Article

Structural basis for substrate selection by the SARS-CoV-2 replicase

https://doi.org/10.1038/s41586-022-05664-3

Received: 9 May 2022

Accepted: 15 December 2022

Published online: 01 February 2023

Check for updates

Brandon F. Malone[1], Jason K. Perry[2], Paul Dominic B. Olinares[3], Hery W. Lee[4], James Chen[1,8], Todd C. Appleby[2], Joy Y. Feng[2], John P. Bilello[2], Honkit Ng[5], Johanna Sotiris[5], Mark Ebrahim[5], Eugene Y. D. Chua[6], Joshua H. Mendez[6], Ed T. Eng[6], Robert Landick[7], Matthias Götte[4], Brian T. Chait[3], Elizabeth A. Campbell[1]✉ & Seth A. Darst[1]✉

The SARS-CoV-2 RNA-dependent RNA polymerase coordinates viral RNA synthesis as part of an assembly known as the replication–transcription complex (RTC)[1]. Accordingly, the RTC is a target for clinically approved antiviral nucleoside analogues, including remdesivir[2]. Faithful synthesis of viral RNAs by the RTC requires recognition of the correct nucleotide triphosphate (NTP) for incorporation into the nascent RNA. To be effective inhibitors, antiviral nucleoside analogues must compete with the natural NTPs for incorporation. How the SARS-CoV-2 RTC discriminates between the natural NTPs, and how antiviral nucleoside analogues compete, has not been discerned in detail. Here, we use cryogenic-electron microscopy to visualize the RTC bound to each of the natural NTPs in states poised for incorporation. Furthermore, we investigate the RTC with the active metabolite of remdesivir, remdesivir triphosphate (RDV-TP), highlighting the structural basis for the selective incorporation of RDV-TP over its natural counterpart adenosine triphosphate[3,4]. Our results explain the suite of interactions required for NTP recognition, informing the rational design of antivirals. Our analysis also yields insights into nucleotide recognition by the nsp12 NiRAN (nidovirus RdRp-associated nucleotidyltransferase), an enigmatic catalytic domain essential for viral propagation[5]. The NiRAN selectively binds guanosine triphosphate, strengthening proposals for the role of this domain in the formation of the 5′ RNA cap[6].

COVID-19, caused by the coronavirus SARS-CoV-2, continues to devastate livelihoods and overwhelm healthcare systems worldwide. Given the urgency of stymieing infection and mitigating disease morbidity, a concerted research effort has begun to explain the molecular details of the viral life cycle and to design therapeutics to disrupt it[1]. The viral RNA-dependent RNA polymerase (RdRp, encoded by non-structural protein 12, or nsp12) functions as part of the holo-RdRp (comprising the RdRp and auxiliary proteins nsp7 and nsp8 as a heterotetramer nsp7/nsp8$_2$/nsp12) in a replication–transcription complex (holo-RdRp + RNA, or RTC) to direct all viral RNA synthesis in conjunction with a group of viral nucleic acid processing enzymes. In less than 2 years, targeting the RTC by the antivirals remdesivir (RDV) and molnupiravir has become an important aspect of clinical care[2,7]. This clinical success underscores the importance of the RTC as a pharmacological target and incentivizes the design of more efficacious nucleotide analogues for the treatment of COVID-19. Furthermore, because of the highly conserved nature of the coronavirus RdRp active site[8], nucleotide analogue inhibitors of the RTC are excellent candidates for pan-viral inhibitors that would be effective against emerging variants and may be repurposed to tackle future pathogenic coronaviruses that could arise.

Biochemical and single-molecule experiments have characterized the RNA synthesis activity of the SARS-CoV-2 RTC, its selectivity for a wide range of nucleotide analogues and their effects on RNA synthesis[3,4,9,10]. These investigations yielded insights into the ability of the RTC to discriminate between natural NTPs and related antiviral analogues. Studies of remdesivir (RDV) revealed that the active metabolite of RDV, RDV triphosphate (RDV-TP), possesses enhanced selectivity for incorporation into nascent RNA over its natural counterpart, adenosine triphosphate (ATP)[3,4]. Rapid, competitive incorporation is probably a critical facet of the clinical effectiveness of RDV. However, the structural basis for NTP recognition and RDV-TP selectivity remains unknown, necessitating structural examinations of stalled RTC ternary complexes with incoming NTP substrates and catalytic metal ions poised for catalysis (Michaelis or pre-incorporation complexes) (Fig. 1a).

Stepwise nucleotide incorporation necessitates transferring the 3′-end of the product-RNA (p-RNA) to the post-translocated site, allowing binding of the incoming NTP substrate along with two catalytic metal ions in the active-site cleft (Fig. 1a)[11–14]. In the subsequent catalytic step(s), the geometry of correct Watson–Crick base pairing of the incoming NTP with the template-RNA (t-RNA) base closes the active site. This conformational change promotes nucleophilic

[1]Laboratory of Molecular Biophysics, The Rockefeller University, New York, NY, USA. [2]Gilead Sciences, Inc., Foster City, CA, USA. [3]Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, New York, NY, USA. [4]Department of Medical Microbiology and Immunology, University of Alberta, Edmonton, Alberta, Canada. [5]The Evelyn Gruss Lipper Cryo-Electron Microscopy Resource Center, The Rockefeller University, New York, NY, USA. [6]National Center for Cryo-EM Access and Training, Simons Electron Microscopy Center, New York Structural Biology Center, New York, NY, USA. [7]Department of Biochemistry, University of Wisconsin-Madison, Madison, WI, USA. [8]Present address: Department of Cell Biology, New York University School of Medicine, New York, NY, USA. ✉e-mail: campbee@rockefeller.edu; darst@rockefeller.edu

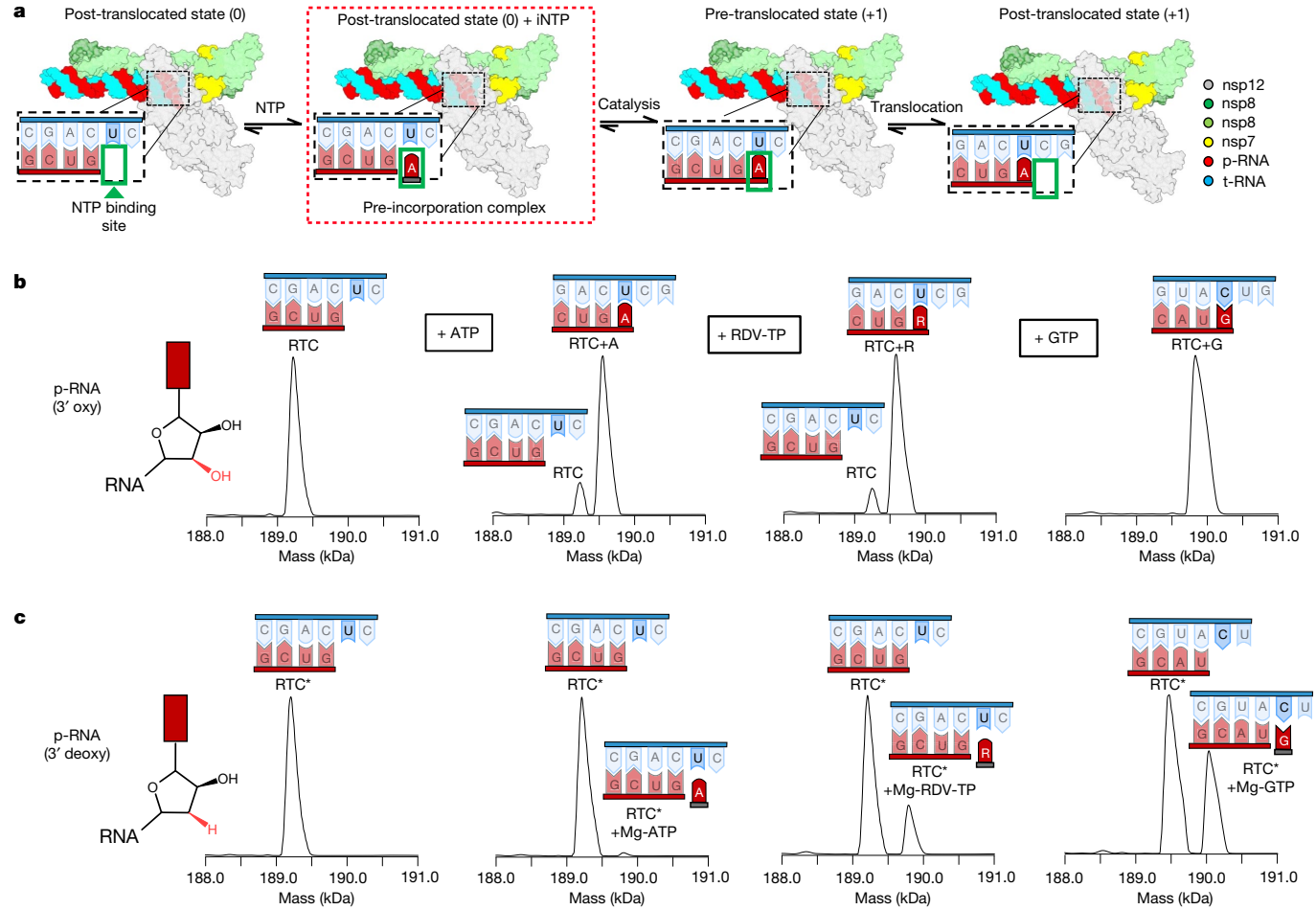


Fig. 1 | Capturing the SARS-CoV-2 RTC ternary complex. a, Schematic depicting the major steps of the nucleotide addition cycle of the RTC. The pre-incorporation complex studied here is highlighted (red dashed box). **b**, nMS analysis of the RTC bound to a 3′ oxy p-RNA in the absence and presence of 300 μM ATP, RDV-TP or GTP, respectively. **c**, Similar nMS analysis to **b** but the RTC was reconstituted using a 3′-deoxy p-RNA (RTC*). p-RNA and t-RNA schematics were created with BioRender.com.

attack of the p-RNA 3′-OH on the NTP α-phosphate, releasing a pyrophosphate in a S_N2 condensation reaction[11]. Available SARS-CoV-2 RTC structures are dominated by apo structures with an empty NTP active site lacking catalytic metals and with an open active-site conformation, hindering structure-based drug design[15–18]. Only two studies document the RTC in the presence of substrates: the antiviral inhibitors favipiravir triphosphate and the activated form of AT-527, AT-9010, but neither investigation yielded views of a catalytically competent RTC because of unproductive binding modes and open conformations of the active sites[19,20]. Therefore, to gain mechanistic insights into RTC recognition of its substrates during RNA synthesis, we determined five cryogenic-electron microscopy (cryo-EM) structures, capturing the RTC with each of the four natural NTPs as well as with RDV-TP. The five cryo-EM maps range in nominal resolutions between roughly 2.6 and 3.3 Å, with the RdRp active sites resolved locally to roughly 2.2–2.9 Å (Extended Data Figs. 1–5 and Supplementary Table 1)[21], enabling near-atomic resolution insights into the mechanism of NTP recognition and RDV-TP selectivity.

Trapping RTC pre-incorporation complexes

Guided by previous mechanistic work on nucleic acid polymerases, we investigated a battery of chemical strategies to block incorporation after substrate binding. These approaches included using nucleotide diphosphate (NDP) substrates[12], alternative metal cofactors (such as Ca^{2+})[22], α-β non-hydrolysable nucleotide analogues[23] and 3′-deoxy p-RNA scaffolds. Although the use of NDP substrates trapped the hepatitis C virus (HCV) RdRp recognition complex[12], the SARS-CoV-2 RTC incorporated guanosine diphosphate (GDP), adenosine diphosphate (ADP) and to a minor extent RDV-diphosphate (RDV-DP) under assay conditions similar to the cryo-EM sample preparation (Extended Data Fig. 6, see also ref. [24]). The ability of the SARS-CoV-2 RTC to incorporate NDPs reflects the relative promiscuity of the RdRp active site, which also retained RNA synthesis activity in the presence of both Ca^{2+} and α-β-imino analogues (Extended Data Fig. 6e). However, RTC synthesis activity was fully ablated with the incorporation of a 3′-deoxy nucleotide into the p-RNA, leading us to use this approach to visualize the RTC pre-incorporation complexes (Extended Data Fig. 6g).

To validate the effectiveness of using 3′-deoxy RNA scaffolds, we used native mass spectrometry (nMS) to probe for the extension of the primer by one base or the formation of a stalled ternary complex (Fig. 1b,c and Supplementary Fig. 2) in the presence of various NTPs. In the event of nucleotide incorporation, we expected a mass-shift corresponding to nucleotide monophosphate addition, whereas non-covalently bound (unincorporated) substrates would either dissociate or yield a mass-shift corresponding to the NTP and bound metal ions. For the control scaffolds containing a 3′-OH, near complete p-RNA extension by a single nucleotide monophosphate occurred (Fig. 1b). In the presence of p-RNA strands lacking a 3′-OH, no incorporation of the next nucleotide (ATP, RDV-TP or guanosine triphosphate (GTP))

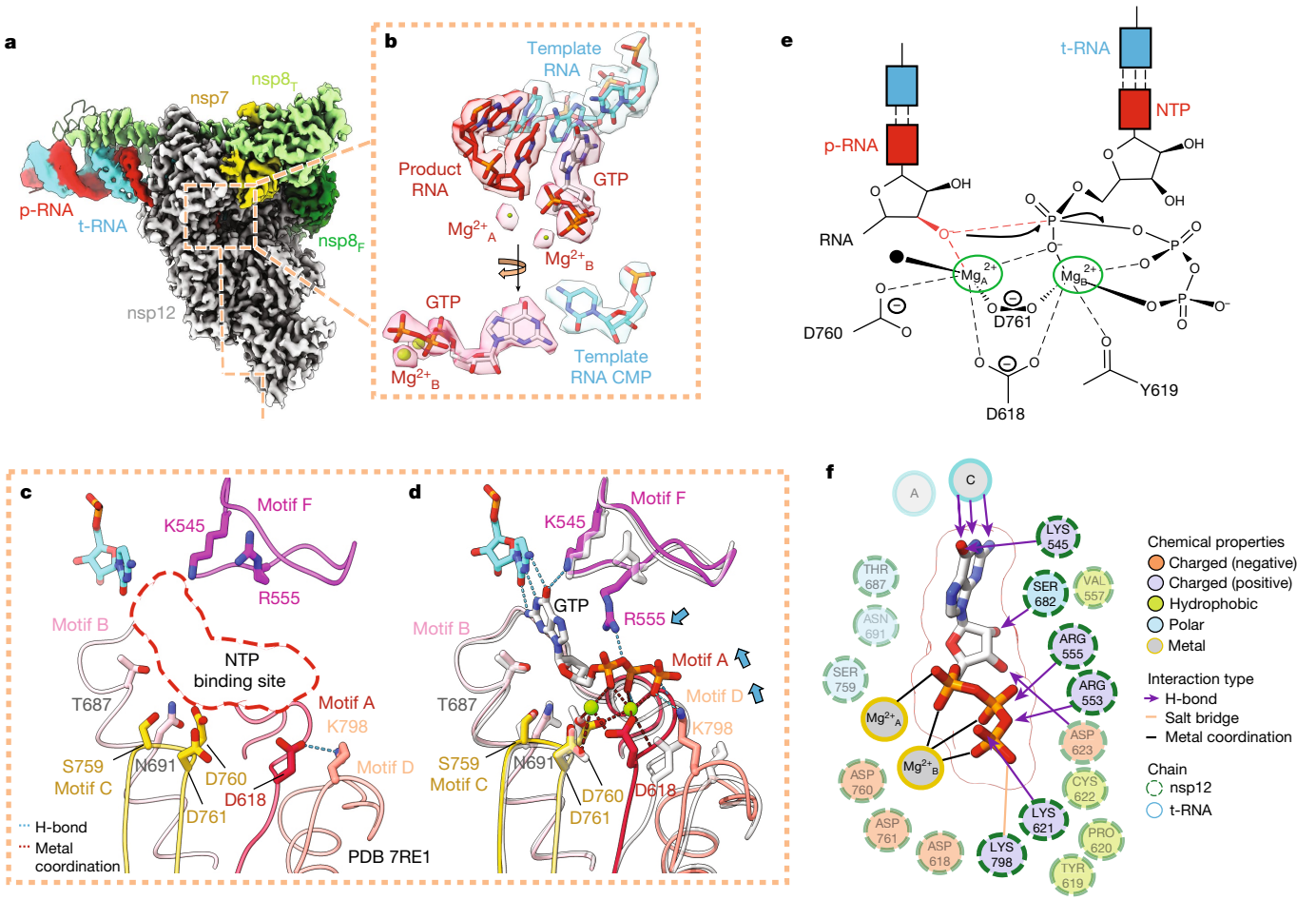


Fig. 2 | Structural basis of nucleotide recognition. a, Cryo-EM density of the 2.7 Å nominal resolution +GTP structure (S4_GTP), coloured according to the fitted model chains. **b**, Cryo-EM density of the bound incoming GTP, two associated metal ions and the nearby bases of the t-RNA and p-RNA strands. **c**, Close-up of the active site of the apo complex PDB 7RE1 (ref. [18]), illustrating the arrangement of the conserved RdRp active site motifs A–D and F (crimson, hot pink, gold, salmon and magenta, respectively) and of key residues when the NTP binding site (dashed red line) is empty. **d**, Comparison of the S4_GTP and apo structure (faded grey) active sites, highlighting observed motif/residue rearrangements (blue arrows) on NTP binding. Movements in motifs A and D close the active site. **e**, Schematic depicting nucleotide addition, in the presence of an intact 3′-OH, based on disposition of magnesiums and NTP in the S4_GTP structure. **f**, 2D schematic of the suite of interactions involved in NTP binding based on distances in the S4_GTP structure. Residues interacting indirectly or weakly with the NTP are faded for clarity.

was observed, confirming the suitability of this strategy for trapping pre-incorporation complexes (Fig. 1c). In the pre-incorporation complexes, we observed peaks corresponding to the RTC with the incoming Mg-NTP (ATP, RDV-TP or GTP) with the relative peak intensity in the order of GTP > RDV-TP > ATP (Fig. 1c).

Structural basis of NTP recognition

Having identified a strategy to capture substrate-bound RTCs stalled in a pre-incorporation state, we prepared cryo-EM samples of RTCs with each of the natural nucleotides, ATP, cytidine triphosphate (CTP), GTP, uridine triphosphate (UTP) or with RDV-TP (Extended Data Figs. 1–5 and Supplementary Table 2). The resulting structures yielded the first views of the RTC poised for catalysis. Two of the structures, the RTC with CTP (+CTP) or GTP (+GTP), yielded nominal resolutions better than 3 Å (2.7 Å), with local resolution around the active site estimated to be roughly 2.2 Å (ref. [21]), enabling near-atomic resolution insights into substrate recognition.

Density features in the +CTP and +GTP cryo-EM maps are consistent with the positions of divalent metal ions observed in many previous nucleic acid polymerase structures[11–13,25]. On the basis of the presence of 5 mM $MgCl_2$ in the cryo-EM buffer and the observed octahedral coordination geometries, we assigned these densities as Mg^{2+}_A and Mg^{2+}_B (Fig. 2b and Supplementary Video 1), consistent with the hypothesized universal two-metal ion mechanism for nucleic acid polymerases[11]. The density peak for Mg^{2+}_A in the +CTP and +GTP structures is weaker than that for Mg^{2+}_B, and density for Mg^{2+}_A in the +ATP, +UTP and +RDV-TP structures is completely absent. Despite the absence of Mg^{2+}_A in the +ATP, +UTP and +RDV-TP structures, most details of the active-site configurations were indistinguishable from the +CTP and +GTP structures.

A consistent feature observed across all five structures is the closure of the active site around the incoming NTP, mediated by a rotation of the RdRp motif A towards the NTP substrate (Fig. 2c,d). Closure of motif A stabilizes substrate binding by: (1) enabling the backbone carbonyl of Y619 to coordinate Mg^{2+}_B and the catalytic residue D618 to coordinate both Mg^{2+}_A and Mg^{2+}_B (Fig. 2e), (2) promoting the formation of a hydrogen-bonding (H-bonding) network through D623 that enables the recognition of the substrate ribose 2′-OH by S682, T687 and N691 (motif B), (3) enabling weak H-bonding interactions between the β- and γ-phosphates and motif A residues K621 or C622 (Fig. 2f, and Extended Data Fig. 7f) and (4) disrupting the polar D618–K798 interaction in the apo-RTC[15,16,18], repositioning D618 for metal coordination and K798 to interact with the NTP γ-phosphate (Fig. 2d).

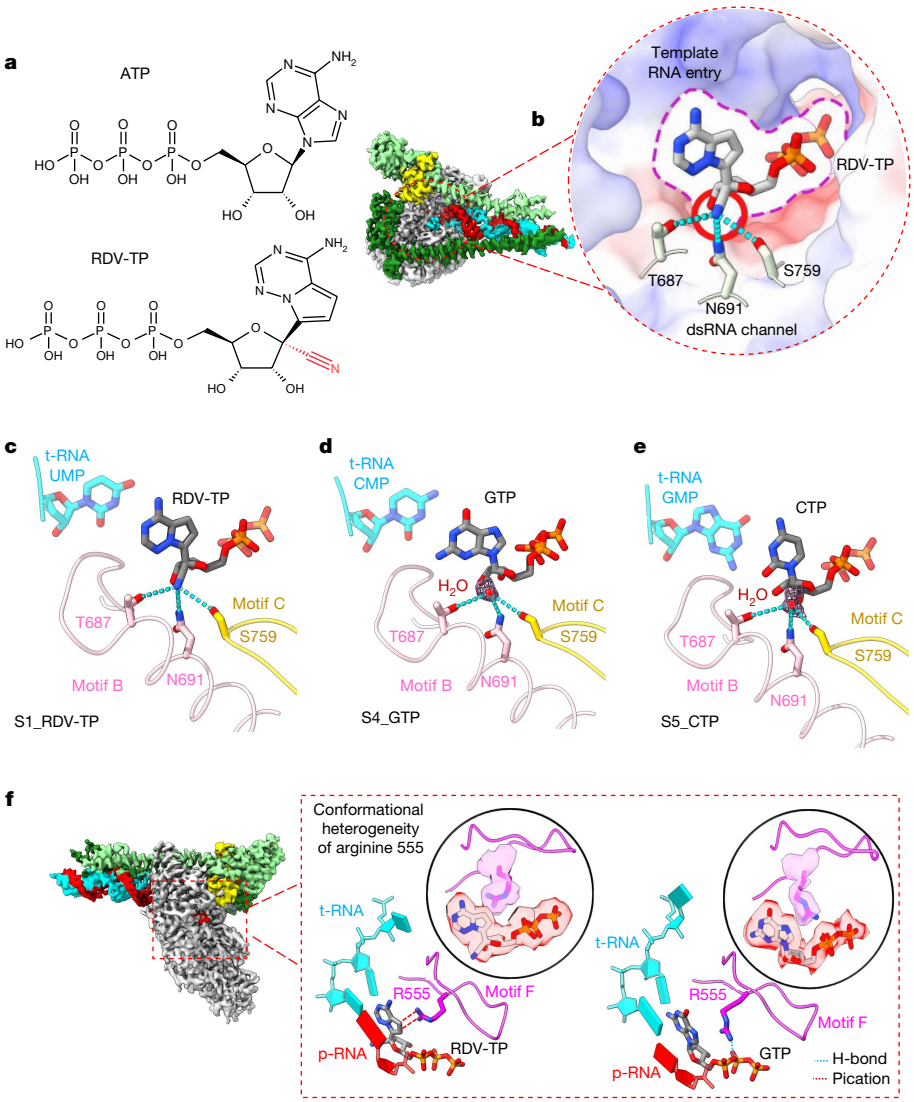


Fig. 3 | Molecular basis of remdesivir's incorporation selectivity. a, Chemical structures of ATP and RDV-TP, highlighting the position of the RDV-TP 1′ cyano group. **b**, Cryo-EM densities of the S2_RDV-TP structure, coloured according to fitted model chains. Zoom-in on the bound RDV-TP illustrates the RDV-TP 1′ cyano group is accommodated in a hydrophilic pocket formed by motif B and C residues. Protein surface is coloured according to electrostatics. **c**–**e**, Comparison of the active sites of the S1_RDV-TP (**c**), S4_GTP (**d**) and S5_CTP (**e**) structures reveals that the RdRp cyano pocket can also bind a water molecule (map density around the water shown in mesh), which needs to be displaced for RDV-TP binding. **f**, A comparison of the S2_RDV-TP and S4_GTP structures reveals two predominant rotamers of R555 that mediate either a pi–pi stacking (red dotted lines) or a H-bond (blue dotted line) interaction with the incoming NTP.

The detailed interactions of nsp12 residues with the four natural NTP substrates were essentially identical except for conserved basic residues of motif F. K545 (motif F) shows selective interactions with the incoming NTPs, forming a H-bond with carbonyl oxygens of UTP (C4) and GTP (C6), interactions that are precluded by the amino group at these positions in the ATP, RDV-TP or CTP bases (Extended Data Fig. 7). The cryo-EM densities also suggest that the side chains of K551 and R555 are dynamic, accessing several conformations in each structure. K551 primarily interacts with the γ-phosphate of the CTP and RDV-TP substrates, but weak density indicates disorder in the other structures. In the +ATP and +RDV-TP structures, R555 predominantly forms a pi–pi stacking interaction with the NTP substrate base. This differs from the predominant conformation in the +CTP, +GTP and +UTP structures where the side chain of R555 forms a H-bond with the β-phosphate oxygens of the NTP substrate (Fig. 3 and Extended Data Fig. 7). In either state, R555 buttresses the incoming nucleotide, arranging the active site for catalysis. We speculate that the observed R555 stacking interaction may first promote the formation of a canonical Watson–Crick base pair by stabilizing the nucleobase opposite the template base, then reorient to interact with the nucleotide β-phosphate to promote subsequent catalysis by reinforcing the correct geometry of the pyrophosphate leaving group, roles that may be conserved across RdRps such as HCV RdRp[26].

The structural observation of SARS-CoV-2 RdRp motif A closure resembles a similar motion of motif A observed in substrate-bound RdRp complexes of polio virus, norovirus and HCV[12–14], illustrating the universal nature of the RdRp palm closure step during NTP recognition. Although the palm-mediated active site closure is well-documented, our results indicate that the structural plasticity of positively charged residues in motif F (K545, K551, R553 and R555) in the fingers domain helps orient the NTP substrate for catalysis through interactions with both the nucleobase and triphosphate moieties. We note that these residues are invariant across the major coronavirus clades and GISAID database of patient SARS-CoV-2 sequences (Extended Data Fig. 8), reflecting the near immutable nature of the coronavirus RdRp active

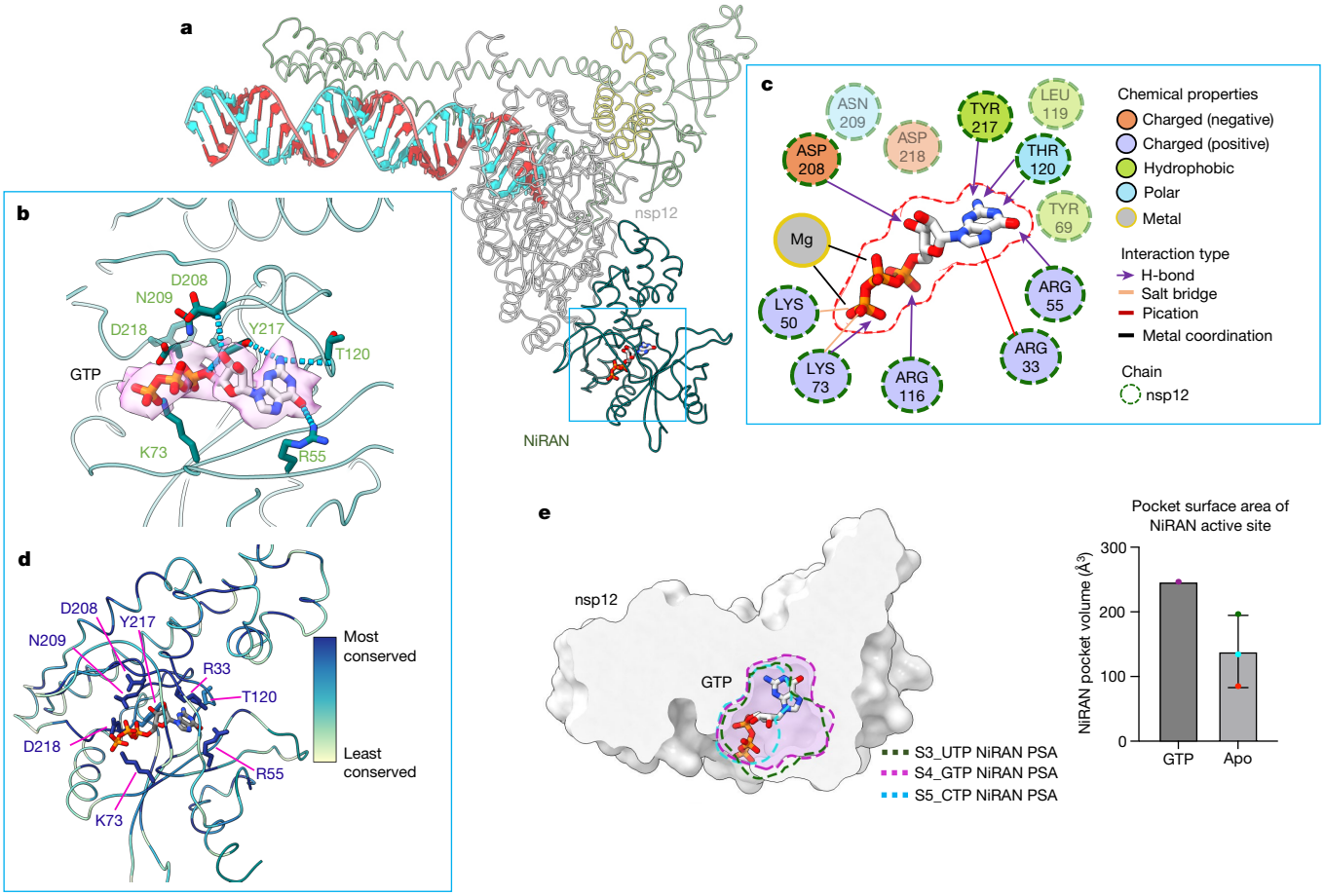


Fig. 4 | NiRAN specific recognition of GTP. a, View of the NiRAN domain of the RTC, bound to GTP, which lies at the amino-terminal end of nsp12. **b**, GTP is selectively recognized in the NiRAN pocket by a series of hydrogen-bonding and electrostatic interactions in which interacting residues are shown as sticks. **c**, 2D schematic illustrating the NiRAN-GTP interactions. **d**, Residues that mediate GTP recognition are shown as sticks and coloured according to their conservation across the α and β coronavirus clades. **e**, Binding of GTP in the NiRAN is mediated through an induced fit that widens the pocket for insertion of the guanine base. Cross-sections (dashed lines) of the pocket surface area (PSA) of the S4_GTP (GTP bound), S3_UTP and S5_CTP (apo-NiRAN) are shown overlaid on a clipped surface view of the nsp12 NiRAN. The NiRAN pocket volume was measured using the Schrodinger Sitemap tool. The mean ± s.d. of the pocket volumes of the S1_RDV-TP, S3_UTP and S5_CTP structures (apo-NiRAN, $n = 3$) is shown. The estimated volume sizes of the datasets are shown as overlaid circles in which the green, cyan and red circles represent the calculated PSA for the S3, S5 and S1 structures, respectively.

site. Previously, we noted that the same residues line the NTP entry channel following nsp13-mediated p-RNA backtracking[27], illustrating the pleiotropic role that motif F plays during RNA synthesis.

Structural basis of RDV-TP selectivity

RDV-TP has been characterized to possess a twofold to threefold greater selectivity for incorporation into elongating RNA compared to ATP[3,4], a property that is thought to be critical for its clinical efficacy. In our structures, both nucleotides are observed base-paired to the cognate t-RNA U; the two nucleotides superimpose with a root-mean-square-deviation of 0.34 Å over 32 common atoms. The key difference between RDV-TP and ATP is the 1′-cyano moiety on the RDV-TP ribose (Fig. 3a), which juts into a hydrophilic pocket formed by residues T687, N691 (motif B) and S759 (motif C) (Fig. 3c), creating a network of polar interactions (Fig. 3c). In the highest resolution cryo-EM maps (+CTP and +GTP, Supplementary Table 1), we observed a stably bound water molecule occupying the hydrophilic pocket in the absence of the RDV-TP cyano group (compare Fig. 3c to 3d,e), suggesting that the enhanced affinity (lower K_m) for RDV-TP over ATP[3,4] may be attributed to both an entropic effect through release of the bound water and an enthalpic effect through formation of new polar interactions between RDV-TP and the surrounding polar residues.

A recent study reports that SARS-CoV-2 acquired phenotypic resistance to RDV through an nsp12-S759A substitution following serial passage in cell culture[28], indicating that binding of the RDV-TP 1′-cyano group in this hydrophilic pocket is an important facet of RDV susceptibility. To investigate whether the polar interactions observed in the 1′-cyano pocket are important for RDV-TP selectivity, we assayed to discover whether RDV-TP incorporation is perturbed in the presence of single alanine substitutions of S759, T687 or N691. Consistent with previous work[3,4], the wild-type nsp12 enzyme possesses an almost twofold greater selectivity for RDV-TP over ATP whereas the S759A mutation renders the enzyme fivefold more selective for ATP (Extended Data Fig. 9). Both the T687A and N691A mutations greatly impair RNA synthesis activity regardless of whether ATP or RDV-TP was the substrate, however, the mutant enzymes strongly prefer to incorporate ATP over RDV-TP (Supplementary Fig. 3). We conclude that accommodation of the RDV-TP 1′-cyano group in the hydrophilic pocket formed from nsp12 S759, T687 and N691 explains how RDV-TP outcompetes ATP as an RdRp substrate[3,4,28].

Structural insights into RNA capping

Apart from the RdRp active site, the RTC possesses an additional catalytic domain known as the NiRAN (nidovirus RdRp-associated

nucleotidyltransferase), which lies at the N terminal end of the RdRp[5,15,29]. Although the NiRAN domain is essential for viral propagation, its functions during the viral life cycle remain enigmatic. Recent studies have suggested that the NiRAN functions in the viral RNA capping pathway in conjunction with an additional viral protein, nsp9 (refs. [5,30]). Others have postulated a role for the NiRAN domain in protein-mediated priming of RNA synthesis[31]. In previous structural studies, ADP[15] or GDP[30] have been observed bound in the NiRAN active site in a 'base-out' pose, where the nucleotide base points out of the active-site pocket and makes few, if any, protein contacts. ADP has not been shown to serve as a substrate for the NiRAN enzymatic activity, and the GDP was bound along with an N-terminally modified (and resultingly inactive) nsp9 (ref. [30]), so the relevance of these nucleotide poses is unclear. In another study, the diphosphate form of the GTP analogue AT-9010 is seen bound to the NiRAN in a 'base-in' pose[20], with the phosphates occupying the same positions as the phosphates in the 'base-out' pose but flipped, and with the guanine base bound in a tight pocket internal to the NiRAN (Extended Data Fig. 10).

In our cryo-EM structures of the RTC with each of the natural NTPs and RDV-TP (Supplementary Table 1), only GTP was observed stably and specifically bound in the NiRAN active site. The cryo-EM density supports the presence of GTP (and not GDP) in the NiRAN (Fig. 4). The GTP is bound in the 'base-in' pose similar to the diphosphate form of AT-9010 (ref. [20]), with the base buried in an apparently guanine-specific pocket that extends into the core of the NiRAN fold, enabling it to make key contacts with hydrophilic residues that line the pocket interior. A central element to this recognition is interactions mediated by R55 and Y217, two conserved residues across the α- and β-coronavirus clades, which provide base specificity for guanosine (Fig. 4b,d). The apo-NiRAN pocket is sterically incompatible with the bound guanine base; GTP association is mediated by an induced fit mechanism that involves expansion of the active-site pocket to accommodate the guanosine base (Fig. 4e).

A recent study proposes that the NiRAN mediates two successive steps in the viral RNA capping pathway that initially entails transfer of the nascent RNA 5′ end to the amino terminus of nsp9, forming a covalent RNA-protein intermediate in a process termed RNAylation[6]. A successive polyribonucleotidyltransferase (PRNTase) reaction uses the RNAylated nsp9 as a substrate and transfers the bound RNA to GDP or GTP to produce the cap structure, GpppN-RNA. The PRNTase activity strictly uses GDP or GTP. Our results indicate that the NiRAN uniquely recognizes GTP (but not other NTPs), explaining the base specificity observed for the PRNTase activity.

Discussion

Antiviral nucleotide analogues are highly effective and versatile therapeutics that can be adapted to treat new coronaviral borne diseases as they emerge. Their suitability for repurposing is exemplified by the use of RDV and molnupiravir for the treatment of COVID-19, two analogues that were initially developed as therapeutics against Ebola virus[32] and influenza virus[33], respectively. Repurposing efforts rely on the conserved active site of the RdRp, which has not been observed to mutate as readily as other COVID-19 therapeutic targets such as the Spike protein (Extended Data Fig. 8). Whereas active-site residues are highly conserved across RdRp-encoding viruses, subtle differences can alter the incorporation selectivity of a nucleotide analogue more than 100-fold[4], highlighting the importance of mechanistic investigations of RdRps across the viral realm.

To probe the mechanism of nucleotide recognition by the SARS-CoV-2 RdRp, we determined structures of stalled RdRp complexes containing an RNA primer template, incoming nucleotide substrates or the antiviral RDV-TP, and catalytic metal ions. Structural views of the RTC with each of the respective nucleotides illustrate how the RdRp structurally adapts to proper Watson–Crick base pairing geometry, closing in around the t-RNA/NTP base pair to facilitate catalytic Mg^{2+} coordination and in-line attack of the p-RNA 3′-OH on the NTP α-phosphate.

Through visualizing these structures, we ascertained that the enhanced selectivity for RDV, which primarily manifests biochemically as an effect on K_m, is mediated by the accommodation of its 1′-cyano group in a conserved hydrophilic pocket near the RdRp active site (Fig. 3). This same RdRp active-site pocket harbours an ordered water molecule in the +CTP and +GTP structures (Fig. 3); the bound water is probably present in the +ATP and +UTP complexes as well but not visualized owing to limited resolution (Extended Data Figs. 2 and 3). A mutation that alters the hydrophilicity of this pocket gives rise to RDV resistance[28] and are naturally found in viral families that show reduced sensitivity to RDV[4]. Critically, structural insights into NTP recognition could be exploited in the design of nucleotide analogues as part of structure-based drug design programs that aim to target the coronavirus RdRp. Our studies also shed light on the presence of a guanosine-specific pocket in the essential NiRAN domain, supporting proposals that the NiRAN is involved in the production of capped messenger RNAs as well as further detailing this pocket as a therapeutic target.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-022-05664-3.

1. Malone, B., Urakova, N., Snijder, E. J. & Campbell, E. A. Structures and functions of coronavirus replication–transcription complexes and their relevance for SARS-CoV-2 drug design. *Nat. Rev. Mol. Cell Biol.* **23**, 21–39 (2022).
2. Beigel, J. H. et al. Remdesivir for the treatment of COVID-19—final report. *N. Engl. J. Med.* **383**, 1813–1826 (2020).
3. Dangerfield, T. L., Huang, N. Z. & Johnson, K. A. Remdesivir is effective in combating COVID-19 because it is a better substrate than ATP for the viral RNA-dependent RNA polymerase. *iScience* **23**, 101849 (2020).
4. Gordon, C. J. et al. 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.* **295**, 6785–6797 (2020).
5. Slanina, H. et al. Coronavirus replication–transcription complex: vital and selective NMPylation of a conserved site in nsp9 by the NiRAN-RdRp subunit. *Proc. Natl Acad. Sci. USA* **118**, e2022310118 (2021).
6. Park, G. J. et al. The mechanism of RNA capping by SARS-CoV-2. *Nature* **609**, 793–800 (2022).
7. Jayk Bernal, A. et al. Molnupiravir for oral treatment of COVID-19 in nonhospitalized patients. *N. Engl. J. Med.* **386**, 509–520 (2021).
8. Bruenn, J. A. A structural and primary sequence comparison of the viral RNA-dependent RNA polymerases. *Nucleic Acids Res.* **31**, 1821–1829 (2003).
9. Bravo, J. P. K., Dangerfield, T. L., Taylor, D. W. & Johnson, K. A. Remdesivir is a delayed translocation inhibitor of SARS CoV-2 replication. *Mol. Cell* https://doi.org/10.1016/j.molcel.2021.01.035 (2021).
10. Seifert, M. et al. Inhibition of SARS-CoV-2 polymerase by nucleotide analogs from a single-molecule perspective. *eLife* **10**, e70968 (2021).
11. Steitz, T. A. DNA polymerases: structural diversity and common mechanisms. *J. Biol. Chem.* **274**, 17395–17398 (1999).
12. Appleby, T. C. et al. Structural basis for RNA replication by the hepatitis C virus polymerase. *Science* **347**, 771 (2015).
13. Shu, B. & Gong, P. Structural basis of viral RNA-dependent RNA polymerase catalysis and translocation. *Proc. Natl Acad. Sci. USA* **113**, E4005–E4014 (2016).
14. Zamyatkin, D. F. et al. Structural insights into mechanisms of catalysis and inhibition in Norwalk virus polymerase. *J. Biol. Chem.* **283**, 7705–7712 (2008).
15. Chen, J. et al. Structural basis for helicase-polymerase coupling in the SARS-CoV-2 replication-transcription complex. *Cell* https://doi.org/10.1016/j.cell.2020.07.033 (2020).
16. Hillen, H. S. et al. Structure of replicating SARS-CoV-2 polymerase. *Nature* **584**, 154–156 (2020).
17. Wang, Q. et al. Structural basis for RNA replication by the SARS-CoV-2 polymerase. *Cell* **182**, 417–428.e413 (2020).
18. Chen, J. et al. Ensemble cryo-EM reveals conformational states of the nsp13 helicase in the SARS-CoV-2 helicase replication–transcription complex. *Nat. Struct. Mol. Biol.* **29**, 250–260 (2022).
19. Naydenova, K. et al. Structure of the SARS-CoV-2 RNA-dependent RNA polymerase in the presence of favipiravir-RTP. *Proc. Natl Acad. Sci. USA* **118**, e2021946118 (2021).
20. Shannon, A. et al. A dual mechanism of action of AT-527 against SARS-CoV-2 polymerase. *Nat. Commun.* **13**, 621 (2022).
21. Cardone, G., Heymann, J. B. & Steven, A. C. One number does not fit all: mapping local variations in resolution in cryo-EM reconstructions. *J. Struct. Biol.* **184**, 226–236 (2013).

22. Salgado, P. S. et al. The structural basis for RNA specificity and Ca^{2+} inhibition of an RNA-dependent RNA polymerase. *Structure* **12**, 307–316 (2004).
23. Temiakov, D. et al. Structural basis for substrate selection by T7 RNA polymerase. *Cell* **116**, 381–391 (2004).
24. Wang, M. et al. SARS-CoV-2 RdRp uses NDPsas a substrate and is able to incorporate NHC into RNA by using diphosphate form molnupiravir. *Int. J. Biol. Macromol.* **226**, 946–955 (2023).
25. Yang, W., Lee, J. Y. & Nowotny, M. Making and breaking nucleic acids: two-Mg2+-ion catalysis and substrate specificity. *Mol. Cell* **22**, 5–13 (2006).
26. Ben Ouirane, K., Boulard, Y. & Bressanelli, S. The hepatitis C virus RNA-dependent RNA polymerase directs incoming nucleotides to its active site through magnesium-dependent dynamics within its F motif. *J. Biol. Chem.* **294**, 7573–7587 (2019).
27. Malone, B. et al. Structural basis for backtracking by the SARS-CoV-2 replication–transcription complex. *Proc. Natl Acad. Sci. USA* **118**, e2102516118 (2021).
28. Stevens, L. J. et al. Mutations in the SARS-CoV-2 RNA dependent RNA polymerase confer resistance to remdesivir by distinct mechanisms. *Sci. Transl. Med.* https://doi.org/10.1126/scitranslmed.abo0718 (2022).
29. Kirchdoerfer, R. N. & Ward, A. B. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nat. Commun.* **10**, 2342 (2019).
30. Yan, L. et al. Cryo-EM structure of an extended SARS-CoV-2 replication and transcription complex reveals an intermediate state in cap synthesis. *Cell* **184**, 184–193.e110 (2021).
31. Shannon, A. et al. Protein-primed RNA synthesis in SARS-CoVs and structural basis for inhibition by AT-527. Preprint at *bioRxiv* https://doi.org/10.1101/2021.03.23.436564 (2021).
32. Mulangu, S. et al. A randomized, controlled trial of Ebola virus disease therapeutics. *N. Engl. J. Med.* **381**, 2293–2303 (2019).
33. Painter, G. R., Natchus, M. G., Cohen, O., Holman, W. & Painter, W. P. Developing a direct acting, orally available antiviral agent in a pandemic: the evolution of molnupiravir as a potential treatment for COVID-19. *Curr. Opin. Virol.* **50**, 17–22 (2021).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Protein expression and purification

SARS-CoV-2 nsp12 was expressed as previously described[34] and purified as follows. Briefly, a pQE-30/pcl-ts ind$^+$ plasmid containing a His_6-small ubiquitin-like modifier (SUMO) SARS-CoV-2 nsp12 and untagged nsp7 and 8 (Addgene no. 160540) was transformed into *Escherichia coli* BL21 cells (Agilent). Cells were grown and protein expression was induced by the addition of 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), 10 ng ml^{-1} tetracycline and 50 μg ml^{-1} nalidixic acid. Cells were collected and lysed in a French press (Avestin). The lysate was cleared by centrifugation and purified on a HiTrap Heparin HP column (Cytiva). The fractions containing nsp12 were loaded onto a HisTrap HP column (Cytiva) for further purification. Eluted nsp12 was dialysed, cleaved with His_6-Ulp1 SUMO protease and passed through a HisTrap HP column to remove the SUMO protease. Flow-through was collected, concentrated by centrifugal filtration (Amicon) and loaded on a Superdex 200 Hiload 16/600 (Cytiva). Glycerol was added to the purified nsp12, aliquoted, flash-frozen with liquid N_2 and stored at −80 °C.

SARS-CoV-2 nsp7/8 was expressed and purified as described[15]. Briefly, the pCDFDuet-1 plasmid containing His_6 SARS-CoV-2 nsp7/8 (Addgene no. 159092) was transformed into *E. coli* BL21 (DE3). Cells were grown and protein expression was induced by the addition of IPTG. Cells were collected and lysed in a French press (Avestin). The lysate was cleared by centrifugation and purified on a HisTrap HP column (Cytiva). Eluted nsp7/8 was dialysed, cleaved with His_6-Prescission Protease to cleave His_6 tag and then passed through a HisTrap HP column to remove the protease (Cytiva). Flow-through was collected, concentrated by centrifugal filtration (Amicon) and loaded onto a Superdex 75 Hiload 16/600 (Cytiva). Glycerol was added to the purified nsp7/8, aliquoted, flash-frozen with liquid N_2 and stored at −80 °C.

Preparation of SARS-CoV-2 RTC for Cryo-EM

Cryo-EM samples of SARS-CoV-2 RTC were prepared as previously described[15,18]. Briefly, purified nsp12 and nsp7/8 were concentrated, mixed in a 1:3 molar ratio and incubated for 20 min at 22 °C. An annealed RNA scaffold (Horizon Discovery, Ltd) was added to the nsp7/8/12 complex and incubated for 10 min at 30 °C. The sample was buffer exchanged into cryo-EM buffer (20 mM HEPES pH 8.0, 100 mM K-Acetate, 5 mM $MgCl_2$, 2 mM DTT) and further incubated for 20 min at 30 °C. The sample was purified over a Superose 6 Increase 10/300 GL column (Cytiva) in cryo-EM buffer. The peak corresponding to nsp7/8/12/RNA complex was pooled and concentrated by centrifugal filtration (Amicon).

Cryo-EM grid preparation

Before grid freezing, beta-octyl glucoside (β-OG) was added to the sample (0.07% w/v final). The final buffer condition for the cryo-EM sample was 20 mM HEPES pH 8.0, 100 mM K-Acetate, 5 mM $MgCl_2$, 2 mM DTT, 0.07% (w/v) β-OG with 300 μM final of the respective NTP added immediately before freezing. C-flat holey carbon grids (CF-1.2/1.3-4Au, EMS) were glow-discharged for 20 s before the application of 3.5 μl of sample. Using a Vitrobot Mark IV (Thermo Fisher Scientific), grids were blotted and plunge-frozen into liquid ethane with 95% chamber humidity at 4 °C.

Cryo-EM data acquisition and processing

Structural biology software was accessed through the SBGrid consortium[35]. The following pertains to each of the respective datasets:

S1_RDV-TP. Grids were imaged using a 300 kV Titan Krios (Thermo Fisher Scientific) equipped with a GIF BioQuantum and K3 camera (Gatan). Images were recorded with Leginon[36] with a pixel size of 1.065 Å per pixel (px) (micrograph dimensions of 5,760 × 4,092 px) over a defocus range of −0.8 to −2.5 μm with a 20 eV energy filter slit. Videos were recorded in 'counting mode' (native K3 camera binning 2) with roughly 25 e$^-$ per px per s in dose-fractionation mode with subframes of 50 ms over a 2.5 s exposure (50 frames) to give a total dose of roughly 54 e$^-$/Å^2. Dose-fractionated videos were gain-normalized, drift-corrected, summed and dose-weighted using MotionCor2 (ref. [37]). The contrast transfer function (CTF) was estimated for each summed image using the Patch CTF module in cryoSPARC v.3.1.0 (ref. [38]). Particles were picked and extracted from the dose-weighted images with box size of 256 px using cryoSPARC blob picker and particle extraction. The entire dataset consisted of 26,130 motion-corrected images with 8,017,151 particles. Particles were sorted using three rounds of cryoSPARC two-dimensional (2D) classification (n = 50, where n equals the number of classes), resulting in 2,473,065 curated particles. Initial models (denoted as monomer and dimer) were generated using cryoSPARC ab initio reconstruction (n = 3) on a subset of the particles. Particles were further curated using these initial models as three-dimensional (3D) templates for iterative cryoSPARC heterogeneous refinement (n = 4), resulting in 314,848 particles in the monomer class (cyan map, Extended Data Fig. 1) and 264,453 particles in the dimer class (pink map, Extended Data Fig. 1). Curated particles in the monomer and dimer classes were re-extracted to a box size of 384 and input to cryoSPARC homogenous and non-uniform refinements[39]. Particles within each class were further processed through two rounds of RELION v.3.1 Bayesian polishing[40]. Polished particles were refined using cryoSPARC local and clobal CTF refinement in combination with cryoSPARC non-uniform refinement, resulting in structures with the following particle counts and nominal resolutions: monomer RTC (251,160 particles; 3.54 Å) and dimer RTC (249,468 particles; 3.87 Å). To facilitate model building of the RTC, particles from the dimer RTC class underwent masked particle subtraction in which both protomers were masked, subtracted from the full map and combined with the original monomer RTC to yield a structure of the RTC at a nominal resolution of 3.38 Å from 613,848 particles (Extended Data Fig. 1e). Local resolution calculations were generated using blocres and blocfilt from the Bsoft package[21] (Extended Data Fig. 1c). The angular distribution of particle orientations (Extended Data Fig. 1b) and directional resolution through the 3DFSC package[41] (Extended Data Fig. 1d) were calculated for the final class.

S2_ATP. Grids were imaged using a C_S (spherical abberation) corrected 300 kV Titan Krios (Thermo Fisher Scientific) equipped with a GIF BioQuantum and K3 camera (Gatan). Images were recorded with Leginon[36] with a pixel size of 1.076 Å per px (micrograph dimensions of 5,760 × 4,092 px) over a defocus range of −0.8 to −2.5 μm with a 20 eV energy filter slit. Videos were recorded in counting mode (native K3 camera binning 2) with roughly 25 e$^-$ per px per s in dose-fractionation mode with subframes of 50 ms over a 2.5 s exposure (50 frames) to give a total dose of roughly 52 e$^-$/Å^2. Dose-fractionated videos were gain-normalized, drift-corrected, summed and dose-weighted using MotionCor2 (ref. [37]). The CTF was estimated for each summed image using the patch CTF module in cryoSPARC v.3.1.0 (ref. [38]). Particles were picked and extracted from the dose-weighted images with box size of 256 px using cryoSPARC blob picker and particle extraction. The entire dataset consisted of 11,657 motion-corrected images with 5,364,858 particles. Particles were sorted using three rounds of cryoSPARC 2D classification (n = 50), resulting in 2,255,856 curated particles. Initial models (denoted as monomer and dimer) were generated using cryoSPARC ab initio reconstruction (n = 4) on a subset of the particles. Particles were further curated using these initial models as 3D templates for iterative cryoSPARC heterogeneous refinement (n = 4), resulting in 102,529 particles in the monomer class (cyan map, Extended Data Fig. 2) and 318,672 particles in the dimer class (pink map, Extended Data Fig. 2). Curated particles in the monomer and dimer classes were

re-extracted to a box size of 384 and input to cryoSPARC homogenous and non-uniform refinements[39]. Particles within each class were further processed through two rounds of RELION v.3.1 Bayesian polishing[40]. Polished particles were refined using cryoSPARC local and clobal CTF refinement in combination with cryoSPARC non-uniform refinement, resulting in structures with the following particle counts and nominal resolutions: monomer RTC (96,868 particles; 3.71 Å) and dimer RTC (299,965 particles; 3.50 Å). To facilitate model building of the RTC, particles from the dimer RTC class underwent masked particle subtraction in which both protomers were masked, subtracted from the full map and combined with the original monomer RTC to yield a structure of the RTC at a nominal resolution of 3.09 Å from 330,442 particles (Extended Data Fig. 2e). This combined RTC class was further refined with masked cryoSPARC local refinement using masks around the RdRp and NiRAN active sites. Locally refined maps were combined into a RTC composite map using PHENIX 'combine focused maps' to aid model building[42]. Local resolution calculations were generated using blocres and blocfilt from the Bsoft package[21] (Extended Data Fig. 2c). The angular distribution of particle orientations (Extended Data Fig. 2b) and directional resolution through the 3DFSC package[41] (Extended Data Fig. 2d) were calculated for the final class.

S3_UTP. Grids were imaged using a 300 kV Titan Krios (Thermo Fisher Scientific) equipped with a GIF BioQuantum and K3 camera (Gatan). Images were recorded with Leginon[36] with a pixel size of 1.065 Å per px (micrograph dimensions of 5,760 × 4,092 px) over a defocus range of −0.8 to −2.5 µm with a 20 eV energy filter slit. Videos were recorded in counting mode (native K3 camera binning 2) with roughly 25 e^- per px per s in dose-fractionation mode with subframes of 50 ms over a 2.5 s exposure (50 frames) to give a total dose of roughly 53 $e^-/Å^2$. Dose-fractionated videos were gain-normalized, drift-corrected, summed and dose-weighted using MotionCor2 (ref. [37]). The CTF was estimated for each summed image using the patch CTF module in cryoSPARC v.3.1.0 (ref. [38]). Particles were picked and extracted from the dose-weighted images with box size of 256 px using cryoSPARC blob picker and particle extraction. The entire dataset consisted of 30,850 motion-corrected images with 14,149,078 particles. Particles were sorted using three rounds of cryoSPARC 2D classification (n = 50), resulting in 3,297,109 curated particles. Initial models (denoted as monomer and dimer) were generated using cryoSPARC ab initio reconstruction (n = 4) on a subset of the particles. Particles were further curated using these initial models as 3D templates for iterative cryoSPARC heterogeneous refinement (n = 6), resulting in 648,814 particles in the monomer class (cyan map, Extended Data Fig. 3) and 773,911 particles in the dimer class (pink map, Extended Data Fig. 3). Curated particles in the monomer and dimer classes were re-extracted to a box size of 384 and input to cryoSPARC homogenous and non-uniform refinements[39]. Particles within each class were further processed through two rounds of RELION v.3.1 Bayesian polishing[40]. Polished particles were refined using cryoSPARC local and global CTF refinement in combination with cryoSPARC non-uniform refinement, resulting in structures with the following particle counts and nominal resolutions: monomer RTC (614,648 particles; 3.34 Å) and dimer RTC (730,009 particles; 3.42 Å). To facilitate model building of the RTC, particles from the dimer RTC class underwent masked particle subtraction in which both protomers were masked, subtracted from the full map and combined with the original monomer RTC to yield a structure of the RTC at a nominal resolution of 3.13 Å from 719,889 particles (Extended Data Fig. 3e). This combined RTC class was further refined with masked cryoSPARC local refinement using masks around the RdRp and NiRAN active sites. Locally refined maps were combined into a RTC composite map using PHENIX 'combine focused maps' to aid model building[42]. Local resolution calculations were generated using blocres and blocfilt from the Bsoft package[21] (Extended Data Fig. 3c). The angular distribution of particle orientations (Extended Data Fig. 3b) and directional resolution through the 3DFSC package[41] (Extended Data Fig. 3d) were calculated for the final class.

S4_GTP. Grids were imaged using a C_S corrected 300 kV Titan Krios (Thermo Fisher Scientific) equipped with a GIF BioQuantum and K3 camera (Gatan). Images were recorded with SerialEM[43] with a pixel size of 1.08 Å per px (micrograph dimensions of 5,760 × 4,092 px) over a defocus range of −0.8 to −3.0 µm with a 20 eV energy filter slit. Videos were recorded in counting mode (native K3 camera binning 2) with roughly 25 e^- per px per s in dose-fractionation mode with subframes of 50 ms over a 2.5 s exposure (50 frames) to give a total dose of roughly 51 $e^-/Å^2$. Dose-fractionated videos were gain-normalized, drift-corrected, summed and dose-weighted using MotionCor2 (ref. [37]). The CTF was estimated for each summed image using the patch CTF module in cryoSPARC v.3.1.0 (ref. [38]). Particles were picked and extracted from the dose-weighted images with box size of 256 px using cryoSPARC blob picker and particle extraction. The entire dataset consisted of 4,527 motion-corrected images with 2,419,929 particles. Particles were sorted using three rounds of cryoSPARC 2D classification (n = 50), resulting in 941,507 curated particles. An initial model of the monomer RTC was generated using cryoSPARC ab initio reconstruction (n = 3) on a subset of the particles. Particles were further curated using this initial model as a 3D template for iterative cryoSPARC heterogeneous refinement (n = 6), resulting in 484,682 particles in the resultant monomer class (cyan map, Extended Data Fig. 4). Curated particles were re-extracted to a box size of 384 and input to cryoSPARC homogenous and non-uniform refinements[39]. Particles were further processed through two rounds of RELION v.3.1 Bayesian polishing[40]. Polished particles were refined using cryoSPARC local and global CTF refinement in combination with cryoSPARC non-uniform refinement, resulting in a structure with the following particle count and nominal resolution: monomer RTC (456,629 particles; 2.68 Å) (Extended Data Fig. 4e). Local resolution calculations were generated using blocres and blocfilt from the Bsoft package[21] (Extended Data Fig. 4c). The angular distribution of particle orientations (Extended Data Fig. 4b) and directional resolution through the 3DFSC package[41] (Extended Data Fig. 4d) were calculated for the final class.

S5_CTP. Grids were imaged using a C_S corrected 300 kV Titan Krios (Thermo Fisher Scientific) equipped with a GIF BioQuantum and K3 camera (Gatan). Images were recorded with SerialEM[43] with a pixel size of 0.515 Å per px (micrograph dimensions of 5,760 × 4,092 px) over a defocus range of −0.8 to −3.0 µm with a 20 eV energy filter slit. Videos were recorded in counting mode (native K3 camera binning 2) with roughly 25 e^- per px per s in dose-fractionation mode with subframes of 50 ms over a 2.5 s exposure (50 frames) to give a total dose of roughly 57 $e^-/Å^2$. Dose-fractionated videos were gain-normalized, drift-corrected, summed and dose-weighted using MotionCor2 (ref. [37]). The CTF was estimated for each summed image using the Patch CTF module in cryoSPARC v.3.1.0 (ref. [38]). Particles were picked and extracted from the dose-weighted images with box size of 512 px using cryoSPARC blob picker and particle extraction. The entire dataset consisted of 13,905 motion-corrected images with 1,986,527 particles. Particles were sorted using three rounds of cryoSPARC 2D classification (n = 50), resulting in 460,232 curated particles. An initial model (denoted a monomer) was generated using cryoSPARC ab initio reconstruction (n = 3) on a subset of the particles. Particles were further curated using these initial models as 3D templates for iterative cryoSPARC heterogeneous refinement (n = 6), resulting in 143,110 particles in the monomer class (cyan map, Extended Data Fig. 5) and 93,597 particles in an extracted dimer class (pink map, Extended Data Fig. 5). Curated particles in the monomer and dimer classes were re-extracted to a box size of 768 and input to cryoSPARC homogenous and non-uniform refinements[39]. Particles within each class were further processed through two rounds of RELION v.3.1 Bayesian polishing[40]. Polished particles were refined using cryoSPARC local and global CTF refinement in combination with

cryoSPARC non-uniform refinement, resulting in structures with the following particle counts and nominal resolutions: monomer RTC (128,484 particles; 3.09 Å) and dimer RTC (83,555 particles; 3.69 Å). To facilitate model building of the RTC, particles from the dimer RTC class underwent masked particle subtraction in which both protomers were masked, subtracted from the full map and combined with the original monomer RTC to yield a structure of the RTC at a nominal resolution of 2.67 Å from 171,107 particles (Extended Data Fig. 5e). This combined RTC class was further refined with masked cryoSPARC local refinement using masks around the RdRp and NiRAN active sites. Locally refined maps were combined into a RTC composite map using PHENIX 'combine focused maps' to aid model building[42]. Local resolution calculations were generated using blocres and blocfilt from the Bsoft package[21] (Extended Data Fig. 5c). The angular distribution of particle orientations (Extended Data Fig. 5b) and directional resolution through the 3DFSC package[41] (Extended Data Fig. 5d) are shown for the final class.

Model building and refinement

An initial model of the RTC was derived from Protein Data Bank (PDB) 7RE1 (ref. [18]). The models were manually fit into the cryo-EM density maps using Chimera[44] and rigid-body and real-space refined using PHENIX real-space-refine[42]. For real-space refinement, rigid-body refinement was followed by all-atom and *B* factor refinement with Ramachandran and secondary structure restraints. Models were inspected and modified in Coot v.0.9.5 (ref. [45]) and the refinement process was repeated iteratively.

nMS analysis

The RTC samples were initially reconstituted as described above in the following buffer: 20 mM HEPES pH 8.0, 80 mM K-Acetate, 5 mM $MgCl_2$ and 2 mM DTT. For the NTP pre-incorporation experiments, the RTCs with RNA scaffolds containing either 3′-oxy p-RNA (RTC) or 3′-deoxy p-RNA (RTC*) were incubated with 300 μM NTP (ATP, RDV-TP or GTP) on ice for 2 min before buffer exchange. For the NDP incubation experiments, the RTC samples reconstituted with the respective RNA scaffolds were incubated with 2 mM nucleotide (ADP, RDV-DP, GDP or ATP) on ice for 2 min before buffer exchange. After nucleotide incubation, all samples were immediately buffer exchanged into an nMS-compatible solution (150 mM ammonium acetate, pH 7.5, 0.01% Tween-20) with a 40 kDa molecular weight cut-off (Thermo Fisher Scientific). For nMS analysis, a 2–3 μl aliquot of the buffer-exchanged sample was loaded into a gold-coated quartz capillary tip prepared in-house and then electrosprayed into an Exactive Plus with an extended mass range instrument (Thermo Fisher Scientific) with a static direct infusion nanospray source[46]. The MS parameters used include: spray voltage, 1.22 kV; capillary temperature, 125–150 °C; in-source dissociation, 0–10 V; S-lens radio-frequency level, 200; resolving power, 8,750 or 17,500 at an *m/z* of 200; AGC target, 1×10^6; maximum injection time, 200 ms; number of microscans was five; injection flatapole, 8 V; interflatapole, 7 V; bent flatapole, 4 V; high energy collision dissociation (HCD), 200 V; ultrahigh vacuum pressure, $5.8–6.1 \times 10^{-10}$ mbar and total number of scans was at least 100. Mass calibration in positive extended mass range mode was performed using caesium iodide. For data processing, the acquired MS spectra were visualized using Thermo Xcalibur Qual Browser (v.4.2.47). Deconvolution was performed either manually or using the software UniDec v.4.2.0 (refs. [47,48]). The following parameters were used for data processing with UniDec: background subtraction (if applied) was a subtract curve of ten; smooth charge state distribution was enabled and peak shape function was Gaussian. Mass accuracies were calculated as the percentage difference between the measured and expected masses relative to the expected mass. The observed mass accuracies (calculated as the percentage difference between the measured and expected masses relative to the expected mass) ranged from 0.005 to 0.06%.

The expected masses for the component proteins are nsp7, 9,137 Da; nsp8 (N-terminal Met lost), 21,881 Da and nsp12 (has two Zn^{2+} ions coordinated with six deprotonated cysteine residues), 106,785 Da (ref. [15]). The RNA scaffolds were also analysed separately, and their sequences were verified by mass measurements using nMS.

In vitro primer elongation assays

Assays were performed using reconstituted template-primer RNA scaffolds (Supplementary Table 1) (Horizon Discover Ltd/Dharmacon) annealed in 10 mM HEPES pH 8.0, 50 mM KCl and 2 mM $MgCl_2$. Reactions (20 μl) containing 100 nM RNA scaffold, 0.75 μM nsp12, 2 μM nsp7/8 and NTPs 300 μM (if present as natural nucleotides, NDPs or α-β analogues) and 1 μl of α-^{32}P-GTP (Perkin-Elmer) were incubated at 30 °C for 10 min before addition of a 2× stop solution (Invitrogen-Gel Loading buffer II). Assay buffer was 100 mM K-acetate, 20 mM HEPES pH 8.0, 5 mM $MgCl_2$ and 2 mM beta-mercaptoethanol in which the $MgCl_2$ was substituted for $CaCl_2$ when monitoring the effects of alternative metal ions. Products of the elongation reactions were separated on 10% acrylamide-8M urea denaturing gels and analysed by phosphorimaging.

Substrate selectivity assays

The selectivity of nucleotide incorporation by SARS-CoV-2 wild-type and mutant RdRps was determined, and materials and reagents acquired as previously described[4,49]. The concentration for all wild-type and mutant nsp12 enzymes per reaction was 150 nM, 750 nM for nsp7 and 1.5 μM for nsp8. Nsp12, nsp7 and nsp8 were assembled in reaction. Reactions were incubated for 10 min to allow for single-nucleotide incorporation. Michaelis–Menten kinetics analyses were applied as described. Nucleotide incorporation efficiencies (catalytic efficiencies) were calculated by taking the ratio between V_{max} and K_m. Selectivity values were determined from the ratio between ATP incorporation efficiency to that of RDV-TP. Finally, the discrimination index was calculated by taking the ratios of the mutant RdRps selectivity values to that of the wild-type one[4,49].

Quantification and statistical analysis

The local resolution of the cryo-EM maps (Extended Data Figs. 1–5) was estimated using blocres[21] with the following parameters: box size 15 and sampling with respective map pixel size. Directional 3DFSCs (Extended Data Figs. 1–5) were calculated using 3DFSC (ref. [41]). The quantification and statistical analyses for model refinement and validation were generated using MolProbity[50] and PHENIX[42].

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All unique/stable reagents generated in this study are available without restriction from the lead contact, S.A.D. (darst@rockefeller.edu). The cryo-EM density maps and atomic coordinates have been deposited in the EMDataBank and PDB as follows: S1_RDV-TP (EMD-26639, PDB 7UO4), S2_ATP (EMD-26641, PDB 7UO7), S3_UTP (EMD-26642, PDB 7UO9), S4_GTP (EMD-26645, PDB 7UOB) and S5_CTP (EMD-26646, PDB 7UOE). The model used for initial building (PDB 7RE1) is available on the PDB. The atomic models used for initial model building and analysis are available from the Protein Data Bank under the accession codes 7RE1, 6XEZ, 7CYQ and 7ED5.

34. Dangerfield, T. L., Huang, N. Z. & Johnson, K. A. Expression and purification of tag-free SARS-CoV-2 RNA-dependent RNA polymerase in *Escherichia coli*. *STAR Protoc.* **2**, 100357 (2021).
35. Morin, A. et al. Collaboration gets the most out of software. *eLife* **2**, e01456 (2013).

36. Suloway, C. et al. Automated molecular microscopy: the new Leginon system. *J. Struct. Biol.* **151**, 41–60 (2005).
37. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
38. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
39. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020).
40. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
41. Tan, Y. Z. et al. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793–796 (2017).
42. Afonine, P. V. et al. New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr. D Struct. Biol.* **74**, 814–840 (2018).
43. Mastronarde, D. N. SerialEM: a program for automated tilt series acquisition on Tecnai microscopes using prediction of specimen position. *Microsc. Microanal.* **9**, 1182–1183 (2003).
44. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
45. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
46. Olinares, P. D. B. & Chait, B. T. Native mass spectrometry analysis of affinity-captured endogenous yeast RNA exosome complexes. *Methods Mol. Biol.* **2062**, 357–382 (2020).
47. Marty, M. T. et al. Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Anal. Chem.* **87**, 4370–4376 (2015).
48. Reid, D. J. et al. MetaUniDec: high-throughput deconvolution of native mass spectra. *J. Am. Soc. Mass. Spectrom.* **30**, 118–127 (2019).
49. Gordon, C. J., Tchesnokov, E. P., Feng, J. Y., Porter, D. P. & Götte, M. The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. *J. Biol. Chem.* **295**, 4773–4779 (2020).
50. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).

Acknowledgements We thank A. Mueller and R. Saecker for helpful discussions. Some of the work reported here was conducted at the Simons Electron Microscopy Center and the National Resource for Automated Molecular Microscopy and National Center for CryoEM Access and Training located at the NYSBC, supported by grants from the National Insitutes of Health (NIH) National Institute of General Medical Sciences (grant no. P41 GM103310), NYSTAR, the Simons Foundation (grant no. SF349247), the NIH Common Fund Transformative High Resolution Cryo-Electron Microscopy programme (grant no. U24 GM129539) and NY State Assembly Majority. This work was supported by NIH grant nos. P41 GM109824 and P41 GM103314 to B.T.C., and NIH grant no. R01 AI161278 and Gilead Sciences (to E.A.C. and S.A.D.). M.G. received funding from the Canadian Institutes of Health Research (CIHR), grant no. 170343, and the Alberta Ministry of Jobs, Economy and Northern Development through the Major Innovation Fund on Antimicrobial Resistance (AMR) - One Health Consortium.

Author contributions B.F.M., J.C., J.K.P., T.C.A., J.Y.F., J.P.B., E.A.C. and S.A.D. conceived and designed this study. B.F.M., H.W.L. and J.C. performed cloning, protein purification and biochemistry. P.D.B.O. conducted mass spectrometry experiments. B.F.M. prepared cryo-EM specimens. Cryo-EM data were collected by B.F.M., E.Y.D.C., J.H.M., E.T.E., M.E., J.S. and H.N. B.F.M. processed all cryo-EM data. B.F.M., J.K.P., E.A.C. and S.A.D. built and analysed atomic models. E.A.C., R.L., B.T.C., M.G. and S.A.D. supervised and acquired financial support. B.F.M. wrote the first draft of the manuscript; all authors contributed to the final version.

Competing interests E.A.C. and S.A.D. received funding from Gilead Sciences, Inc. in support of this study. J.K.P., T.C.A., J.Y.F. and J.P.B. are Gilead employees. M.G. received funding from Gilead Sciences Inc. for studies on the mechanism of action of remdesivir. All other authors declare no competing interests.

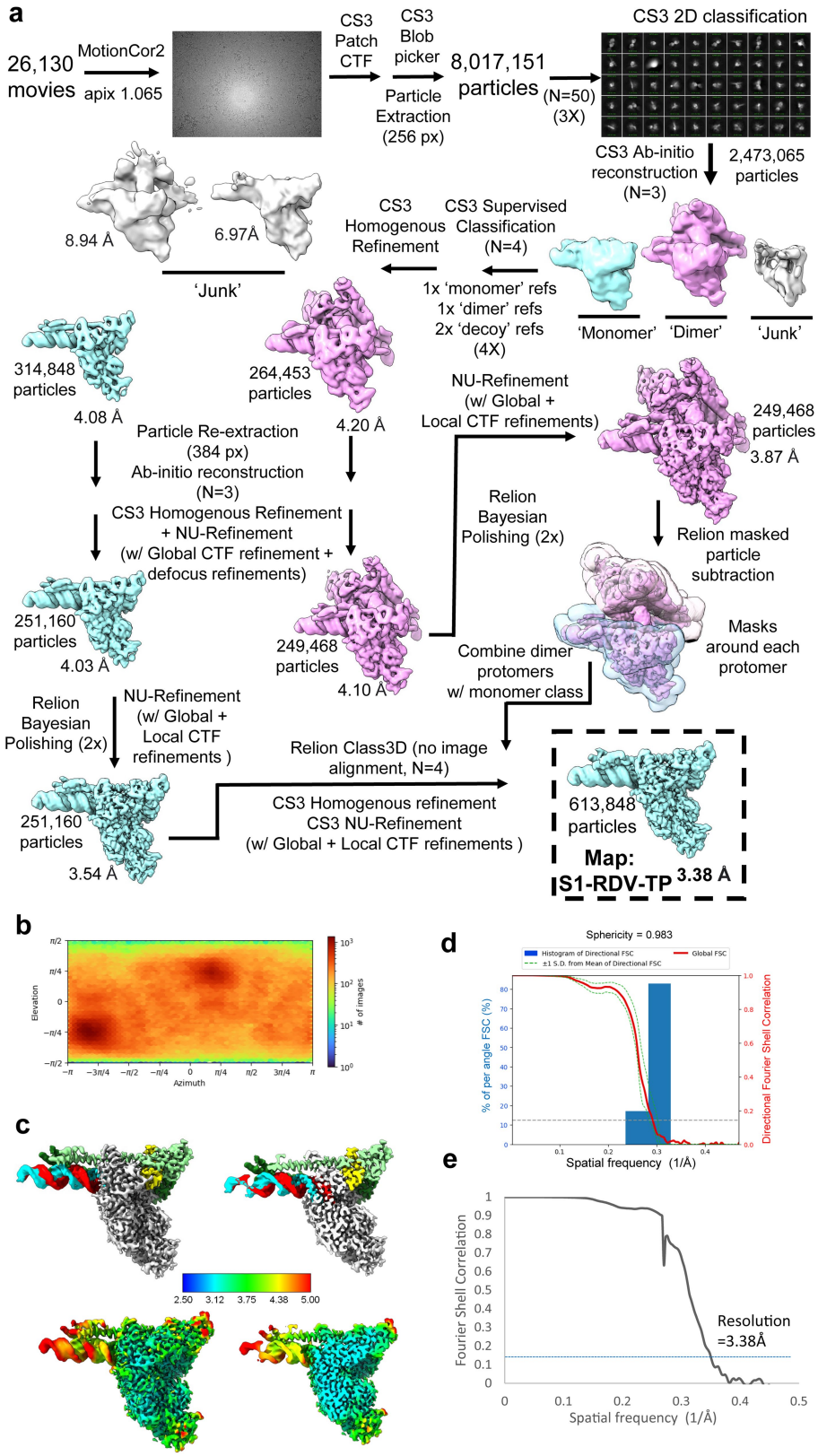
Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-022-05664-3.

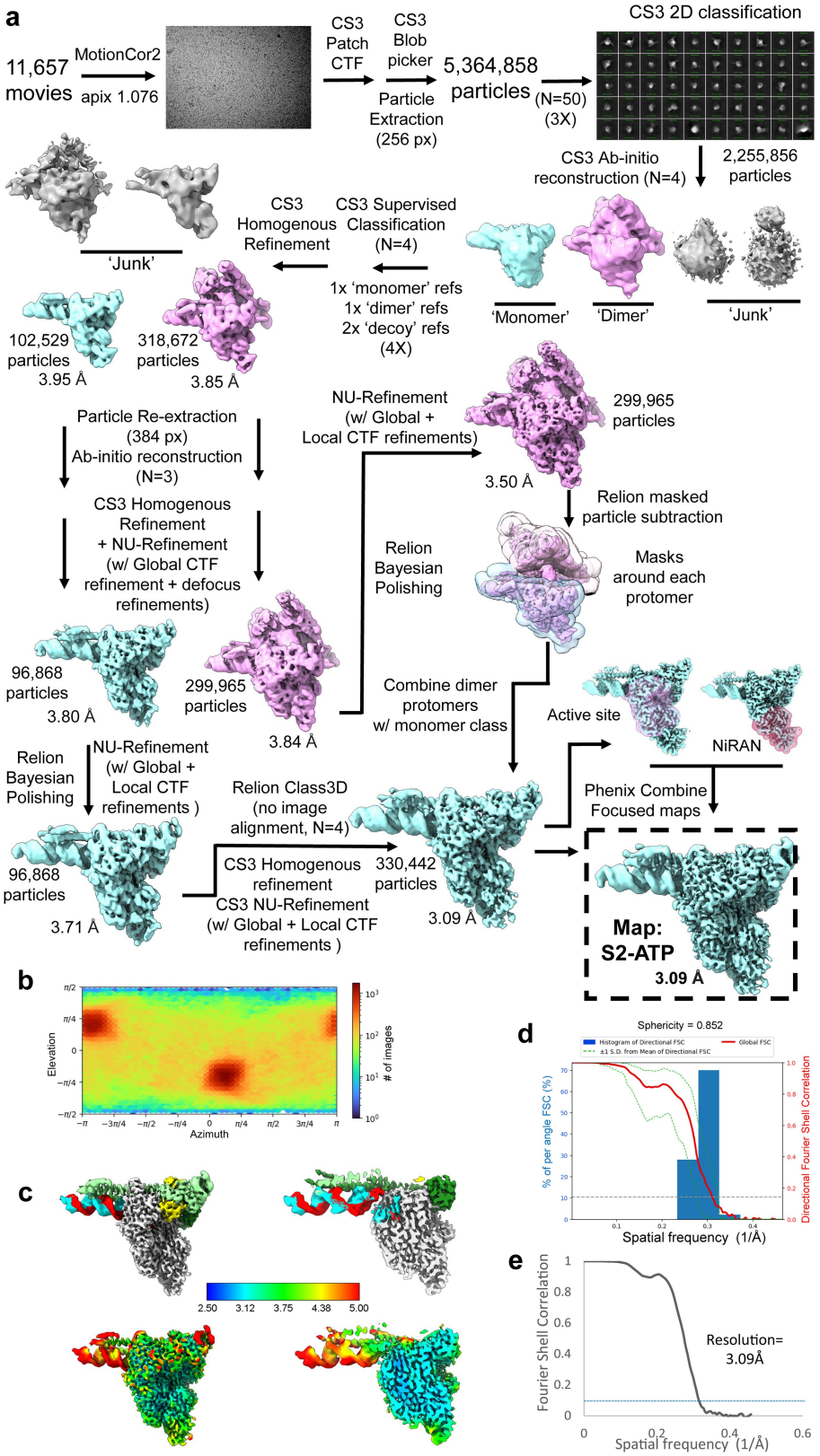
Correspondence and requests for materials should be addressed to Elizabeth A. Campbell or Seth A. Darst.

Peer review information *Nature* thanks Dong Wang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

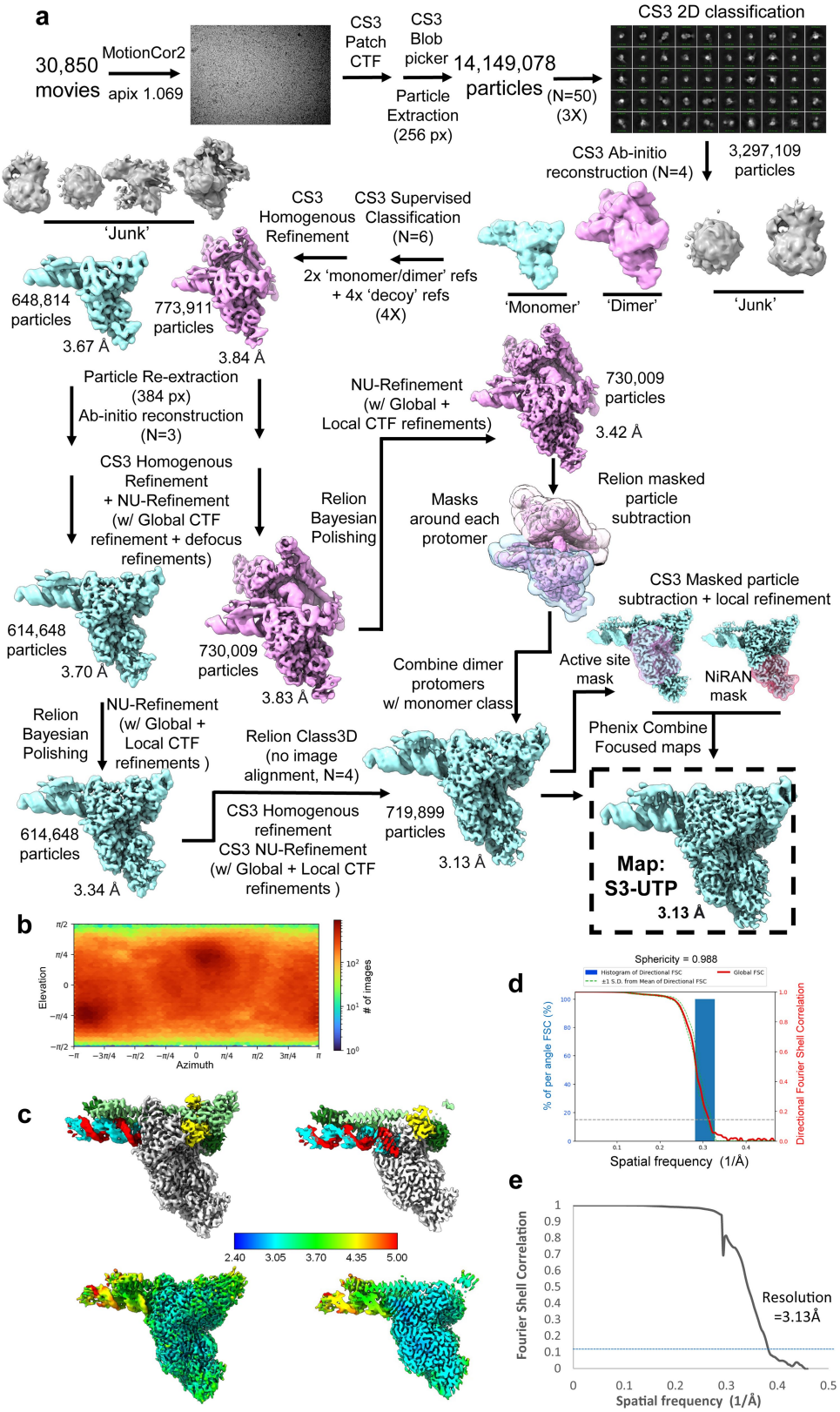
Reprints and permissions information is available at http://www.nature.com/reprints.



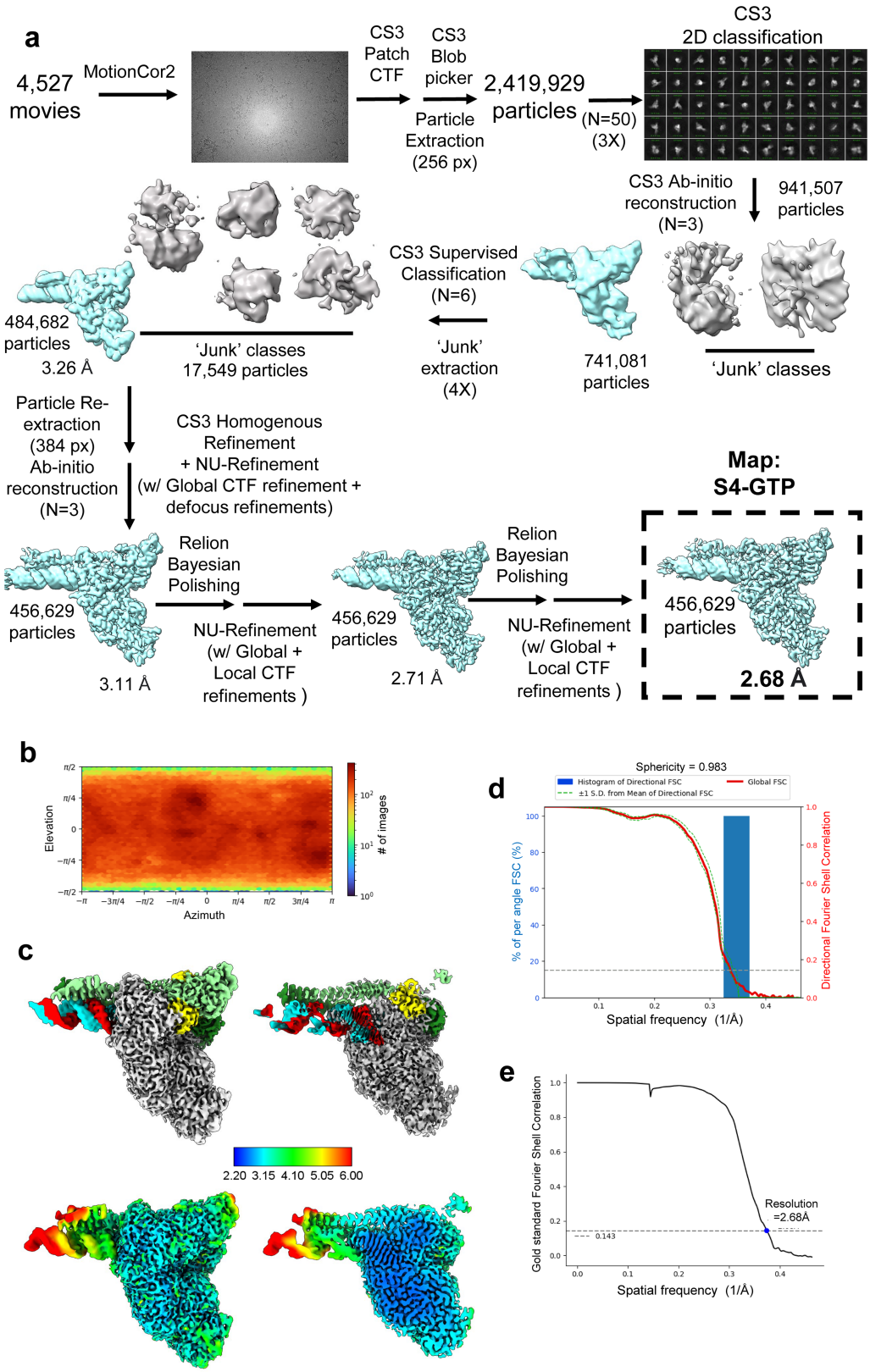
Extended Data Fig. 1 | Cryo-EM processing pipeline and analysis for S1_RDV-TP dataset. (**a**) Cryo-EM processing pipeline for the S1_RDV-TP dataset. (**b**) Angular distribution plot for the S1_RDV-TP dataset, calculated in cryoSPARC. Scale depicts number of particles assigned to a specific angular bin. (**c**) Nominal 3.38 Å resolution cryo-EM reconstruction filtered by local resolution and colored according to fitted model chain. Right panel is clipped to reveal RTC active site. (**d**) Directional 3D FSC for S1_RDV-TP, determined with 3DFSC. (**e**) Gold-standard FSC plot for the S1_RDV-TP dataset, calculated by comparing two half maps from cryoSPARC. The blue dotted line represents the 0.143 FSC cutoff.



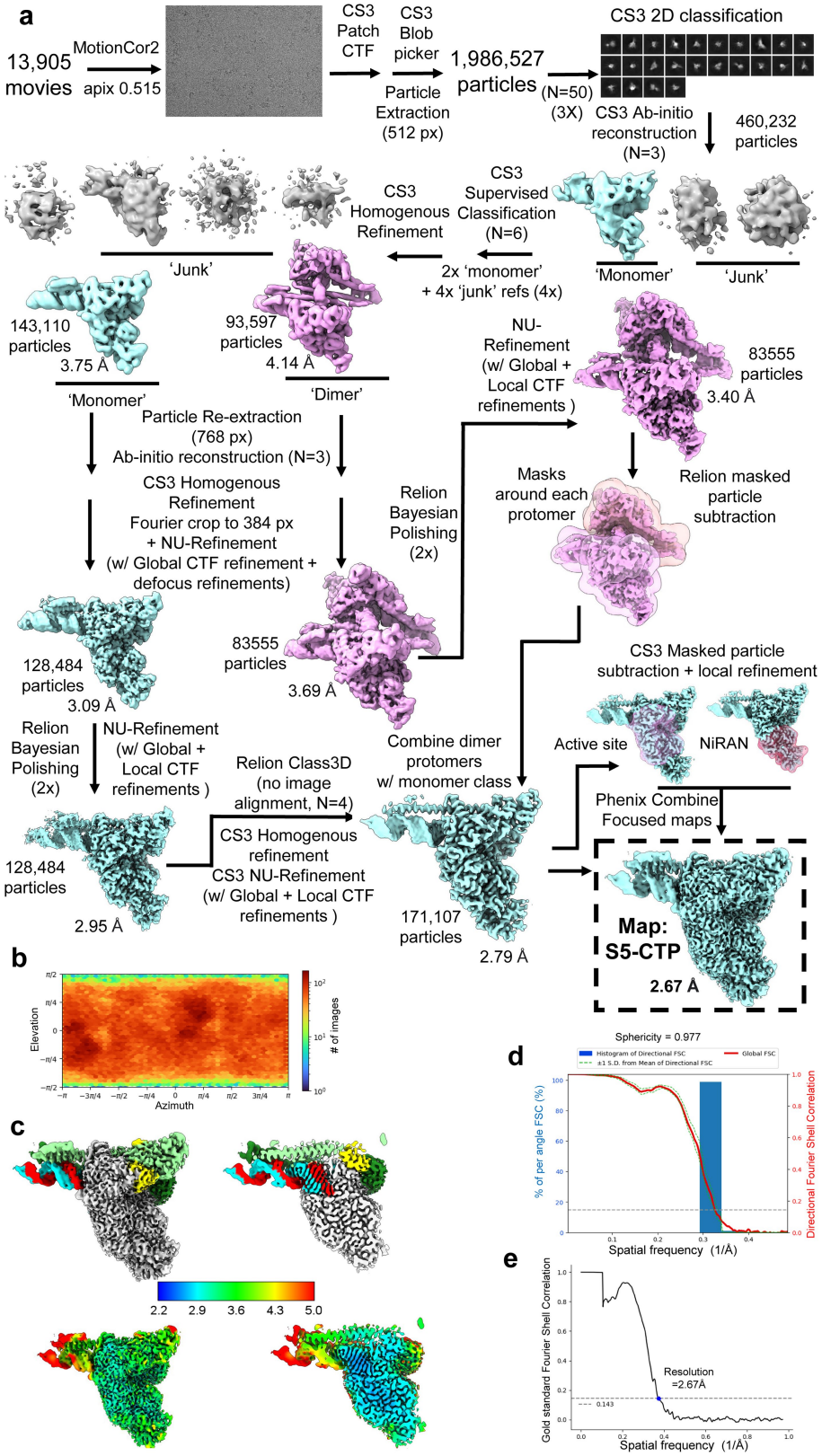
Extended Data Fig. 2 | Cryo-EM processing pipeline and analysis for S2_ATP dataset. (**a**) Cryo-EM processing pipeline for the S2_ATP dataset. (**b**) Angular distribution plot for the S2_ATP dataset, calculated in cryoSPARC. Scale depicts number of particles assigned to a specific angular bin. (**c**) Nominal 3.09 Å resolution cryo-EM reconstruction filtered by local resolution and colored according to fitted model chain. Right panel is clipped to reveal RTC active site. (**d**) Directional 3D FSC for S2_ATP, determined with 3DFSC. (**e**) Gold-standard FSC plot for the S2_ATP dataset, calculated by comparing two half maps from cryoSPARC. The blue dotted line represents the 0.143 FSC cutoff.



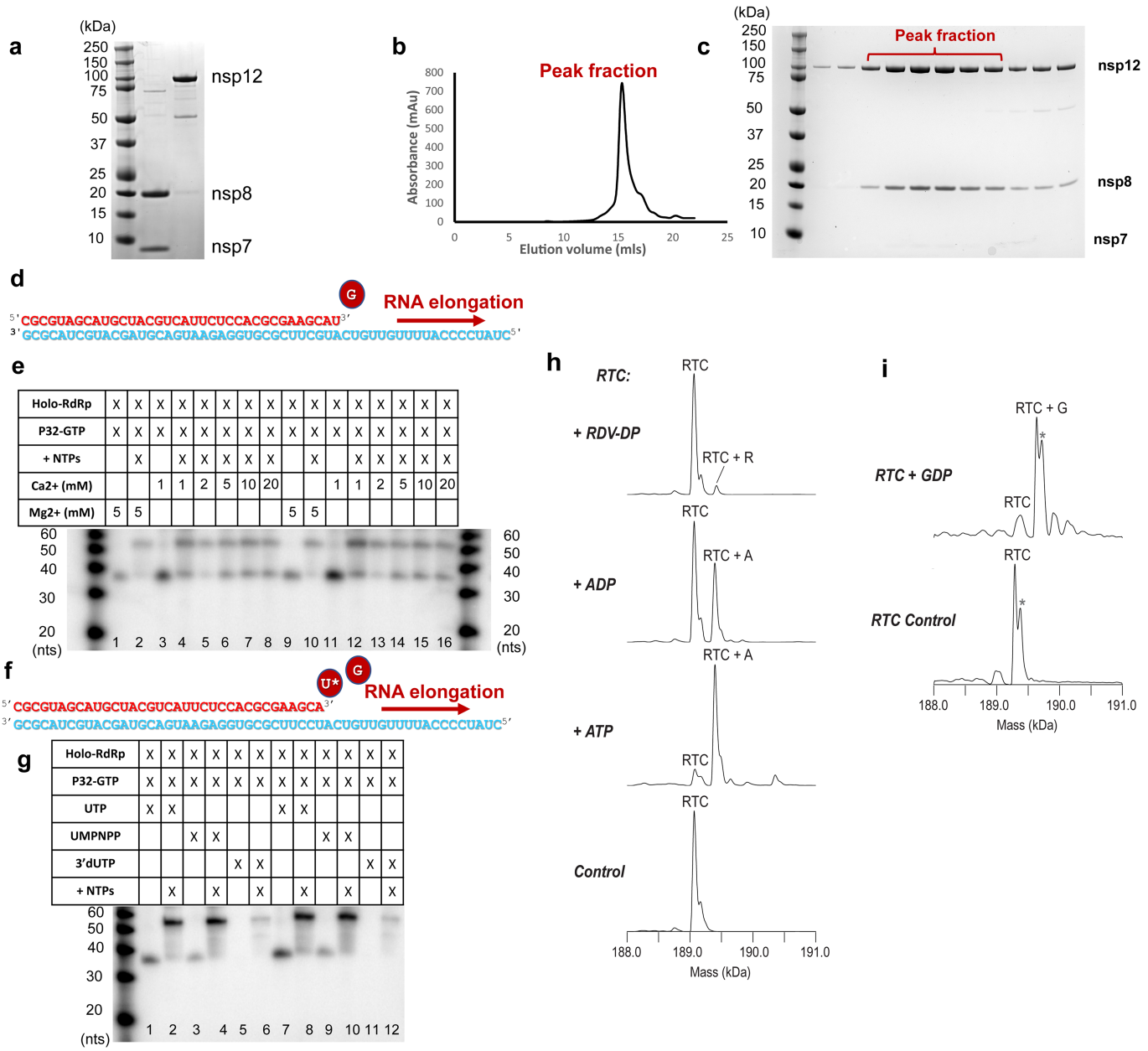
Extended Data Fig. 3 | Cryo-EM processing pipeline and analysis for S3_UTP dataset. (**a**) Cryo-EM processing pipeline for the S3_UTP dataset. (**b**) Angular distribution plot for the S3_UTP dataset, calculated in cryoSPARC. Scale depicts number of particles assigned to a specific angular bin. (**c**) Nominal 3.13 Å resolution cryo-EM reconstruction filtered by local resolution and colored according to fitted model chain. Right panel is clipped to reveal RTC active site. (**d**) Directional 3D FSC for S3_UTP, determined with 3DFSC. (**e**) Gold-standard FSC plot for the S3_UTP dataset, calculated by comparing two half maps from cryoSPARC. The blue dotted line represents the 0.143 FSC cutoff.



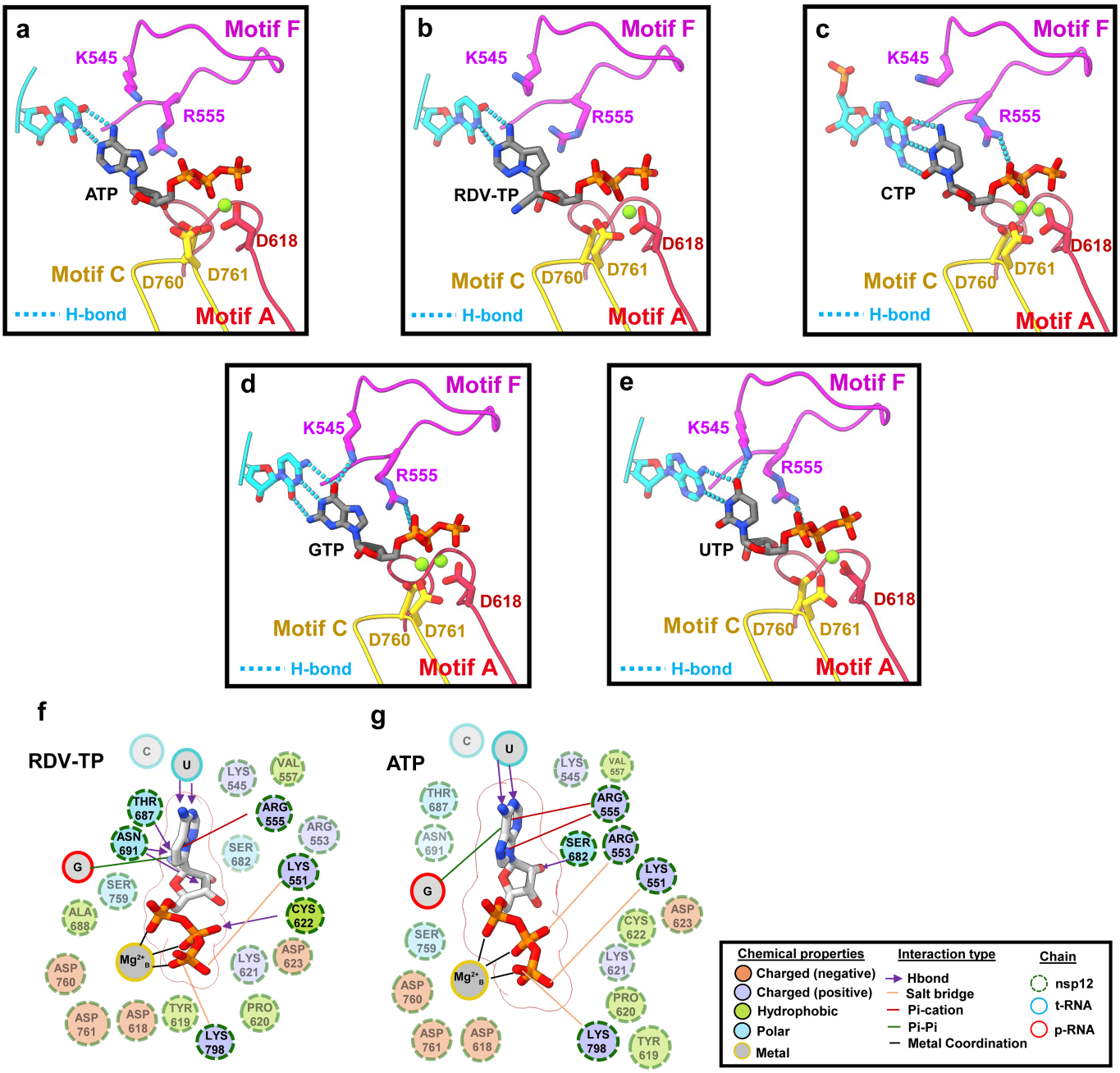
Extended Data Fig. 4 | Cryo-EM processing pipeline and analysis for S4_GTP dataset. (**a**) Cryo-EM processing pipeline for the S4_GTP dataset. (**b**) Angular distribution plot for the S4_GTP dataset, calculated in cryoSPARC. Scale depicts number of particles assigned to a specific angular bin. (**c**) Nominal 2.68 Å resolution cryo-EM reconstruction filtered by local resolution and colored according to fitted model chain. Right panel is clipped to reveal RTC active site. (**d**) Directional 3D FSC for S4_GTP, determined with 3DFSC. (**e**) Gold-standard FSC plot for the S4_GTP dataset, calculated by comparing two half maps from cryoSPARC. The black dotted line represents the 0.143 FSC cutoff.



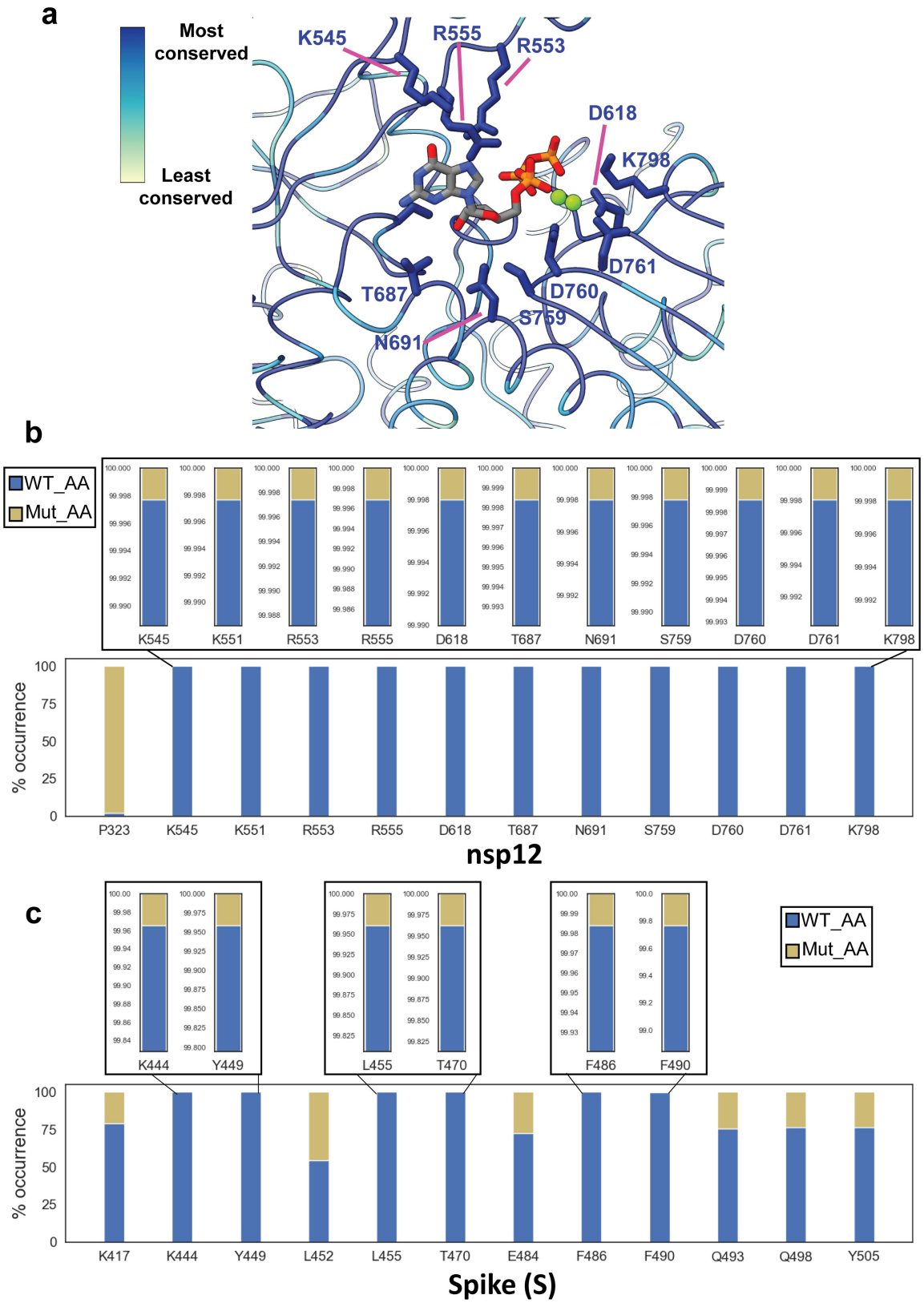
Extended Data Fig. 5 | Cryo-EM processing pipeline and analysis for S5_CTP dataset. (**a**) Cryo-EM processing pipeline for the S5_CTP dataset. (**b**) Angular distribution plot for the S5_CTP dataset, calculated in cryoSPARC. Scale depicts number of particles assigned to a specific angular bin. (**c**) Nominal 2.67 Å resolution cryo-EM reconstruction filtered by local resolution and colored according to fitted model chain. Right panel is clipped to reveal RTC active site. (**d**) Directional 3D FSC for S5_CTP, determined with 3DFSC. (**e**) Gold-standard FSC plot for the S5_CTP dataset, calculated by comparing two half maps from cryoSPARC. The black dotted line represents the 0.143 FSC cutoff.



Extended Data Fig. 6 | Biochemical assays probing stabilization of pre-incorporation complex. (**a**) SDS-PAGE of purified SARS-CoV-2 nsp7/8 & nsp12. The experiment was performed once. (**b**) Size exclusion chromatography for the purified RTC complex, composed of nsp7/8_2/12 bound to the reconstituted product/template-RNA scaffold (chromatogram trace using the S4 scaffold is depicted). (**c**) SDS-PAGE of assembled RTC complex following size exclusion. Gel (n = 1) illustrates RTC assembly on the S4 scaffold. (**d**) S4 RNA scaffold utilized for the S4_GTP structure and incorporation assays. (**e**) Gel-based primer elongation assay in presence of physiological metal, Mg^{2+}, and non-physiological metal, Ca^{2+}, to investigate use of Ca^{2+} for stabilization of the pre-incorporation complex (gel depicts n = 2 with n = 3 experiments performed). Products are visualized alongside Decade RNA ladder (Invitrogen) showing separation in nucleotides (nts). (**f**) S3 RNA scaffold used for the S3_UTP structure and incorporation assays. (**g**) Gel-based primer elongation assay in presence of UTP, UMPNPP and 3′deoxy UTP (gel depicts n = 2 with n = 3 experiments performed). Products are visualized alongside Decade RNA ladder (Invitrogen) showing separation in nucleotides (nts). (**h**) Native mass spectrometry analysis of RNA extension in presence of ADP, RDV-DP & ATP using the S1/S2 RNA scaffold. (**i**) Native mass spectrometry analysis of RNA extension in presence of GDP using the S4 RNA scaffold.

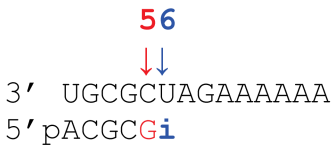


Extended Data Fig. 7 | Motif F residue interactions with the incoming NTPs. (**a**–**e**) Analysis of the active sites of structures (S1–S5) illustrate that K545 and R555 interact differentially with each of the incoming nucleotides. H-bonds are depicted as dotted-cyan lines. (**f**,**g**) 2D schematics illustrating the set of interactions between RDV-TP (**f**) and ATP (**g**) and the active site.

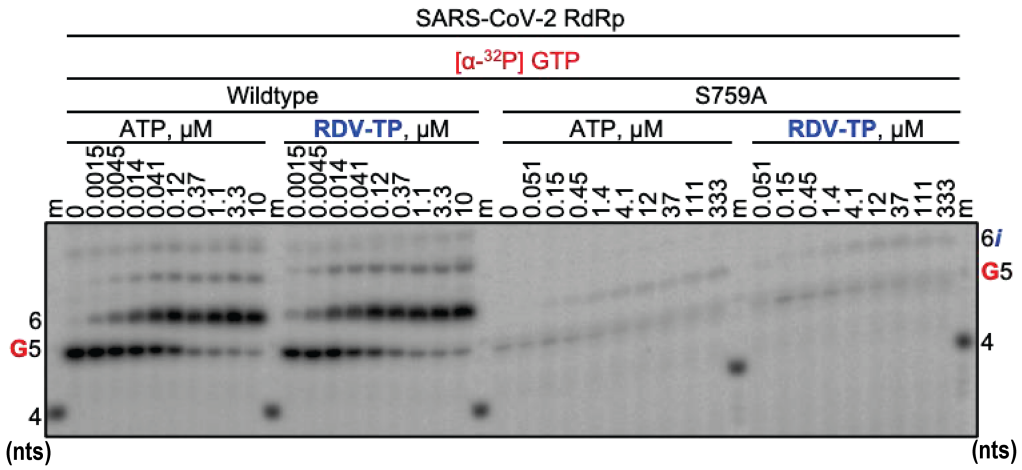


Extended Data Fig. 8 | Conservation of the CoV active site across clades and SARS-CoV-2 strains. (**a**) Zoom-in on the active site of S4_GTP, highlighting residues (sticks) which interact with the bound GTP/2xMg in which ribbon/sticks are colored according to the amino acid conservation across a representative list of viruses found in the α & β coronavirus clades (Supplementary Table 3). (**b**) Bar plot showing the frequency of occurrence of the wild-type amino acid (reference strain Wuhan/Hu-1/2019) as well as their summed mutations according to the GISAID database, as of April 2022, for the nsp12 active site residues and the nsp12 residue P323. (**c**) Bar plot showing the frequency of occurrence of the wild-type amino acid (reference strain Wuhan/Hu-1/2019) as well as their summed mutations according to the GISAID database, as of April 2022, for Spike (S) residues found in the ACE2 binding region.

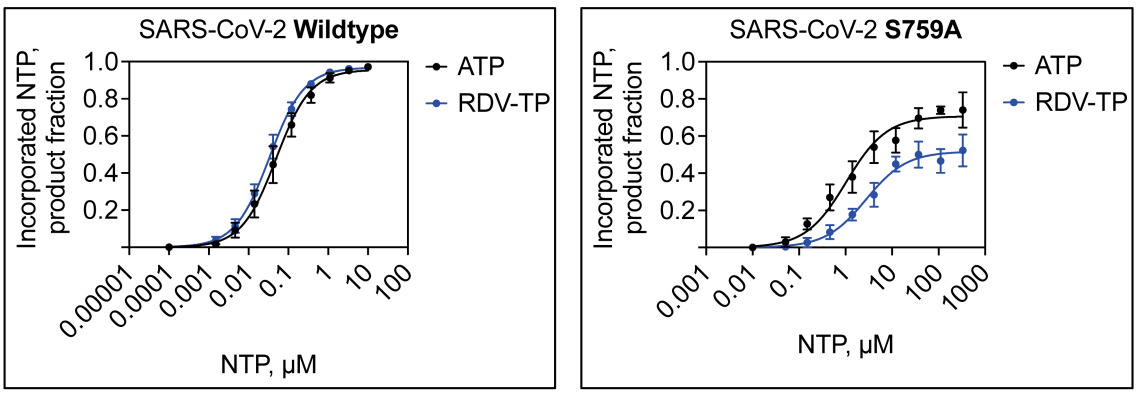
a.



b.



c.

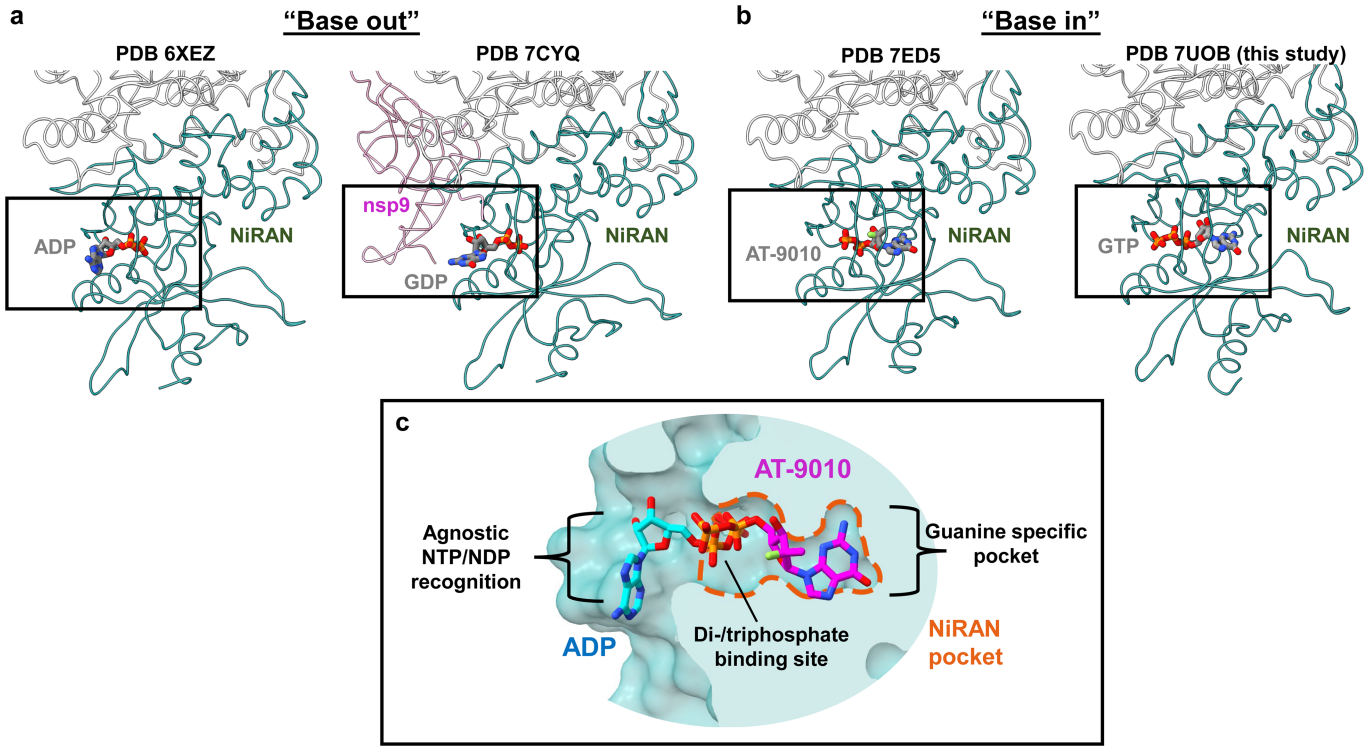


d.

	Wildtype (*n* = 4)		**S759A** (*n* = 3)	
	ATP	**RDV-TP**	**ATP**	**RDV-TP**
V_{max}	0.96 ± 0.013	0.97 ± 0.0085	0.71 ± 0.020	0.52 ± 0.018
K_m	0.049 ± 0.0037	0.033 ± 0.0016	1.0 ± 0.17	2.8 ± 0.48
V_{max}/K_m	20	29	0.71	0.19
Selectivity	**0.69**		**3.7**	
Discrimination	**Reference**		**5.4**	

Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Biochemical analysis of the role of residue S759 in RDV-TP recognition. (**a**) RNA template sequence used to determine the efficiency of adenosine triphosphate (ATP) or RDV-TP incorporation at position 6 (i). G indicates incorporation of [α32P] guanosine triphosphate (GTP) at position 5 (red). (**b**) Migration patterns of the products of ATP or RDV-TP incorporation reactions with wild-type and S759A SARS-CoV-2 RdRp complexes are shown. Main products emerge at position 6. The 5′- 32P-labeled 4-nucleotide primer (4) is used as a size marker (m), illustrating sizing in nucleotides (nts) (**c**) Graphical representation of ATP or RDV-TP single-nucleotide incorporation during RNA synthesis as a function of their respective concentrations shown in (**b**). Best-fit lines illustrate fitting of the data points to Michaelis-Menten kinetics function using GraphPad Prism 7.0. Data shows mean values with error bars illustrating the SDs of the data. All data represent at least three independent experiments (n = 4 for wild-type and n = 3 for S759A SARS-CoV-2 RdRp complexes). (**d**) V_{max} refers to the maximum velocity required to convert substrates into products in a substrate-saturated system, reported as a product fraction of incorporated nucleotide. K_m is a parameter indicating the concentration of substrate at one-half V_{max}, in μM. The standard error of the linear fit is denoted (±). The V_{max}/K_m ratio reflects catalytic efficiency and is used here to determine the selectivity of RDV-TP by taking the ratio between the catalytic efficiency of ATP and of RDV-TP. The discrimination index is defined as the ratio between the mutant RdRp's selectivity values in relation to that of the wild-type.



Extended Data Fig. 10 | Comparison of NiRAN binding poses. (**a**) Models (PDBs 6XEZ & 7CYQ) of the nsp12 NiRAN with a bound base in the 'Base-out' pose. (**b**) Models (PDBs 7ED5 & 7UOB) of the nsp12 NiRAN with a bound base in the 'Base-in' pose. (**c**) Surface representation detailing the NiRAN active site pocket bound to a nucleotide in the 'base-out' pose (PDB 6XEZ) aligned with a structure with a bound nucleotide in the 'base-in' pose (PDB 7ED5).

nature portfolio

Corresponding author(s): Elizabeth A. Campbell, Seth A. Darst

Last updated by author(s): May 19, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Leginon (ref. 36) SerialEM (ref. 43)
Data analysis	Structural biology software was accessed through the SBGrid consortium (ref. 35). MotionCor2 (ref. 37) CryoSPARC v3.1.0 (ref. 38) RELION 3.1 (ref. 40) Bsoft package (ref. 20) 3DFSC package (ref. 41) ChimeraX (ref. 44) Phenix (ref. 42) Thermo Xcalibur Qual Browser v.4.2.47 UniDec v.4.2.0 (ref. 47, 48) Molprobity (ref. 49)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Data availability. All unique/stable reagents generated in this study are available without restriction from the corresponding authors, Seth A. Darst (darst@rockefeller.edu) and E.A. Campbell (campbee@rockefeller.edu) . The cryo-EM density maps and atomic coordinates have been deposited in the EMDataBank and Protein Data Bank as follows: S1_RDV-TP (EMD-26639, PDB 7UO4), S2_ATP (EMD-26641, PDB 7UO7), S3_UTP (EMD-26642, PDB 7UO9), S4_GTP (EMD-26645, 7UOB), S5_CTP (EMD-26646, PDB 7UOE).

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	n/a
Population characteristics	n/a
Recruitment	n/a
Ethics oversight	n/a

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We did not perform power analysis for sample size calculation.
Data exclusions	We excluded some particles from our cryo-EM data based on their 2-D projections and 3D reconstruction due to their shape not resembling our complex such as contaminations and/or low-resolution features.
Replication	Experiments were performed 2-3 biological replicates, each with three technical replicates (see figure legends for detail). Biological replicates were performed on different days within 2 weeks.
Randomization	Randomization does not apply to our study as it does not involve clinical data.
Blinding	Blinding does not apply in our study as it does not involve clinical data.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging