating concentrations (0.2 μM nsp12-polymerase, 1.8 μM nsp7, and 1.8 μM nsp8), as well as purified nsp13-helicase at the indicated concentration (Supplementary Fig. S10A and B, "Materials and methods" section). Structural studies have reported that two nsp13-helicases bind the core RTC, one of them also binding the template RNA [15]. Accordingly, we see a reduction in activation times upon addition of nsp13-helicase (Supplementary Fig. S10C–E). While an increase in tension lowered the affinity of the core RTC to the RNA (Fig. 2F), addition of nsp13-helicase increased its affinity. We further note that in the presence of saturating amounts of nsp13-helicase, i.e. >10 nM, the duration of the activation phase is still set by one characteristic timescale (Supplementary Fig. S10C). Moreover, the addition of nsp13-helicases only minimally increased the success rate of the primer-extension reaction (Supplementary Fig. S10F). In terms of our kinetic model, the helicase increases both the RNA-binding affinity (K_{EQ}^{bnd}) and transition rate into elongation-competency (k_{conf}). Regardless, an average activation time of ∼360 s remains at a saturating concentration of 20 nM nsp13-helicase (Supplementary Fig. S10C–E). This demonstrates that the time needed for a conformational change identified for the core RTC remains relevant when viewed in the context of the extended complex.

Discussion

The CoV RTC, in particular the "core RTC," fulfills the essential role of replicating the viral genome [13] and is therefore a key target for antiviral drugs. There is a wealth of structural [8, 10, 18, 20] and biochemical studies focused on understanding and inhibiting the CoV polymerase's RNA synthesis capabilities [9, 11, 12, 26, 28, 44]. The assembly of the core RTC was earlier proposed to be the rate-limiting step in the RNA synthesis reaction [23, 24]. The present study reveals that an already assembled core RTC must undergo a conformational change that represents the true rate-limiting step towards processive elongation.

We expanded the single-molecule high-throughput magnetic tweezers assay we previously developed [26] to assess the lag time between injection of the RTC proteins into the flow chamber and the start of RNA synthesis (Fig. 1A). We investigated how varying the stoichiometry and concentration of the core RTC proteins, as well as their mode of expression, impacted both activation and elongation phases. We showed that the approximately ten minutes of activation time (Fig. 1) did not result from either the time required for the core RTC to assemble or to bind the RNA (Fig. 2), but rather from a conformational change within the complex (Figs. 3 and 5).

Despite both reconstituted core RTCs investigated display long activation times that are most likely explained by a conformational change, we note the differently expressed nsp12-polymerases displayed quantitative differences in their activation times. We consistently observe lower activation times using the nsp12-polymerase expressed in *Sf9* (Figs. 1D and 2D). Differences may arise from expression system specific post-translational modifications made to nsp12-polymerase. Interestingly, it was the *E. coli* derived nsp12 that when exposed to its co-factors for sufficient time adopted its optimal conformer (from an activation time standpoint) (Fig. 3A). We note that when expressing nsp12-polymerase in *E. coli*, the gene sequence was codon optimized not for total protein production but to maximize activity in bulk [27] ("Materials and methods" section). Rare codons were intentionally maintained within the coding sequence with the idea of retaining natural pause sites for the ribosome to allow the protein more time to fold into a functional conformer during translation. While clearly not an elongation-competent conformer from the onset (Fig. 1C and D), an altered free-energy landscape for folding the nsp12-polymerase can facilitate access to the elongation-competent conformer. We suspect that the *Sf9* expressed nsp12-polymerase, whose coding sequence was codon-optimized for protein yield, i.e. lacking the rare codons, adopts a conformer capable of activating faster initially but has reduced access to the optimal conformation. This likely explains why pre-incubating the *E. coli* expressed nsp12-polymerase with nsp7 and nsp8 resulted in rapid activation, while repeating the same experiment with the *Sf9* expressed nsp12-polymerase did not (Fig. 3A). It also supports our hypothesis that nsp7 and nsp8 stabilize a conformation in which nsp12-polymerase properly exposes its active site to the RNA template.

Activation of the SARS-CoV-2 core RTC requires the co-factors nsp7 and nsp8 (Figs 3 and 5). Cryogenic-electron microscopy (cryo-EM) studies of apo-core RTC and RNA bound core RTC revealed nsp8 must undergo a large conformational change allowing its N-terminal tails to form rigid contacts ∼28 bp upstream of the active site [18–20]. While this is clearly a conformational change required for RNA synthesis, our data do not support the interpretation that the rotation of the nsp8s constitutes the activation step. Namely, using a pu-

rified core RTC or pre-incubating the individual nsps allowed for rapid activation (Figs 1D and 3B). In these core RTCs, the nsp8 monomers must still unfold their N-terminal tails prior to starting RNA synthesis [17, 19]. Given the nsp8 tails align along the RNA, establishing the contacts should be tension dependent. Furthermore, we saw a reduction in the mean activation time when including nsp13-helicase (Supplementary Fig. S10), suggesting that nsp13-helicase successfully bound the extended nsp8-tails [15]. Finally, varying the tension applied to the template RNA did not reduce the (effective) activation time to seconds (Fig. 2F). Therefore, the activation step we reveal here is a distinct conformational change that, at least in part, rearranges protein domains that either do not interact with the RNA or move orthogonal to the force. We hypothesize that the conformational change is mainly in nsp12-polymerase, with the co-factors stabilizing a conformer in which the RNA is positioned properly for RNA synthesis. In the absence of either nsp7 or nsp8, nsp12-polymerase is forced into a conformation that is inactive and is harder to escape from (Supplementary Fig. S4).

Previous studies have identified viral genomic RNA–polymerase interactions to initiate either replication or transcription in RNA viruses. The dengue virus polymerase (NS5) interacts with the stem loop 5 in the viral genome, which enables its transition into an "elongation complex" [45]. Alphaviruses have a promotor sequence in their genome that activates the RdRp to transcribe a subgenomic RNA encoding for structural proteins [46]. Our results also indicate an important role for the RNA in activating the core RTC, which was orders of magnitudes faster with RNA as opposed to without (Fig. 3). The rich structure of the SARS-CoV-2 genome [47] may contain a motif yet to be discovered to recruit and activate the RTC.

Recent studies demonstrated that a P323L mutation in the nsp12-polymerase, present in most SARS-CoV-2 variants of concern [48], increases the core RTC's primer-extension rate [49] and the viruses' fitness [50]. The mutated residue stabilizes the nsp12/nsp8 interface [49]. Based on our results, we suspect the mutation modulates the RTC's activation time. One may speculate how such a mutation could be beneficial to the virus. Seminal work on CoV murine hepatitis virus showed that stopping translation in infected cells rapidly resulted in a drop of (-)RNA synthesis [51], indicating that a continuous production of nsps was necessary for viral RNA synthesis during infection. Hence, the virus is seemingly assembling and activating new RTCs throughout the infectious cycle. During a viral infection, the extended RTC incorporates additional components beyond the core. These include nsp13-helicases [15], an nsp10–nsp14 complex with exonuclease and methyltransferase activities [16, 52], which may also associate with nsp16 [53], and the capping co-factor nsp9 [27], among others [16]. The experiments of Supplementary Fig. S10 show that the polymerase-helicase complex must still undergo a post-assembly conformational change to become elongation-competent. The extended activation time of the core RTC may enable the recruitment of all essential nsps before replication begins. The association of additional components could influence the kinetics of this activation, a possibility that warrants future investigations.

Our results provide a mechanistic insight into the assembly and activation of the core RTC, establishing a platform to recruit other nsps and assemble a complete and functional CoV RTC to reveal the molecular determinants of CoV replication.

Acknowledgements

We thank Gonzalo Cosa for fruitful discussions.

Author contributions: conceptualisation: D.D, S.C.B, A.D, M.K, C.E.C, and J.J.A formal analysis: M.K, A.D, and M.D investigation: A.D, S.C.B, and M.K resources: B.W, I.A., T.K.A, R.N.K, D.K, H.W.L, M.G, J.C.M, and K.D.R. supervision: D.D writing-original draft: M.K, D.D, and M.D writing – review & editing: all authors.

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

The authors declare that they have no competing interests.

Funding

D.D. was supported by the Interdisciplinary Center for Clinical Research (IZKF) at the University Hospital of the University of Erlangen-Nuremberg, the German Research Foundation grant DFG-DU-1872/4–1, BaSyC – Building a Synthetic Cell" Gravitation grant (024.003.019) of the Netherlands Ministry of Education, Culture and Science (OCW) and the Netherlands Organization for Scientific Research (NWO), NIH fundings R01 AI161841-01, U19 AI171292 and U19 AI171421, and NWO funding OCENW.XL21.XL21.115. IA was supported by NIH funding R01 GM067153. K.D.R. was supported by NIGMS funding R35-GM12260. M.G. was supported by the Alberta Ministry of Technology and Innovation through SPP-ARC (Striving for Pandemic Preparedness – The Alberta Research Consortium). Funding to pay the Open Access publication charges for this article was provided by OCENW.XL21.XL21.115.

Data availability

Data described are available via https://doi.org/10.48338/vu01-cstzpy.

References

1. Beigel JH, Tomashek KM, Dodd LE *et al.* Remdesivir for the Treatment of Covid-19—final report. *N Engl J Med* 2020;**383**:1813–26. https://doi.org/10.1056/NEJMoa2007764
2. Jayk Bernal A, Gomes Da Silva MM, Musungaie DB *et al.* Molnupiravir for oral treatment of Covid-19 in nonhospitalized patients. *N Engl J Med* 2022;**386**:509–20. https://doi.org/10.1056/NEJMoa2116044
3. Sanderson T, Hisner R, Donovan-Banfield I *et al.* A molnupiravir-associated mutational signature in global SARS-CoV-2 genomes. *Nature* 2023;**623**:594–600. https://doi.org/10.1038/s41586-023-06649-6
4. Martinot M, Jary A, Fafi-Kremer S *et al.* Emerging RNA-dependent RNA polymerase mutation in a remdesivir-treated B-cell immunodeficient patient with protracted coronavirus disease 2019. *Clin Infect Dis* 2021;**73**:e1762–5. https://doi.org/10.1093/cid/ciaa1474
5. Stevens LJ, Pruijssers AJ, Lee HW *et al.* Mutations in the SARS-CoV-2 RNA-dependent RNA polymerase confer resistance to remdesivir by distinct mechanisms. *Sci Transl*

Med 2022;**14**:eabo0718.
https://doi.org/10.1126/scitranslmed.abo0718
6. Chien M, Anderson TK, Jockusch S *et al.* Nucleotide analogues as inhibitors of SARS-CoV-2 polymerase, a key drug target for COVID-19. *J Proteome Res* 2020;**19**:4690–7.
https://doi.org/10.1021/acs.jproteome.0c00392
7. Chinthapatla R, Sotoudegan M, Srivastava P *et al.* Interfering with nucleotide excision by the coronavirus 3′-to-5′ exoribonuclease. *Nucleic Acids Res* 2023;**51**:315–36.
https://doi.org/10.1093/nar/gkac1177
8. Feng JY, Du Pont V, Babusis D *et al.* The nucleoside/nucleotide analogs tenofovir and emtricitabine are inactive against SARS-CoV-2. *Molecules* 2022;**27**:4212.
https://doi.org/10.3390/molecules27134212
9. Gordon CJ, Tchesnokov EP, Feng JY *et al.* The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. *J Biol Chem* 2020;**295**:4773–9.
https://doi.org/10.1074/jbc.AC120.013056
10. Malone BF, Perry JK, Olinares PDB *et al.* Structural basis for substrate selection by the SARS-CoV-2 replicase. *Nature* 2023;**614**:781–7.
https://doi.org/10.1038/s41586-022-05664-3
11. Seifert M, Bera SC, van Nies P *et al.* Inhibition of SARS-CoV-2 polymerase by nucleotide analogs from a single-molecule perspective. *eLife* 2021;**10**:e70968.
https://doi.org/10.7554/eLife.70968
12. Tchesnokov EP, Gordon CJ, Woolner E *et al.* Template-dependent inhibition of coronavirus RNA-dependent RNA polymerase by Remdesivir reveals a second mechanism of action. *J Biol Chem* 2020;**295**:16156–65.
https://doi.org/10.1074/jbc.AC120.015720
13. Steiner S, Kratzel A, Barut GT *et al.* SARS-CoV-2 biology and host interactions. *Nat Rev Micro* 2024;**22**:206–25.
https://doi.org/10.1038/s41579-023-01003-z
14. Narwal M, Armache J-P, Edwards TJ *et al.* SARS-CoV-2 polyprotein substrate regulates the stepwise Mpro cleavage reaction. *J Biol Chem* 2023;**299**:206–25.
https://doi.org/10.1016/j.jbc.2023.104697
15. Chen J, Malone B, Llewellyn E *et al.* Structural basis for helicase-polymerase coupling in the SARS-CoV-2 replication-transcription complex. *Cell* 2020;**182**:1560–73.e13.
https://doi.org/10.1016/j.cell.2020.07.033
16. Yan L, Yang Y, Li M *et al.* Coupling of N7-methyltransferase and 3′-5′ exoribonuclease with SARS-CoV-2 polymerase reveals mechanisms for capping and proofreading. *Cell* 2021;**184**:3474–85.
https://doi.org/10.1016/j.cell.2021.05.033
17. Gao Y, Yan L, Huang Y *et al.* Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* 2020;**368**:779–82.
https://doi.org/10.1126/science.abb7498
18. Hillen HS, Kokic G, Farnung L *et al.* Structure of replicating SARS-CoV-2 polymerase. *Nature* 2020;**584**:154–6.
https://doi.org/10.1038/s41586-020-2368-8
19. Kirchdoerfer RN, Ward AB. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nat Commun* 2019;**10**:2342.
https://doi.org/10.1038/s41467-019-10280-3
20. Wang Q, Wu J, Wang H *et al.* Structural basis for RNA replication by the SARS-CoV-2 polymerase. *Cell* 2020;**182**:417–28.
https://doi.org/10.1016/j.cell.2020.05.034
21. Anderson TK, Hoferle PJ, Chojnacki KJ *et al.* An alphacoronavirus polymerase structure reveals conserved replication factor functions. *Nucleic Acids Res* 2024;**52**:5975–86.
https://doi.org/10.1093/nar/gkae153
22. Biswal M, Diggs S, Xu D *et al.* Two conserved oligomer interfaces of NSP7 and NSP8 underpin the dynamic assembly of SARS-CoV-2 RdRP. *Nucleic Acids Res* 2021;**49**:5956–66.
https://doi.org/10.1093/nar/gkab370
23. Campagnola G, Govindarajan V, Pelletier A *et al.* The SARS-CoV nsp12 polymerase active site is tuned for large-genome replication. *J Virol* 2022;**96**:e0067122. https://doi.org/10.1128/jvi.00671-22
24. Subissi L, Posthuma CC, Collet A *et al.* One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. *Proc Natl Acad Sci USA* 2014;**111**:E3900–9.
https://doi.org/10.1073/pnas.1323705111
25. Wilamowski M, Hammel M, Leite W *et al.* Transient and stabilized complexes of Nsp7, Nsp8, and Nsp12 in SARS-CoV-2 replication. *Biophys J* 2021;**120**:3152–65.
https://doi.org/10.1016/j.bpj.2021.06.006
26. Bera SC, Seifert M, Kirchdoerfer RN *et al.* The nucleotide addition cycle of the SARS-CoV-2 polymerase. *Cell Rep* 2021;**36**:109650.
https://doi.org/10.1016/j.celrep.2021.109650
27. Wang B, Svetlov V, Wolf YI *et al.* Allosteric activation of SARS-CoV-2 RNA-dependent RNA polymerase by remdesivir triphosphate and other phosphorylated nucleotides. *mBio* 2021;**12**:e0142321. https://doi.org/10.1128/mbio.01423-21
28. Gordon CJ, Lee HW, Tchesnokov EP *et al.* Efficient incorporation and template-dependent polymerase inhibition are major determinants for the broad-spectrum antiviral activity of remdesivir. *J Biol Chem* 2022;**298**:101529.
https://doi.org/10.1016/j.jbc.2021.101529
29. Berger I, Fitzgerald DJ, Richmond TJ. Baculovirus expression system for heterologous multiprotein complexes. *Nat Biotechnol* 2004;**22**:1583–7. https://doi.org/10.1038/nbt1036
30. Bieniossek C, Richmond TJ, Berger I. MultiBac: multigene baculovirus-based eukaryotic protein complex production. *CP Protein Science* 2008;**51**:5.20.1–5.20.26.
https://doi.org/10.1002/0471140864.ps0520s51
31. Papini FS, Seifert M, Dulin D. High-yield fabrication of DNA and RNA constructs for single molecule force and torque spectroscopy experiments. *Nucleic Acids Res* 2019;**47**:e144.
https://doi.org/10.1093/nar/gkz851
32. Quack S, Dulin D. Surface functionalization, nucleic acid tether characterization, and force calibration for a magnetic tweezers assay. *Methods Mol Biol* 2024;**2694**:403–20.
https://doi.org/10.1007/978-1-0716-3377-9_19
33. Cnossen JP, Dulin D, Dekker NH. An optimized software framework for real-time, high-throughput tracking of spherical beads. *Rev Sci Instrum* 2014;**85**:103712.
https://doi.org/10.1063/1.4898178
34. Dulin D, Vilfan ID, Berghuis BA *et al.* Elongation-competent pauses govern the fidelity of a viral RNA-dependent RNA Polymerase. *Cell Rep* 2015;**10**:983–92.
https://doi.org/10.1016/j.celrep.2015.01.031
35. Dulin D, Arnold JJ, van Laar T *et al.* Signatures of nucleotide analog incorporation by an RNA-dependent RNA polymerase revealed using high-throughput magnetic tweezers. *Cell Rep* 2017;**21**:1063–76.
https://doi.org/10.1016/j.celrep.2017.10.005
36. Eslami-Mosallam B, Katechis I, Depken M . 2019; Fitting in the age of single-molecule experiments: a guide to maximum-likelihood estimation and its advantages. In: Joo C, Rueda D (eds.), *Biophysics of RNA-Protein Interactions, Biological and Medical Physics, Biomedical Engineering*. New York, NY: Springer New York, 85–105.
37. Kirkpatrick S, Gelatt CD, Vecchi MP. Optimization by simulated annealing. *Science* 1983;**220**:671–80.
https://doi.org/10.1126/science.220.4598.671
38. Eslami-Mosallam B, Klein M, Smagt CVD *et al.* A kinetic model predicts SpCas9 activity, improves off-target classification, and reveals the physical basis of targeting fidelity. *Nat Commun* 2022;**13**:1367.
https://doi.org/10.1038/s41467-022-28994-2

39. Dulin D, Berghuis BA, Depken M *et al*. Untangling reaction pathways through modern approaches to high-throughput single-molecule force-spectroscopy experiments. *Curr Opin Struct Biol* 2015;**34**:116–22. https://doi.org/10.1016/j.sbi.2015.08.007
40. Dulin D . 2024; An introduction to magnetic tweezers. In Heller I, Dulin D, Peterman EJG (eds.), *Single Molecule Analysis : Methods and Protocols, Methods in Molecular Biology*. New York, NY: Springer US, 375–401. https://doi.org/10.1007/978-1-0716-3377-9_18
41. Ostrofet E, Papini FS, Dulin D. Correction-free force calibration for magnetic tweezers experiments. *Sci Rep* 2018;**8**:15920. https://doi.org/10.1038/s41598-018-34360-4
42. Dulin D, Vilfan ID, Berghuis BA *et al*. Backtracking behavior in viral RNA-dependent RNA polymerase provides the basis for a second initiation site. *Nucleic Acids Res* 2015;**43**:10421–9. https://doi.org/10.1093/nar/gkv1098
43. Yan L, Zhang Y, Ge J *et al*. Architecture of a SARS-CoV-2 mini replication and transcription complex. *Nat Commun* 2020;**11**:5874. https://doi.org/10.1038/s41467-020-19770-1
44. Maheden K, Todd B, Gordon CJ *et al*. Inhibition of viral RNA-dependent RNA polymerases with clinically relevant nucleotide analogs. *Enzymes* 2021;**49**:315–54. https://doi.org/10.1016/bs.enz.2021.07.002
45. Osawa T, Aoki M, Ehara H *et al*. Structures of dengue virus RNA replicase complexes. *Mol Cell* 2023;**83**:2781–91. https://doi.org/10.1016/j.molcel.2023.06.023
46. Strauss JH, Strauss EG. The alphaviruses: gene expression, replication, and evolution. *Microbiol Rev* 1994;**58**:491–562. https://doi.org/10.1128/mr.58.3.491-562.1994
47. Cao C, Cai Z, Xiao X *et al*. The architecture of the SARS-CoV-2 RNA genome inside virion. *Nat Commun* 2021;**12**:3917. https://doi.org/10.1038/s41467-021-22785-x
48. Kim S-M, Kim E-H, Casel MAB *et al*. SARS-CoV-2 variants with NSP12 P323L/G671S mutations display enhanced virus replication in ferret upper airways and higher transmissibility. *Cell Rep* 2023;**42**:113077. https://doi.org/10.1016/j.celrep.2023.113077
49. Ferrer-Orta C, Vázquez-Monteagudo S, Ferrero DS *et al*. Point mutations at specific sites of the nsp12–nsp8 interface dramatically affect the RNA polymerization activity of SARS-CoV-2. *Proc Natl Acad Sci USA* 2024;**121**:e2317977121. https://doi.org/10.1073/pnas.2317977121
50. Goldswain H, Dong X, Penrice-Randal R *et al*. The P323L substitution in the SARS-CoV-2 polymerase (NSP12) confers a selective advantage during infection. *Genome Biol* 2023;**24**:47. https://doi.org/10.1186/s13059-023-02881-5
51. Sawicki SG, Sawicki DL. Coronavirus minus-strand RNA synthesis and effect of cycloheximide on coronavirus RNA synthesis. *J Virol* 1986;**57**:328–34. https://doi.org/10.1128/jvi.57.1.328-334.1986
52. Moeller NH, Shi K, Demir Ö *et al*. Structure and dynamics of SARS-CoV-2 proofreading exoribonuclease ExoN. *Proc Natl Acad Sci USA* 2022;**119**:e2106379119. https://doi.org/10.1073/pnas.2106379119
53. Matsuda A, Plewka J, Rawski M *et al*. Despite the odds: formation of the SARS-CoV-2 methylation complex. *Nucleic Acids Res* 2024;**52**:6441–6458. https://doi.org/10.1093/nar/gkae165

Received: October 1, 2024. **Revised:** April 14, 2025. **Editorial Decision:** May 6, 2025. **Accepted:** May 16, 2025