

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

3

4 **Brandon F. Malone<sup>1</sup>, Jason K. Perry<sup>2</sup>, Paul Dominic B. Olinares<sup>3</sup>,  
5 James Chen<sup>1,7</sup>, Todd K. Appelby<sup>2</sup>, Joy Y. Feng<sup>2</sup>, John P. Bilello<sup>2</sup>,  
6 Honkit Ng<sup>4</sup>, Johanna Sotiris<sup>4</sup>, Mark Ebrahim<sup>4</sup>, Eugene Y.D. Chua<sup>5</sup>,  
7 Joshua H. Mendez<sup>5</sup>, Ed T. Eng<sup>5</sup>, Robert Landick<sup>6</sup>, Brian T. Chait<sup>3</sup>,  
8 Elizabeth A. Campbell<sup>1,8</sup>, Seth A. Darst<sup>1,8,9</sup>**

9

10 <sup>1</sup>Laboratory of Molecular Biophysics, The Rockefeller University, New York, NY 10065  
11 USA.

12 <sup>2</sup>Gilead Sciences, Inc., Foster City, CA 94404 USA.

13 <sup>3</sup>Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller  
14 University, New York, NY 10065 USA.

15 <sup>4</sup>The Evelyn Gruss Lipper Cryo-Electron Microscopy Resource Center, The Rockefeller  
16 University, New York, NY 10065, USA

17 <sup>5</sup>National Center for Cryo-EM Access and Training, Simons Electron Microscopy  
18 Center, New York Structural Biology Center, New York, NY 10027 USA.

19 <sup>6</sup>Department of Biochemistry, University of Wisconsin-Madison, Madison, WI, USA

20

21 <sup>7</sup>Current address: Department of Cell Biology, New York University School of Medicine,  
22 New York, NY 10016 USA.

23 <sup>8</sup>Correspondence to: [darst@rockefeller.edu](mailto:darst@rockefeller.edu), campbee@rockefeller.edu

24 <sup>9</sup>Lead contact: [darst@rockefeller.edu](mailto:darst@rockefeller.edu)

25

26

27

28

29

30

31

32

33

34 **The SARS-CoV-2 RNA-dependent RNA polymerase coordinates viral RNA**  
35 **synthesis as part of an assembly known as the replication-transcription complex**  
36 **(RTC)<sup>1</sup>. Accordingly, the RTC is a target for clinically approved antiviral**  
37 **nucleoside analogs, including remdesivir<sup>2</sup>. Faithful synthesis of viral RNAs by the**  
38 **RTC requires recognition of the correct nucleotide triphosphate (NTP) for**  
39 **incorporation into the nascent RNA. To be effective inhibitors, antiviral**  
40 **nucleoside analogs must compete with the natural NTPs for incorporation. How**  
41 **the SARS-CoV-2 RTC discriminates between the natural NTPs, and how antiviral**  
42 **nucleoside analogs compete, has not been discerned in detail. Here, we use cryo-**  
43 **electron microscopy to visualize the RTC bound to each of the natural NTPs in**  
44 **states poised for incorporation. Furthermore, we investigate the RTC with the**  
45 **active metabolite of remdesivir, remdesivir triphosphate (RDV-TP), highlighting**  
46 **the structural basis for the selective incorporation of RDV-TP over its natural**  
47 **counterpart ATP<sup>3,4</sup>. Our results elucidate the suite of interactions required for NTP**  
48 **recognition, informing the rational design of antivirals. Our analysis also yields**  
49 **insights into nucleotide recognition by the nsp12 NiRAN, an enigmatic catalytic**  
50 **domain essential for viral propagation<sup>5</sup>. The NiRAN selectively binds GTP,**  
51 **strengthening proposals for the role of this domain in the formation of the 5' RNA**  
52 **cap<sup>6</sup>.**

53

54 COVID-19, caused by the coronavirus SARS-CoV-2, continues to devastate livelihoods,  
55 and overwhelm healthcare systems worldwide. Given the urgency of stymieing infection  
56 and mitigating disease morbidity, a concerted research effort has begun to elucidate the  
57 molecular details of the viral lifecycle and to design therapeutics to disrupt it <sup>1</sup>. The viral  
58 RNA-dependent RNA polymerase (RdRp, encoded by non-structural protein 12, or  
59 nsp12) functions as part of the holo-RdRp (comprising the RdRp and auxiliary proteins  
60 nsp7 and nsp8 as a heterotetramer nsp7/nsp8<sub>2</sub>/nsp12) in a replication–transcription  
61 complex (holo-RdRp + RNA, or RTC) to direct all viral RNA synthesis in conjunction with  
62 a coterie of viral nucleic acid processing enzymes. In less than two years, targeting the  
63 RTC by the antivirals remdesivir (RDV) and molnupiravir has become the staple of  
64 clinical care <sup>2,7</sup>. This clinical success underscores the importance of the RTC as a  
65 pharmacological target and incentivizes the design of more efficacious nucleotide  
66 analogues for the treatment of COVID-19. Furthermore, due to the highly conserved  
67 nature of the coronavirus RdRp active site <sup>8</sup>, nucleotide analogue inhibitors of the RTC  
68 are excellent candidates for pan-viral inhibitors that would be effective against emerging  
69 variants and may be repurposed to tackle future pathogenic coronaviruses that could  
70 arise.

71 Biochemical and single-molecule experiments have characterized the RNA  
72 synthesis activity of the SARS-CoV-2 RTC, its selectivity for a wide range of nucleotide  
73 analogs, and their effects on RNA synthesis <sup>3,4,9,10</sup>. These investigations yielded insights  
74 into the ability of the RTC to discriminate between natural NTPs and related antiviral  
75 analogs. Studies of remdesivir (RDV) revealed that the active metabolite of RDV,  
76 RDV triphosphate (RDV-TP), possesses enhanced selectivity for incorporation into  
77 nascent RNA over its natural counterpart, ATP <sup>3,4</sup>. Rapid, competitive incorporation is  
78 likely a critical facet of the clinical effectiveness of RDV. However, the structural basis

79 for NTP recognition and RDV-TP selectivity remains unknown, necessitating structural  
80 examinations of stalled RTC ternary complexes with incoming NTP substrates and  
81 catalytic metal ions poised for catalysis (Michaelis or pre-incorporation complexes) (Fig.  
82 1a).

83 Stepwise nucleotide incorporation necessitates transferring the 3'-end of the  
84 product-RNA (p-RNA) to the post-translocated site, allowing binding of the incoming  
85 NTP substrate along with two catalytic metal ions in the active-site cleft (Fig. 1a)<sup>11,12,13</sup>.  
86 In the subsequent catalytic step(s), the geometry of correct Watson-Crick base pairing  
87 of the incoming NTP with the template-RNA (t-RNA) base closes the active site. This  
88 conformational change promotes nucleophilic attack of the p-RNA 3'-OH on the NTP  $\alpha$ -  
89 phosphate, releasing a pyrophosphate in a  $S_N2$  condensation reaction<sup>11</sup>. Available  
90 SARS-CoV-2 RTC structures are dominated by apo structures with an empty NTP  
91 active site lacking catalytic metals and with an open active-site conformation, hindering  
92 structure-based drug design<sup>14,15,16,17</sup>. Only two studies document the RTC in the  
93 presence of substrates: the antiviral inhibitors favipiravir triphosphate and the activated  
94 form of AT-527, AT-9010, but neither investigation yielded views of a catalytically  
95 competent RTC due to unproductive binding modes and open conformations of the  
96 active sites<sup>18,19</sup>. Therefore, to gain mechanistic insights into RTC recognition of its  
97 substrates during RNA synthesis, we determined five cryo-electron microscopy (cryo-  
98 EM) structures, capturing the RTC with each of the four natural NTPs as well as with  
99 RDV-TP. The five cryo-EM maps range in nominal resolutions between ~2.6-3.3 Å, with  
100 the RdRp active sites resolved locally to ~2.2-2.9 Å (Extended Data Fig. 1-5, Extended  
101 Data Table 1)<sup>20</sup>, enabling near-atomic resolution insights into the mechanism of NTP  
102 recognition and RDV-TP selectivity.

103

#### 104 **Trapping SARS-CoV-2 RTC pre-incorporation complexes**

105 Guided by previous mechanistic work on nucleic acid polymerases, we investigated a  
106 battery of chemical strategies to block incorporation after substrate binding. These  
107 approaches included using nucleotide diphosphate (NDP) substrates<sup>13</sup>, alternative  
108 metal co-factors (such as  $Ca^{2+}$ )<sup>21</sup>,  $\alpha$ - $\beta$  non-hydrolysable nucleotide analogues<sup>22</sup>, and  
109 3'-deoxy p-RNA scaffolds. Although the use of NDP substrates trapped the hepatitis C  
110 virus (HCV) RdRp recognition complex<sup>13</sup>, the SARS-CoV-2 RTC efficiently incorporated  
111 GDP, less efficiently ADP, and to a minor extent RDV-diphosphate (RDV-DP)  
112 (Extended Data Fig. 6, see also<sup>23</sup>). The ability of the SARS-CoV-2 RTC to incorporate  
113 NDPs reflects the relative promiscuity of the RdRp active site, which also retained RNA  
114 synthesis activity in the presence of both  $Ca^{2+}$  and  $\alpha$ - $\beta$ -imino analogues (Extended Data  
115 Fig. 6e). However, RTC synthesis activity was fully ablated with the incorporation of a  
116 3'-deoxy nucleotide into the p-RNA, leading us to utilize this approach to visualize the  
117 RTC pre-incorporation complexes (Extended Data Fig. 6g).

118 To validate the effectiveness of using 3'-deoxy RNA scaffolds, we used native  
119 mass spectrometry (nMS) to probe for the extension of the primer by one base or the  
120 formation of a stalled ternary complex (Fig 1b, c, Extended Data Fig. 7) in the presence  
121 of various NTPs. In the event of nucleotide incorporation, we expected a mass-shift  
122 corresponding to NMP addition, whereas non-covalently bound (unincorporated)

123 substrates would either dissociate or yield a mass-shift corresponding to the NTP and  
124 bound metal ions. For the control scaffolds containing a 3'-OH, near complete p-RNA  
125 extension by a single NMP occurred (Fig. 1b). In the presence of p-RNA strands lacking  
126 a 3'-OH, no incorporation of the next nucleotide (ATP, RDV-TP, or GTP) was observed,  
127 confirming the suitability of this strategy for trapping pre-incorporation complexes (Fig.  
128 1c). In the pre-incorporation complexes, we observed peaks corresponding to the RTC  
129 with the incoming Mg-NTP (ATP, RDV-TP, or GTP) with the relative peak intensity in  
130 the order of GTP > RDV-TP > ATP (Fig. 1c).

131

## 132 **Structural basis of NTP recognition**

133 Having identified a strategy to capture substrate-bound RTCs stalled in a pre-  
134 incorporation state, we prepared cryo-EM samples of RTCs with each of the natural  
135 nucleotides, ATP, CTP, GTP, UTP, or with RDV-TP (Extended Data Fig. 1-5). The  
136 resulting structures yielded the first views of the RTC poised for catalysis. Two of the  
137 structures, the RTC with CTP (+CTP) or GTP (+GTP), yielded nominal resolutions  
138 better than 3 Å (2.7 Å), with local resolution around the active site estimated to be ~2.2  
139 Å<sup>20</sup>, enabling near-atomic resolution insights into substrate recognition.

140 Density features in the +CTP and +GTP cryo-EM maps are consistent with the  
141 positions of divalent metal ions observed in many previous nucleic acid polymerase  
142 structures<sup>11-13, 24</sup>. Based on the presence of 5 mM MgCl<sub>2</sub> in the cryo-EM buffer and the  
143 observed octahedral coordination geometries, we assigned these densities as Mg<sup>2+</sup><sub>A</sub>  
144 and Mg<sup>2+</sup><sub>B</sub> (Fig. 2b, Supplementary Video 1), consistent with the hypothesized universal  
145 two-metal ion mechanism for nucleic acid polymerases<sup>11</sup>. The density peak for Mg<sup>2+</sup><sub>A</sub> in  
146 the +CTP and +GTP structures is weaker than that for Mg<sup>2+</sup><sub>B</sub>, and density for Mg<sup>2+</sup><sub>A</sub> in  
147 the +ATP, +UTP, and +RDV-TP structures is completely absent. Despite the absence of  
148 Mg<sup>2+</sup><sub>A</sub> in the +ATP, +UTP, and +RDV-TP structures, most details of the active-site  
149 configurations were indistinguishable from the +CTP and +GTP structures.

150 A consistent feature observed across all five structures is the closure of the  
151 active site around the incoming NTP, mediated by a rotation of the RdRp motif A  
152 towards the NTP substrate (Fig. 2c, d). Closure of motif A stabilizes substrate binding  
153 by: (I) enabling the backbone carbonyl of Y619 to coordinate Mg<sup>2+</sup><sub>B</sub> and the catalytic  
154 residue D618 to coordinate both Mg<sup>2+</sup><sub>A</sub> and Mg<sup>2+</sup><sub>B</sub> (Fig. 2e), (II) promoting the formation  
155 of a hydrogen-bonding (H-bonding) network through D623 that enables the recognition  
156 of the substrate ribose 3'-OH by S682 and N691 (motif B), (III) enabling weak H-  
157 bonding interactions between the β- and γ-phosphates and motif A residues K621 or  
158 C622 (Fig. 2f, Extended Data Fig. 9f); and (IV) disrupting the polar D618-K798  
159 interaction in the apo-RTC<sup>14, 15, 17</sup>, repositioning D618 for metal coordination, and K798  
160 to interact with the NTP γ-phosphate (Fig. 2d).

161 The detailed interactions of nsp12 residues with the four natural NTP substrates  
162 were essentially identical except for conserved basic residues of motif F. K545 (motif F)  
163 displays selective interactions with the incoming NTPs, forming a H-bond with carbonyl  
164 oxygens of UTP (C4) and GTP (C6), interactions that are precluded by the amino group  
165 at these positions in the ATP, RDV-TP, or CTP bases (Extended Data Fig. 9). The cryo-  
166 EM densities also suggest that the side chains of K551 and R555 are dynamic,

167 accessing multiple conformations in each structure. K551 primarily interacts with the  $\gamma$ -  
168 phosphate of the CTP and RDV-TP substrates, but weak density indicates disorder in  
169 the other structures. In the +ATP and +RDV-TP structures, R555 predominantly forms a  
170 pi-pi stacking interaction with the NTP substrate base. This differs from the predominant  
171 conformation in the +CTP, +GTP, and +UTP structures where the side chain of R555  
172 forms a H-bond with the  $\beta$ -phosphate oxygens of the NTP substrate (Fig. 3, Extended  
173 data Fig. 9). In either state, R555 buttresses the incoming nucleotide, arranging the  
174 active site for catalysis. We speculate that the observed R555 stacking interaction may  
175 first promote the formation of a canonical Watson-Crick base pair by stabilizing the  
176 nucleobase opposite the template base, then reorient to interact with the nucleotide  $\beta$ -  
177 phosphate to promote subsequent catalysis by reinforcing the correct geometry of the  
178 pyrophosphate leaving group, roles that may be conserved across RdRps such as HCV  
179 RdRp<sup>25</sup>.

180 The structural observation of SARS-CoV-2 RdRp motif A closure resembles a  
181 similar motion of motif A observed in substrate-bound RdRp complexes of polio virus  
182 and HCV<sup>12, 13</sup>, illustrating the universal nature of the RdRp palm closure step during  
183 NTP recognition. Although the palm-mediated active site closure is well-documented,  
184 our results indicate that the structural plasticity of positively charged residues in motif F  
185 (K545, K551, R553, and R555) in the fingers domain helps orient the NTP substrate for  
186 catalysis through interactions with both the nucleobase and triphosphate moieties. We  
187 note that these residues are invariant across the major coronavirus clades and GISAID  
188 database of patient SARS-CoV-2 sequences (Extended Data Fig. 8), reflecting the near  
189 immutable nature of the coronavirus RdRp active site. Previously, we noted that the  
190 same residues line the NTP entry channel following nsp13-mediated p-RNA  
191 backtracking<sup>26</sup>, illustrating the pleiotropic role that motif F plays during RNA synthesis.

192

### 193 **Structural basis of RDV-TP selectivity**

194 RDV-TP has been characterized in biochemical studies to possess an almost threefold  
195 greater selectivity for incorporation into elongating RNA compared to ATP<sup>3, 4</sup>, a property  
196 which is thought to improve its ability to inhibit RNA synthesis. In our structures, both  
197 nucleotides are observed base-paired to the cognate t-RNA U; the two nucleotides  
198 superimpose with a root-mean-square-deviation (RMSD) of 0.34 Å over 32 common  
199 atoms. The key difference between RDV-TP and ATP is the 1'-cyano moiety on the  
200 RDV-TP ribose (Fig. 3a), which juts into a hydrophilic pocket formed by residues T687,  
201 N691 (motif B), and S759 (motif C) (Fig. 3c), creating a network of polar interactions  
202 (Fig. 3c). In the highest resolution cryo-EM maps (+CTP and +GTP; Extended Data  
203 Table 1), we observed a stably bound water molecule occupying the hydrophilic pocket  
204 in the absence of the RDV-TP cyano-group (compare Fig. 3c to 3d, e), suggesting that  
205 the enhanced affinity (lower  $K_m$ ) for RDV-TP over ATP<sup>3, 4</sup> may be attributed to both an  
206 entropic effect through release of the bound water and an enthalpic effect through  
207 formation of new polar interactions between RDV-TP and the surrounding polar  
208 residues. A recent study reports that SARS-CoV-2 acquired phenotypic resistance to  
209 RDV through an nsp12-S759A substitution following serial passage in cell culture<sup>27</sup>,  
210 indicating that binding of the RDV-TP 1'-cyano group in this hydrophilic pocket is an  
211 important facet of RDV susceptibility.

212

213 **Structural insights into RNA capping**

214 Apart from the RdRp active site, the RTC possesses an additional catalytic domain  
215 known as the NiRAN (Nidovirus RdRp-associated nucleotidyltransferase), which lies N-  
216 terminal of the RdRp<sup>5, 14, 28</sup>. Although the NiRAN domain is essential for viral  
217 propagation, its functions during the viral life cycle remain enigmatic. Recent studies  
218 have suggested that the NiRAN functions in the viral RNA capping pathway in  
219 conjunction with an additional viral protein, nsp9<sup>5, 29</sup>. Others have postulated a role for  
220 the NiRAN domain in protein-mediated priming of RNA synthesis<sup>30</sup>. In previous  
221 structural studies, ADP<sup>14</sup> or GDP<sup>29</sup> have been observed bound in the NiRAN active  
222 site in a 'base-out' pose, where the nucleotide base points out of the active site pocket  
223 and makes few, if any, protein contacts. ADP has not been shown to serve as a  
224 substrate for the NiRAN enzymatic activity, and the GDP was bound along with an N-  
225 terminally modified (and resultingly inactive) nsp9<sup>29</sup>, so the relevance of these  
226 nucleotide poses is unclear. In another study, the diphosphate form of the GTP analog  
227 AT-9010 is seen bound to the NiRAN in a 'base-in' pose<sup>19</sup>, with the phosphates  
228 occupying the same positions as the phosphates in the 'base-out' pose but flipped, and  
229 with the guanine base bound in a tight pocket internal to the NiRAN (Extended Data  
230 Fig. 10).

231 In our cryo-EM structures of the RTC with each of the natural NTPs and RDV-TP  
232 (Extended Data Table 1), only GTP was observed stably and specifically bound in the  
233 NiRAN active site. The cryo-EM density supports the presence of GTP (and not GDP) in  
234 the NiRAN (Fig. 4). The GTP is bound in the 'base-in' pose similar to the diphosphate  
235 form of AT-9010<sup>19</sup>, with the base buried in an apparently guanine-specific pocket that  
236 extends into the core of the NiRAN fold, enabling it to make key contacts with  
237 hydrophilic residues that line the pocket interior. A central element to this recognition is  
238 interactions mediated by R55 and Y217, two conserved residues across the  $\alpha$ - and  $\beta$ -  
239 coronavirus clades, which provide base specificity for guanosine (Fig. 4 b, d). The apo-  
240 NiRAN pocket is sterically incompatible with the bound guanine base; GTP association  
241 is mediated by an induced fit mechanism that involves expansion of the active-site  
242 pocket to accommodate the guanosine base (Fig. 4e).

243 A recent preprint proposes that the NiRAN mediates two successive steps in the  
244 viral RNA capping pathway that initially entails transfer of the nascent RNA 5'-end to the  
245 amino terminus of nsp9, forming a covalent RNA-protein intermediate in a process  
246 termed RNAylation<sup>6</sup>. A successive polyribonucleotidyltransferase (PRNTase) reaction  
247 utilizes the RNAylated nsp9 as a substrate and transfers the bound RNA to GDP/GTP  
248 to produce the cap structure, GpppN-RNA. The PRNTase activity strictly utilizes GDP or  
249 GTP. Although the authors report slightly higher PRNTase activity when GDP is the  
250 substrate, GTP may be the more likely physiological substrate given that concentrations  
251 of GTP are much higher than GDP in the cellular milieu<sup>31</sup>. Our results indicate that the  
252 NiRAN uniquely recognizes GTP (but not other NTPs), explaining the base specificity  
253 observed for the PRNTase activity.

254

255

256 **Discussion**

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

268 To probe the mechanism of nucleotide recognition by the SARS-CoV-2 RdRp,  
269 we determined structures of stalled RdRp complexes containing an RNA primer-  
270 template, incoming nucleotide substrates or the antiviral RDV-TP, and catalytic metal  
271 ions. Structural views of the RTC with each of the respective nucleotides illustrate how  
272 the RdRp structurally adapts to proper Watson-Crick base pairing geometry, closing in  
273 around the t-RNA/NTP base pair to facilitate catalytic Mg<sup>2+</sup> coordination and in-line  
274 attack of the p-RNA 3'-OH on the NTP  $\alpha$ -phosphate.

275 Through visualizing these structures, we ascertained that the enhanced  
276 selectivity for RDV, which primarily materializes biochemically as an effect on  $K_m$ , is  
277 mediated by the accommodation of its 1'-cyano group in a conserved hydrophilic pocket  
278 near the RdRp active site (Fig. 3). This same RdRp active site pocket harbors an  
279 ordered water molecule in the +CTP and +GTP structures (Fig. 3); the bound water is  
280 likely present in the +ATP and +UTP complexes as well but not visualized due to limited  
281 resolution (Extended Data Figs. 2, 3). Mutations that alter the hydrophilicity of this  
282 pocket give rise to RDV resistance <sup>27</sup> and are naturally found in viral families that exhibit  
283 reduced sensitivity to RDV <sup>4</sup>. Critically, structural insights into NTP recognition could be  
284 exploited in the design of nucleotide analogues as part of structure-based drug design  
285 programs that aim to target the coronavirus RdRp. Our studies also shed light on the  
286 presence of a guanosine-specific pocket in the essential NiRAN domain, supporting  
287 proposals that the NiRAN is involved in the production of capped mRNAs as well as  
288 further detailing this pocket as a therapeutic target.

289  
290  
291

292 **REFERENCES**

293

294 (1) Malone, B.; Urakova, N.; Snijder, E. J.; Campbell, E. A. Structures and functions of  
295 coronavirus replication–transcription complexes and their relevance for SARS-CoV-2 drug  
296 design. *Nature Reviews Molecular Cell Biology* 2022, 23 (1), 21-39. DOI: 10.1038/s41580-021-  
297 00432-z.

298 (2) Beigel, J. H.; Tomashek, K. M.; Dodd, L. E.; Mehta, A. K.; Zingman, B. S.; Kalil, A. C.;  
299 Hohmann, E.; Chu, H. Y.; Luetkemeyer, A.; Kline, S.; et al. Remdesivir for the Treatment of  
300 Covid-19 — Final Report. *New England Journal of Medicine* 2020, 383 (19), 1813-1826. DOI:  
301 10.1056/NEJMoa2007764 (acccesed 2021/02/11).

302 (3) Dangerfield, T. L.; Huang, N. Z.; Johnson, K. A. Remdesivir Is Effective in Combating  
303 COVID-19 because It Is a Better Substrate than ATP for the Viral RNA-Dependent RNA  
304 Polymerase. *iScience* 2020, 23 (12), 101849. DOI: <https://doi.org/10.1016/j.isci.2020.101849>.

305 (4) Gordon, C. J.; Tchesnokov, E. P.; Woolner, E.; Perry, J. K.; Feng, J. Y.; Porter, D. P.; Götte,  
306 M. Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from  
307 severe acute respiratory syndrome coronavirus 2 with high potency. *J Biol Chem* 2020, 295 (20),  
308 6785-6797. DOI: 10.1074/jbc.RA120.013679 From NLM.

309 (5) Slanina, H.; Madhugiri, R.; Bylapudi, G.; Schultheiß, K.; Karl, N.; Gulyaeva, A.;  
310 Gorbalyena, A. E.; Linne, U.; Ziebuhr, J. Coronavirus replication–transcription complex: Vital  
311 and selective NMPylation of a conserved site in nsp9 by the NiRAN-RdRp subunit. *Proceedings  
312 of the National Academy of Sciences* 2021, 118 (6), e2022310118. DOI:  
313 10.1073/pnas.2022310118.

314 (6) Park, G. J.; Osinski, A.; Hernandez, G.; Eitson, J. L.; Majumdar, A.; Tonelli, M.; Henzler-  
315 Wildman, K.; Pawłowski, K.; Chen, Z.; Li, Y.; et al. The mechanism of RNA capping by SARS-  
316 CoV-2. *bioRxiv* 2022, 2022.2002.2007.479471. DOI: 10.1101/2022.02.07.479471.

317 (7) Jayk Bernal, A.; Gomes da Silva, M. M.; Musungaie, D. B.; Kovalchuk, E.; Gonzalez, A.;  
318 Delos Reyes, V.; Martín-Quirós, A.; Caraco, Y.; Williams-Diaz, A.; Brown, M. L.; et al.  
319 Molnupiravir for Oral Treatment of Covid-19 in Nonhospitalized Patients. *New England Journal  
320 of Medicine* 2021, 386 (6), 509-520. DOI: 10.1056/NEJMoa2116044 (acccesed 2022/04/01).

321 (8) Bruenn, J. A. A structural and primary sequence comparison of the viral RNA-dependent  
322 RNA polymerases. *Nucleic Acids Research* 2003, 31 (7), 1821-1829. DOI: 10.1093/nar/gkg277  
323 (acccesed 8/18/2020).

324 (9) Bravo, J. P. K.; Dangerfield, T. L.; Taylor, D. W.; Johnson, K. A. Remdesivir is a delayed  
325 translocation inhibitor of SARS CoV-2 replication. *Molecular Cell*. DOI:  
326 10.1016/j.molcel.2021.01.035 (acccesed 2021/02/04).

327 (10) Seifert, M.; Bera, S. C.; van Nies, P.; Kirchdoerfer, R. N.; Shannon, A.; Le, T.-T.-N.; Meng,  
328 X.; Xia, H.; Wood, J. M.; Harris, L. D.; et al. Inhibition of SARS-CoV-2 polymerase by  
329 nucleotide analogs from a single-molecule perspective. *eLife* 2021, 10, e70968. DOI:  
330 10.7554/eLife.70968.

331 (11) Steitz, T. A. DNA Polymerases: Structural Diversity and Common Mechanisms \*. *Journal  
332 of Biological Chemistry* 1999, 274 (25), 17395-17398. DOI: 10.1074/jbc.274.25.17395  
333 (acccesed 2021/02/15).

334 (12) Shu, B.; Gong, P. Structural basis of viral RNA-dependent RNA polymerase catalysis and  
335 translocation. *Proceedings of the National Academy of Sciences* 2016, 113 (28), E4005-E4014.  
336 DOI: 10.1073/pnas.1602591113 (acccesed 2022/04/01).

337 (13) Appleby, T. C.; Perry, J. K.; Murakami, E.; Barauskas, O.; Feng, J.; Cho, A.; Fox, D.;  
338 Wetmore, D. R.; McGrath, M. E.; Ray, A. S.; et al. Structural basis for RNA replication by the  
339 hepatitis C virus polymerase. *Science* 2015, 347 (6223), 771. DOI: 10.1126/science.1259210.  
340 (14) Chen, J.; Malone, B.; Llewellyn, E.; Grasso, M.; Shelton, P. M. M.; Olinares, P. D. B.;  
341 Maruthi, K.; Eng, E. T.; Vatandaslar, H.; Chait, B. T.; et al. Structural Basis for Helicase-  
342 Polymerase Coupling in the SARS-CoV-2 Replication-Transcription Complex. *Cell* 2020. DOI:  
343 10.1016/j.cell.2020.07.033 From NLM.  
344 (15) Hillen, H. S.; Kokic, G.; Farnung, L.; Dienemann, C.; Tegunov, D.; Cramer, P. Structure of  
345 replicating SARS-CoV-2 polymerase. *Nature* 2020, 584 (7819), 154-156. DOI: 10.1038/s41586-  
346 020-2368-8.  
347 (16) Wang, Q.; Wu, J.; Wang, H.; Gao, Y.; Liu, Q.; Mu, A.; Ji, W.; Yan, L.; Zhu, Y.; Zhu, C.; et  
348 al. Structural Basis for RNA Replication by the SARS-CoV-2 Polymerase. *Cell* 2020, 182 (2),  
349 417-428.e413. DOI: <https://doi.org/10.1016/j.cell.2020.05.034>.  
350 (17) Chen, J.; Wang, Q.; Malone, B.; Llewellyn, E.; Pechersky, Y.; Maruthi, K.; Eng, E. T.;  
351 Perry, J. K.; Campbell, E. A.; Shaw, D. E.; et al. Ensemble cryo-EM reveals conformational  
352 states of the nsp13 helicase in the SARS-CoV-2 helicase replication-transcription complex.  
353 *Nature Structural & Molecular Biology* 2022, 29 (3), 250-260. DOI: 10.1038/s41594-022-  
354 00734-6.  
355 (18) Naydenova, K.; Muir, K. W.; Wu, L.-F.; Zhang, Z.; Coscia, F.; Peet, M. J.; Castro-  
356 Hartmann, P.; Qian, P.; Sader, K.; Dent, K.; et al. Structure of the SARS-CoV-2 RNA-dependent  
357 RNA polymerase in the presence of favipiravir-RTP. *Proceedings of the National Academy of  
358 Sciences* 2021, 118 (7), e2021946118. DOI: 10.1073/pnas.2021946118.  
359 (19) Shannon, A.; Fattorini, V.; Sama, B.; Selisko, B.; Feracci, M.; Falcou, C.; Gauffre, P.; El  
360 Kazzi, P.; Delpal, A.; Decroly, E.; et al. A dual mechanism of action of AT-527 against SARS-  
361 CoV-2 polymerase. *Nature Communications* 2022, 13 (1), 621. DOI: 10.1038/s41467-022-  
362 28113-1.  
363 (20) Cardone, G.; Heymann, J. B.; Steven, A. C. One number does not fit all: mapping local  
364 variations in resolution in cryo-EM reconstructions. *Journal of structural biology* 2013, 184 (2),  
365 226-236. DOI: 10.1016/j.jsb.2013.08.002 PubMed.  
366 (21) Salgado, P. S.; Makeyev, E. V.; Butcher, S. J.; Bamford, D. H.; Stuart, D. I.; Grimes, J. M.  
367 The Structural Basis for RNA Specificity and Ca2+ Inhibition of an RNA-Dependent RNA  
368 Polymerase. *Structure* 2004, 12 (2), 307-316. DOI: <https://doi.org/10.1016/j.str.2004.01.012>.  
369 (22) Temiakov, D.; Patlan, V.; Anikin, M.; McAllister, W. T.; Yokoyama, S.; Vassylyev, D. G.  
370 Structural Basis for Substrate Selection by T7 RNA Polymerase. *Cell* 2004, 116 (3), 381-391.  
371 DOI: [https://doi.org/10.1016/S0092-8674\(04\)00059-5](https://doi.org/10.1016/S0092-8674(04)00059-5).  
372 (23) Wang, M.; Wu, C.; Liu, N.; Zhang, F.; Dong, H.; Wang, S.; Chen, M.; Jiang, X.; Gu, L.  
373 SARS-CoV-2 RdRp is a versatile enzyme with proofreading activity and ability to incorporate  
374 NHC into RNA by using diphosphate form molnupiravir as a substrate. *bioRxiv* 2021,  
375 2021.2011.2015.468737. DOI: 10.1101/2021.11.15.468737.  
376 (24) Yang, W.; Lee, J. Y.; Nowotny, M. Making and breaking nucleic acids: two-Mg2+-ion  
377 catalysis and substrate specificity. *Mol Cell* 2006, 22 (1), 5-13. DOI:  
378 10.1016/j.molcel.2006.03.013 From NLM.  
379 (25) Ben Ouirane, K.; Boulard, Y.; Bressanelli, S. The hepatitis C virus RNA-dependent RNA  
380 polymerase directs incoming nucleotides to its active site through magnesium-dependent  
381 dynamics within its F motif. *Journal of Biological Chemistry* 2019, 294 (19), 7573-7587. DOI:  
382 10.1074/jbc.RA118.005209 (acccesed 2021/02/15).

383 (26) Malone, B.; Chen, J.; Wang, Q.; Llewellyn, E.; Choi, Y. J.; Olinares, P. D. B.; Cao, X.;  
384 Hernandez, C.; Eng, E. T.; Chait, B. T.; et al. Structural basis for backtracking by the SARS-  
385 CoV-2 replication–transcription complex. *Proceedings of the National Academy of Sciences*  
386 2021, *118* (19), e2102516118. DOI: 10.1073/pnas.2102516118.

387 (27) Stevens Laura, J.; Pruijssers Andrea, J.; Lee Hery, W.; Gordon Calvin, J.; Tchesnokov Egor,  
388 P.; Gribble, J.; George Amelia, S.; Hughes Tia, M.; Lu, X.; Li, J.; et al. Mutations in the SARS-  
389 CoV-2 RNA dependent RNA polymerase confer resistance to remdesivir by distinct  
390 mechanisms. *Science Translational Medicine* *0* (0), eab00718. DOI:  
391 10.1126/scitranslmed.ab00718 (acccesed 2022/04/30).

392 (28) Kirchdoerfer, R. N.; Ward, A. B. Structure of the SARS-CoV nsp12 polymerase bound to  
393 nsp7 and nsp8 co-factors. *Nature Communications* 2019, *10* (1), 2342. DOI: 10.1038/s41467-  
394 019-10280-3.

395 (29) Yan, L.; Ge, J.; Zheng, L.; Zhang, Y.; Gao, Y.; Wang, T.; Huang, Y.; Yang, Y.; Gao, S.; Li,  
396 M.; et al. Cryo-EM Structure of an Extended SARS-CoV-2 Replication and Transcription  
397 Complex Reveals an Intermediate State in Cap Synthesis. *Cell* 2021, *184* (1), 184-193.e110.  
398 DOI: <https://doi.org/10.1016/j.cell.2020.11.016>.

399 (30) Shannon, A.; Fattorini, V.; Sama, B.; Selisko, B.; Feracci, M.; Falcou, C.; Gauffre, P.;  
400 Kazzi, P. E.; Decroly, E.; Rabah, N.; et al. Protein-primed RNA synthesis in SARS-CoVs and  
401 structural basis for inhibition by AT-527. *bioRxiv* 2021, 2021.2003.2023.436564. DOI:  
402 10.1101/2021.03.23.436564.

403 (31) Traut, T. W. Physiological concentrations of purines and pyrimidines. *Mol Cell Biochem*  
404 1994, *140* (1), 1-22. DOI: 10.1007/bf00928361 From NLM.

405 (32) Mulangu, S.; Dodd, L. E.; Davey, R. T., Jr.; Tshiani Mbaya, O.; Proschan, M.; Mukadi, D.;  
406 Lusakibanza Manzo, M.; Nzolo, D.; Tshomba Oloma, A.; Ibanda, A.; et al. A Randomized,  
407 Controlled Trial of Ebola Virus Disease Therapeutics. *N Engl J Med* 2019, *381* (24), 2293-2303.  
408 DOI: 10.1056/NEJMoa1910993 From NLM.

409 (33) Painter, G. R.; Natchus, M. G.; Cohen, O.; Holman, W.; Painter, W. P. Developing a direct  
410 acting, orally available antiviral agent in a pandemic: the evolution of molnupiravir as a potential  
411 treatment for COVID-19. *Current opinion in virology* 2021, *50*, 17-22. DOI:  
412 10.1016/j.coviro.2021.06.003 PubMed.

413 (34) Dangerfield, T. L.; Huang, N. Z.; Johnson, K. A. Expression and purification of tag-free  
414 SARS-CoV-2 RNA-dependent RNA polymerase in Escherichia coli. *STAR Protoc* 2021, *2* (1),  
415 100357. DOI: 10.1016/j.xpro.2021.100357 From NLM.

416 (35) Morin, A.; Eisenbraun, B.; Key, J.; Sanschagrin, P. C.; Timony, M. A.; Ottaviano, M.; Sliz,  
417 P. Collaboration gets the most out of software. *eLife* 2013, *2*, e01456. DOI:  
418 10.7554/eLife.01456.

419 (36) Suloway, C.; Pulokas, J.; Fellmann, D.; Cheng, A.; Guerra, F.; Quispe, J.; Stagg, S.; Potter,  
420 C. S.; Carragher, B. Automated molecular microscopy: the new Leginon system. *J. Struct. Biol.*  
421 2005, *151* (1), 41-60. DOI: 10.1016/j.jsb.2005.03.010.

422 (37) Zheng, S. Q.; Palovcak, E.; Armache, J. P.; Verba, K. A.; Cheng, Y.; Agard, D. A.  
423 MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron  
424 microscopy. *Nat. Methods* 2017, *14* (4), 331-332. DOI: 10.1038/nmeth.4193.

425 (38) Punjani, A.; Rubinstein, J. L.; Fleet, D. J.; Brubaker, M. A. cryoSPARC: algorithms for  
426 rapid unsupervised cryo-EM structure determination. *Nat. Methods* 2017, *14* (3), 290-296. DOI:  
427 10.1038/nmeth.4169.

428 (39) Punjani, A.; Zhang, H.; Fleet, D. J. Non-uniform refinement: adaptive regularization  
429 improves single-particle cryo-EM reconstruction. *Nat Methods* 2020, *17* (12), 1214-1221. DOI:  
430 10.1038/s41592-020-00990-8 From NLM.

431 (40) Zivanov, J.; Nakane, T.; Forsberg, B. O.; Kimanis, D.; Hagen, W. J.; Lindahl, E.; Scheres,  
432 S. H. New tools for automated high-resolution cryo-EM structure determination in RELION-3.  
433 *eLife* 2018, *7*. DOI: 10.7554/eLife.42166.

434 (41) Tan, Y. Z.; Baldwin, P. R.; Davis, J. H.; Williamson, J. R.; Potter, C. S.; Carragher, B.;  
435 Lyumkis, D. Addressing preferred specimen orientation in single-particle cryo-EM through  
436 tilting. *Nat. Methods* 2017, *14* (8), 793-796. DOI: 10.1038/nmeth.4347.

437 (42) Afonine, P. V.; Klaholz, B. P.; Moriarty, N. W.; Poon, B. K.; Sobolev, O. V.; Terwilliger, T.  
438 C.; Adams, P. D.; Urzhumtsev, A. New tools for the analysis and validation of cryo-EM maps  
439 and atomic models. *Acta Crystallogr D Struct Biol* 2018, *74* (Pt 9), 814-840. DOI:  
440 10.1107/s2059798318009324 From NLM.

441 (43) Mastronarde, D. N. SerialEM: A Program for Automated Tilt Series Acquisition on Tecnai  
442 Microscopes Using Prediction of Specimen Position. *Microscopy and Microanalysis* 2003, *9*  
443 (S02), 1182-1183. DOI: 10.1017/S1431927603445911 From Cambridge University Press  
444 Cambridge Core.

445 (44) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E.  
446 C.; Ferrin, T. E. UCSF Chimera—A visualization system for exploratory research and analysis.  
447 *Journal of Computational Chemistry* 2004, *25* (13), 1605-1612. DOI: 10.1002/jcc.20084.

448 (45) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr  
449 D Biol Crystallogr* 2004, *60* (Pt 12 Pt 1), 2126-2132. DOI: 10.1107/s0907444904019158 From  
450 NLM.

451 (46) Olinares, P. D. B.; Chait, B. T. Native Mass Spectrometry Analysis of Affinity-Captured  
452 Endogenous Yeast RNA Exosome Complexes. *Methods Mol Biol* 2020, *2062*, 357-382. DOI:  
453 10.1007/978-1-4939-9822-7\_17 From NLM.

454 (47) Marty, M. T.; Baldwin, A. J.; Marklund, E. G.; Hochberg, G. K. A.; Benesch, J. L. P.;  
455 Robinson, C. V. Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary  
456 Interactions to Polydisperse Ensembles. *Analytical Chemistry* 2015, *87* (8), 4370-4376. DOI:  
457 10.1021/acs.analchem.5b00140.

458 (48) Reid, D. J.; Diesing, J. M.; Miller, M. A.; Perry, S. M.; Wales, J. A.; Montfort, W. R.;  
459 Marty, M. T. MetaUniDec: High-Throughput Deconvolution of Native Mass Spectra. *J Am Soc  
460 Mass Spectrom* 2019, *30* (1), 118-127. DOI: 10.1007/s13361-018-1951-9 From NLM.

461 (49) Chen, V. B.; Arendall, W. B., 3rd; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral,  
462 G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: all-atom structure  
463 validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 2010, *66* (Pt  
464 1), 12-21. DOI: 10.1107/s0907444909042073 From NLM.

465 (50) Adams, P. D.; Afonine, P. V.; Bunkóczki, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd,  
466 J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; et al. PHENIX: a comprehensive  
467 Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*  
468 2010, *66* (Pt 2), 213-221. DOI: 10.1107/s0907444909052925 From NLM.

469

470

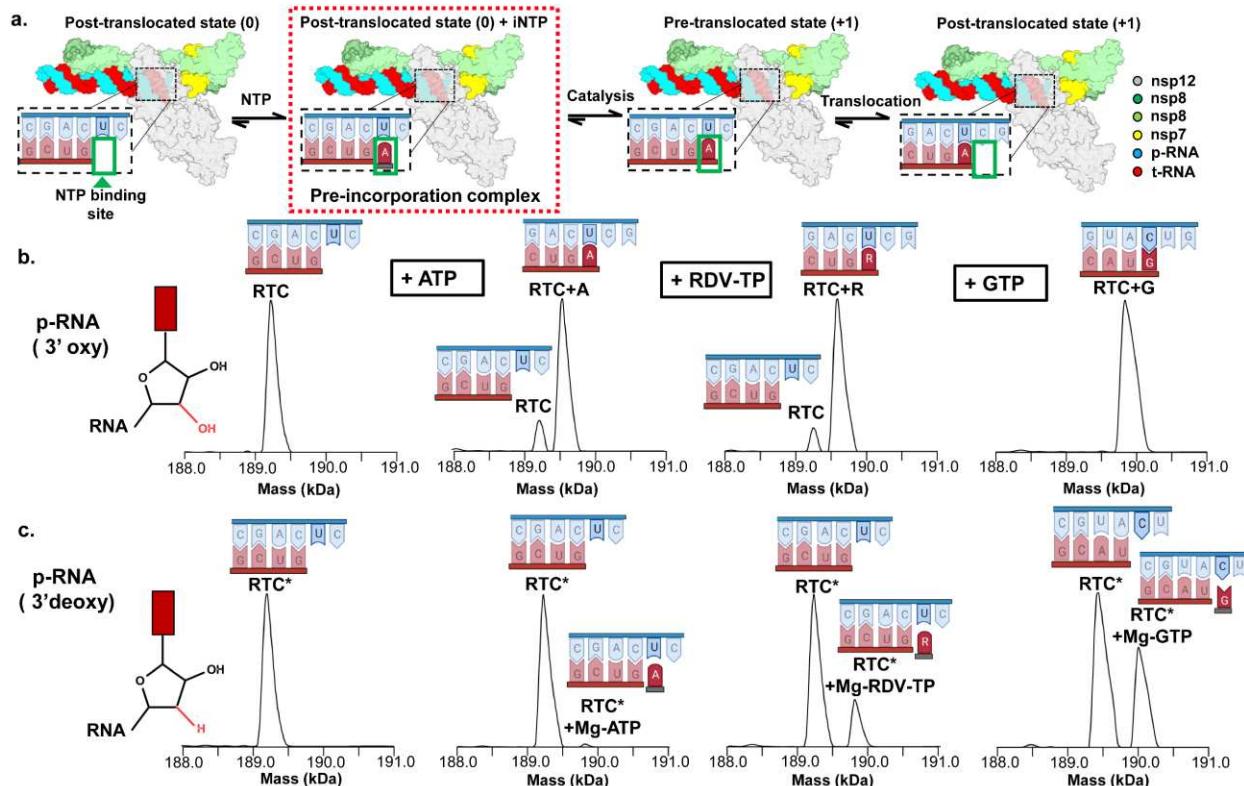
471

472

473

474 **Fig. 1 | Capturing the SARS-CoV-2 RTC ternary complex**

475



476

477

478

479 **Fig 1. (a)** Schematic depicting the major steps of the nucleotide addition cycle of the  
480 replication-transcription complex. The pre-incorporation complex studied here is  
481 highlighted (red dashed box). **(b)** Native mass spectrometry (nMS) analysis of the RTC  
482 bound to a 3'oxy product-RNA (p-RNA) in the absence and presence of 300 μM ATP,  
483 RDV-TP, or GTP respectively. **(c)** Similar nMS analysis to **(b)** but the RTC was  
484 reconstituted using a 3'-deoxy p-RNA (RTC\*).

485

486

487

488

489

490

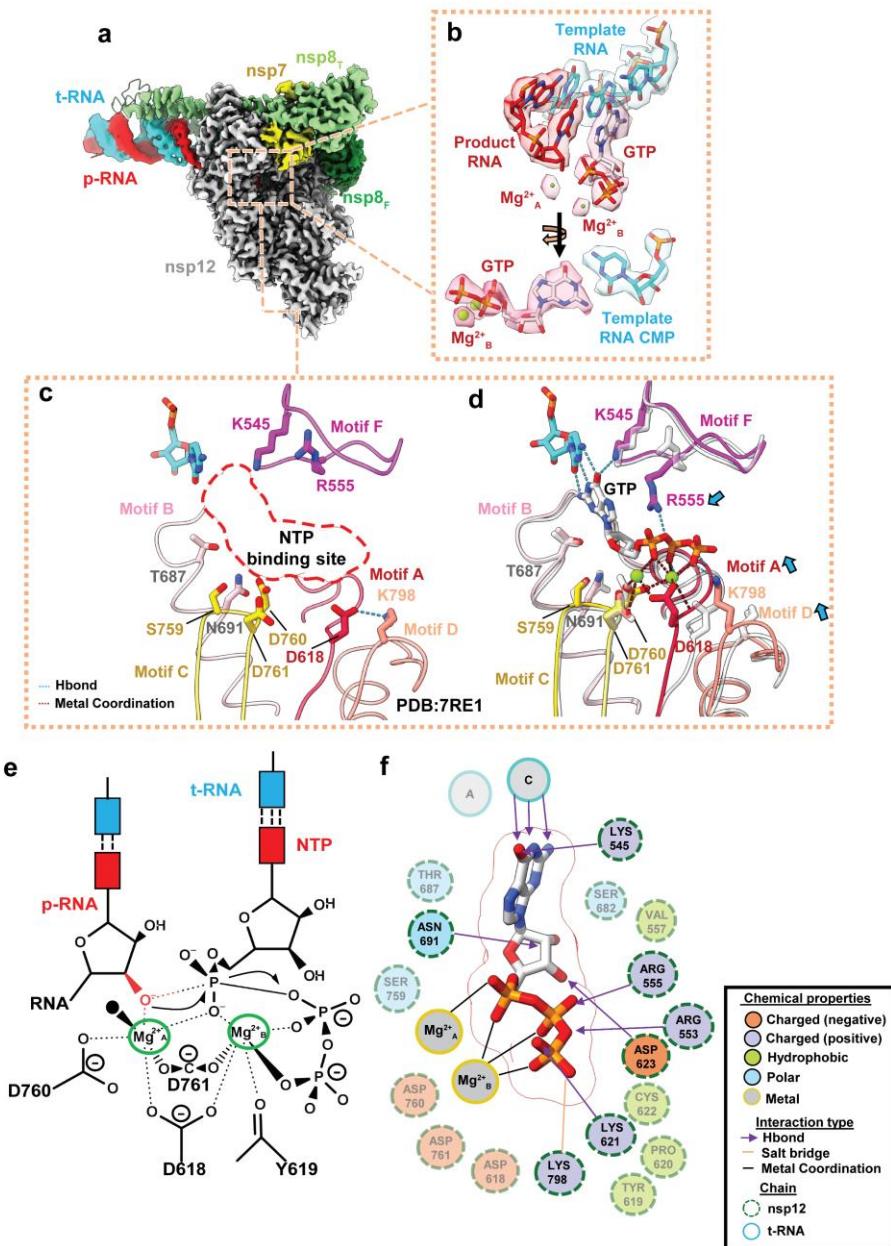
491

492

493

494 **Fig. 2 | Structural basis of nucleotide recognition**

495

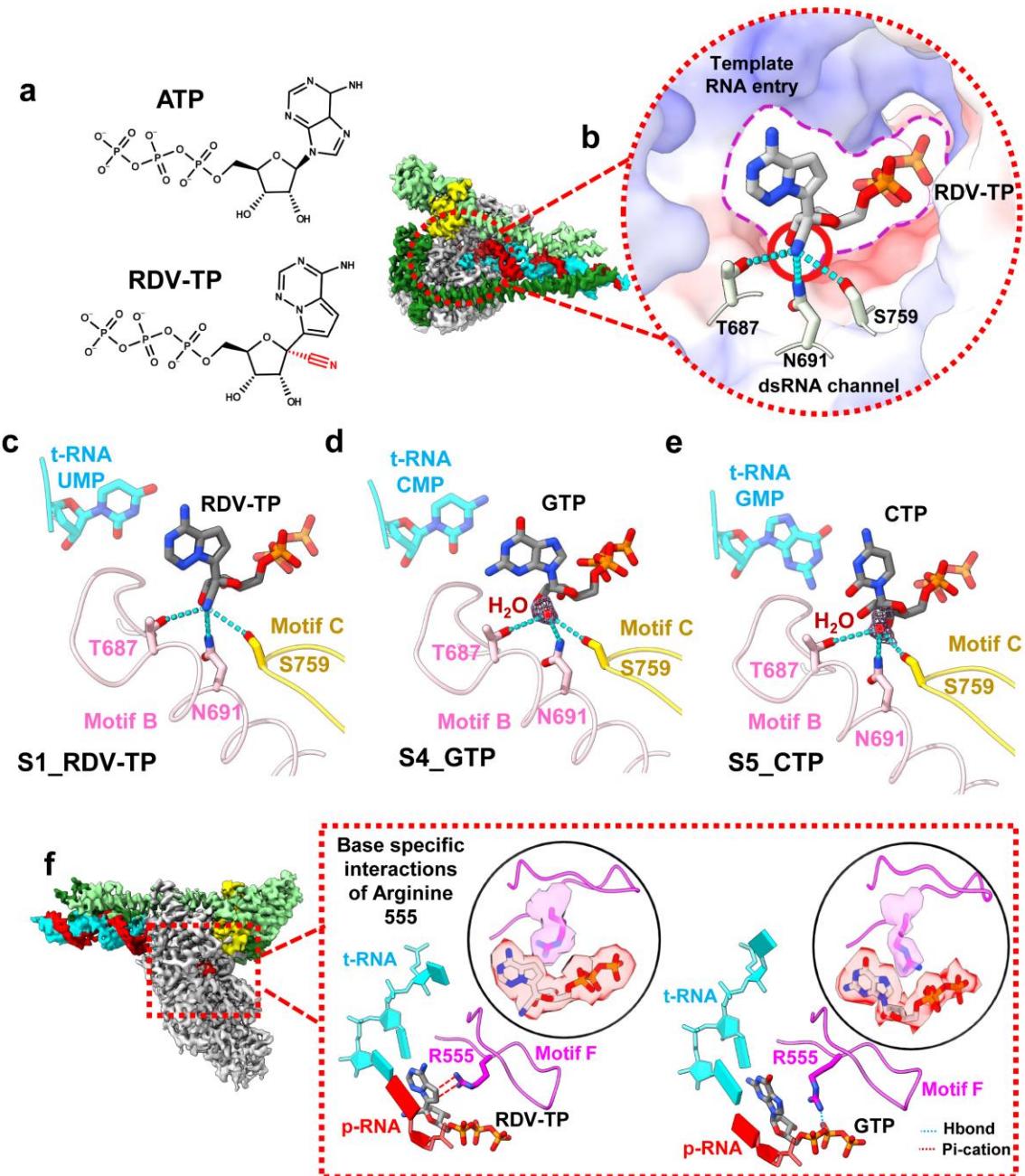


496

497 **Fig 2. (a)** Cryo-EM density of the 2.7 Å nominal resolution +GTP structure (S4\_GTP),  
498 colored according to the fitted model chains **(b)** Cryo-EM density of the bound incoming  
499 GTP, two associated metal ions, and the nearby bases of the template-RNA (t-RNA)  
500 and product-RNA (p-RNA) strands. **(c)** Close-up of the active site of the apo complex  
501 PDB 7RE1<sup>(17)</sup>, illustrating the arrangement of the conserved RdRp active site motifs A-  
502 D and F (crimson, hot pink, gold, salmon, magenta, respectively) and of key residues  
503 when the NTP binding site (dashed red line) is empty. **(d)** Comparison of the S4\_GTP  
504 and apo structure (faded grey) active sites, highlighting observed motif/residue  
505 rearrangements (blue arrows) on NTP binding. Movements in motifs A and D close the

506 active site. **(e)** Schematic depicting nucleotide addition, in the presence of an intact 3'-  
507 OH, based on disposition of magnesiums and NTP in the S4\_GTP structure **(f)** 2D  
508 schematic of the suite of interactions involved in NTP binding based on distances in the  
509 S4\_GTP structure. Residues interacting indirectly with the NTP are faded for clarity.  
510

511 **Fig. 3 | Molecular basis of Remdesivir's incorporation selectivity**



512

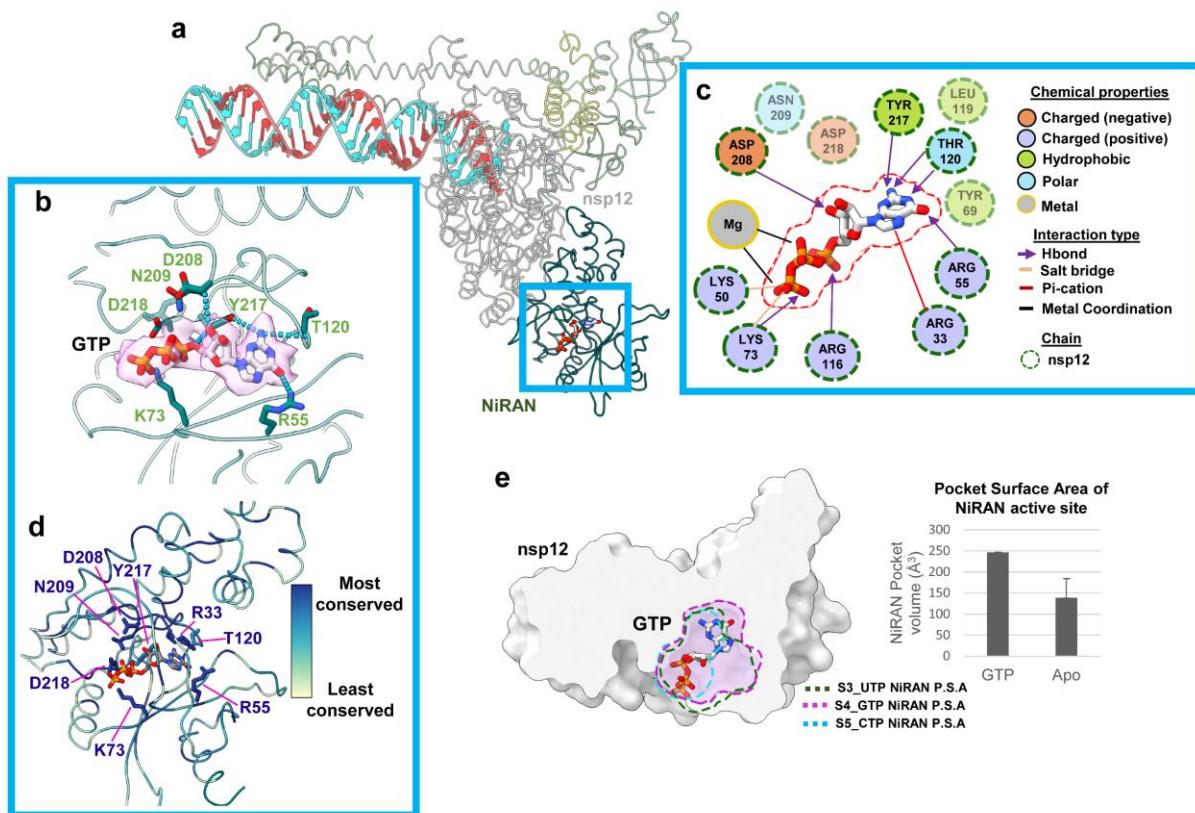
513 **Fig 3. (a)** Chemical structures of ATP and RDV-TP, highlighting the position of the  
514 RDV-TP 1' cyano group. **(b)** Cryo-EM densities of the S2\_RDV-TP structure, colored  
515 according to fitted model chains. Zoom-in on the bound RDV-TP illustrates the RDV-TP

516 1' cyano group is accommodated in a hydrophilic pocket formed by motif B and C  
517 residues. Protein surface is colored according to electrostatics. (c-e) Comparison of the  
518 active sites of the S1\_RDV-TP (c), S4\_GTP (d), S5\_CTP (e) structures reveals that the  
519 RdRp cyano pocket can also bind a water molecule (map density around the water  
520 shown in mesh) which needs to be displaced for RDV-TP binding (f) A comparison of  
521 the S2\_RDV-TP and S4\_GTP structures reveals two predominant rotamers of R555  
522 which mediate either a pi-pi stacking (red dotted lines) or a H-bond (blue dotted line)  
523 interaction with the incoming NTP.

524

525 **Fig. 4 | NiRAN specific recognition of GTP**

526



527

528 **Fig 4.** (a) View of the NiRAN domain of the RTC, bound to GTP, which lies at the  
529 amino-terminal end of nsp12. (b) GTP is selectively recognized in the NiRAN pocket by  
530 a series of hydrogen bonding and electrostatic interactions in which interacting residues  
531 are shown as sticks. (c) 2D schematic illustrating the NiRAN-GTP interactions. (d)  
532 Residues that mediate GTP recognition are shown as sticks and colored according to  
533 their conservation across the  $\alpha$  &  $\beta$  coronavirus clades. (e) Binding of GTP in the NiRAN  
534 is mediated through an induced fit that widens the pocket for insertion of the guanine  
535 base. Cross-sections (dashed lines) of the pocket surface area (P.S.A) of the S4\_GTP  
536 (GTP bound), S3\_UTP and S5\_CTP (apo-NiRAN) are shown overlaid on a clipped  
537 surface view of the nsp12 NiRAN. The NiRAN pocket volume was measured using the  
538 Schrodinger Sitemap tool.

539

## 540 METHODS

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

544

## 545 Protein expression and purification

546 *SARS-CoV-2 nsp12* was expressed as previously described <sup>34</sup> and purified as follows.  
547 Briefly, a pQE-30/pcl-ts ind+ plasmid containing a His<sub>6</sub>-SUMO SARS-CoV-2 nsp12 and  
548 untagged nsp7 & 8 (Addgene #160540) was transformed into *E. coli* BL21 cells  
549 (Agilent). Cells were grown and protein expression was induced by the addition of 0.2  
550 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), 10 ng/mL tetracycline and 50 ug/mL  
551 nalidixic acid. Cells were collected and lysed in a French press (Avestin). The lysate  
552 was cleared by centrifugation and purified on a HiTrap Heparin HP column (Cytiva). The  
553 fractions containing nsp12 were loaded onto a HisTrap HP column (Cytiva) for further  
554 purification. Eluted nsp12 was dialyzed, cleaved with His<sub>6</sub>-Ulp1 SUMO protease, and  
555 passed through a HisTrap HP column to remove the SUMO protease. Flow-through was  
556 collected, concentrated by centrifugal filtration (Amicon), and loaded on a Superdex 200  
557 Hiload 16/600 (Cytiva). Glycerol was added to the purified nsp12, aliquoted, flash-frozen  
558 with liquid N<sub>2</sub>, and stored at -80°C.

559 *SARS-CoV-2 nsp7/8* was expressed and purified as described <sup>14</sup>. Briefly, the  
560 pCDFDuet-1 plasmid containing His<sub>6</sub> SARS-CoV-2 nsp7/8 (Addgene #159092) was  
561 transformed into *E. coli* BL21 (DE3). Cells were grown and protein expression was  
562 induced by the addition of IPTG. Cells were collected and lysed in a French press  
563 (Avestin). The lysate was cleared by centrifugation and purified on a HisTrap HP  
564 column (Cytiva). Eluted nsp7/8 was dialyzed, cleaved with His<sub>6</sub>-Prescission Protease to  
565 cleave His<sub>6</sub> tag, and then passed through a HisTrap HP column to remove the protease  
566 (Cytiva). Flow-through was collected, concentrated by centrifugal filtration (Amicon), and  
567 loaded onto a Superdex 75 Hiload 16/600 (Cytiva). Glycerol was added to the purified  
568 nsp7/8, aliquoted, flash-frozen with liquid N<sub>2</sub>, and stored at -80°C.

## 569 Preparation of SARS-CoV-2 replication/transcription complex (RTC) for Cryo-EM.

570 Cryo-EM samples of SARS-CoV-2 RTC were prepared as previously described <sup>14, 17</sup>.  
571 Briefly, purified nsp12 and nsp7/8 were concentrated, mixed in a 1:3 molar ratio, and  
572 incubated for 20 min at 22°C. An annealed RNA scaffold (Horizon Discovery, Ltd.) was  
573 added to the nsp7/8/12 complex and incubated for 10 min at 30°C. Sample was buffer  
574 exchanged into cryo-EM buffer [20 mM HEPES pH 8.0, 100 mM K-Acetate, 5 mM  
575 MgCl<sub>2</sub>, 2 mM DTT] and further incubated for 20 min at 30°C. The sample was purified  
576 over a Superose 6 Increase 10/300 GL column (Cytiva) in cryo-EM buffer. The peak  
577 corresponding to nsp7/8/12/RNA complex was pooled and concentrated by centrifugal  
578 filtration (Amicon).

579

580 **Cryo-EM grid preparation.** Prior to grid freezing, beta-octyl glucoside (β-OG) was  
581 added to the sample (0.07 % w/v final). The final buffer condition for the cryo-EM

582 sample was 20 mM HEPES pH 8.0, 100 mM K-Acetate, 5 mM MgCl<sub>2</sub>, 2 mM DTT,  
583 0.07% (w/v) β-OG with 300 μM final of the respective NTP added immediately before  
584 freezing. C-flat holey carbon grids (CF-1.2/1.3-4Au, EMS) were glow-discharged for 20  
585 seconds prior to the application of 3.5 μL of sample. Using a Vitrobot Mark IV (Thermo  
586 Fisher Scientific), grids were blotted and plunge-frozen into liquid ethane with 95%  
587 chamber humidity at 4°C.

588

589 **Cryo-EM data acquisition and processing.** Structural biology software was accessed  
590 through the SBGrid consortium <sup>35</sup>. The following pertains to each of the respective  
591 datasets:

592 **S1\_RDV-TP:** Grids were imaged using a 300 kV Titan Krios (Thermo Fisher Scientific)  
593 equipped with a GIF BioQuantum and K3 camera (Gatan). Images were recorded with  
594 Leginon <sup>36</sup> with a pixel size of 1.065 Å/px (micrograph dimension of 5760 × 4092 px)  
595 over a defocus range of −0.8 μm to −2.5 μm with a 20 eV energy filter slit. Movies were  
596 recorded in “counting mode” (native K3 camera binning 2) with ~25 e-/px/s in dose-  
597 fractionation mode with subframes of 50 ms over a 2.5 s exposure (50 frames) to give a  
598 total dose of ~54 e-/Å<sup>2</sup>. Dose-fractionated movies were gain-normalized, drift-corrected,  
599 summed, and dose-weighted using MotionCor2 <sup>37</sup>. The contrast transfer function (CTF)  
600 was estimated for each summed image using the Patch CTF module in cryoSPARC  
601 v3.1.0 <sup>38</sup>. Particles were picked and extracted from the dose-weighted images with box  
602 size of 256 px using cryoSPARC Blob Picker and Particle Extraction. The entire dataset  
603 consisted of 26,130 motion-corrected images with 8,017,151 particles. Particles were  
604 sorted using three rounds of cryoSPARC 2D classification (N=50, where N equals the  
605 number of classes), resulting in 2,473,065 curated particles. Initial models (denoted as  
606 monomer & dimer) were generated using cryoSPARC *Ab initio* Reconstruction (N=3) on  
607 a subset of the particles. Particles were further curated using these initial models as 3D  
608 templates for iterative cryoSPARC Heterogeneous Refinement (N=4), resulting in  
609 314,848 particles in the monomer class (cyan map, Extended Data Fig. 1) and 264,453  
610 particles in the dimer class (pink map, Extended Data Fig. 1). Curated particles in the  
611 monomer and dimer classes were re-extracted to a box size of 384 and input to  
612 cryoSPARC Homogenous and Non-uniform refinements <sup>39</sup>. Particles within each class  
613 were further processed through two rounds of RELION 3.1 Bayesian Polishing <sup>40</sup>.  
614 Polished particles were refined using cryoSPARC Local and Global CTF Refinement in  
615 combination with cryoSPARC Non-uniform Refinement, resulting in structures with the  
616 following particle counts and nominal resolutions: monomer RTC (251,160 particles;  
617 3.54 Å) & dimer RTC (249,468 particles; 3.87 Å). To facilitate model building of the  
618 RTC, particles from the dimer RTC class underwent masked particle subtraction in  
619 which both protomers were masked, subtracted from the full map, and combined with  
620 the original monomer RTC to yield a structure of the RTC at a nominal resolution of 3.38  
621 Å from 613,848 particles (Extended Data Fig. 1e). Local resolution calculations were  
622 generated using blocres and blocfilt from the Bsoft package <sup>20</sup> (Extended Data Fig. 1c).  
623 The angular distribution of particle orientations (Extended Data Fig. 1b) and directional  
624 resolution through the 3DFSC package <sup>41</sup> (Extended Data Fig. 1d) were calculated for  
625 the final class.

626 **S2\_ATP:** Grids were imaged using a CS corrected 300 kV Titan Krios (Thermo Fisher  
627 Scientific) equipped with a GIF BioQuantum and K3 camera (Gatan). Images were  
628 recorded with Leginon <sup>36</sup> with a pixel size of 1.076 Å/px (micrograph dimension of 5760  
629 × 4092 px) over a defocus range of −0.8 μm to −2.5 μm with a 20 eV energy filter slit.  
630 Movies were recorded in “counting mode” (native K3 camera binning 2) with ~25 e-/px/s  
631 in dose-fractionation mode with subframes of 50 ms over a 2.5 s exposure (50 frames)  
632 to give a total dose of ~52 e-/Å<sup>2</sup>. Dose-fractionated movies were gain-normalized, drift-  
633 corrected, summed, and dose-weighted using MotionCor2 <sup>37</sup>. The CTF was estimated  
634 for each summed image using the Patch CTF module in cryoSPARC v3.1.0 <sup>38</sup>. Particles  
635 were picked and extracted from the dose-weighted images with box size of 256 px using  
636 cryoSPARC Blob Picker and Particle Extraction. The entire dataset consisted of 11,657  
637 motion-corrected images with 5,364,858 particles. Particles were sorted using three  
638 rounds of cryoSPARC 2D classification (N=50), resulting in 2,255,856 curated particles.  
639 Initial models (denoted as monomer & dimer) were generated using cryoSPARC *Ab*  
640 *initio* Reconstruction (N=4) on a subset of the particles. Particles were further curated  
641 using these initial models as 3D templates for iterative cryoSPARC Heterogeneous  
642 Refinement (N=4), resulting in 102,529 particles in the monomer class (cyan map,  
643 Extended Data Fig. 2) and 318,672 particles in the dimer class (pink map, Extended  
644 Data Fig. 2). Curated particles in the monomer and dimer classes were re-extracted to a  
645 box size of 384 and input to cryoSPARC Homogenous and Non-uniform refinements <sup>39</sup>.  
646 Particles within each class were further processed through two rounds of RELION 3.1  
647 Bayesian Polishing <sup>40</sup>. Polished particles were refined using cryoSPARC Local and  
648 Global CTF Refinement in combination with cryoSPARC Non-uniform Refinement,  
649 resulting in structures with the following particle counts and nominal resolutions:  
650 monomer RTC (96,868 particles; 3.71 Å) & dimer RTC (299,965 particles; 3.50 Å). To  
651 facilitate model building of the RTC, particles from the dimer RTC class underwent  
652 masked particle subtraction in which both protomers were masked, subtracted from the  
653 full map, and combined with the original monomer RTC to yield a structure of the RTC  
654 at a nominal resolution of 3.09 Å from 330,442 particles (Extended Data Fig. 2e). This  
655 combined RTC class was further refined with masked cryoSPARC Local Refinement  
656 using masks around the RdRp and NiRAN active sites. Locally refined maps were  
657 combined into a RTC composite map using PHENIX ‘Combine Focused Maps’ to aid  
658 model building <sup>42</sup>. Local resolution calculations were generated using blocres and blocfilt  
659 from the Bsoft package <sup>20</sup> (Extended Data Fig. 2c). The angular distribution of particle  
660 orientations (Extended Data Fig. 2b) and directional resolution through the 3DFSC  
661 package <sup>41</sup> (Extended Data Fig. 2d) were calculated for the final class.

662 **S3\_UTP:** Grids were imaged using a 300 kV Titan Krios (Thermo Fisher Scientific)  
663 equipped with a GIF BioQuantum and K3 camera (Gatan). Images were recorded with  
664 Leginon <sup>36</sup> with a pixel size of 1.065 Å/px (micrograph dimension of 5760 × 4092 px)  
665 over a defocus range of −0.8 μm to −2.5 μm with a 20 eV energy filter slit. Movies were  
666 recorded in “counting mode” (native K3 camera binning 2) with ~25 e-/px/s in dose-  
667 fractionation mode with subframes of 50 ms over a 2.5 s exposure (50 frames) to give a  
668 total dose of ~53 e-/Å<sup>2</sup>. Dose-fractionated movies were gain-normalized, drift-corrected,  
669 summed, and dose-weighted using MotionCor2 <sup>37</sup>. The CTF was estimated for each  
670 summed image using the Patch CTF module in cryoSPARC v3.1.0 <sup>38</sup>. Particles were  
671 picked and extracted from the dose-weighted images with box size of 256 px using

672 cryoSPARC Blob Picker and Particle Extraction. The entire dataset consisted of 30,850  
673 motion-corrected images with 14,149,078 particles. Particles were sorted using three  
674 rounds of cryoSPARC 2D classification (N=50), resulting in 3,297,109 curated particles.  
675 Initial models (denoted as monomer & dimer) were generated using cryoSPARC *Ab*  
676 *initio* Reconstruction (N=4) on a subset of the particles. Particles were further curated  
677 using these initial models as 3D templates for iterative cryoSPARC Heterogeneous  
678 Refinement (N=6), resulting in 648,814 particles in the monomer class (cyan map,  
679 Extended Data Fig. 3) and 773,911 particles in the dimer class (pink map, Extended  
680 Data Fig. 3). Curated particles in the monomer and dimer classes were re-extracted to a  
681 box size of 384 and input to cryoSPARC Homogenous and Non-uniform refinements<sup>39</sup>.  
682 Particles within each class were further processed through two rounds of RELION 3.1  
683 Bayesian Polishing<sup>40</sup>. Polished particles were refined using cryoSPARC Local and  
684 Global CTF Refinement in combination with cryoSPARC Non-uniform Refinement,  
685 resulting in structures with the following particle counts and nominal resolutions:  
686 monomer RTC (614,648 particles; 3.34 Å) & dimer RTC (730,009 particles; 3.42 Å). To  
687 facilitate model building of the RTC, particles from the dimer RTC class underwent  
688 masked particle subtraction in which both protomers were masked, subtracted from the  
689 full map, and combined with the original monomer RTC to yield a structure of the RTC  
690 at a nominal resolution of 3.13 Å from 719,889 particles (Extended Data Fig. 3e). This  
691 combined RTC class was further refined with masked cryoSPARC Local Refinement  
692 using masks around the RdRp and NiRAN active sites. Locally refined maps were  
693 combined into a RTC composite map using PHENIX ‘Combine Focused Maps’ to aid  
694 model building<sup>42</sup>. Local resolution calculations were generated using blocres and blocfilt  
695 from the Bsoft package<sup>20</sup> (Extended Data Fig. 3c). The angular distribution of particle  
696 orientations (Extended Data Fig. 3b) and directional resolution through the 3DFSC  
697 package<sup>41</sup> (Extended Data Fig. 3d) were calculated for the final class.

698 **S4\_GTP:** Grids were imaged using a CS corrected 300 kV Titan Krios (Thermo Fisher  
699 Scientific) equipped with a GIF BioQuantum and K3 camera (Gatan). Images were  
700 recorded with SerialEM<sup>43</sup> with a pixel size of 1.08 Å/px (micrograph dimension of 5760  
701 × 4092 px) over a defocus range of −0.8 μm to −3.0 μm with a 20 eV energy filter slit.  
702 Movies were recorded in “counting mode” (native K3 camera binning 2) with ~25 e-/px/s  
703 in dose-fractionation mode with subframes of 50 ms over a 2.5 s exposure (50 frames)  
704 to give a total dose of ~51 e-/Å<sup>2</sup>. Dose-fractionated movies were gain-normalized, drift-  
705 corrected, summed, and dose-weighted using MotionCor2<sup>37</sup>. The CTF was estimated  
706 for each summed image using the Patch CTF module in cryoSPARC v3.1.0<sup>38</sup>. Particles  
707 were picked and extracted from the dose-weighted images with box size of 256 px using  
708 cryoSPARC Blob Picker and Particle Extraction. The entire dataset consisted of 4,527  
709 motion-corrected images with 2,419,929 particles. Particles were sorted using three  
710 rounds of cryoSPARC 2D classification (N=50), resulting in 941,507 curated particles.  
711 An initial model of the monomer RTC was generated using cryoSPARC *Ab initio*  
712 Reconstruction (N=3) on a subset of the particles. Particles were further curated using  
713 this initial model as a 3D template for iterative cryoSPARC Heterogeneous Refinement  
714 (N=6), resulting in 484,682 particles in the resultant monomer class (cyan map,  
715 Extended Data Fig. 4). Curated particles were re-extracted to a box size of 384 and  
716 input to cryoSPARC Homogenous and Non-uniform refinements<sup>39</sup>. Particles were  
717 further processed through two rounds of RELION 3.1 Bayesian Polishing<sup>40</sup>. Polished

718 particles were refined using cryoSPARC Local and Global CTF Refinement in  
719 combination with cryoSPARC Non-uniform Refinement, resulting in a structure with the  
720 following particle count and nominal resolution: monomer RTC (456,629 particles;  
721 2.68 Å) (Extended Data Fig. 4e). Local resolution calculations were generated using  
722 blocres and blocfilt from the Bsoft package <sup>20</sup> (Extended Data Fig. 4c). The angular  
723 distribution of particle orientations (Extended Data Fig. 4b) and directional resolution  
724 through the 3DFSC package <sup>41</sup> (Extended Data Fig. 4d) were calculated for the final  
725 class.

726

727 **S5\_CTP:** Grids were imaged using a CS corrected 300 kV Titan Krios (Thermo Fisher  
728 Scientific) equipped with a GIF BioQuantum and K3 camera (Gatan). Images were  
729 recorded with SerialEM <sup>43</sup> with a pixel size of 0.515 Å/px (micrograph dimension of 5760  
730 × 4092 px) over a defocus range of −0.8 μm to −3.0 μm with a 20 eV energy filter slit.  
731 Movies were recorded in “counting mode” (native K3 camera binning 2) with ~25 e-/px/s  
732 in dose-fractionation mode with subframes of 50 ms over a 2.5 s exposure (50 frames)  
733 to give a total dose of ~57 e-/Å<sup>2</sup>. Dose-fractionated movies were gain-normalized, drift-  
734 corrected, summed, and dose-weighted using MotionCor2 <sup>37</sup>. The CTF was estimated  
735 for each summed image using the Patch CTF module in cryoSPARC v3.1.0 <sup>38</sup>. Particles  
736 were picked and extracted from the dose-weighted images with box size of 512 px using  
737 cryoSPARC Blob Picker and Particle Extraction. The entire dataset consisted of 13,905  
738 motion-corrected images with 1,986,527 particles. Particles were sorted using three  
739 rounds of cryoSPARC 2D classification (N=50), resulting in 460,232 curated particles.  
740 An initial model (denoted as monomer) was generated using cryoSPARC *Ab initio*  
741 Reconstruction (N=3) on a subset of the particles. Particles were further curated using  
742 these initial models as 3D templates for iterative cryoSPARC Heterogeneous  
743 Refinement (N=6), resulting in 143,110 particles in the monomer class (cyan map,  
744 Extended Data Fig. 5) and 93,597 particles in an extracted dimer class (pink map,  
745 Extended Data Fig. 5). Curated particles in the monomer and dimer classes were re-  
746 extracted to a box size of 768 and input to cryoSPARC Homogenous and Non-uniform  
747 refinements <sup>39</sup>. Particles within each class were further processed through two rounds of  
748 RELION 3.1 Bayesian Polishing <sup>40</sup>. Polished particles were refined using cryoSPARC  
749 Local and Global CTF Refinement in combination with cryoSPARC Non-uniform  
750 Refinement, resulting in structures with the following particle counts and nominal  
751 resolutions: monomer RTC (128,484 particles; 3.09 Å) & dimer RTC (83,555 particles;  
752 3.69 Å). To facilitate model building of the RTC, particles from the dimer RTC class  
753 underwent masked particle subtraction in which both protomers were masked,  
754 subtracted from the full map, and combined with the original monomer RTC to yield a  
755 structure of the RTC at a nominal resolution of 2.67 Å from 171,107 particles (Extended  
756 Data Fig. 5e). This combined RTC class was further refined with masked cryoSPARC  
757 Local Refinement using masks around the RdRp and NiRAN active sites. Locally  
758 refined maps were combined into a RTC composite map using PHENIX ‘Combine  
759 Focused Maps’ to aid model building <sup>42</sup>. Local resolution calculations were generated  
760 using blocres and blocfilt from the Bsoft package <sup>20</sup> (Extended Data Fig. 5c). The  
761 angular distribution of particle orientations (Extended Data Fig. 5b) and directional  
762 resolution through the 3DFSC package <sup>41</sup> (Extended Data Fig. 5d) are shown for the  
763 final class.

764

765 **Model building and refinement.** An initial model of the RTC was derived from PDB  
766 7RE1<sup>17</sup>. The models were manually fit into the cryo-EM density maps using Chimera<sup>44</sup>  
767 and rigid-body and real-space refined using Phenix real-space-refine<sup>42</sup>. For real-space  
768 refinement, rigid body refinement was followed by all-atom and B-factor refinement with  
769 Ramachandran and secondary structure restraints. Models were inspected and  
770 modified in Coot 0.9.5<sup>45</sup> and the refinement process was repeated iteratively.

771

772 **Native mass spectrometry (nMS) analysis.** The RTC samples were initially  
773 reconstituted as described above in the following buffer: 20 mM HEPES pH 8.0, 80 mM  
774 K-Acetate, 5 mM MgCl<sub>2</sub>, 2 mM DTT. For the NTP pre-incorporation experiments, the  
775 replication-transcription complexes with RNA scaffolds containing either 3'-oxy p-RNA  
776 (RTC) or 3'-deoxy p-RNA (RTC\*) were incubated with 300 μM NTP (ATP, RDV-TP or  
777 GTP) on ice for 2 min prior to buffer exchange. For the NDP incubation experiments, the  
778 RTC samples reconstituted with the respective RNA scaffolds were incubated with 2 mM  
779 nucleotide (ADP, RDV-DP, GDP or ATP) on ice for 2 min before buffer exchange. After  
780 nucleotide incubation, all samples were immediately buffer exchanged into an nMS-  
781 compatible solution (150 mM ammonium acetate, pH 7.5, 0.01% Tween-20) with a 40  
782 kDa MWCO (ThermoFisher Scientific). For nMS analysis, a 2 – 3 μL aliquot of the buffer-  
783 exchanged sample was loaded into a gold-coated quartz capillary tip prepared in-house  
784 and then electrosprayed into an Exactive Plus with extended mass range (EMR)  
785 instrument (Thermo Fisher Scientific) with a static direct infusion nanospray source<sup>46</sup>.  
786 The MS parameters used include: spray voltage, 1.22 kV; capillary temperature, 125 -  
787 150 °C; in-source dissociation, 0 - 10 V; S-lens RF level, 200; resolving power, 8,750 or  
788 17,500 at *m/z* of 200; AGC target, 1 x 10<sup>6</sup>; maximum injection time, 200 ms; number of  
789 microscans, 5; injection flatapole, 8 V; interflatapole, 7 V; bent flatapole, 4 V; high energy  
790 collision dissociation (HCD), 200 V; ultrahigh vacuum pressure, 5.8 – 6.1 × 10<sup>-10</sup> mbar;  
791 total number of scans, at least 100. Mass calibration in positive EMR mode was performed  
792 using cesium iodide. For data processing, the acquired MS spectra were visualized using  
793 Thermo Xcalibur Qual Browser (v. 4.2.47). Deconvolution was performed either manually  
794 or using the software UniDec v. 4.2.0<sup>47,48</sup>. The following parameters were used for data  
795 processing with UniDec: background subtraction (if applied), subtract curve 10; smooth  
796 charge state distribution, enabled; peak shape function, Gaussian. Mass accuracies were  
797 calculated as the % difference between the measured and expected masses relative to  
798 the expected mass. The observed mass accuracies (calculated as the % difference  
799 between the measured and expected masses relative to the expected mass) ranged from  
800 0.005 – 0.06%.

801

802 The expected masses for the component proteins are nsp7: 9,137 Da; nsp8 (N-  
803 terminal Met lost): 21,881 Da, and nsp12 (has two Zn<sup>2+</sup> ions coordinated with 6  
804 deprotonated cysteine residues): 106,785 Da<sup>14</sup>. The RNA scaffolds were also analyzed  
separately, and their sequences were verified by mass measurements using nMS.

805

806 **In-vitro primer elongation assays.** Assays were performed using reconstituted  
807 template-primer RNA scaffolds (Table S1) (Horizon Discover Ltd./Dharmacon) annealed  
808 in 10 mM HEPES pH 8.0, 50 mM KCl, 2 mM MgCl<sub>2</sub>. Reactions (20 μL) containing 100

809 nM RNA scaffold, 0.75  $\mu$ M nsp12, 2  $\mu$ M nsp7/8 and NTPs 300  $\mu$ M (if present as natural  
810 nucleotides, NDPs or  $\alpha$ - $\beta$  analogues) and 1  $\mu$ L  $\alpha$ - $^{32}$ P-GTP (Perkin-Elmer) were  
811 incubated at 30°C for 10 minutes prior to addition of a 2x stop solution (Invitrogen-Gel  
812 Loading buffer II). Assay buffer was 100 mM K-acetate, 20 mM HEPES pH 8.0, 5 mM  
813 MgCl<sub>2</sub> & 2 mM BME in which the MgCl<sub>2</sub> was substituted with CaCl<sub>2</sub> when monitoring the  
814 effects of alternative metal ions. Products of the elongation reactions were separated on  
815 10% acrylamide-8M urea denaturing gels and analyzed by phosphorimaging.

816

817 **Quantification and statistical analysis.** The local resolution of the cryo-EM maps  
818 (Extended Data Figs. 1-5) was estimated using blocres <sup>20</sup> with the following parameters:  
819 box size 15 & sampling with respective map pixel size. Directional 3DFSCs (Extended  
820 Data Figs. 1-5) were calculated using 3DFSC <sup>41</sup>. The quantification and statistical  
821 analyses for model refinement and validation were generated using MolProbity <sup>49</sup> and  
822 PHENIX <sup>50</sup>.

823

824 **Data and code availability.** All unique/stable reagents generated in this study are  
825 available without restriction from the corresponding authors, Seth A. Darst  
826 ([darst@rockefeller.edu](mailto:darst@rockefeller.edu)) and E.A. Campbell ([campbee@rockefeller.edu](mailto:campbee@rockefeller.edu)) . The cryo-EM  
827 density maps and atomic coordinates have been deposited in the EMDataBank and  
828 Protein Data Bank as follows: S1\_RDV-TP (EMD-26639, PDB 7UO4), S2\_ATP (EMD-  
829 26641, PDB 7UO7), S3\_UTP (EMD-26642, PDB 7UO9), S4\_GTP (EMD-26645, 7UOB),  
830 S5\_CTP (EMD-26646, PDB 7UOE).

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852 **Acknowledgments.** We thank Andreas Mueller and Ruth Saecker for helpful  
853 discussions. Some of the work reported here was conducted at the Simons Electron  
854 Microscopy Center (SEMC) and the National Resource for Automated Molecular  
855 Microscopy (NRAMM) and National Center for CryoEM Access and Training (NCCAT)  
856 located at the NYSBC, supported by grants from the NIH National Institute of General  
857 Medical Sciences (P41 GM103310), NYSTAR, the Simons Foundation (SF349247), the  
858 NIH Common Fund Transformative High Resolution Cryo-Electron Microscopy program  
859 (U24 GM129539) and NY State Assembly Majority. This work was supported by  
860 NIH P41 GM109824 and P41 GM103314 to B.T.C, and NIH R01 AI161278 and  
861 Gilead Sciences (to E.A.C. and S.A.D).

862

863 **Author contributions.** B.F.M., J.C., J.K.P., T.K.A., J.Y.F., E.A.C. and S.A.D. conceived  
864 and designed this study. B.F.M. and J.C. performed cloning, protein purification and  
865 biochemistry. P.D.B.O. conducted mass spectrometry experiments. B.F.M. prepared  
866 cryo-EM specimens. Cryo-EM data were collected by B.F.M., E.C., J.M., E.T.E., M.E.,  
867 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  
868 and analyzed atomic models. E.A.C., R.L., B.T.C. and S.A.D supervised and acquired  
869 financial support. B.F.M. wrote the first draft of the manuscript; all authors contributed to  
870 the final version.

871

872 **Competing interests.** E.A.C. and S.A.D. received funding from Gilead Sciences, Inc. in  
873 support of this study. J.K.P., T.K.A., J.Y.F., and J.P.B. are Gilead employees.

874

875

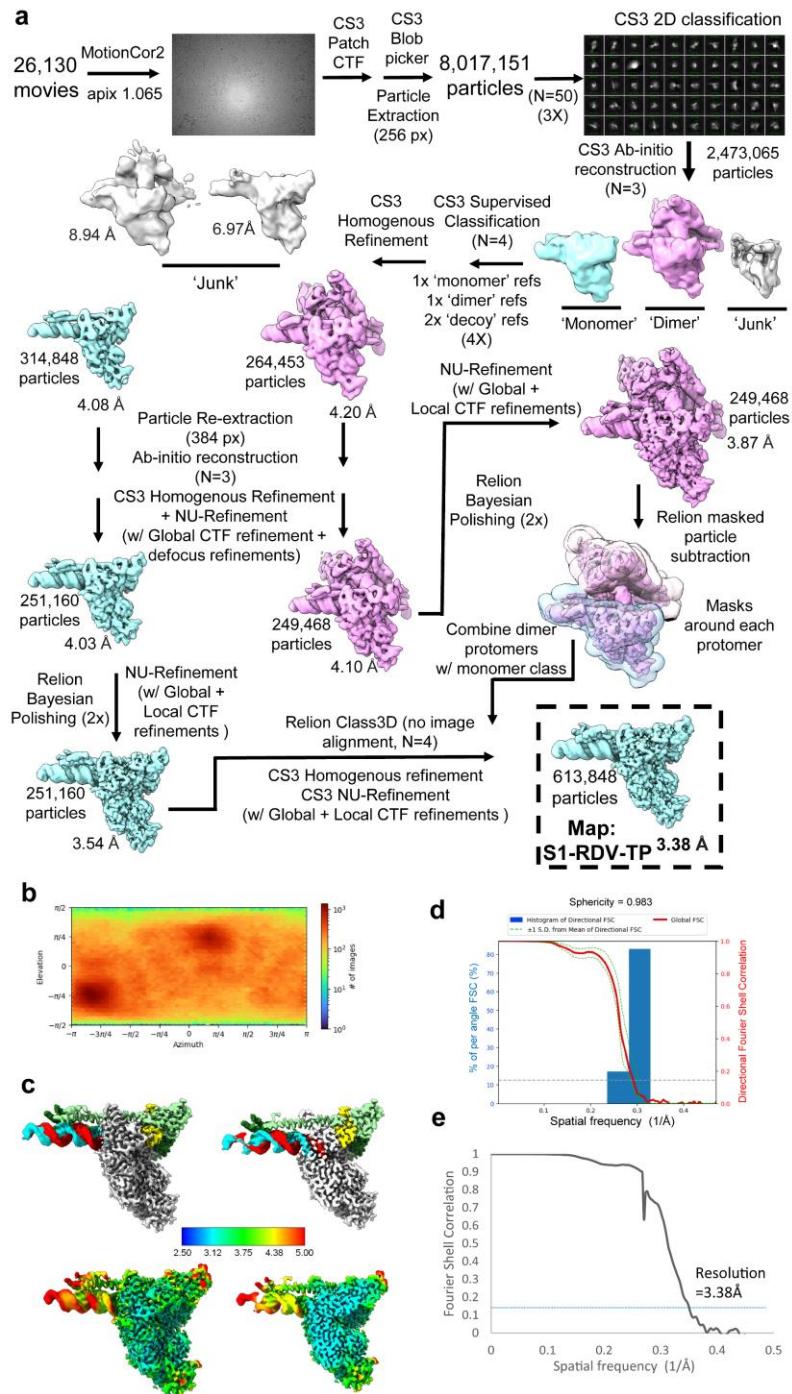
876 **Table 1. Cryo-EM data collection, refinement and validation statistics**

877  
878

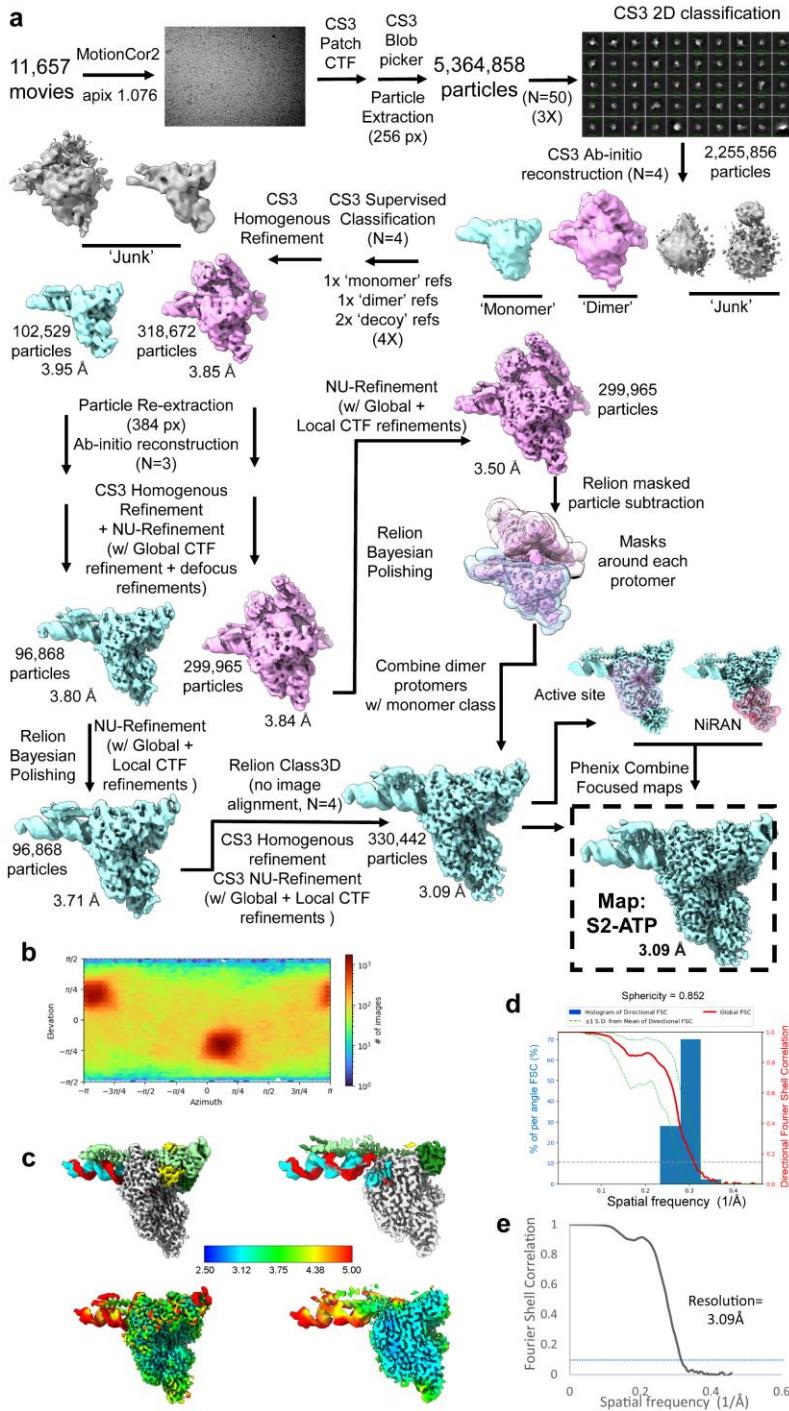
	<b>S1_RDV-TP</b>	<b>S2_ATP</b>	<b>S3_UTP</b>	<b>S4_GTP</b>	<b>S5_CTP</b>
	EMDB-26639	EMD-26641	EMD-26642	EMDB-26645	EMDB-26646
	PDB 7UO4	PDB 7UO7	PDB 7UO9	PDB 7UOB	PDB 7UOE
<b>Data collection and processing</b>					
<b>Magnification</b>	81,000	81,000	81,000	64,000	130,000
<b>Voltage (kV)</b>	300	300	300	300	300
<b>Electron exposure (e- /Å<sup>2</sup>)</b>	54.58	51.89	53.34	51.44	56.767
<b>Defocus range (μm)</b>	-0.8 μm to -2.5 μm	-0.8 μm to -2.5 μm	-0.8 μm to -2.5 μm	-0.8 μm to -3.0 μm	-0.8 μm to -3.0 μm
<b>Pixel size (Å)</b>	1.065	1.076	1.069	1.08	0.515
<b>Symmetry imposed</b>	C1	C1	C1	C1	C1
<b>Initial particle images (no.)</b>	8,017,151	5,363,858	14,149,078	2,419,929	1,986,527
<b>Final particle images (no.)</b>	613,848	330,442	719,899	456,629	171,107
<b>Map resolution (Å)</b> <b>FSC threshold 0.143</b>	3.38	3.09	3.13	2.68	2.67
<b>Map resolution range (Å)</b>	2.91 - 6.68	2.82 - 7.53	2.74 - 6.81	2.29 - 7.82	2.34 - 8.24
<b>Refinement</b>					
<b>Initial models used (PDB code)</b>	7RE1	7RE1	7RE1	7RE1	7RE1
<b>Model resolution (Å)</b> <b>FSC threshold 0.5</b>	3.5	3.4	3.2	2.8	2.9
<b>Map sharpening B factor (Å<sup>2</sup>)</b>	144.9	100.7	141.4	106.7	90.6
<b>Model composition</b>					
<b>Non-hydrogen atoms</b>	12,470	12,262	12,560	12,718	12,705
<b>Protein residues</b>	1,374	1,371	1,375	1,375	1,376
<b>Nucleic acid residues (RNA)</b>	70	66	67	69	70

<b>Ligands</b>	4	4	4	7	5
<b>B factors (Å<sup>2</sup>)</b>					
<b>Protein</b>	23.34	74.21	37.93	16.96	57.92
<b>Nucleic acids</b>	116.61	153.2	117.39	82.17	143.63
<b>Ligand</b>	16.71	58.48	25.68	13.75	48.02
<b>R.m.s. deviations</b>					
<b>Bond lengths (Å)</b>	0.004	0.005	0.005	0.005	0.004
<b>Bond angles (°)</b>	0.503	0.643	0.508	0.605	516
<b>Validation</b>					
<b>MolProbity score</b>	1.93	1.8	1.65	1.45	1.44
<b>Clashscore</b>	7.28	8.32	4.94	3.27	4.95
<b>Poor rotamers (%)</b>	1.27	0.76	0.25	1.52	1.18
<b>Ramachandran plot</b>					
<b>Favored (%)</b>	92.83	95.01	94.15	96.78	97.37
<b>Allowed (%)</b>	7.17	4.99	5.85	3	2.41
<b>Disallowed (%)</b>	0	0	0	0.22	0.22

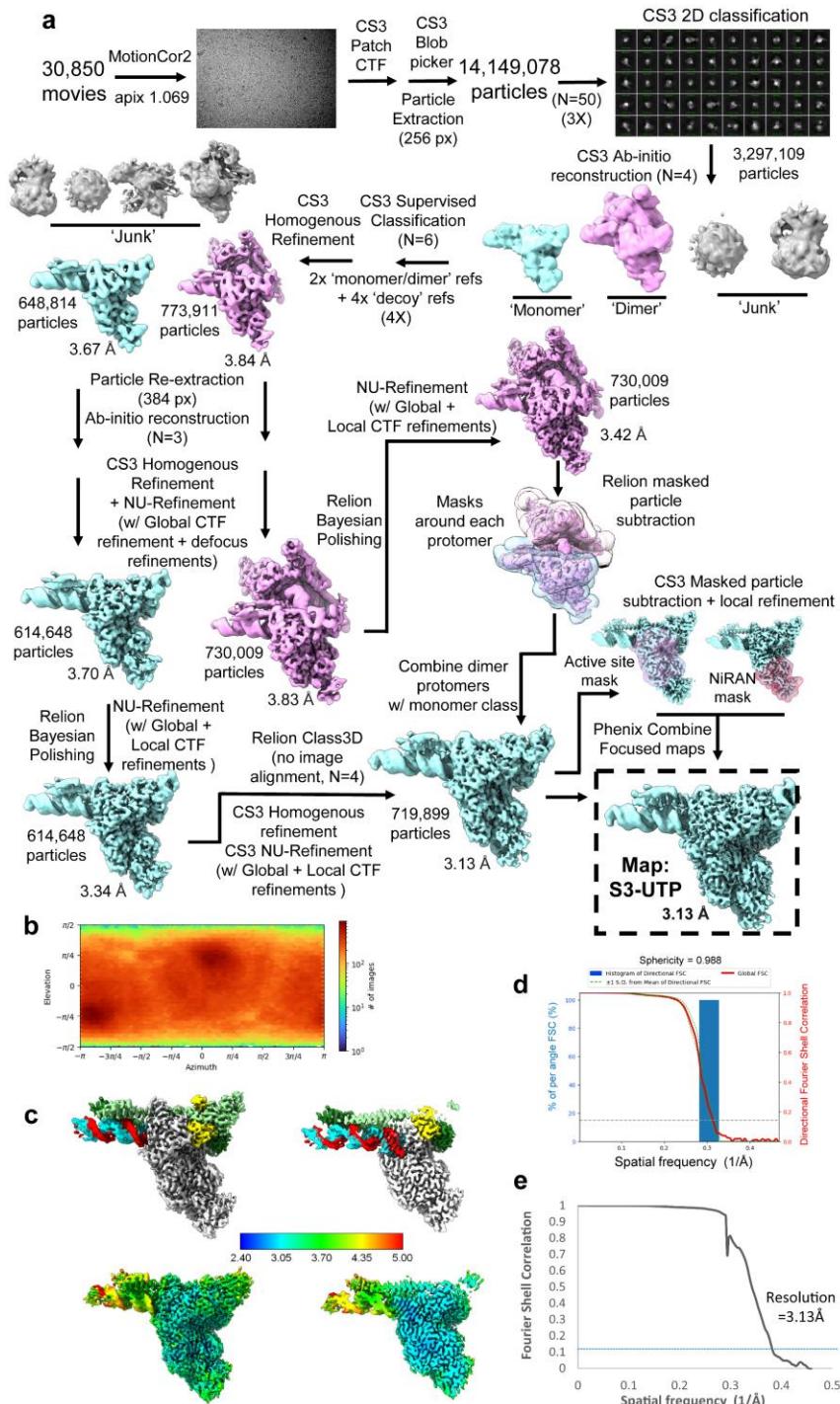
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895



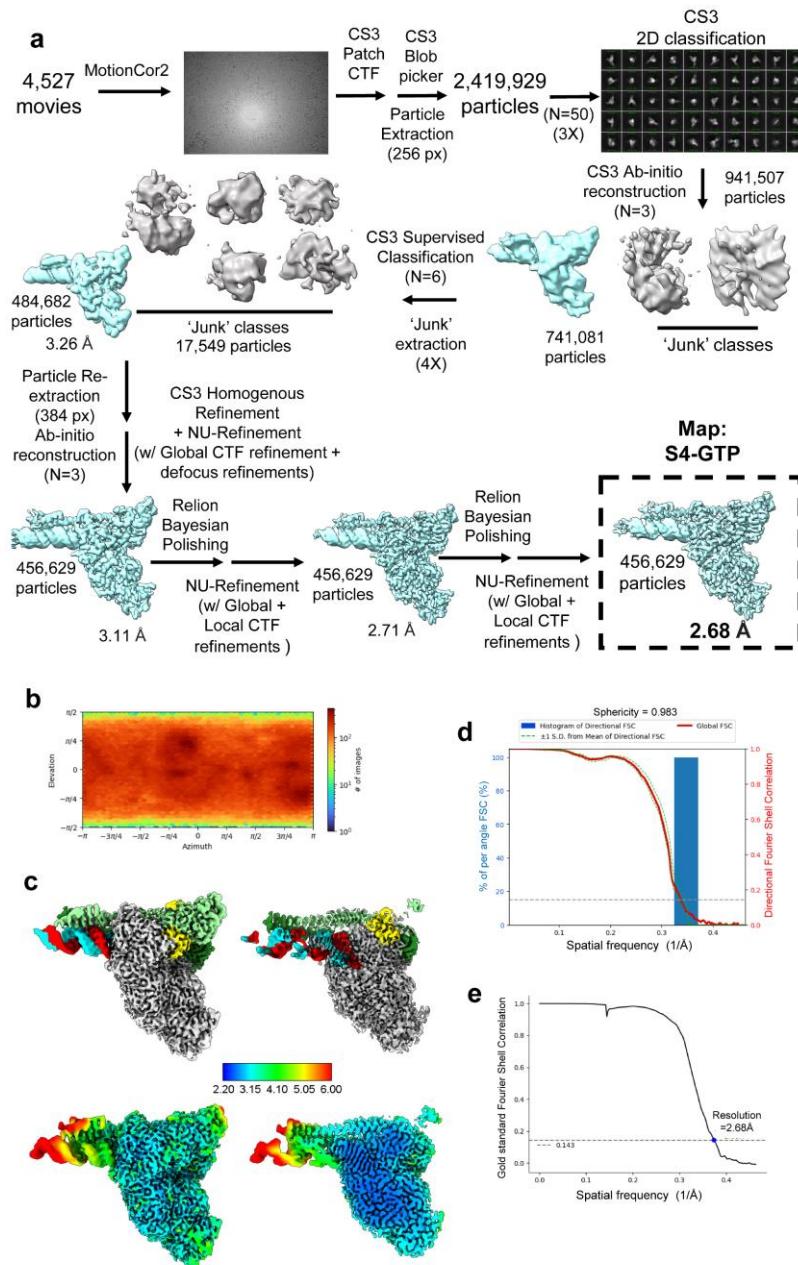
896  
897 **Extended Data Fig 1. (a)** Cryo-EM processing pipeline for the S1\_RDV-TP dataset. **(b)**  
898 Angular distribution plot for the S1\_RDV-TP dataset, calculated in cryoSPARC. Scale  
899 depicts number of particles assigned to a specific angular bin. **(c)** Nominal 3.38 Å  
900 resolution cryo-EM reconstruction filtered by local resolution and colored according to  
901 fitted model chain. Right panel is clipped to reveal RTC active site. **(d)** Directional 3D  
902 FSC for S1\_RDV-TP, determined with 3DFSC. **(e)** Gold-standard FSC plot for the  
903 S1\_RDV-TP dataset, calculated by comparing two half maps from cryoSPARC. The  
904 blue dotted line represents the 0.143 FSC cutoff.



905  
906 **Extended Data Fig 2. (a)** Cryo-EM processing pipeline for the S2\_ATP dataset. **(b)**  
907 Angular distribution plot for the S2\_ATP dataset, calculated in cryoSPARC. Scale  
908 depicts number of particles assigned to a specific angular bin. **(c)** Nominal 3.09 Å  
909 resolution cryo-EM reconstruction filtered by local resolution and colored according to  
910 fitted model chain. Right panel is clipped to reveal RTC active site. **(d)** Directional 3D  
911 FSC for S2\_ATP, determined with 3DFSC. **(e)** Gold-standard FSC plot for the S2\_ATP  
912 dataset, calculated by comparing two half maps from cryoSPARC. The blue dotted line  
913 represents the 0.143 FSC cutoff.

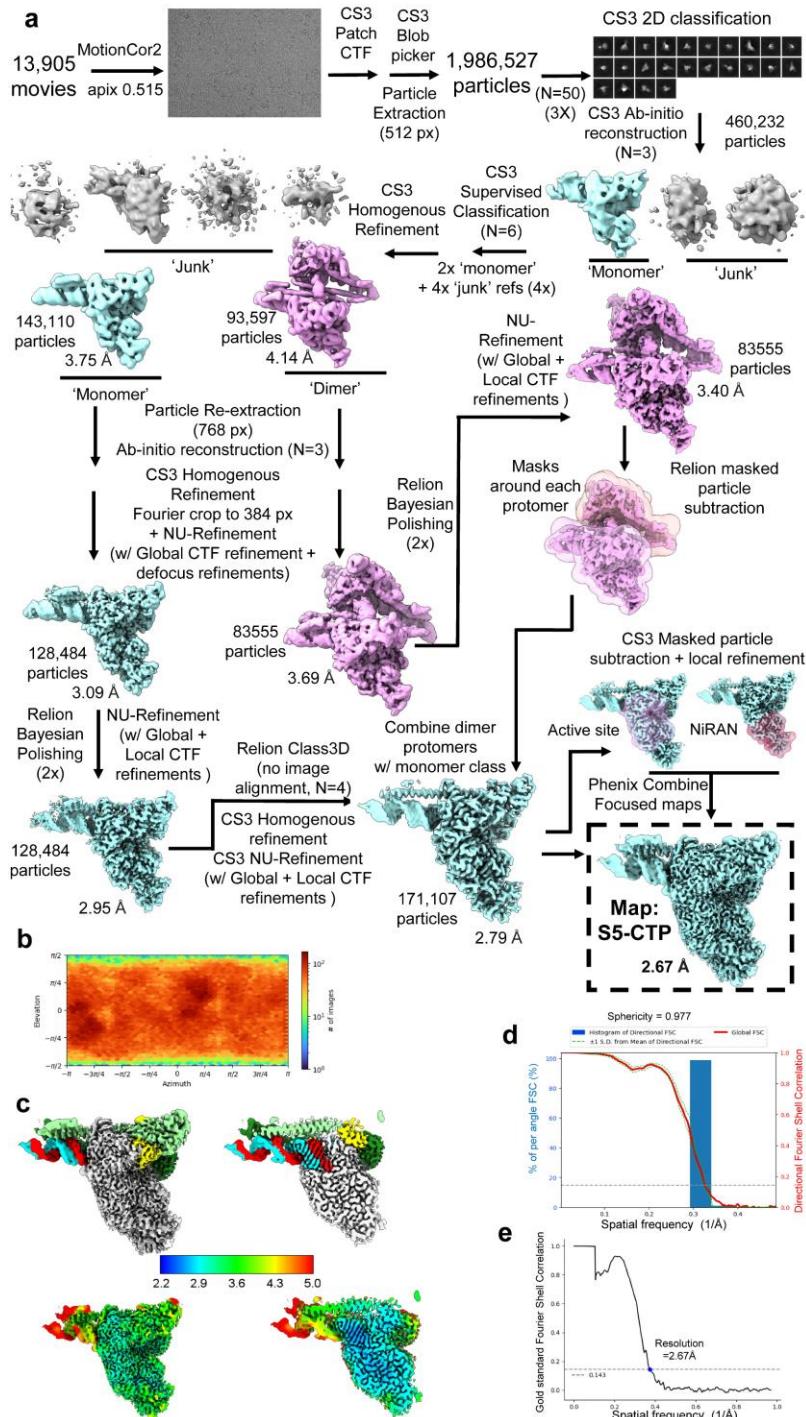


914  
915 **Extended Data Fig 3. (a)** Cryo-EM processing pipeline for the S3\_UTP dataset. **(b)**  
916 Angular distribution plot for the S3\_UTP dataset, calculated in cryoSPARC. Scale  
917 depicts number of particles assigned to a specific angular bin. **(c)** Nominal 3.13 Å  
918 resolution cryo-EM reconstruction filtered by local resolution and colored according to  
919 fitted model chain. Right panel is clipped to reveal RTC active site. **(d)** Directional 3D  
920 FSC for S3\_UTP, determined with 3DFSC. **(e)** Gold-standard FSC plot for the S3\_UTP  
921 dataset, calculated by comparing two half maps from cryoSPARC. The blue dotted line  
922 represents the 0.143 FSC cutoff.



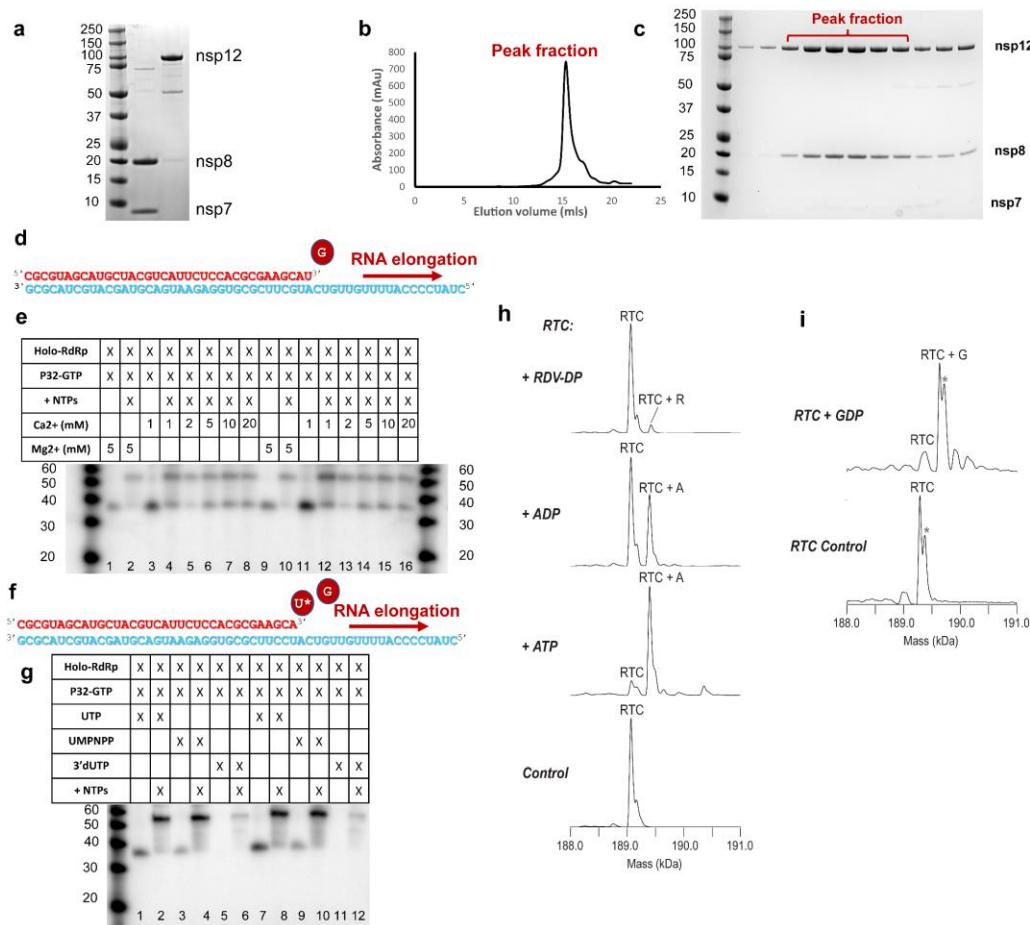
923  
924  
925  
926  
927  
928  
929  
930  
931

**Extended Data Fig 4. (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.



932  
933 **Extended Data Fig 5. (a)** Cryo-EM processing pipeline for the S5\_CTP dataset. **(b)**  
934 Angular distribution plot for the S5\_CTP dataset, calculated in cryoSPARC. Scale  
935 depicts number of particles assigned to a specific angular bin. **(c)** Nominal 2.67 Å  
936 resolution cryo-EM reconstruction filtered by local resolution and colored according to  
937 fitted model chain. Right panel is clipped to reveal RTC active site. **(d)** Directional 3D  
938 FSC for S5\_CTP, determined with 3DFSC. **(e)** Gold-standard FSC plot for the S5\_CTP  
939 dataset, calculated by comparing two half maps from cryoSPARC. The black dotted line  
940 represents the 0.143 FSC cutoff.

941



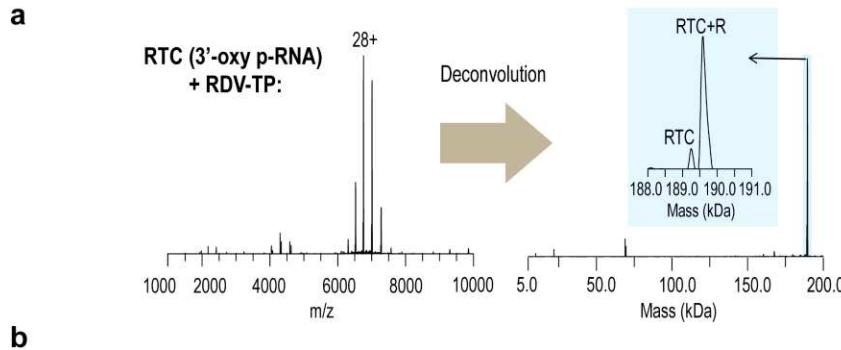
942

943

944 **Extended Data Fig 6. (a)** SDS-PAGE of purified SARS-CoV-2 nsp7/8 & nsp12. **(b)** Size  
 945 exclusion chromatography for the purified RTC complex, composed of nsp7/8/12  
 946 bound to the reconstituted product/template-RNA scaffold. **(c)** SDS-PAGE of assembled  
 947 RTC complex following size exclusion. **(d)** S4 RNA scaffold utilized for the S4\_GTP  
 948 structure and incorporation assays. **(e)** Gel-based primer elongation assay in presence  
 949 of physiological metal, Mg<sup>2+</sup>, and non-physiological metal, Ca<sup>2+</sup>, to investigate use of  
 950 Ca<sup>2+</sup> for stabilization of the pre-incorporation complex (gel depicts n=2 with n=3  
 951 experiments performed). **(f)** S3 RNA scaffold used for the S3\_UTP structure and  
 952 incorporation assays. **(g)** Gel-based primer elongation assay in presence of UTP,  
 953 UMPNPP and 3'deoxy UTP (gel depicts n=2 with n=3 experiments performed). **(h)**  
 954 Native mass spectrometry analysis of RNA extension in presence of ADP, RDV-DP &  
 955 ATP using the S1/S2 RNA scaffold. **(i)** Native mass spectrometry analysis of RNA  
 956 extension in presence of GDP using the S4 RNA scaffold.

957

958



**b**

Native MS mass measurements of RTC incubated with or without NTPs.

Sample Set	Sample	Protein Complex	Mass (Da)			% Mass Error <sup>c</sup>
			Measured <sup>a</sup>	Expected	Δ <sup>b</sup>	
RTC +/- ATP or RDV-TP (3'-oxy)	RTC alone	RTC	189,219	189,183	36	0.02
	RTC + ATP	RTC	189,238	189,183	55	0.03
		RTC + A	189,559	189,512	47	0.02
	RTC + RDV-TP	RTC	189,248	189,183	65	0.03
		RTC + R	189,588	189,536	52	0.03
RTC +/- GTP (3'-oxy)	RTC alone	RTC	189,451	189,403	48	0.03
	RTC4 + GTP	RTC + G	189,834	189,748	86	0.05
RTC* +/- ATP or RDV-TP (3'-deoxy)	RTC* alone	RTC*	189,206	189,167	39	0.02
	RTC* + ATP	RTC*	189,211	189,167	44	0.02
		RTC* + MgATP	189,804	189,699	105	0.06
	RTC* + RDV-TP	RTC*	189,209	189,167	42	0.02
		RTC* + Mg-RDV-TP	189,788	189,723	65	0.03
RTC* +/- GTP (3'-deoxy)	RTC* alone	RTC*	189,442	189,387	55	0.03
	RTC* + GTP	RTC*	189,471	189,387	84	0.04
		RTC* + Mg-GTP	190,033	189,935	98	0.05

<sup>a</sup>Obtained from UniDec software deconvolution.

<sup>b</sup>Mass difference between the expected and measured masses.

<sup>c</sup>Mass accuracy calculation as the percent mass difference relative to the expected mass.

959

960

**Extended Data Fig. 7. (a)** Data processing and deconvolution of native MS spectra using the UniDec software. Analysis of a representative MS spectrum (sample: RTC + RDV-TP) is shown. **(b)** The table of native MS mass measurements obtained from UniDec deconvolution of the RTC and RTC\* samples used in the NTP incorporation experiments

966

967

968

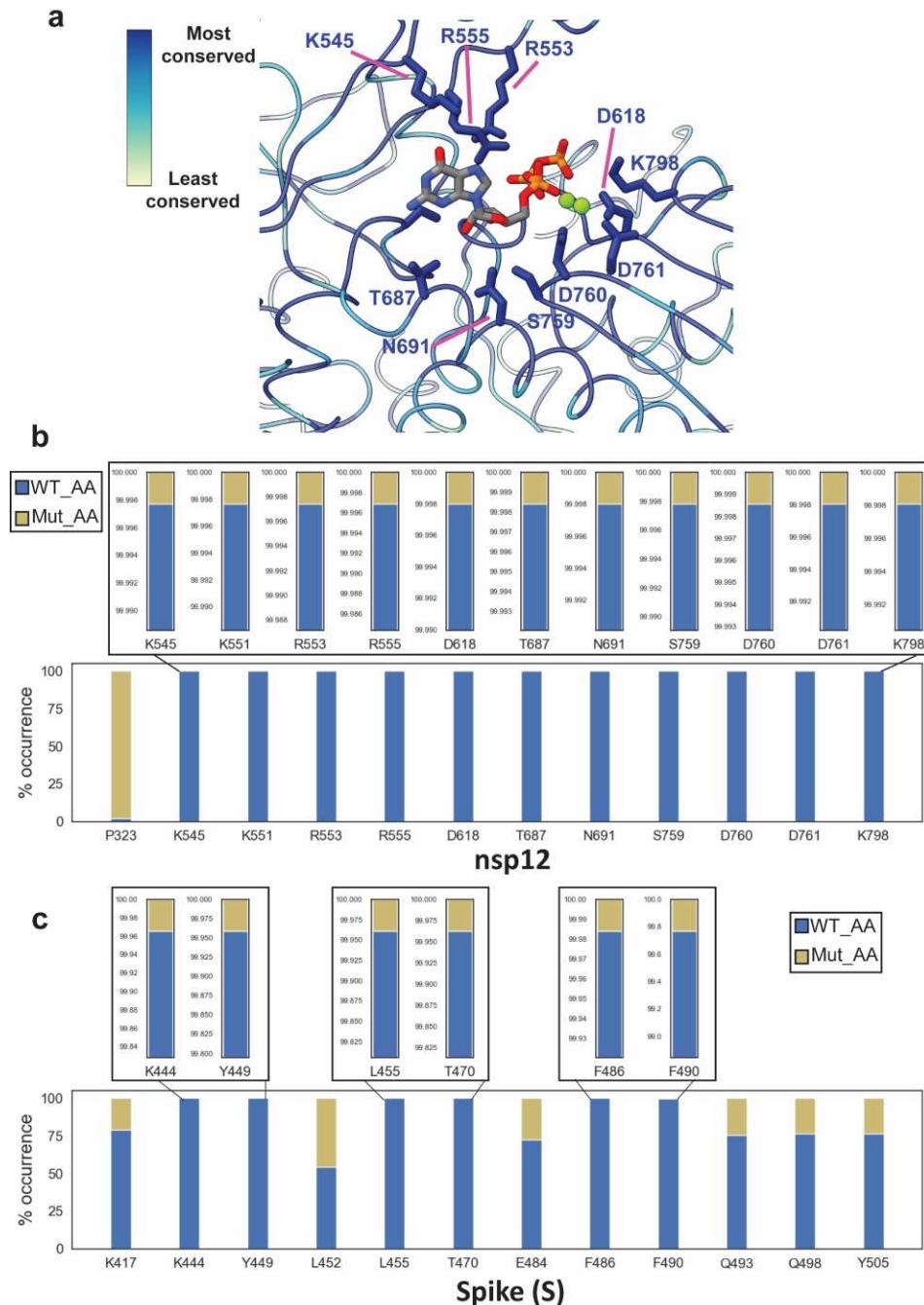
969

970

971

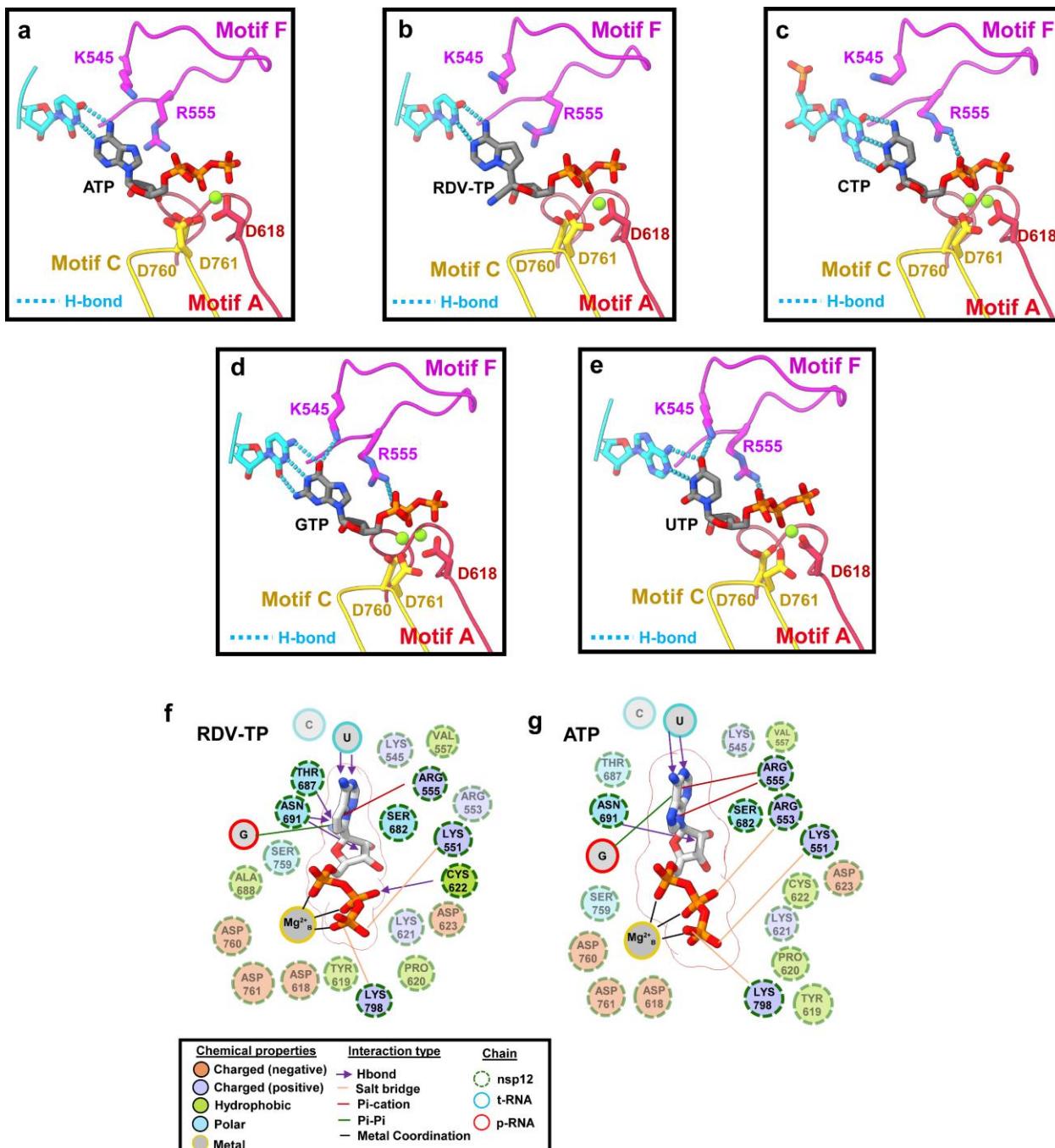
972

973



974

975 **Extended Data Fig 8. (a)** Zoom-in on the active site of S4\_GTP, highlighting residues  
 976 (sticks) which interact with the bound GTP/2xMg in which ribbon/sticks are colored  
 977 according to the amino acid conservation across a representative list of viruses found in  
 978 the  $\alpha$  &  $\beta$  coronavirus clades (Supplementary information). (b) Bar plot showing the  
 979 frequency of occurrence of the wild-type amino acid (reference strain Wuhan/Hu-  
 980 1/2019) as well as their summed mutations according to the GISAID database, as of  
 981 April 2022, for the nsp12 active site residues and the nsp12 residue P323. (c) Bar plot  
 982 showing the frequency of occurrence of the wild-type amino acid (reference strain  
 983 Wuhan/Hu-1/2019) as well as their summed mutations according to the GISAID  
 984 database, as of April 2022, for Spike (S) residues found in the ACE2 binding region.



985

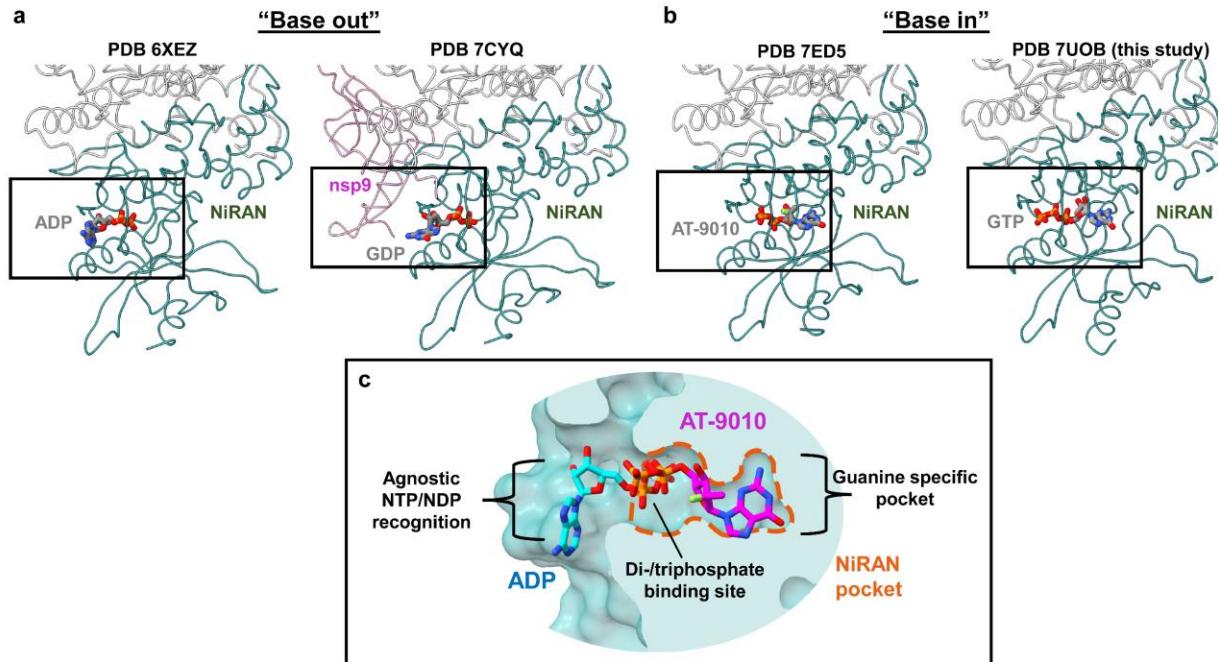
986

987

988 **Extended Data Fig 9. (a - e)** Analysis of the active sites of structures (S1-S5) illustrate  
 989 that K545 and R555 interact differentially with each of the incoming nucleotides. H-  
 990 bonds are depicted as dotted-cyan lines. (f, g) 2D schematics illustrating the set of  
 991 interactions between RDV-TP (f) and ATP (g) and the active site,

992

993



994  
995 **Extended Data Fig 10.** (a) Models (PDBs 6XEZ & 7CYQ) of the nsp12 NiRAN with a  
996 bound base in the 'Base-out' pose. (b) Models (PDBs 7ED5 & 7UOB) of the nsp12  
997 NiRAN with a bound base in the 'Base-in' pose. (c) Surface representation detailing the  
998 NiRAN active site pocket bound to a nucleotide in the 'base-out' pose (PDB 6XEZ)  
999 aligned with a structure with a bound nucleotide in the 'base-in' pose (PDB 7ED5).

1000

1001

1002

1003

1004

1005

1006

1007

1008

1009

1010

1011

1012

1013

1014

1015

1016 **Supplementary information**

**S1/S2 RNA scaffold**

<sup>5</sup>'CGCGUAGCAUGCUACGUCAUUCUCCUAAGAAGCUG\*<sup>3</sup>'

<sup>3</sup>'GCGCAUCGUACGAUGCAGUAAGAGGAUUCGUACUCGGCGAGGUACCCUAUC<sup>5</sup>'

**S3 RNA scaffold**

<sup>5</sup>'CGCGUAGCAUGCUACGUCAUUCUCCACGCGAAGCA\*<sup>3</sup>'

<sup>3</sup>'GCGCAUCGUACGAUGCAGUAAGAGGAUUCGUACUGUUGUUUACCCUAUC

**S4 RNA scaffold**

<sup>5</sup>'CGCGUAGCAUGCUACGUCAUUCUCCACGCGAAGCAU\*<sup>3</sup>'

<sup>3</sup>'GCGCAUCGUACGAUGCAGUAAGAGGAUUCGUACUGUUGUUUACCCUAUC

**S5 RNA scaffold**

<sup>5</sup>'CGCGUAGCAUGCUACGUCAUUCUCCACGCGAAGCAU\*<sup>3</sup>'

<sup>3</sup>'gcgcaucguacgaugCAGUAAGAGGugcgcuucguAGuguuguuUACCCUAUC

\* Denotes 3'-deoxy

1017

1018