

The mechanism of RNA capping by SARS-CoV-2

Vincent Tagliabracci (✉ Vincent.Tagliabracci@UTSouthwestern.edu)

HHMI/UT Southwestern Medical Center <https://orcid.org/0000-0002-9735-4678>

Gina Park

UT Southwestern Medical Center

Adam Osinski

UT Southwestern <https://orcid.org/0000-0003-0010-0733>

Genaro Hernandez

UT Southwestern Medical Center

Jennifer Eitson

Department of Microbiology, University of Texas Southwestern Medical Center

Abir Majumdar

UT Southwestern Medical Center <https://orcid.org/0000-0002-3516-5855>

Marco Tonelli

Katherine Henzler-Wildman

henzlerwildm@wisc.edu <https://orcid.org/0000-0002-5295-2121>

Krzysztof Pawłowski

Warsaw University of Life Sciences

Zhe Chen

UT Southwestern Medical Center

Yang Li

UT Southwestern Medical Center

John Schoggins

Department of Microbiology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.

<https://orcid.org/0000-0002-7944-6800>

Biological Sciences - Article

Keywords:

Posted Date: February 15th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1336910/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Nature on August 9th, 2022. See the published version at <https://doi.org/10.1038/s41586-022-05185-z>.

The mechanism of RNA capping by SARS-CoV-2

Gina J. Park^{1†}, Adam Osinski^{1†}, Genaro Hernandez¹, Jennifer L. Eitson², Abir Majumdar¹,
Marco Tonelli³, Katie Henzler-Wildman³, Krzysztof Pawłowski^{1,4}, Zhe Chen⁵, Yang Li⁵, John
W. Schoggins², Vincent S. Tagliabracci^{1,6,7,8*}

Affiliations:

¹Department of Molecular Biology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.

²Department of Microbiology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.

³Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706, USA.

⁴Department of Biochemistry and Microbiology, Institute of Biology, Warsaw University of Life Sciences, Warsaw 02-776, Poland.

⁵Department of Biophysics, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.

⁶Harold C. Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA

⁷Hamon Center for Regenerative Science and Medicine, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA

⁸Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA

*Correspondence to: Vincent S. Tagliabracci (vincent.tagliabracci@utsouthwestern.edu)

†These authors contributed equally to this work.

This PDF file includes:

Abstract

Main text

Figs. 1-6

Extended Data Figs. 1-12

Extended Data Tables 1-3

Methods

Acknowledgements

Author contributions

Competing interests

Materials & Correspondence

References

Preliminary PDB Validation Reports

35 **Abstract**

36 The SARS-CoV-2 RNA genome contains a 5'-cap that facilitates translation of viral proteins,
37 protection from exonucleases and evasion of the host immune response¹⁻⁴. How this cap is made
38 is not completely understood. Here, we reconstitute the SARS-CoV-2 ⁷MeGpppA_{2'-O-Me}-RNA cap
39 using virally encoded non-structural proteins (nsps). We show that the kinase-like NiRAN domain⁵
40 of nsp12 transfers RNA to the amino terminus of nsp9, forming a covalent RNA-protein
41 intermediate (a process termed RNylation). Subsequently, the NiRAN domain transfers RNA to
42 GDP, forming the cap core structure GpppA-RNA. The nsp14⁶ and nsp16⁷ methyltransferases then
43 add methyl groups to form functional cap structures. Structural analyses of the replication-
44 transcription complex bound to nsp9 identified key interactions that mediate the capping reaction.
45 Furthermore, we demonstrate in a reverse genetics system⁸ that the N-terminus of nsp9 and the
46 kinase-like active site residues in the NiRAN domain are required for successful SARS-CoV-2
47 replication. Collectively, our results reveal an unconventional mechanism by which SARS-CoV-
48 2 caps its RNA genome, thus exposing a new target in the development of antivirals to treat
49 COVID-19.

50

51 **Main Text**

52 Coronaviruses (CoVs) are a family of positive-sense, single-stranded RNA viruses that cause
53 disease in humans, ranging from mild common colds to more severe respiratory infections⁹. The
54 most topical of these is severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the
55 etiological agent of the ongoing COVID-19 pandemic, which to date has resulted in over 5.7-
56 million deaths and almost 400-million cases globally¹⁰.

57 The SARS-CoV-2 RNA genome contains two open reading frames (ORF1a and ORF1ab), which
58 are translated by host ribosomes to form two large polyproteins². These polyproteins are
59 subsequently cleaved by viral proteases to form 16 non-structural proteins (nsp1-16), some of
60 which make up the Replication-Transcription Complex (RTC)². At the core of the RTC is the
61 nsp12 RNA-dependent RNA polymerase (RdRp), which is the target of several promising
62 antivirals used to treat COVID-19 including remdesivir¹¹ and molnupiravir¹². In addition to the
63 RdRp domain, nsp12 contains an N-terminal Nidovirus RdRp-Associated Nucleotidyltransferase
64 (NiRAN) domain (**Fig. 1a**)⁵. The NiRAN domain shares sequence and structural similarity with
65 the pseudokinase selenoprotein-O (SelO), which transfers AMP from ATP to protein substrates (a
66 process termed AMPylation)¹³⁻¹⁵. Notably, the active site kinase-like residues of the NiRAN
67 domain are highly conserved in *Nidovirales* (**Extended Data Fig. 1**) and are required for equine
68 arteritis virus (EAV) and SARS-CoV-1 replication in cell culture⁵. Several hypotheses for the
69 function of the NiRAN domain have been proposed, including roles in protein-primed RNA
70 synthesis, RNA ligation, and mRNA capping^{5,16}.

71 The CoV RNA genome, like eukaryotic mRNAs, contains a methylated guanosine linked to the
72 first nucleotide of the RNA via a reverse 5' to 5' triphosphate linkage (**Extended Data Fig. 2**)^{1,4}.
73 This 5' cap is important for RNA stability, initiation of mRNA translation, and protection from
74 exonucleases¹⁷. Methylation of the ribose 2'-OH position of the first nucleotide completes the cap
75 and protects the RNA from the host immune system^{18,19}. Thus, formation of the RNA cap is crucial
76 for successful replication and transcription of the viral genome.

77 All eukaryotes share a conserved co-transcriptional capping mechanism (**Extended Data Fig. 2**)
78 involving: **1**) an RNA triphosphatase (RTPase), which removes the γ -phosphate from the nascent
79 5'-triphosphorylated RNA (5'-pppRNA) to yield a 5'-diphosphorylated RNA (5'-ppRNA); **2**) a
80 guanylyltransferase (GTase), which transfers GMP from GTP to 5'-ppRNA to form the cap core

81 structure GpppN-RNA; **3**) a (guanine-N7)-methyltransferase (N7-MTase), which methylates the
82 cap guanine at the N7 position; and **4**) a (nucleoside-2'-O)-methyltransferase (2'-O-MTase), which
83 methylates the ribose-2'-OH position on the first nucleotide of the RNA. In CoVs, the nsp13,
84 nsp14, and nsp16 proteins have RTPase²⁰, N7-MTase⁶, and 2'-O-MTase⁷ activities, respectively.
85 Thus, it was presupposed that the CoV capping mechanism occurs in a similar fashion to the
86 eukaryotic capping pathway, with the NiRAN domain functioning as the GTase^{3,5,21}. However,
87 evidence to support this claim has been lacking.

88 In this study, we discover that the NiRAN domain transfers monophosphorylated RNA (5'-pRNA)
89 from 5'-pppRNA to the N-terminus of nsp9 as an intermediate step in cap synthesis. The NiRAN
90 domain then transfers 5'-pRNA from RNAylated nsp9 to GDP to form the cap core structure
91 GpppA-RNA. We then reconstitute cap-0 and cap-1 structures using the nsp14 and nsp16
92 methyltransferases. Furthermore, we present a cryo-EM structure of the SARS-CoV-2 RTC with
93 the native N-terminus of nsp9 bound in the NiRAN active site. Finally, we demonstrate in a
94 reverse genetics system that the N-terminus of nsp9 and the kinase-like active site residues in the
95 NiRAN domain are required for SARS-CoV-2 replication.

96 **The NiRAN domain NMPylates the N-terminus of nsp9**

97 The NiRAN domain has been shown to transfer nucleotide monophosphates (NMPs) from
98 nucleotide triphosphates (NTPs) (referred to as NMPylation) to protein substrates, including
99 nsp9¹⁶ and the nsp12 co-factors, nsp7²² and nsp8²³. We observed NiRAN-dependent NMPylation
100 of native nsp9, but not native nsp7 or nsp8 (**Fig. 1b, Extended Data Fig. 3**). Quantification of ³²P
101 incorporation and intact mass analyses suggests stoichiometric incorporation of NMPs into nsp9
102 (**Extended Data Fig. 3g-j**). Mutation of nsp9 Asn1 to Ala or Asp reduced NMPylation of nsp9
103 (**Fig. 1c, Extended Data Fig. 4a**), consistent with previous work that suggested NMPylation
104 occurs on the backbone nitrogen of nsp9 Asn1¹⁶. To provide direct evidence that the amino
105 terminus of nsp9 is NMPylated by the NiRAN domain, we performed nuclear magnetic resonance
106 (NMR) spectroscopy of AMPylated nsp9. The 2D ¹H,³¹P HSQC and 2D HSQC-TOCSY spectra
107 confirm that the AMP is attached to the nitrogen backbone atom of Asn1 via a phosphoramidate
108 linkage (**Fig. 1d-f, Extended Data 4b, c**).

109

110

111 **The NiRAN domain RNAYlates nsp9**

112 Given the ability of the NiRAN domain to transfer NMPs to nsp9 using NTPs as substrates, we
113 wondered whether the NiRAN domain could also utilize 5'-pppRNA in a similar fashion (**Fig. 2a**).
114 We synthesized a 5'-pppRNA 10-mer corresponding to the first 10 bases in the leader sequence
115 (LS10) of the SARS-CoV-2 genome (hereafter referred to as 5'-pppRNA^{LS10}) (**Extended Data**
116 **Table 1**). We incubated 5'-pppRNA^{LS10} with nsp9 and nsp12 and analysed the reaction products
117 by SDS-PAGE. Remarkably, we observed an electrophoretic mobility shift in nsp9 that was time-
118 dependent, sensitive to RNase A treatment and required an active NiRAN domain, but not an
119 active RdRp domain (**Fig. 2b**). Intact mass analyses of the reaction products confirmed the
120 incorporation of monophosphorylated RNA^{LS10} (5'-pRNA^{LS10}) into nsp9 (**Fig. 2c**). The reaction
121 was dependent on Mn²⁺ (**Extended Data Fig. 5a**) and required a triphosphate at the 5'-end of the
122 RNA (**Extended Data Fig. 5b**). Substituting Ala for Asn1 reduced the incorporation of RNA^{LS10}
123 into nsp9 (**Fig. 2d**). We also observed NiRAN-dependent RNAYlation of nsp9 using LS RNAs
124 ranging from 2 to 20 nucleotides (**Fig. 2e**). Mutation of the first A to any other nucleotide markedly
125 reduced RNAYlation (**Fig. 2f**). Thus, the NiRAN domain RNAYlates the N-terminus of nsp9 in a
126 substrate-selective manner.

127 **The NiRAN domain transfers 5'-pRNA from nsp9 to GDP forming the cap core structure** 128 **GpppA-RNA**

129 Negative-sense RNA viruses of the order *Mononegavirales*, including vesicular stomatitis virus
130 (VSV), have an unconventional capping mechanism in which a polyribonucleotidyltransferase
131 (PRNTase) transfers 5'-pRNA from 5'-pppRNA to GDP via a covalent enzyme-RNA intermediate
132 (**Extended Data Fig. 6a**)^{24,25}. Because the NiRAN domain transfers 5'-pRNA to nsp9, we
133 hypothesized that this protein-RNA species may be an intermediate in a similar reaction
134 mechanism to that of the VSV system. To test this hypothesis, we purified the nsp9-pRNA^{LS10}
135 species by ion exchange and gel filtration chromatography and incubated it with GDP in the
136 presence of nsp12. Treatment with GDP deRNAYlated nsp9 in a NiRAN-dependent manner, as
137 judged by the nsp9 electrophoretic mobility on SDS-PAGE (**Fig. 3a**) and its molecular weight
138 based on intact mass analysis (**Fig. 3b**). The reaction was time-dependent, (**Fig 3c**), preferred Mg²⁺
139 over Mn²⁺ (**Extended Data Fig. 6b**) and was specific for GDP—and to some extent GTP—but
140 not the other nucleotides tested (**Fig. 3d**). Interestingly, although inorganic pyrophosphate (PP_i)

141 was able to deAMPylate nsp9-AMP, it was unable to deRNAylate nsp9-pRNA^{LS10} (**Fig. 3e**). (See
142 Discussion)

143 We used Urea-PAGE to analyse the fate of the RNA^{LS10} during the deRNAylation reaction.
144 Treatment of nsp9-pRNA^{LS10} with nsp12 and [α -³²P]GDP resulted in a [³²P]-labelled RNA species
145 that migrated similarly to GpppA-RNA^{LS10} produced by the Vaccinia capping enzyme (**Fig. 3f**).
146 The reaction was dependent on a functional NiRAN domain but not an active RdRp domain. To
147 confirm the presence of a GpppA-RNA cap, we digested the RNA produced from the nsp12
148 reaction with P1 nuclease and detected GpppA by high performance liquid chromatography/mass
149 spectrometry (HPLC/MS) analysis (**Fig. 3g**). Thus, the NiRAN domain is a GDP
150 polyribonucleotidyltransferase (GDP-PRNTase) that mediates the transfer of 5'-pRNA from nsp9
151 to GDP.

152 In our attempts to generate GpppA-RNA^{LS10} in a “one pot” reaction, we found that GDP inhibited
153 the RNAylation reaction (**Extended Data Fig. 6c**). However, the formation of GpppA-RNA^{LS10}
154 could be generated in one pot provided that the RNAylation occurs prior to the addition of GDP
155 (**Extended Data Fig. 6c, d**).

156 **Nsp14 and nsp16 catalyse the formation of the cap-0 and cap-1 structures**

157 The SARS-CoV-2 genome encodes an N7-MTase domain within nsp14⁶ and a 2'-O-MTase in
158 nsp16, the latter of which requires nsp10 for activity⁷. Nsp14 and the nsp10/16 complex use S-
159 adenosyl methionine (SAM) as the methyl donor. To test whether NiRAN-synthesized GpppA-
160 RNA^{LS10} can be methylated, we incubated ³²P-labelled GpppA-RNA^{LS10} with nsp14 and/or the
161 nsp10/16 complex in the presence of SAM and separated the reaction products by Urea-PAGE
162 (**Fig. 4a**). We extracted RNA from the reaction, treated it with P1 nuclease and CIP, and then
163 analysed the products by thin layer chromatography (TLC) (**Fig. 4b**). As expected, the NiRAN-
164 synthesized cap migrated similarly to the GpppA standard and the products from the Vaccinia
165 capping enzyme reaction (compare lanes 1 and 4). Likewise, reactions that included SAM and
166 nsp14 migrated similarly to the ⁷MeGpppA standard and to the products from the Vaccinia capping
167 enzyme reaction following the addition of SAM (compare lanes 2 and 6). Furthermore, treatment
168 of ⁷MeGpppA-RNA^{LS10}, but not unmethylated GpppA-RNA^{LS10}, with nsp10/16 produced the
169 ⁷MeGpppA_{2'-O-Me}-RNA cap-1 structure (compare lanes 3, 8 and 9). In parallel experiments, we
170 incubated NiRAN-synthesized GpppA-RNA^{LS10} with nsp14 and/or the nsp10/16 complex in the

171 presence of [¹⁴C]-labelled SAM (¹⁴C on the donor methyl group) and separated the reaction
172 products by Urea-PAGE. As expected, nsp14 and the nsp10/16 complex incorporated ¹⁴C into
173 GpppA-RNA^{LS10} to form the cap-0 and cap-1 structures, respectively (**Fig 4c**). Thus, the SARS-
174 CoV-2 ⁷MeGpppA_{2'-O-Me}-RNA capping mechanism can be reconstituted in vitro using virally
175 encoded proteins.

176 Efficient translation of mRNAs is dependent on eIF4E binding to the ⁷MeGpppA-RNA cap²⁶. To
177 test whether the SARS-CoV-2 RNA cap is functional, we incubated [³²P]-labelled ⁷MeGpppA-
178 RNA^{LS10} with GST-tagged eIF4E. We observed [³²P]-labelled RNA in GST pulldowns of
179 [³²P]⁷MeGpppA-RNA but not the unmethylated derivative (**Fig. 4d**). Thus, the ⁷MeGpppA-RNA cap
180 generated by SARS-CoV-2 encoded proteins is a substrate for eIF4E in vitro, suggesting that the
181 cap is functional.

182 **Structural insights into RNA capping by the NiRAN domain**

183 We determined a cryo-EM structure of the nsp7/8/9/12 complex and observed a nsp9 monomer
184 bound in the NiRAN active site (**Fig. 5a, Extended Data Fig. 7-9, Extended Data Table 2**). The
185 native N-terminus of nsp9 occupies a similar position to previously reported structures using a
186 non-native N-terminus of nsp9 (**Fig. 5b, c**)²¹. Our cryo-EM analysis was hindered by the preferred
187 orientation of the complex and sample heterogeneity, yielding final maps with high levels of
188 anisotropy, with distal portions of nsp9 missing, and weak density for the N-lobe of the NiRAN
189 domain (**Extended Data Fig. 7, 8**). Therefore, we used our model and the complex structure by
190 Yan et al.²¹ (PDBID: 7CYQ) to study the structural basis of NiRAN-mediated RNA capping.

191 The first four residues of nsp9 extend into the NiRAN active site, forming electrostatic and
192 hydrophobic contacts in and around a groove near the kinase-like active site (**Fig. 5d**). Asn1 of
193 nsp9 is positioned inside of the active site, primed for transfer of 5'-pppRNA onto its N-terminus.
194 Although the terminal NH₂ group of nsp9 is the substrate for RNylation, the local quality of the
195 structures is not high enough to distinguish its exact position. We have modelled the nsp9 acceptor
196 NH₂ pointing towards what appears to be the phosphates of the nucleotide analogue UMP-NPP in
197 the active site (**Fig 5b**). In the structure by Yan et al.²¹, Asn1 was assigned an opposite
198 conformation and there are unmodeled residues (non-native N-terminus; NH₂-Gly-Ser-) visible in
199 the density maps, distorting local structural features (**Fig. 5c, arrow**)²⁷.

200 Asn2 of nsp9 is in a negatively charged cleft around the NiRAN active site, and contacts Arg733,
201 which extends from the polymerase domain and is partially responsible for positioning nsp9 (**Fig.**
202 **5e**). Both Leu4 and the C-terminal helix of nsp9 form hydrophobic interactions with a β -sheet (β 8-
203 β 9- β 10) in the N-lobe of the NiRAN domain (**Fig. 5e, f**). The N-terminal cap of the nsp9 C-terminal
204 helix also forms electrostatic interactions with a negatively charged pocket on the surface of the
205 NiRAN domain (**Fig. 5f**). Nsp12 lacking the RdRp domain (Δ RdRp; 1-326) neither RNAylates
206 nsp9 nor processes nsp9-pRNA^{LS10} to form GpppA-RNA (**Fig. 5g**). Likewise, deleting the C-
207 terminal helix on nsp9 (Δ C; 1-92) and Ala substitutions of Asn1 and Asn2 abolished RNAylation
208 (**Fig. 5h**).

209 The NiRAN domain resembles SelO, with an RMSD of 5.7 Å over 224 Ca atoms (PDB ID:
210 6EAC¹³, **Extended Data Fig. 10a**). Lys73 (PKA nomenclature; K72) forms a salt bridge with
211 Glu83 (PKA; E91) from the α C (α 2) helix and contacts the phosphates (GDP in 7CYQ, or UMP-
212 NPP in our structure; **Fig. 5i**). As expected, the “DFG” Asp218 (PKA; D184) binds a divalent
213 cation. Interestingly, the NiRAN domain lacks the catalytic Asp (**Extended Data Fig. 1**), (PKA;
214 D166); however, like in SelO, Asp208 is next to the metal binding Asn209 (PKA; N171) and may
215 act as a catalytic base to activate the NH₂ group on the N-terminus of nsp9 (**Fig. 5i**).

216 In canonical kinases, the β 1- β 2 G-loop stabilizes the phosphates of ATP²⁸. In contrast, the NiRAN
217 domain contains a β -hairpin insert (β 2- β 3) where the β 1- β 2 G-loop should be (**Extended Data**
218 **Fig. 10b**). This insertion not only makes contacts with the N-terminus of nsp9, but also contains a
219 conserved Lys (K50) that extends into the active site and stabilizes the phosphates of the bound
220 nucleotide. Likewise, Arg116 also contacts the phosphates of the nucleotide. SelO contains a
221 similar set of basic residues pointing into the active site that accommodate the flipped orientation
222 of the nucleotide to facilitate AMPylation (**Extended Data Fig. 10b**). Notably, Lys73, Arg116
223 and Asp218 in SARS-CoV-1 nsp12 are required for viral replication⁵.

224 **The kinase-like residues of the NiRAN domain and the N-terminus of nsp9 are essential for** 225 **SARS-CoV-2 replication**

226 To determine the importance of the NiRAN domain and the N-terminus of nsp9 in viral replication,
227 we used a DNA-based reverse genetics system that can rescue infectious SARS-CoV-2 (Wuhan-
228 Hu-1/2019 isolate) expressing a fluorescent reporter⁸ (**Extended Data Fig. 11a**). We introduced
229 single point mutations in nsp9 (N1A, N1D and N2A) and nsp12 (K73A, D218A and D760A) and

230 quantified the virus in supernatants of producer cells by RT-qPCR to detect the viral N gene. We
231 observed a 400 to 4000-fold reduction in viral load for all the mutants compared to WT (**Fig. 5j**,
232 **Extended Data Fig. 11a**). To account for the possibility of a proteolytic defect in the mutant viral
233 polyprotein, we tested whether the main viral protease nsp5 (M^{Pro}) can cleave a nsp8-nsp9 fusion
234 protein containing the Asn1/Asn2 mutations in nsp9. The N1D mutant failed to be cleaved by
235 nsp5, suggesting that the replication defect observed for this mutant is a result of inefficient
236 processing of the viral polyprotein. However, the N1A and N2A mutants were efficiently cleaved
237 by nsp5 (**Extended Data Fig. 11b, c**). Collectively, these data provide genetic evidence that the
238 residues involved in capping of the SARS-CoV-2 genome are essential for viral replication.

239 Discussion

240 We propose the following mechanism of RNA capping by CoV: during transcription, the nascent
241 5'-pppRNA binds to the NiRAN active site, in either a *cis* (**Fig. 6a**) or *trans* (**Fig. 6b**) manner and
242 5'-pRNA is subsequently transferred to the N-terminus of nsp9 forming a phosphoramidate bond
243 (**Fig. 6c, panels 1 and 2**). The nsp13 protein produces GDP from GTP, which binds the NiRAN
244 active site and attacks RNAylated nsp9, releasing capped RNA and regenerating unmodified nsp9
245 (**Fig. 6c, panels 3 and 4**). Subsequently, nsp14 and nsp16 perform sequential N7 and 2'-O
246 methylations, forming a fully functional ⁷MeGpppA_{2'-O-Me}-RNA cap.

247 SARS-CoV-2 nsp12 is thought to initiate transcription/replication starting with an NTP, or a short
248 5'-pppRNA primer²⁹. Cryo-EM structures of the RTC suggest that the dsRNA product makes its
249 way out of the RdRp active site in a straight line, supported by the nsp8 helical stalks^{21,30,31}. In a
250 *cis* capping model, the helical duplex with nascent 5'-pppRNA would then need to unwind, flex
251 90°, and extend into the NiRAN active site ~70 Å away (**Fig. 6a**). More likely, a separate RTC
252 complex could perform capping in *trans* (**Fig. 6b**). Notably, Perry et al.³² propose that the nascent
253 RNA strand is separated from the template upon passing through the proof-reading ExoN domain
254 of nsp14 on a neighbouring RTC and threaded towards the NiRAN domain.

255 SARS-CoV-1 nsp13 has RNA helicase, nucleotide triphosphatase (NTPase), and RNA 5'-
256 triphosphatase (RTPase) activities²⁰. The RTPase activity implicated nsp13 in the first step of the
257 capping mechanism; however, while nsp13 can act on 5'-pppRNA, this reaction is inhibited in the
258 presence of cellular concentrations of ATP²⁰. Thus, we favour the idea that the physiological
259 functions for nsp13 are: **1**) to utilize the energy from ATP hydrolysis to unwind double-stranded

260 RNA (helicase), and **2**) to hydrolyse GTP to GDP, which can then act as an acceptor for 5'-pRNA
261 in the NiRAN-catalysed capping reaction.

262 The SARS-CoV-2 capping mechanism is reminiscent of the capping mechanism used by VSV,
263 although there are some differences. The VSV large (L) protein is a multifunctional enzyme that
264 carries out RdRp, PRNTase, and methyltransferase activities to form the cap^{24,33,34}. During the
265 reaction, 5'-pRNA is transferred to a conserved His within the PRNTase domain, which adopts a
266 unique α -helical fold that is distinct from that of protein kinases²⁵. The presence of two different
267 enzymatic mechanisms of capping, proceeding via covalent protein-RNA intermediates, in
268 *Mononegavirales* and in *Nidovirales* is an example of convergent evolution.

269 Consistent with other reports^{16,27}, we observed NiRAN-catalysed NMPylation of nsp9 (**Fig. 1b**,
270 **Extended Data Fig. 3**). While our results do not necessarily preclude a biologically relevant
271 function for nsp9 NMPylation, it is worth noting that this modification is reversible in the presence
272 of PP_i²⁷ (**Fig. 3e**). PP_i is produced during the RdRp reaction, making the stability of NMPylated
273 nsp9 difficult to envision in vivo. By contrast, RNAylated nsp9 was not reversible in the presence
274 of PP_i. Thus, RNAylation is likely the physiologically relevant modification of nsp9 during viral
275 RNA capping.

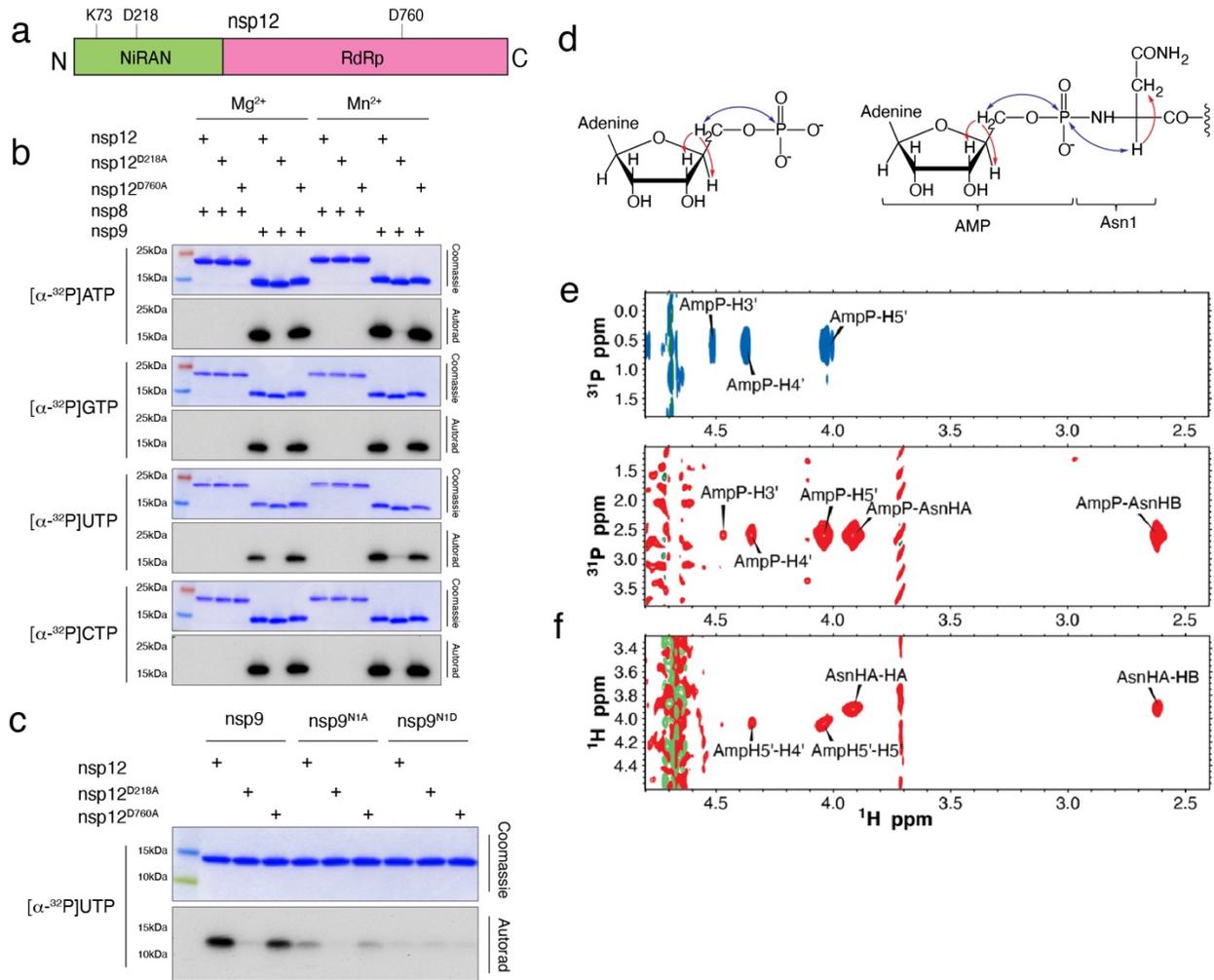
276 Recent work suggested that the NiRAN domain is a GTase that transfers GMP from GTP to 5'-
277 ppRNA, forming a GpppA-RNA cap intermediate^{3,21}. In our efforts to reproduce these results, we
278 failed to detect nsp12-dependent GpppA cap formation by TLC (**Extended Data Fig. 12a**) or by
279 Urea-PAGE analysis of the RNA (**Extended Data Fig. 12b**), in contrast to our control, in which
280 the Vaccinia capping enzyme efficiently generated GpppA-RNA. Because nsp13 and the Vaccinia
281 capping enzyme can hydrolyse GTP to GDP²⁰, the cap reported previously²¹ appears to be GDP
282 formed from nsp13- and Vaccinia capping enzyme-dependent hydrolysis of GTP.

283 In summary, we have defined the mechanism by which SARS-CoV-2 caps its genome and have
284 reconstituted this reaction in vitro using non-structural proteins encoded by SARS-CoV-2. Our
285 results uncover new targets for the development of antivirals to treat COVID-19 and highlight the
286 catalytic adaptability of the kinase domain.

287

288 **Main Figures**

289
290

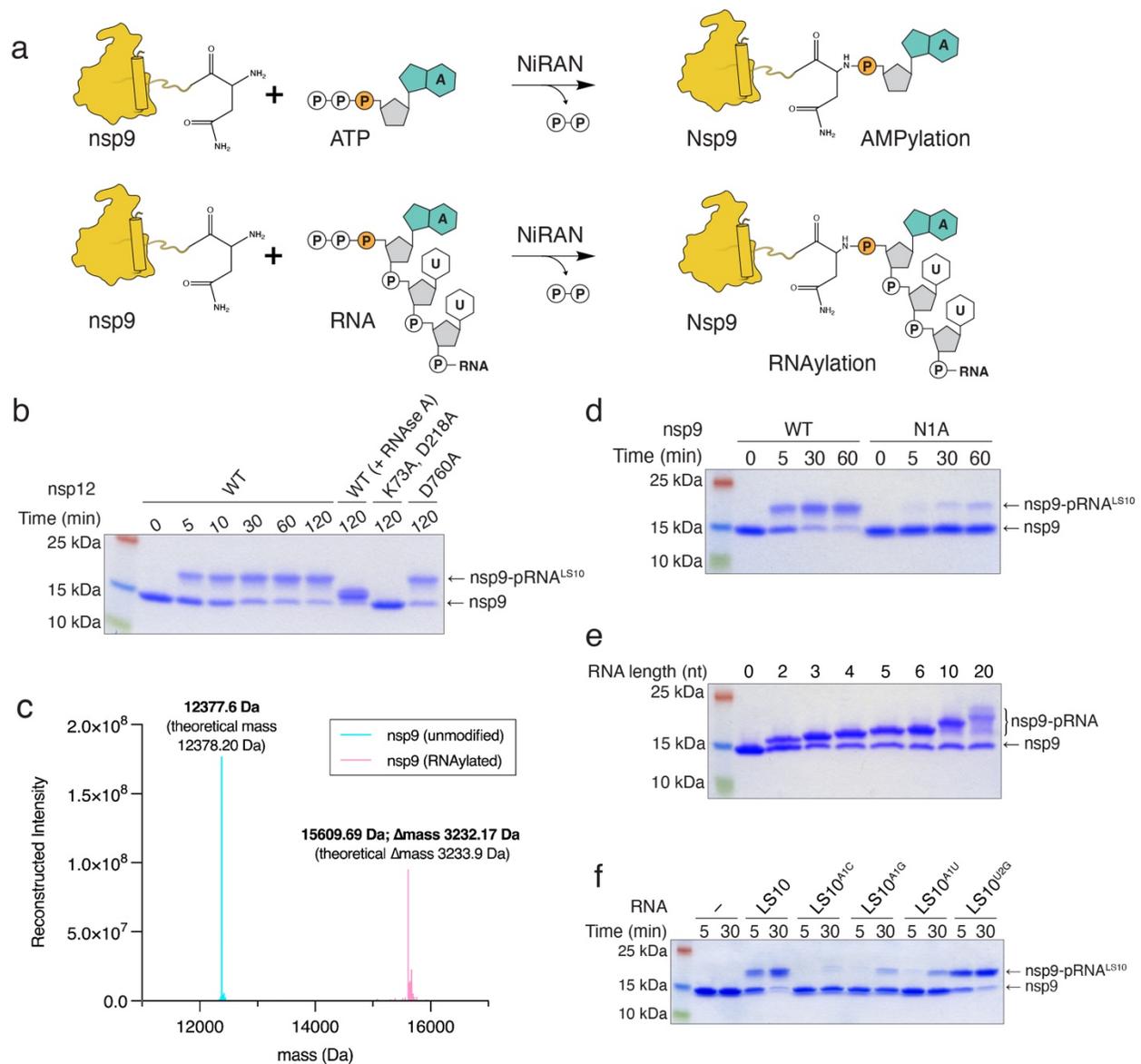


291
292
293

294 **Figure 1. The NiRAN domain NMPylates the N-terminus of nsp9.** **a.** Domain architecture of
295 nsp12 depicting the SelO -like NiRAN domain (*green*) and the RNA-dependent RNA polymerase
296 domain (RdRp; *magenta*), annotated with the predicted catalytic residues. **b.** Incorporation of α-
297 ³²P from [α-³²P]ATP, GTP, UTP, or CTP into nsp8 or nsp9 by WT nsp12, the NiRAN mutant
298 (D218A), or the polymerase mutant (D760A). Reactions were performed in the presence of Mg²⁺
299 or Mn²⁺ and the products were resolved by SDS-PAGE and visualized by Coomassie staining (*top*)
300 and autoradiography (*bottom*). **c.** Incorporation of α-³²P from [α-³²P]UTP into nsp9 or the indicated
301 mutants by the NiRAN domain. Reaction products were analysed as in **b.** **d.** Structure of AMP
302 (*left*) and AMP-nsp9 (*right*) with arrows to indicate the magnetization transfer steps that result in

303 the peaks observed in the 2D NMR spectra. The blue arrows indicate the transfer steps that yield
304 the peaks in the HSQC spectra, while the red arrows show the additional magnetization transfer
305 during the TOCSY that result in the additional peaks found in the HSQC-TOCSY spectra. **e.** 2D
306 ¹H,³¹P-HSQC-TOCSY spectra of AMP (*top, blue*) and AMP-nsp9 (*bottom, red*). **f.** 2D ¹H,¹H-
307 HSQC-TOCSY spectra of AMP-nsp9. Results shown in **b** and **c** are representative of at least 3
308 independent experiments.

309



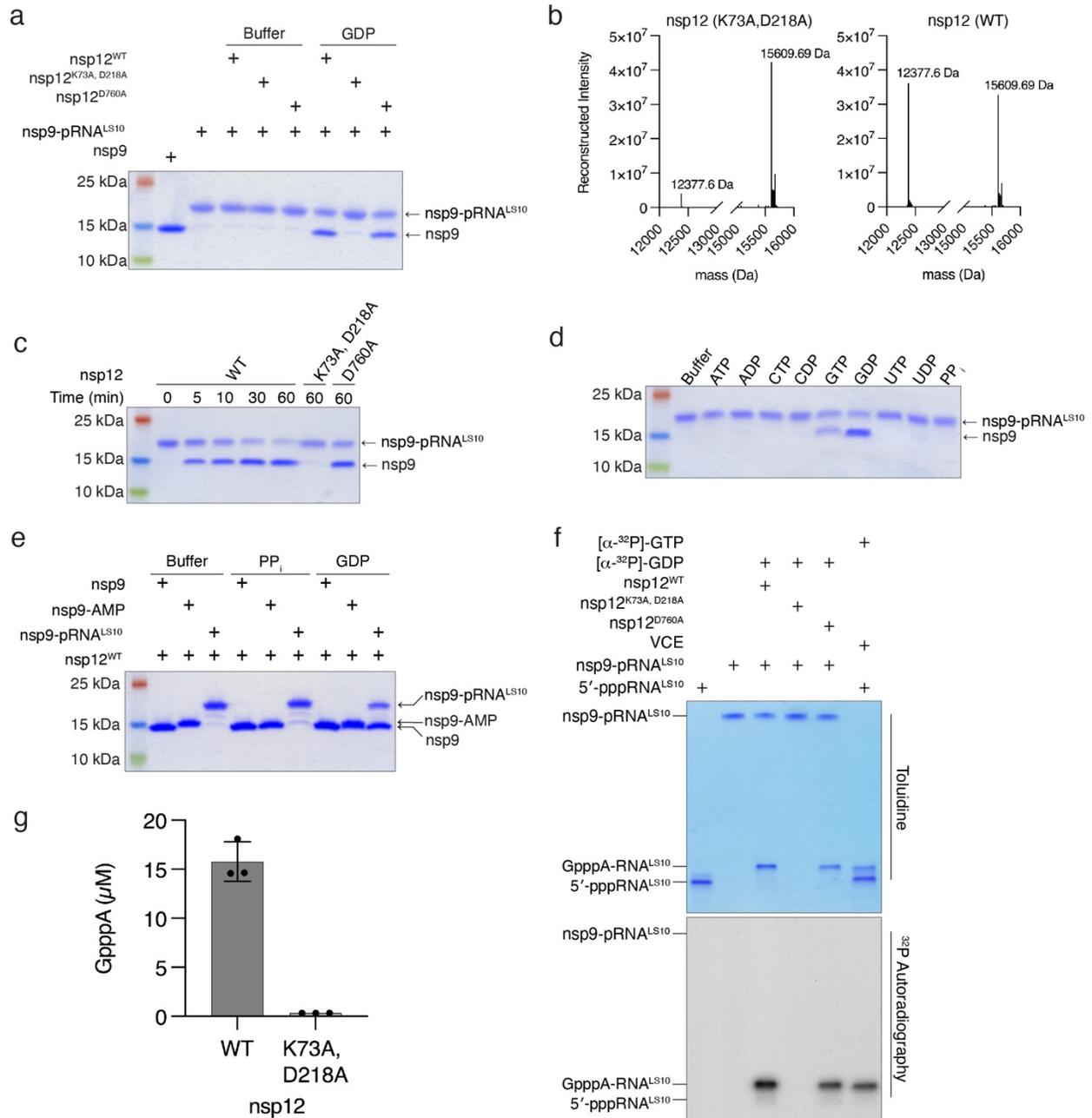
310

311

312 **Figure 2. The NiRAN domain RNAylates nsp9.** **a.** Schematic depicting the nsp9 AMPylation
 313 reaction (*top*) and the proposed nsp9 RNAylation reaction (*bottom*). **b.** Time-dependent
 314 incorporation of RNA into nsp9 by WT nsp12, the NiRAN mutant (K73A, D218A), or the
 315 polymerase mutant (D760A). Reaction products were analysed by SDS-PAGE and Coomassie
 316 staining. Samples were also treated with RNase A. **c.** Intact mass LC/MS spectra (overlaid) of
 317 unmodified nsp9 (*cyan*) or nsp9 after incubation with 5'-pppRNA^{LS10} and WT nsp12 (*pink*). The
 318 theoretical and observed masses are shown in the insets. The Δmass of 3233.17 Da corresponds
 319 to monophosphorylated RNA^{LS10} (5'-pRNA^{LS10}). **d.** Time-dependent incorporation of 5'-pRNA^{LS10}

320 into nsp9 or the nsp9 N1A mutant. Reaction products were analysed as in **b. e.** Incorporation of
321 different lengths of 5'-pppRNAs into nsp9 by the NiRAN domain. Reaction products were
322 analysed as in **b. f.** Time-dependent incorporation of RNAs with substitutions in the first and
323 second base into nsp9 by the NiRAN domain. Reaction products were analysed as in **b.** Results
324 shown are representative of at least 2 independent experiments.

325



326

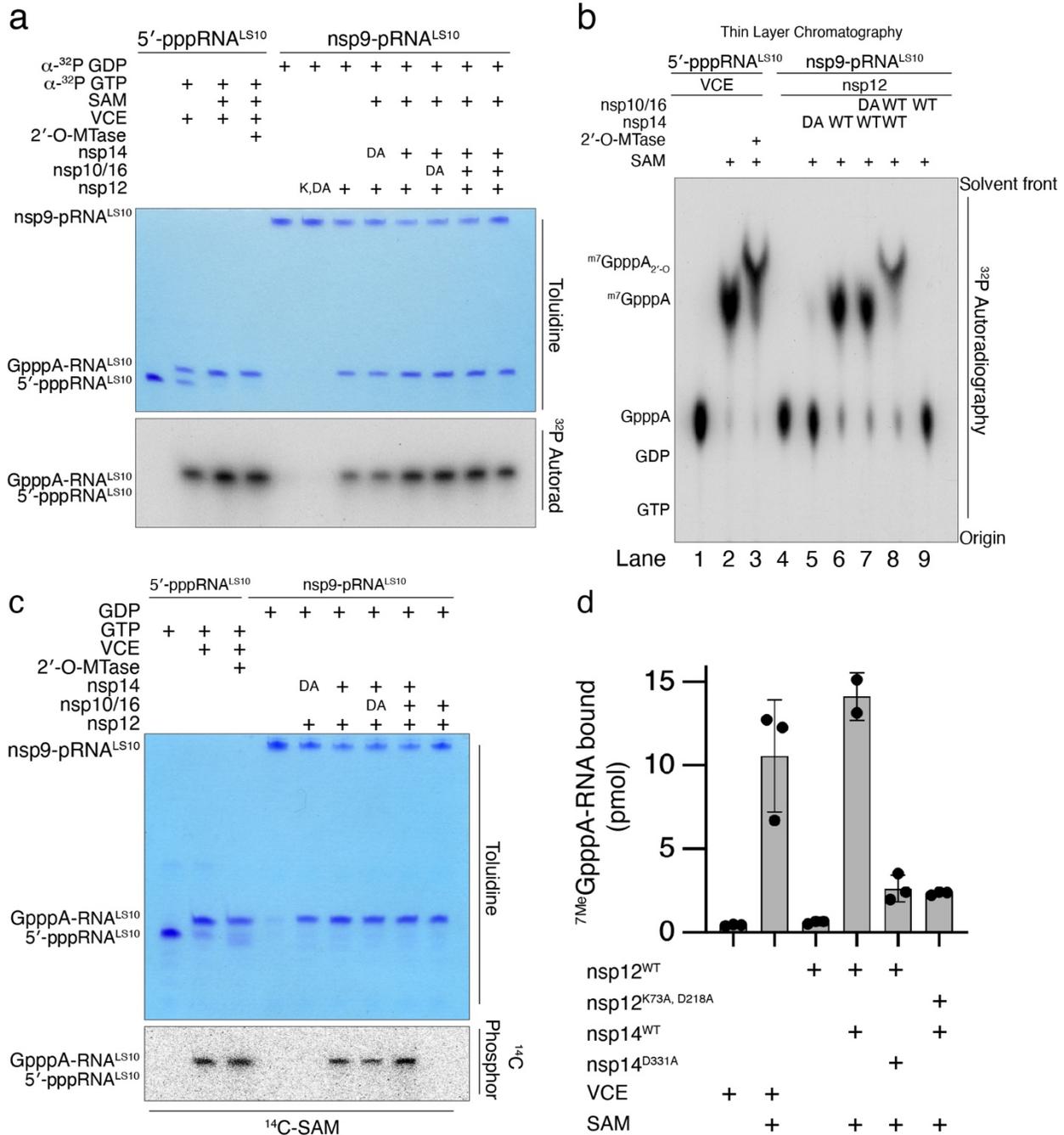
327

328 **Figure 3. The NiRAN domain catalyses the transfer of 5'-pRNA from nsp9 to GDP to form**
 329 **the cap core structure GpppA-RNA.** **a.** DeRNAylation of the covalent nsp9-RNA^{LS10} species by
 330 WT nsp12, the NiRAN mutant (K73A, D218A), or the polymerase mutant (D760A) when
 331 incubated with buffer or GDP. Reaction products were analysed as in **Fig. 2b.** **b.** Intact mass
 332 LC/MS spectra of nsp9-pRNA^{LS10} after incubation with GDP and WT nsp12 (*right*) or the NiRAN
 333 mutant (*left*). The theoretical mass of nsp9 is 12378.2 Da and the theoretical mass of nsp9-

334 pRNA^{LS10} is 15611.5 Da. **c.** Time-dependent deRNAylation of nsp9-pRNA^{LS10} by WT nsp12, the
335 NiRAN mutant (K73A, D218A), or the polymerase mutant (D760A). Reaction products were
336 analysed as in **Fig. 2b. d.** DeRNAylation of nsp9-pRNA^{LS10} by nsp12 in the presence of different
337 NTPs, NDPs or PP_i. Reaction products were analysed as in **Fig. 2b. e.** NiRAN-dependent
338 deAMPylation or deRNAylation of nsp9 in the presence of PP_i or GDP. Reaction products were
339 analysed as in **Fig. 2b. f.** Incorporation of α -³²P from [α -³²P]GDP into nsp9-pRNA^{LS10} by WT
340 nsp12, the NiRAN mutant (K73A, D218A), or the polymerase mutant (D760A). Vaccinia capping
341 enzyme (VCE) was used as a control but incubated with [α -³²P]GTP. Reaction products were
342 resolved by Urea-PAGE and visualized by toluidine blue O staining (upper) and autoradiography
343 (lower). **g.** HPLC/MS quantification of GpppA formed during the NiRAN-catalysed
344 deRNAylation of nsp9-pRNA^{LS10}. Reaction products were digested with nuclease P1 prior to
345 HPLC analysis. Reactions were performed in triplicate and error bars represent the standard
346 deviation. Results shown are representative of at least 2 independent experiments.

347

348



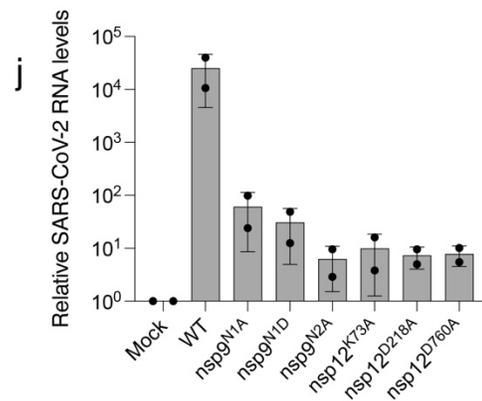
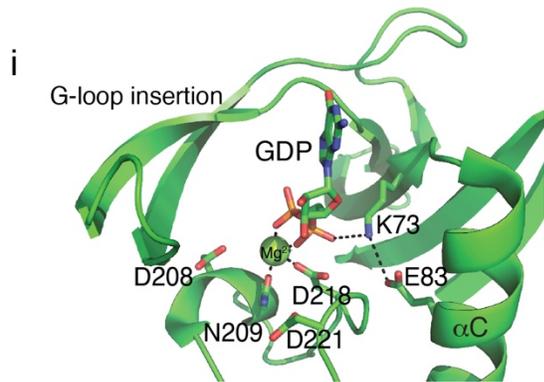
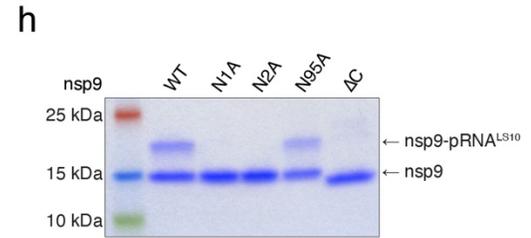
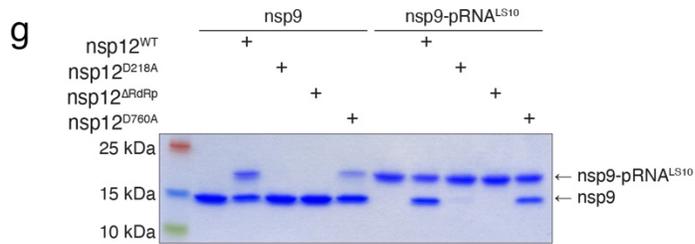
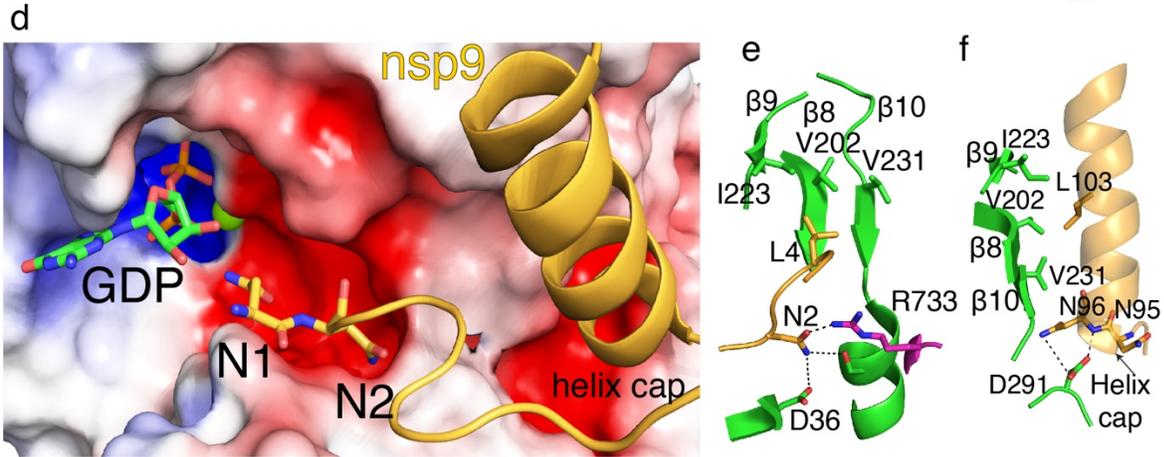
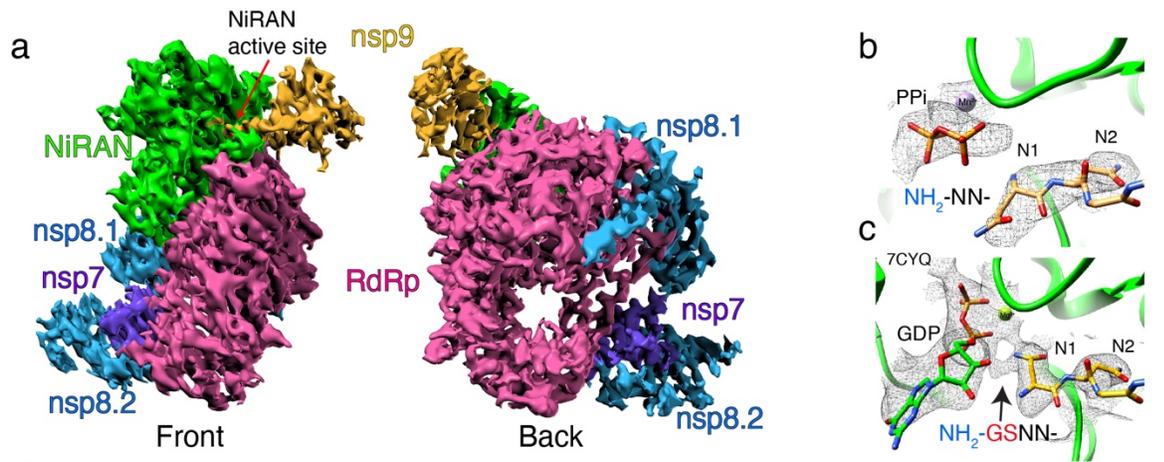
349

350

351 **Figure 4. Nsp14 and nsp16 catalyse the formation of the cap-0 and cap-1 structures. a.**
 352 Incorporation of α -³²P from [α -³²P]GDP into nsp9-pRNA^{LS10} by WT nsp12, or the NiRAN K73A,
 353 D218A mutant (K, DA). Reactions were subsequently incubated with SAM, nsp14 (or the D331A
 354 mutant; DA), and nsp10/16 (or the D130A nsp16 mutant; DA). The Vaccinia capping enzyme
 355 (VCE) was used as a positive control but incubated with [α -³²P]GTP and the Vaccinia 2'-O-MTase.

356 Reaction products were analysed as in **Fig 3f. b.** Thin layer chromatograms depicting the reaction
357 products from **Fig. 4a** following extraction from the Urea PAGE gel and treatment with PI nuclease
358 and CIP. Location of the cold standards (left) was visualized by UV fluorescence and the ^{32}P by
359 autoradiography. **c.** Incorporation of ^{14}C from [methyl- ^{14}C]SAM into GpppA-RNA^{LS10} by nsp14
360 (or the D331A mutant; DA), and nsp10/16 (or the D130A nsp16 mutant; DA). The VCE and the
361 Vaccinia 2'-O-MTase were used as positive controls. Reaction products were analysed as in **Fig**
362 **3f.** The ^{14}C signal was detected by phosphorimaging. **d.** Pull-down assays depicting the binding of
363 [^{32}P]^{7Me}GpppA-RNA^{LS10} to GST-eIF4E. [^{32}P]^{7Me}GpppA-RNA^{LS10} was produced using SARS-
364 CoV-2 virally encoded proteins or the VCE. Radioactivity in GST pull-downs was quantified by
365 scintillation counting. Results represent three independent experiments. Error bars represent the
366 standard deviation (SD). Results shown are representative of at least 2 independent experiments.

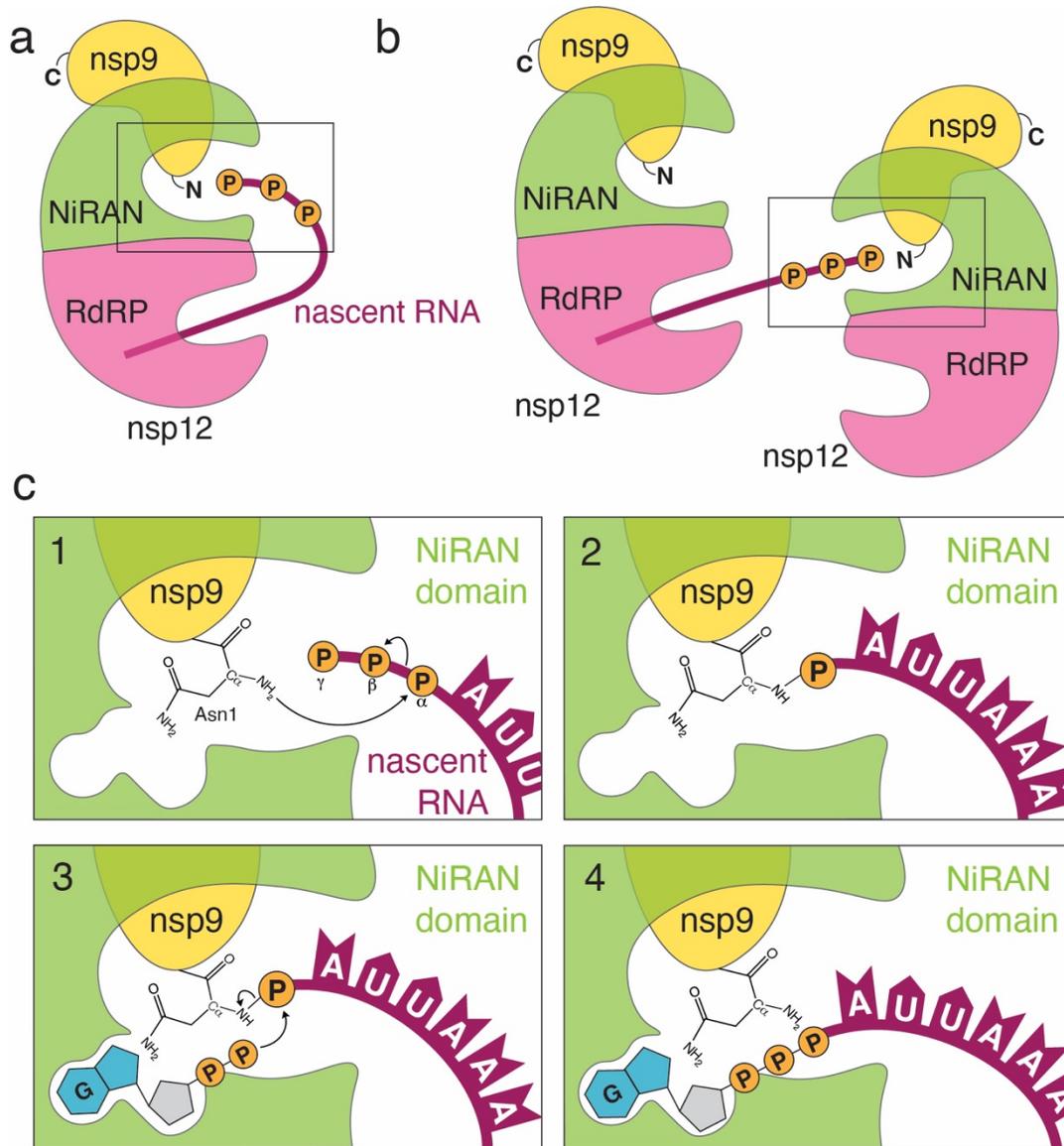
367



370 **Figure 5. Structural and genetic insights into RNA capping by the kinase domain. a.** Front
371 and back views of nsp12/7/8/9 cryo-EM maps, with respect to the NiRAN domain. The NiRAN
372 domain is in green, the RdRp in magenta, nsp7 in violet, nsp8 in light blue and nsp9 in gold. **b, c.**
373 Coulomb density maps of the N-terminus of nsp9 from this study (**b**) and by Yan et al. (**c**)²¹ (PDB
374 ID:7CYQ). The NiRAN domain is shown in green and nsp9 is in gold. The arrow in (**c**) indicates
375 additional density that likely corresponds to unmodeled Gly-Ser residues at the non-native N-
376 terminus of nsp9. **d.** Electrostatic surface view of the NiRAN active site from 7CYQ bound to
377 nsp9 (gold). The N-terminus and the C-terminal helix of nsp9 are shown. Electrostatic surface of
378 nsp12 is contoured at 5 kT. **e.** Cartoon representation depicting the interactions between the nsp9
379 N-terminus (gold) with the β 8- β 9- β 10 sheet in the NiRAN domain (green). Asn2 in nsp9 forms
380 electrostatic interactions with Asp36 in the NiRAN domain and Arg733 in the RdRp domain
381 (magenta). PDB ID 7CYQ was used. **f.** Cartoon representation depicting the interactions between
382 the nsp9 C-terminal helix (gold) and the β 8- β 9- β 10 sheet in the NiRAN domain (green).
383 Interactions between Asn95/96 in nsp9 and D291 in the NiRAN domain are indicated. PDB ID
384 7CYQ was used. **g.** Incorporation of 5'-pRNA^{LS10} into nsp9 and deRNAylation of nsp9-pRNA^{LS10}
385 by WT nsp12, the NiRAN mutant (D218A), the polymerase mutant (D760A), or the isolated
386 NiRAN domain (residues 1-326; Δ RdRP). Reaction products were analysed as in **Fig. 2b. h.**
387 Incorporation of 5'-pRNA^{LS10} into nsp9 (or the indicated mutants) by nsp12. Reaction products
388 were analysed as in **Fig. 2b. i.** Cartoon representation of the NiRAN active site. Catalytic residues
389 and GDP are shown as sticks, Mg²⁺ is a green sphere, and interactions are denoted by dashed lines.
390 **j.** Relative viral yields from WT or mutant SARS-CoV-2 viruses bearing indicated mutations in
391 nsp9 and nsp12. Data represent averages of two biological replicates. Error bars, SD. Results
392 shown in **g** and **h** are representative of at least 2 independent experiments.

393

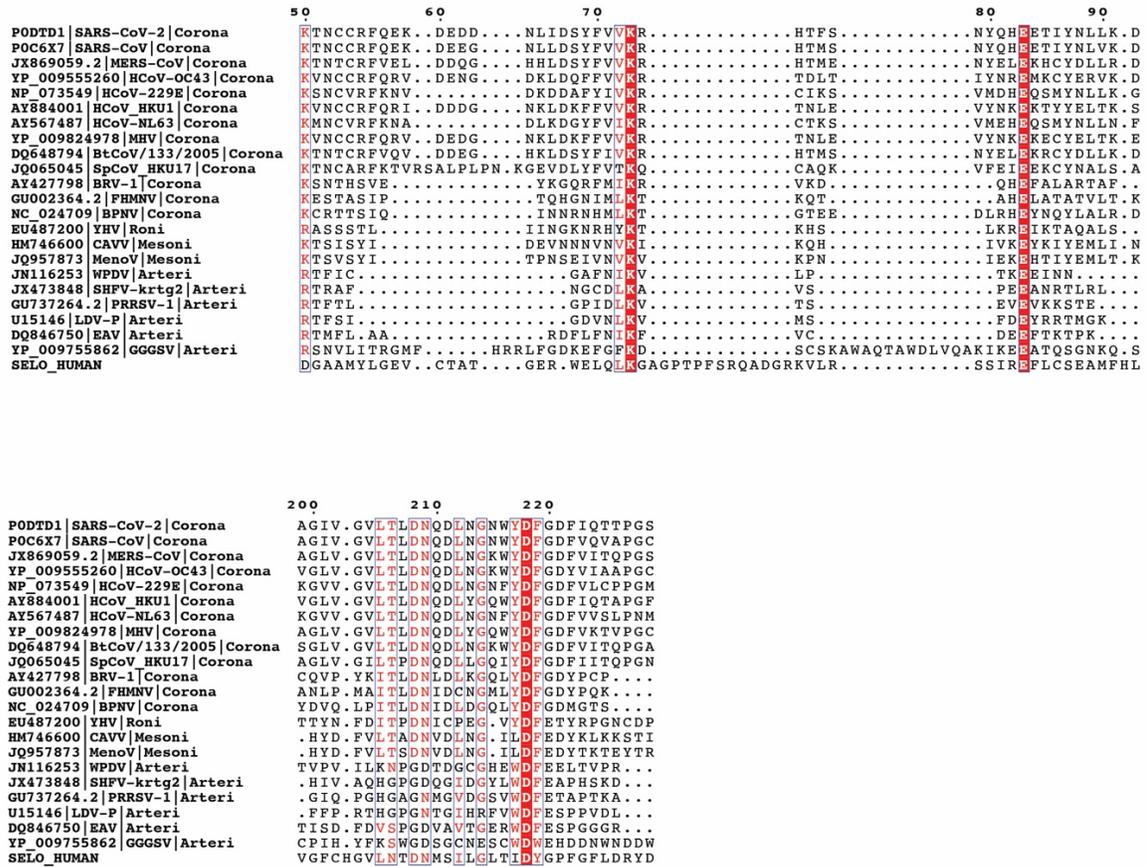
394



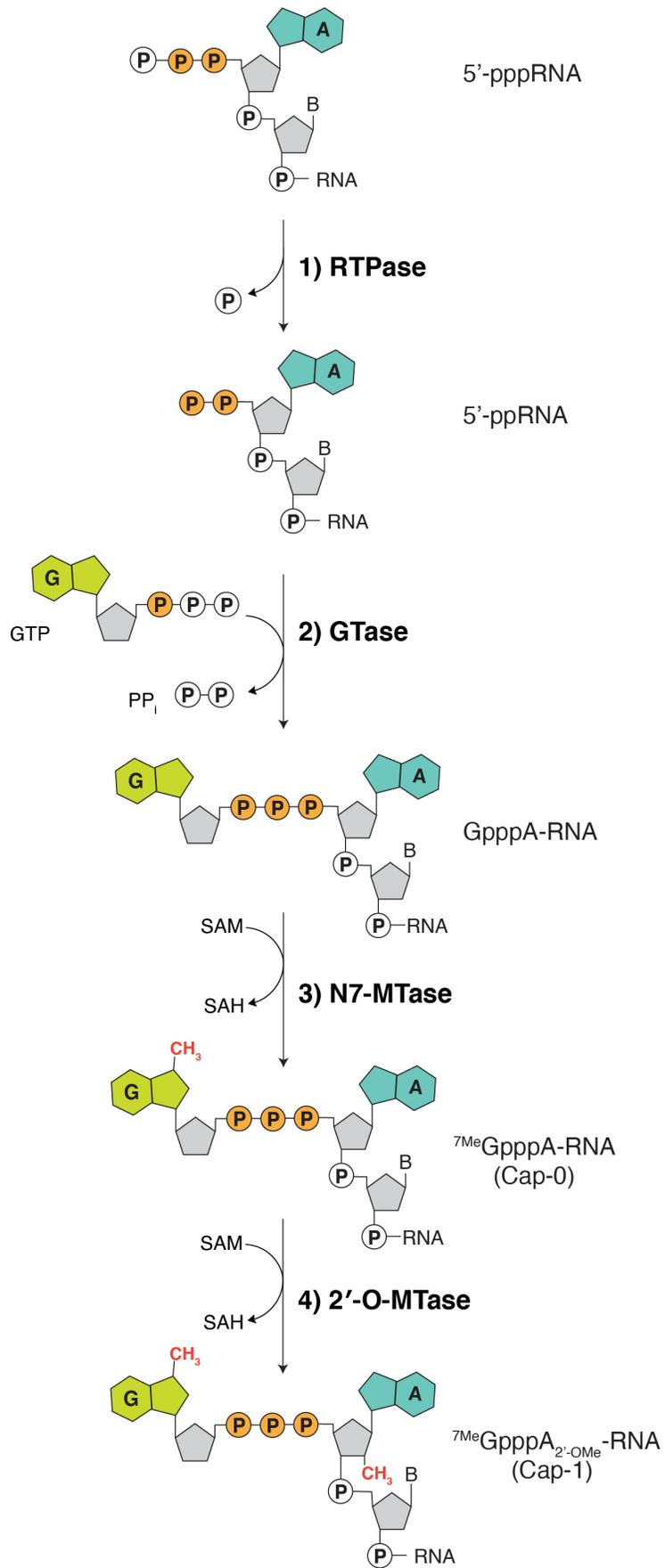
395

396

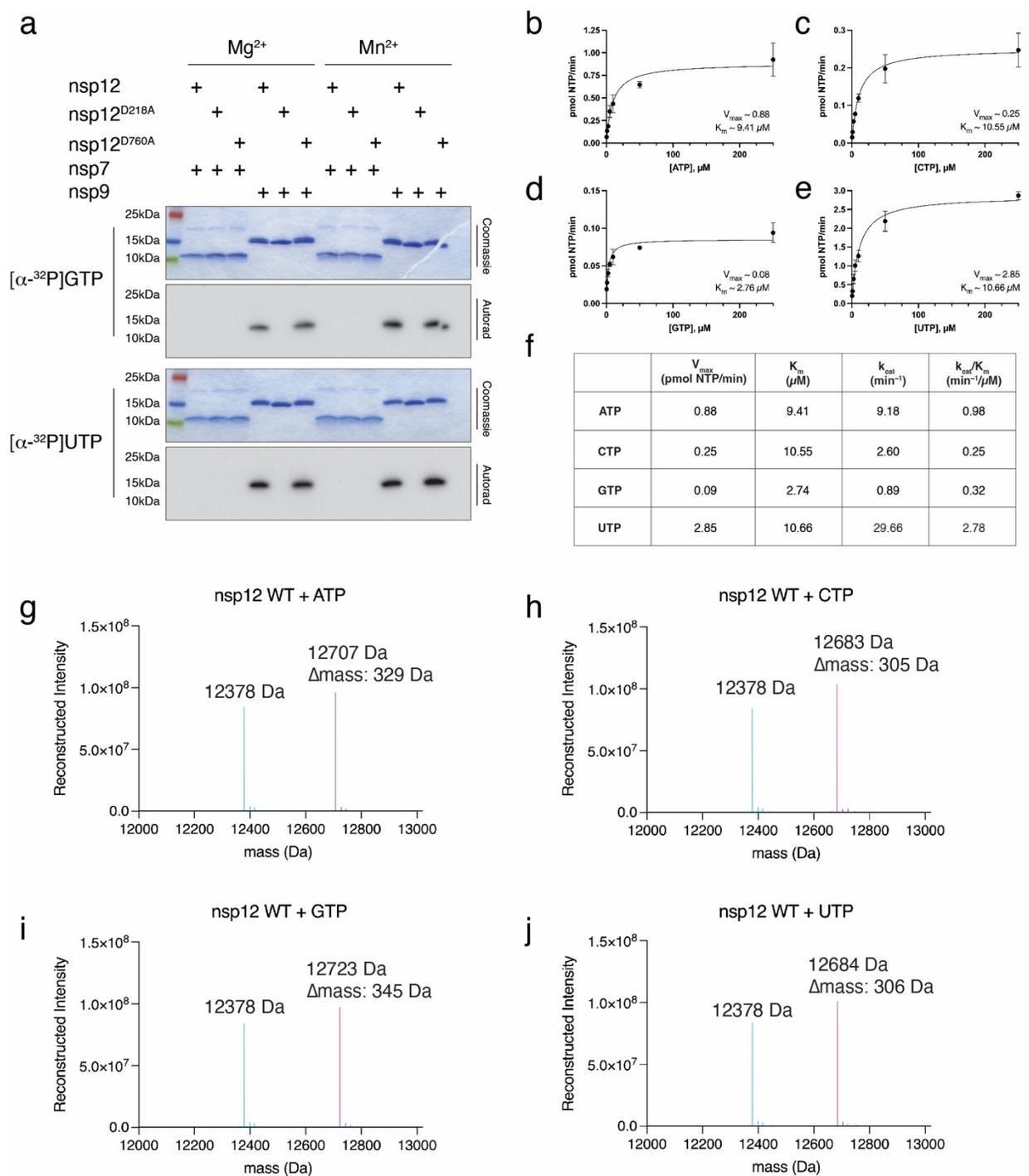
397 **Figure 6. Proposed model of the SARS-CoV-2 RNA capping mechanism** a, b. During
 398 transcription, the nascent 5'-pppRNA binds to the NiRAN active site in either a *cis* (a) or a *trans*
 399 (b) manner. c. Upon binding, the N-terminus of nsp9 attacks the α -phosphate of the nascent 5'-
 400 pppRNA (1), forming the covalent nsp9-pRNA species and releasing PP_i (2). Upon GDP binding
 401 to the NiRAN active site, the β -phosphate of the nsp13-generated GDP attacks the 5'-phosphate
 402 on the nsp9-pRNA (3), releasing capped RNA and regenerating unmodified nsp9 (4). Subsequent
 403 methylation events are carried out by nsp14 and nsp16 to generate the ⁷MeGpppA_{2'-O-Me}-RNA cap.



405
 406 **Extended Data Fig. 1. Sequence alignment of the NiRAN domain reveals similarity to the**
 407 **pseudokinase selenoprotein-O (SeLO).** Multiple sequence alignment highlighting conserved
 408 kinase-like active site residues in the NiRAN domain among several CoVs, other selected
 409 *Nidovirales* (Arteri-, Mesoni- and Roniviruses) and the human SeLO pseudokinase. Top: amino
 410 acid sequence surrounding the Lys-Glu ion pair. Bottom: and the amino acid sequence
 411 surrounding the active site and the “DFG” motif.



413 **Extended Data Fig. 2. The canonical eukaryotic mRNA capping mechanism.** The
414 $^7\text{MeGpppA}_{2'\text{OMe}}$ cap on eukaryotic RNA is formed co-transcriptionally by four enzymes: **1)** an RNA
415 triphosphatase (RTPase), which removes the γ -phosphate from the nascent 5'-triphosphorylated
416 RNA (5'-pppRNA) to yield a 5'-diphosphorylated RNA (5'-ppRNA); **2)** a guanylyltransferase
417 (GTase), which transfers the GMP moiety from GTP to the 5'-ppRNA to form the core cap
418 structure GpppN-RNA; **3)** a (guanine-N7)-methyltransferase (N7-MTase), which methylates the
419 cap guanine at the N7 position; and **4)** a (nucleoside-2'-O)-methyltransferase (2'-O-MTase), which
420 methylates the ribose-2'-OH position on the first nucleotide of the RNA. B denotes any base; GTP,
421 Guanosine triphosphate; GDP, Guanosine diphosphate; PP_i, pyrophosphate; SAM, S-Adenosyl
422 methionine.



424

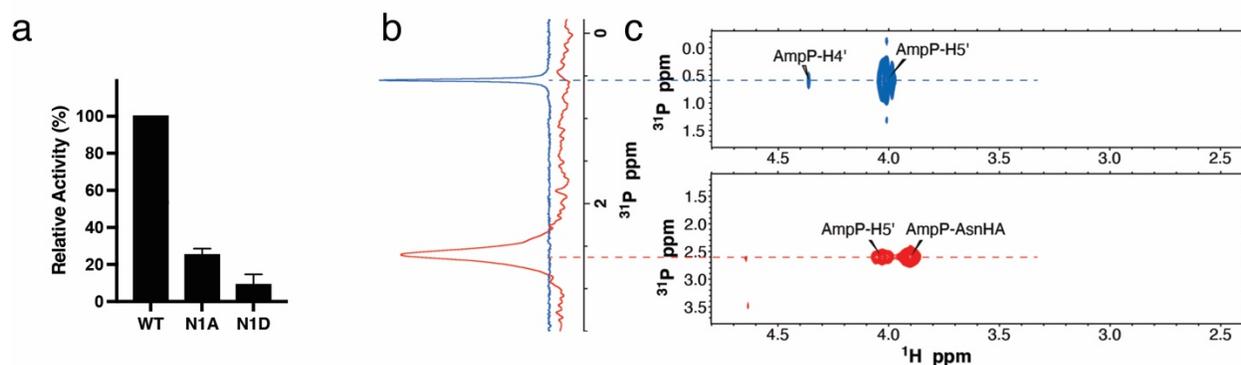
425

426 **Extended Data Fig. 3. The NiRAN domain NMPylates nsp9.** **a.** Incorporation of α -³²P from
 427 $[\alpha$ -³²P]GTP or $[\alpha$ -³²P]UTP into nsp7 or nsp9 by WT nsp12, the NiRAN mutant (K73A, D218A),

428 or the polymerase mutant (D760A). Reactions were performed in the presence of Mg^{2+} or Mn^{2+}
429 and the products were resolved by SDS-PAGE and visualized by Coomassie staining (top) and
430 autoradiography (bottom). **b-e**. Kinetic analysis depicting the concentration dependence of **(b)**
431 ATP, **(c)** CTP, **(d)** GTP, or **(e)** UTP on the rate of nsp9 NMPylation by the NiRAN domain. K_m
432 and V_{max} are indicated on the insets. Plots shown are the mean and SD of triplicate reactions. **f**.
433 Summary of K_m , V_{max} , k_{cat} , and k_{cat}/K_m values for each NTP. **g-j**. Intact mass LC/MS spectra of
434 unmodified nsp9 (*cyan*) overlaid with NMPylated nsp9 (*pink*) following incubation with WT
435 nsp12 and **(g)** ATP, **(h)** CTP, **(i)** GTP, or **(j)** UTP. The observed masses are shown in the insets.
436 The theoretical mass of unmodified nsp9 is 12378.2 Da and the theoretical increase in mass with
437 the addition of each NMP is as follows: AMP, 329 Da; CMP, 305 Da; GMP, 345 Da; UMP, 306
438 Da.

439

440



441

442 **Extended Data Fig. 4. The NiRAN domain NMPylates nsp9 on the N-terminus. a.**

443 Quantification of reaction products from **Fig. 1c** depicting the relative NiRAN-dependent

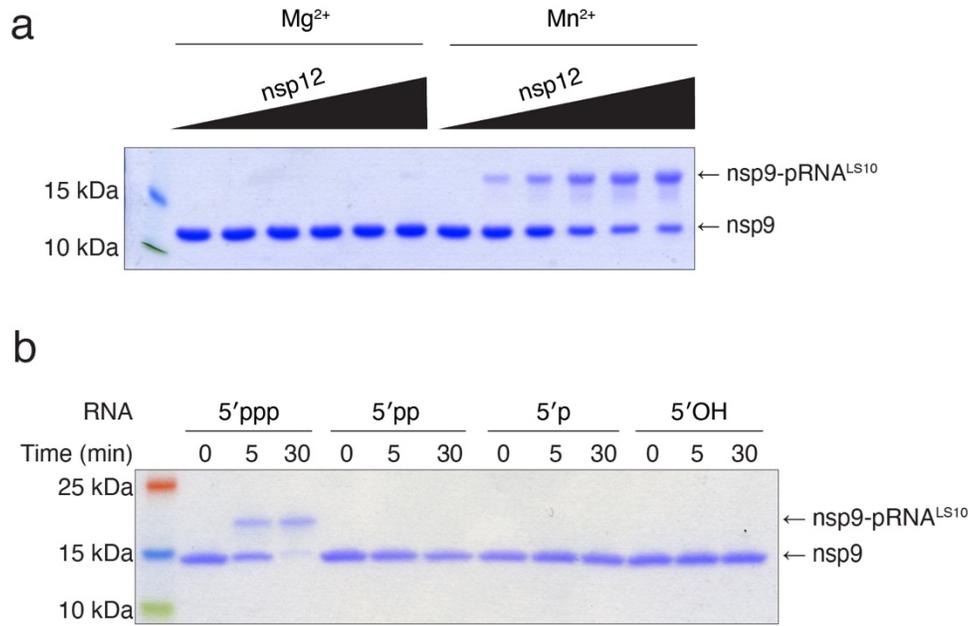
444 UMPylation activity towards nsp9 or the indicated mutants. Radioactive gel bands were excised

445 and quantified by scintillation counting. **b.** 1D ^{31}P spectrum of AMP-nsp9 (*red*) and AMP (*blue*)

446 recorded in the same buffer as reference. **c.** 2D ^1H , ^{31}P -HSQC spectra of AMP (*top, blue*) and

447 AMP-nsp9 (*bottom, red*).

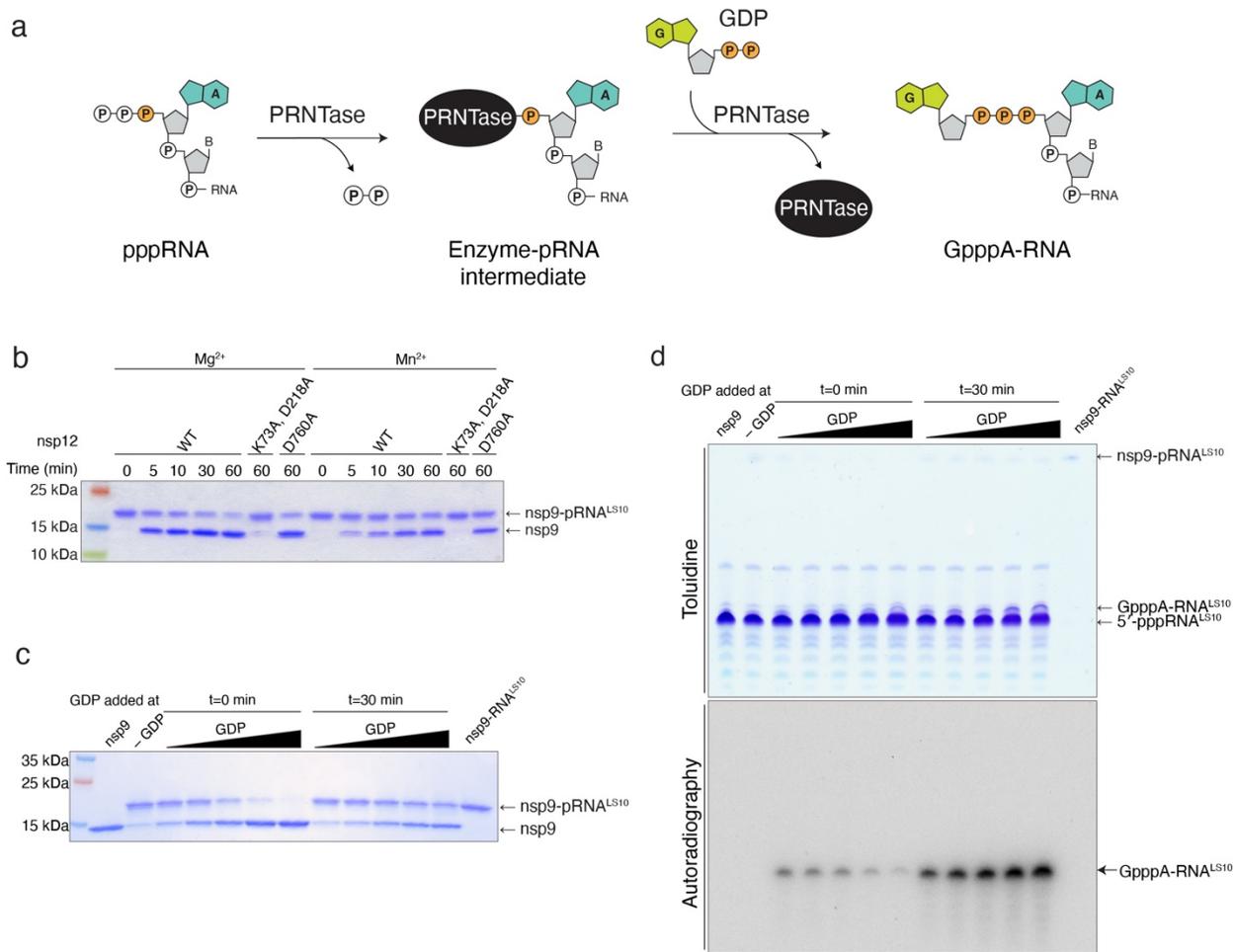
448



449

450 **Extended Data Fig. 5. Characterization of NiRAN RNAylation activity.** **a.** Incorporation of
 451 RNA into nsp9 by nsp12 (0-4 μ M) in the presence of Mg²⁺ or Mn²⁺. Reaction products were
 452 analysed as in **Fig. 2b.** **b.** Incorporation of RNA with the indicated 5' ends into nsp9 by nsp12.
 453 Reaction products were analysed as in **Fig. 2b.**

454



455

456 **Extended Data Fig. 6. Characterization of nsp12 NiRAN GDP-PRNTase activity.**

457 Schematic representation depicting the mechanism of GpppA-RNA formation by vesicular

458 stomatitis virus (VSV) polyribonucleotidyltransferase (PRNTase) enzyme. **b.** Time-dependent

459 deRNAylation of nsp9-pRNA^{LS10} by WT nsp12, the NiRAN mutant (K73A, D218A), or the

460 polymerase mutant (D760A) in the presence of GDP and either Mg²⁺ or Mn²⁺. Reaction products

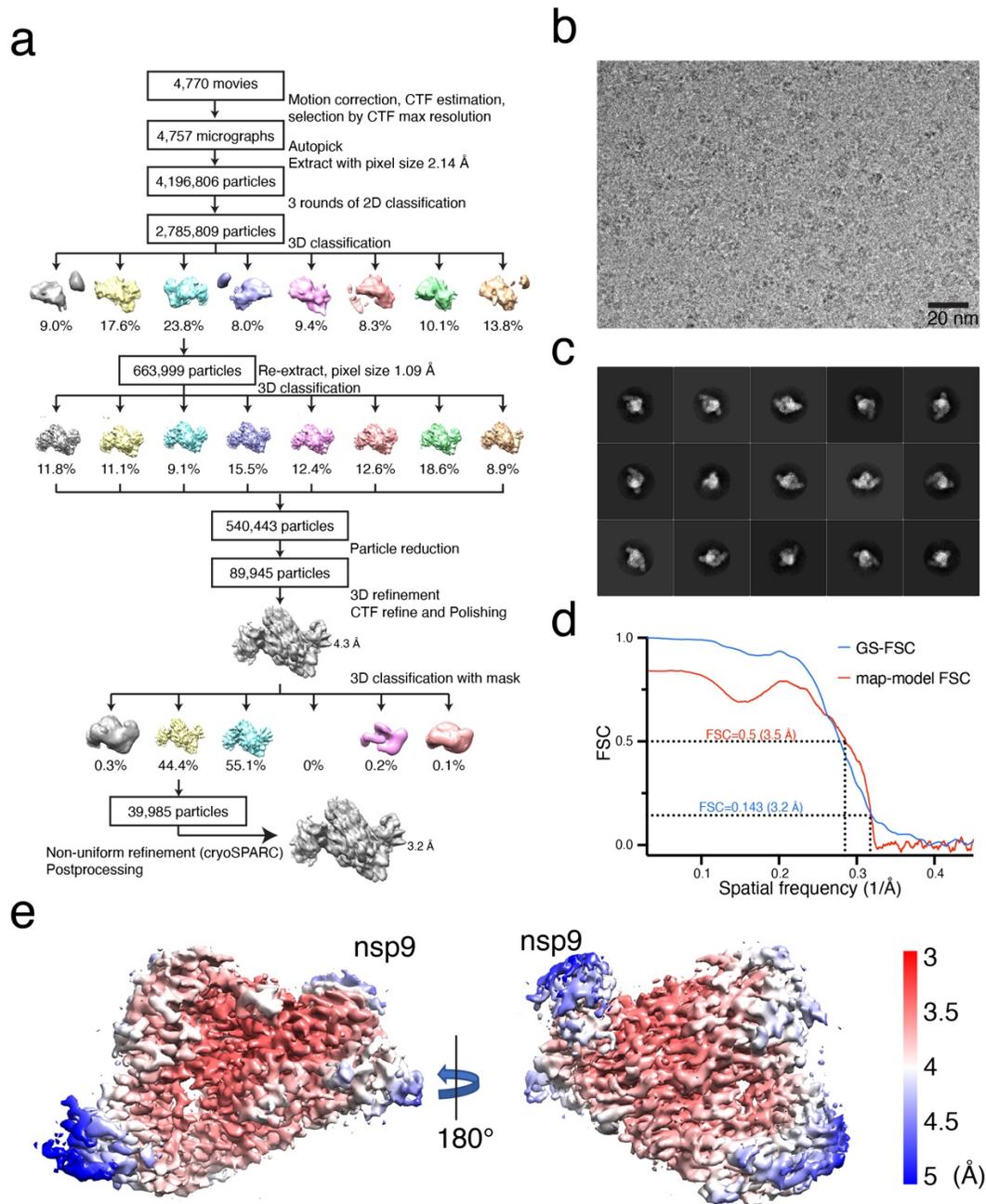
461 were analysed as in **Fig. 2b.** **c, d.** NiRAN-catalysed capping reactions depicting the inhibitory

462 effect of GDP on RNAylation. Nsp9 was incubated with excess 5'-pppRNA^{LS10} in presence of

463 nsp12 with no GDP (-GDP), or increasing concentrations (6.25-100 μM) of [³²P]GDP added either

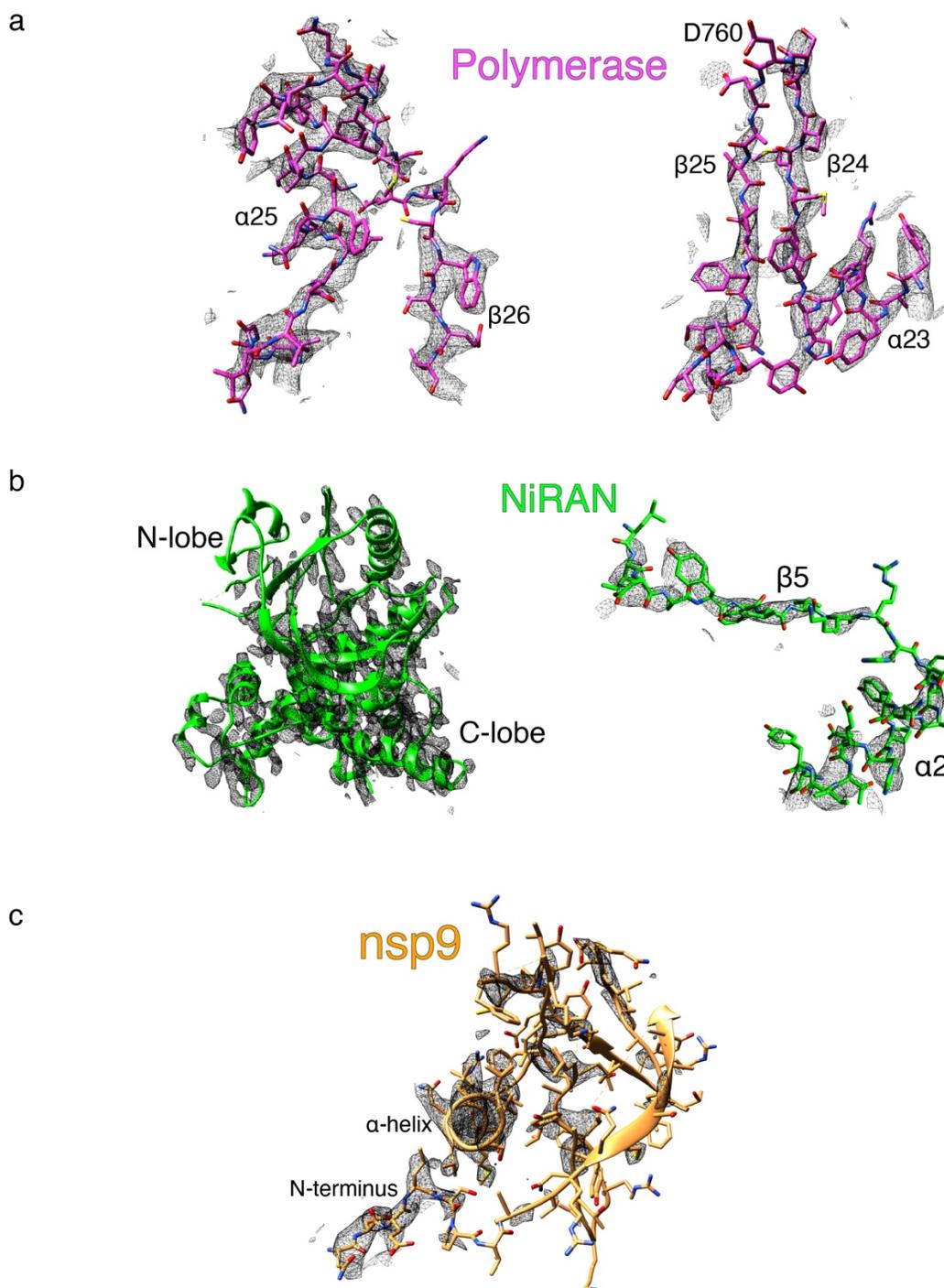
464 at time zero (t =0 min), or after the RNAylation reaction was allowed to proceed for 30 minutes

465 (t=30 min). Reaction products were analysed by SDS-PAGE (**c**), and Urea-PAGE (**d**).



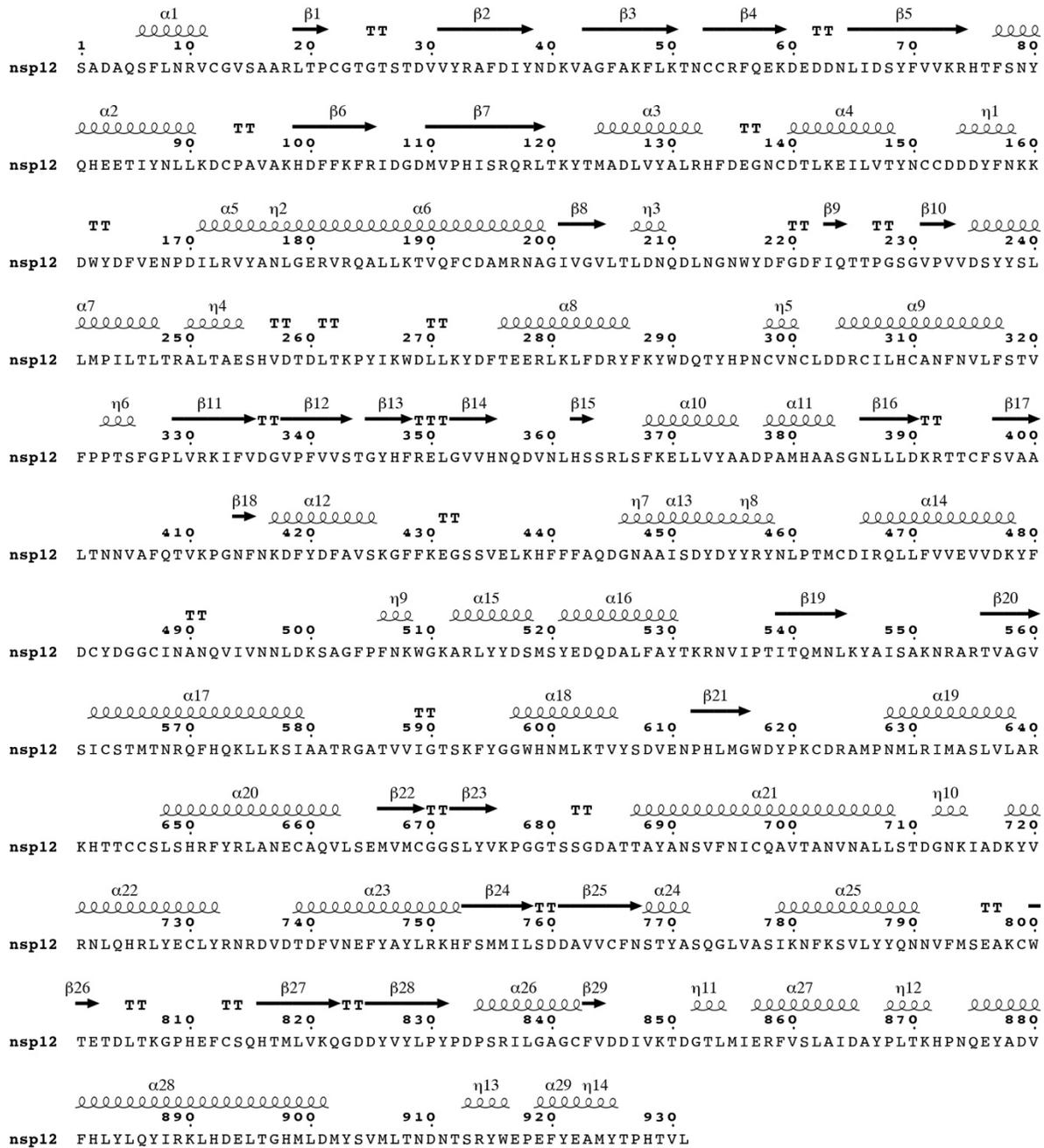
466

467 **Extended Data Fig 7. Cryo-EM analysis of the nsp7/8/9/12 complex. a.** Flow chart representing
 468 data processing for the nsp7/8/9/12 complex. **b.** A representative micrograph of the nsp7/8/9/12
 469 complex grids. **c.** Representative 2D classes generated by RELION 2D-classification. **d.** Gold-
 470 standard FSC curve (blue), and map-model FSC curve (red). Curves were generated by
 471 cryoSPARC and Phenix suite, respectively. **e.** Local resolution of the nsp7/8/9/12 complex
 472 calculated by RELION from final cryoSPARC half-maps. Position of nsp9 is indicated.



474

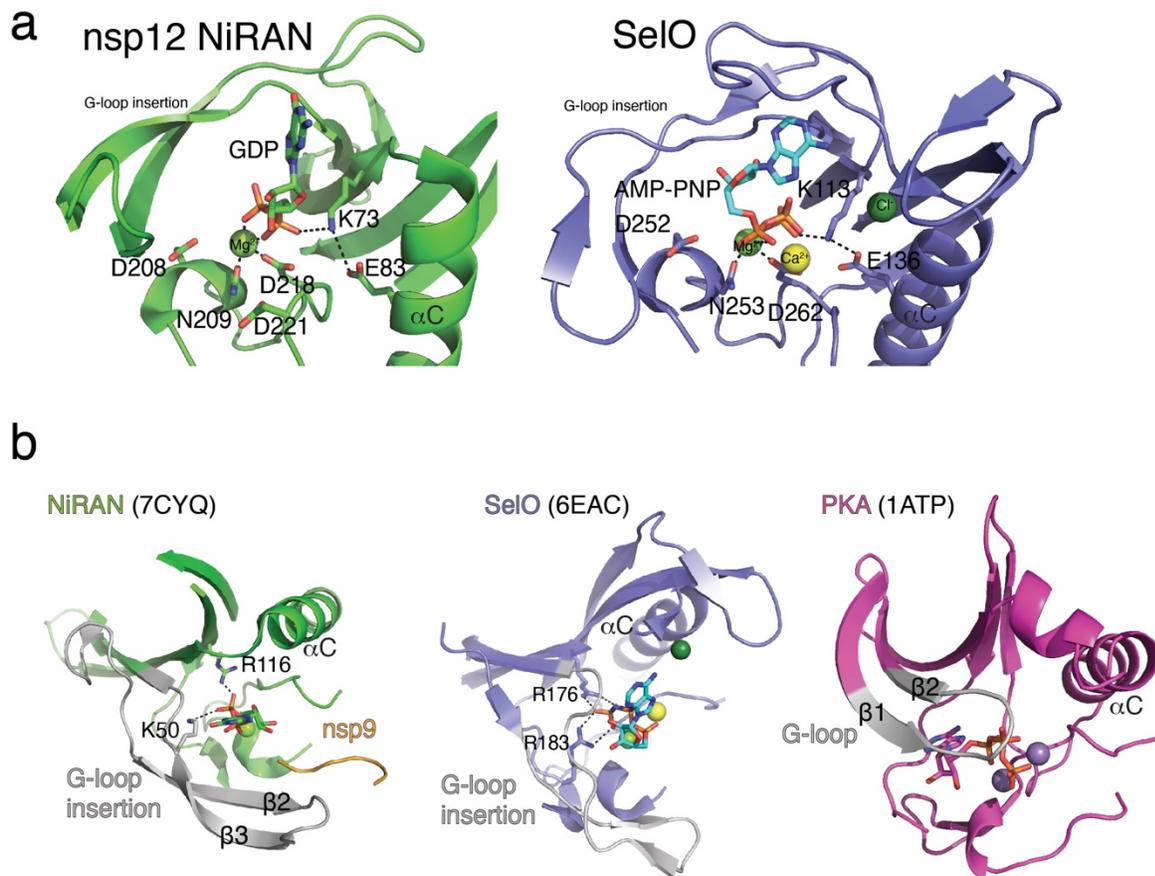
475 **Extended Data Fig 8. Exemplary cryo-EM density (black mesh) on (a) Polymerase (Magenta),**
 476 **(b) NiRAN (green) and (c) nsp9 (gold).** Note weaker density in the N-lobe of the kinase-like
 477 NiRAN domain (left panel of **b**, compare top to bottom of the image), and poor density in nsp9
 478 **(c)**, in areas not in direct contact with the NiRAN domain.



479

480 **Extended Data Fig. 9. Secondary structure of nsp12.** The secondary structural elements in
481 nsp12 (from PDB ID: 7CYQ) are shown.

482



483

484 **Extended Data Fig. 10. Comparison of the kinase-like domains of nsp12 and SelO. a.**

485 Cartoon representation comparing the NiRAN active site catalytic residues (green) to the active

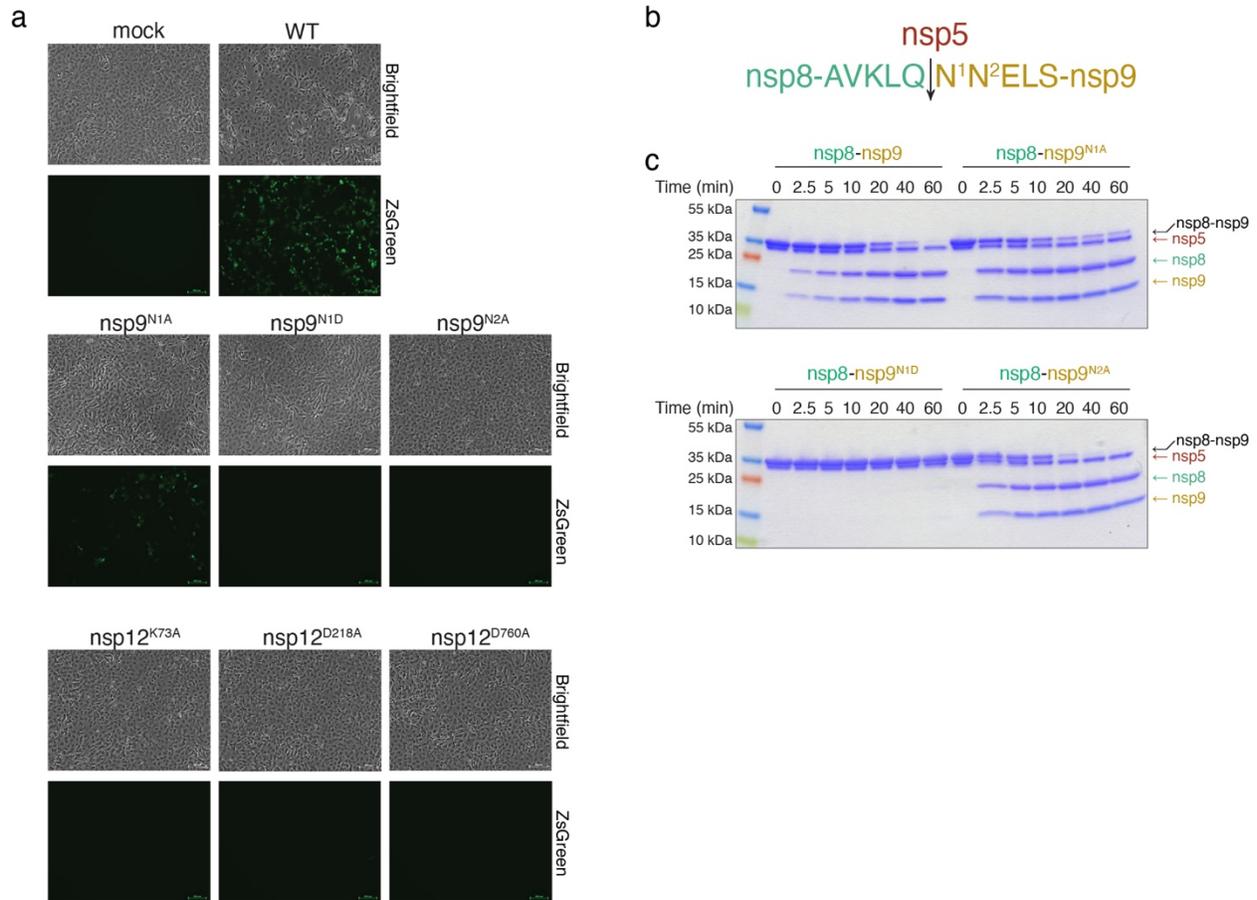
486 site residues in SelO (purple). The divalent cations are shown as spheres. **b.** Comparison of the

487 Gly-rich loop regions in NiRAN (PDB ID: 7CYQ, left), SelO (PDB ID: 6EAC), and the

488 canonical kinase PKA (PDB ID: 1ATP, right). Green sphere – Mg^{2+} , dark green sphere –

489 Chloride, yellow sphere – calcium, violet sphere – Mn^{2+} .

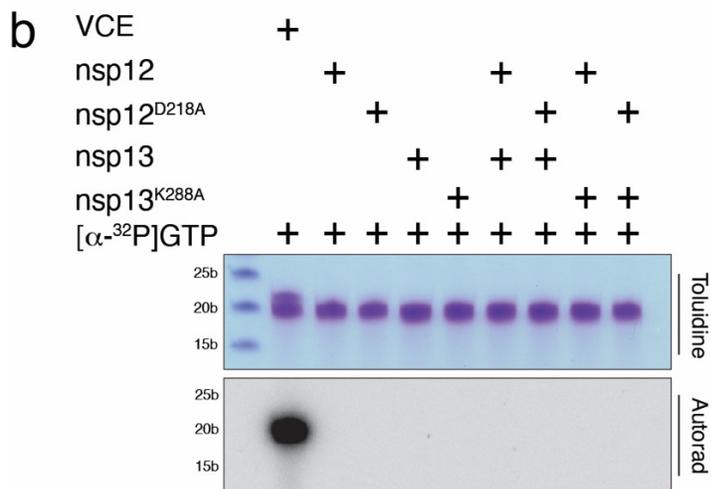
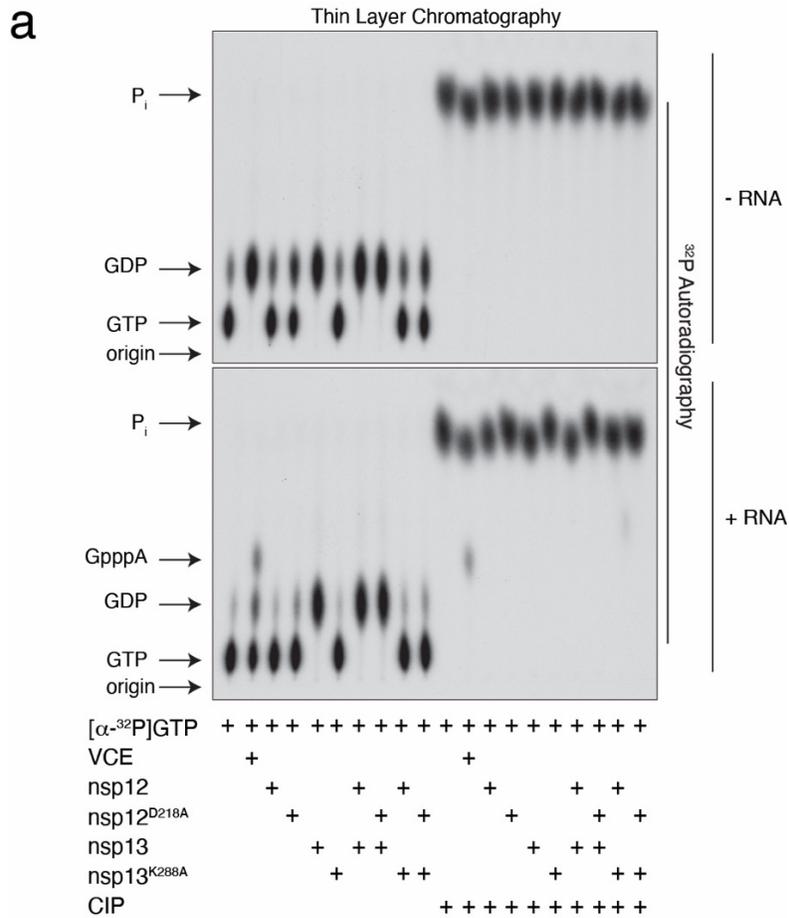
490



491

492 **Extended Data Fig. 11. Genetic insights into RNA capping by the NiRAN domain. a.**
 493 Microscopy images showing brightfield (upper) or fluorescence-based images (ZsGreen; lower)
 494 of SARS-CoV-2-ZsGreen production in VeroE6-C1008-TMPRSS2 cells. Mock-transfected
 495 panels were incubated with transfection reagents lacking DNA. The mutations engineered into
 496 either nsp9 or nsp12 are indicated above each set of images. Data represent one set of images
 497 from two independent biological replicates. Scale bars, 100 μ m. **b.** The amino acid sequence
 498 between nsp8 (green) and nsp9 (gold) depicting the cleavage site for the nsp5 (dark red)
 499 protease. N1 and N2 of nsp9 are highlighted. The arrow denotes the location of cleavage. **c.**
 500 Time-dependent proteolysis of the nsp8-nsp9 fusion protein by nsp5. Reaction products were
 501 separated by SDS PAGE and visualized by Coomassie staining.

502



503
504

505 **Extended Data Fig. 12. The NiRAN domain does not act as a GTase.** **a.** Thin-layer
 506 chromatograms depicting the reaction products resulting from the incubation of [α - ^{32}P]GTP with
 507 nsp12 or the inactive NiRAN mutant (D218A). Reactions were performed as described in ²¹ with

508 (*lower*) or without (*upper*) 5'-pppRNA^{A19C} and included nsp13 or the inactive mutant (K288A)
509 as indicated. Vaccinia capping enzyme (VCE) was used as a positive control. Reaction products
510 were digested with nuclease P1, then treated with or without calf intestinal alkaline phosphatase
511 (CIP) and analysed by PEI-cellulose thin-layer chromatography (TLC) followed by
512 autoradiography. The positions of the origin and standard marker compounds are indicated. **b.**
513 RNA products from (**a**) were analysed by TBE Urea-PAGE and visualized by toluidine blue O
514 staining (*upper*) and autoradiography (*lower*). Markers indicate RNA size by base length.

515 **Extended Data Table 1. RNAs used in this study**

RNA	Sequence
5'-pppRNA ^{A19C 21}	[ppp]ACCCCCCCCCCCCCCCCCCC
5'-pppRNA ^{LS10}	[ppp]AUUAAAGGUU
5'-pppRNA ^{LS2}	[ppp]AU
5'-pppRNA ^{LS3}	[ppp]AUU
5'-pppRNA ^{LS4}	[ppp]AUUA
5'-pppRNA ^{LS5}	[ppp]AUUAA
5'-pppRNA ^{LS6}	[ppp]AUUAAG
5'-pppRNA ^{LS20}	[ppp]AUUAAAGGUUUAUACCUUCC
5'-pppRNA ^{LS10_A1C}	[ppp]CUUAAAGGUU
5'-pppRNA ^{LS10_A1G}	[ppp]GUUAAAGGUU
5'-pppRNA ^{LS10_A1U}	[ppp]UUUAAAGGUU
5'-pppRNA ^{LS10_U2G}	[ppp]AGUAAAGGUU

516

517

518 **Extended Data Table 2. Data collection and refinement statistics.**

519

Magnification	81,000
Voltage (kV)	300
Electron exposure ($e^{-}\text{\AA}^{-2}$)	54
Defocus range (μm)	-1.0 to -2.5
Pixel size (\AA)	1.09
Symmetry imposed	C1
Initial particle images (no.)	4,196,806
Final particle images (no.)	39,985
Map resolution (\AA)	3.2
FSC threshold	0.143
Refinement	
Initial model used (PDB code)	-
Model resolution (\AA)	3.5
FSC threshold	0.5
Map sharpening B factor (\AA^{-2})	-37
Nonhydrogen atoms	9,215
Protein residues	1150
Ligands	4
B factors (\AA^{-2})	
Protein	67.4
Ligands	75.9
R.m.s. deviations	
Bond lengths (\AA)	0.002
Bond angles ($^{\circ}$)	0.453
MolProbity score	1.50
Clashscore	3.12
Poor rotamers (%)	0
Favored (%)	94.1
Allowed (%)	5.9
Disallowed (%)	0

520

521

522

523
524

Extended Data Table 3. Oligonucleotides used in the SARS-CoV-2 infection experiments

Oligo	Sequence	Use
PacI Forward	GGTTGAAGCAGTTAATTAAAGTTACACTTGTG	Fragment 1 and 3
1NtoA Reverse	GCAACAGGACTAAGCTCATTAGCCTGTAATTTGACAGC	Fragment 1
1NtoD Reverse	GCAACAGGACTAAGCTCATTGTCCTGTAATTTGACAGC	Fragment 1
2NtoA Reverse	GCAACAGGACTAAGCTCAGCATTCTGTAATTTGACAGC	Fragment 1
73KtoA Reverse	GAGAAAGTGTGTCTCGCAACTACAAAGTAAG	Fragment 1
MluI Reverse	CCTAAGTTGGCGTATACGCGTAATATATCTGGG	Fragment 2 and 3
1NtoA Forward	GCTGTCAAATTACAGGCTAATGAGCTTAGTCCTGTTGC	Fragment 2
1NtoD Forward	GCTGTCAAATTACAGGACAATGAGCTTAGTCCTGTTGC	Fragment 2
2NtoA Forward	GCTGTCAAATTACAGAATGCTGAGCTTAGTCCTGTTGC	Fragment 2
73KtoA Forward	CTTACTTTGTAGTTGCGAGACACACTTTCTC	Fragment 2

525
526

527 **Acknowledgments**

528
529 We thank members of the Tagliabracci laboratory for discussions, David Karlin for notifying us
530 of the similarity between SelO and the NiRAN domain, Brenden Park for help with kinetics,
531 Andrew Lemoff for help with intact mass analysis, and Jessica Kilgore, Noelle Williams, and the
532 resources available in the UTSW Preclinical Pharmacology Core for detection and quantitation
533 of GpppA. We thank Sam Wilson and Suzannah Rihn for the SARS-CoV-2 infectious clone and
534 for technical guidance and the Structural Biology Laboratory and the Cryo-Electron Microscopy
535 Facility at UT Southwestern Medical Center which are partially supported by grant RP170644
536 from the Cancer Prevention & Research Institute of Texas (CPRIT) for cryo-EM studies. A
537 portion of this research was supported by the W. M. Keck Foundation Medical Research Grant
538 (VST, KP, JS), the National Institutes of Health grant R01GM135189 (VST), 1DP1AI158124
539 (JWS), a Welch Foundation Grant I-1911 (VST), a Life Sciences Research Foundation
540 Fellowship (GH), a Polish National Agency for Scientific Exchange scholarship
541 PPN/BEK/2018/1/00431 (KP). J.W.S. is a Burroughs Wellcome Fund Investigator in the
542 Pathogenesis of Infectious Disease. V.S.T. is a Howard Hughes Medical Institute Investigator,
543 the Michael L. Rosenberg Scholar in Medical Research, a CPRIT Scholar (RR150033), and a
544 Searle Scholar.

545 **Author contributions**

546 G.J.P., A.O., G.H., J.W.S., and V.S.T. designed the experiments; G.J.P., A.O., G.H., J.L.E., A.M.,
547 and V.S.T. conducted experiments; G.J.P. discovered the RNylation and GDP-PRNTase activity
548 of the NiRAN domain; A.O. Z.C. and Y.L. performed the cryo-EM; A.O. performed methylation
549 experiments; G.H. performed GTase and nsp9 NMPylation experiments; A.M. performed intact
550 mass data analysis; M.T. and K.H.W. performed NMR experiments; K.P. performed the
551 bioinformatics; G.J.P., A.O., and G.H. purified proteins; G.J.P., A.O. and V.S.T. performed
552 cloning and site directed mutagenesis; J.L.E. performed infectious SARS-CoV-2 experiments; and
553 G.J.P., A.O., K.P., J.W.S., and V.S.T. wrote the manuscript with input from all authors.

554 **Competing interests**

555 The authors declare no competing interests. Data and materials availability:

556 **Materials & Correspondence**

557 Correspondence and material requests should be addressed to V.S.T.
558 The atomic coordinates have been deposited in the Protein Data Bank with accession code
559 7THM.
560

561 **Methods**

562 Chemicals and reagents

563 Ampicillin sodium (A9518), ATP (A2383), ADP (A2754), chloramphenicol (C0378), CTP
564 (C1506), CDP (C9755), dithiothreitol (DTT; D0632), EDTA (E5134), GTP (G8877), GDP
565 (G7127), imidazole (I2399), IPTG (I5502), kanamycin sulfate (K1377), 2-mercaptoethanol (BME,
566 M3148), Brilliant blue R (B0149), magnesium chloride (MgCl₂; M2670), manganese (II) chloride
567 tetrahydrate (MnCl₂; M3634), PEI-cellulose TLC plates (Z122882), potassium chloride (KCl;
568 P9541), pyrophosphate (221368), Urea (U6504), UTP (U6625), UDP (94330), were obtained from
569 MilliporeSigma (St. Louis, MO). Q5 DNA polymerase (M0492L), all restriction enzymes used
570 for cloning, Proteinase K (P8107S), Yeast Inorganic Pyrophosphatase (M2403), Quick CIP
571 (M0525S), Nuclease P1 (M0660S), Vaccinia Capping System (M2080S), mRNA Cap 2'-O-
572 Methyltransferase (M0366S), G(5')ppp(5')A RNA Cap Structure Analog (GpppA; S1406L), and
573 m7G(5')ppp(5')A RNA Cap Structure Analog (m7GpppA; S1405S) were all obtained from New
574 England Biolabs (Ipswich, MA). Acetic acid (A38-212), RNase inhibitor (N8080119), 2X TBE-
575 Urea Sample Buffer (LC6876), and isopropanol (42383) were all obtained from Thermo Fisher
576 Scientific (Waltham, MA). [α -³²P]-ATP (BLU003H250UC), [α -³²P]-CTP (BLU008H250UC), [α -
577 ³²P]-GTP (BLU006H250UC), [α -³²P]-UTP (BLU007H250UC), and S-[methyl-¹⁴C]-Adenosyl-L-
578 Methionine (NEC363010UC) were all obtained from PerkinElmer (Waltham, MA). All 5'-
579 triphosphorylated RNAs were custom synthesized by ChemGenes Corporation (Wilmington,
580 MA). Phenylmethylsulfonyl fluoride (PMSF; 97064-898) was obtained from VWR (Radnor, PA).

581 4–20% Mini-PROTEAN® TGX Stain-Free™ Protein Gels (4568096) were obtained from Bio-
582 Rad. Uridine-5'-[(α,β)-imido]triphosphate (UMP-NPP; NU-930L) was obtained from Sapphire
583 North America (Ann Arbor, MI).

584 Plasmids

585 SARS-CoV-2 nsp7, nsp8, nsp12, nsp13, nsp14, and nsp16 coding sequences (CDS) were codon-
586 optimized for bacterial expression and synthesized as gBlocks (Integrative DNA Technologies,
587 Coralville, IA). The CDS for nsp9 and nsp10 were amplified from mammalian expression vectors
588 (a generous gift Nevan Krogen)³⁵. The CDS were cloned into modified pET28a bacterial
589 expression vectors containing N-terminal 6/8/10xHis tags followed by the yeast Sumo (smt3) CDS
590 Amino acid mutations were introduced via QuikChange site-directed mutagenesis. Briefly,

591 primers were designed using the Agilent QuikChange primer design program to generate the
592 desired mutation and used in PCR reactions with PfuTurbo DNA polymerase. Reaction products
593 were digested with Dpn1, transformed in DH5 α cells and mutations were confirmed by Sanger
594 sequencing.

595 For protein expression in *Escherichia coli* (*E. coli*), ppSumo-SARS-CoV-2 nsps and mutants were
596 cloned into a BamHI site at the 5' end, which introduced a Ser residue following the diGly motif
597 in smt3. To make native N-termini, the codon encoding the Ser was deleted via QuickChange
598 mutagenesis. Thus, following cleavage with the ULP protease (after the diGly motif), the proteins
599 contained native N-termini.

600 pGEX-2T-GST-eIF4E K119A³⁶ was obtained from Addgene (plasmid # 112818)

601 Protein purification

602 *Nsp5, nsp7, nsp8, nsp10, nsp14 and nsp16*

603 6xHis-Sumo-nsp5/7/8/10/14/16 and corresponding mutant plasmids (with native N-termini
604 following the diGly motif in Sumo) were transformed into Rosetta (DE3) *E. coli* or LOBSTR-
605 BL21(DE3)-RIL cells under 50 μ g/ml kanamycin exposure. 5 ml LB Miller growth medium starter
606 cultures containing 50 μ g/ml kanamycin and 34 μ g/ml chloramphenicol were grown for 2-4 h at
607 37°C and then transferred to growth medium containing the same antibiotics. Typically, 2 L were
608 grown for each protein. Protein expression was induced at O.D. 0.7-1.0 by adding 0.4 mM IPTG
609 and overnight incubation (16 hours) at 18°C. Cultures were centrifuged at 3,000 x g for 10 min
610 and the bacterial pellet was resuspended in lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, 17.4
611 μ g/ml PMSF, 15 mM imidazole pH 8.0 and 5 mM β -ME) and lysed by sonication. Lysates were
612 centrifuged for 30 min at 30,000-35,000 x g and the supernatants incubated with Ni-NTA resin for
613 1-2 h at 4°C. The Ni-NTA resin was washed with 50 mM Tris pH 8.0, 300 mM NaCl, 30 mM
614 imidazole pH 8.0, 1 mM DTT and the protein was eluted in 50 mM Tris pH 8.0, 300 mM NaCl,
615 300 mM imidazole pH 8.0 and 1 mM DTT. The eluted proteins were incubated with 5 μ g/ml Ulp1
616 protease overnight at 4°C. Nsp7 was separated from 6xHis-Sumo by anion exchange (Capto HiRes
617 Q 5/50 column (Cytiva) equilibrated in 50 mM Tris 8.0, 50 mM NaCl, 1 mM DTT, eluted with 0-
618 50% gradient of buffer containing 1 M NaCl). 6xHis-SUMO-nsp8 was treated with Ulp1 on the
619 Ni-NTA resin, to separate nsp8 and 6xHis-SUMO in buffer with no imidazole. Proteins were

620 further purified by size exclusion chromatography using Superdex 200 10/300 increase, Superdex
621 200 16/600, Superdex 75 10/300 increase or Superdex 75 16/600 columns in 50 mM Tris pH 7.5-
622 8.0, 150-300 mM NaCl, 1 mM DTT, depending on yield and size. Fractions containing proteins
623 of interest were pooled, concentrated in an Amicon Ultra-15 with a 3-50 kDa pore size centrifugal
624 filters. The nsp16 protein was incubated with Ni-NTA resin post SEC to remove 6His-SUMO.
625 Purified proteins were aliquoted and stored at -80°C.

626 *Nsp9*

627 6xHis-Sumo-nsp9 and respective mutant plasmids (with native N-termini following the diGly
628 motif in Sumo) were transformed into Rosetta (DE3) *E. coli* cells under 50 µg/ml kanamycin
629 exposure. 5 ml LB Miller growth medium starter cultures containing 50 µg/ml kanamycin and 34
630 µg/ml chloramphenicol were grown for 2-4 h at 37°C and then transferred to Terrific Broth (TB)
631 growth medium containing the same antibiotics and several drops of Antifoam B emulsion (Sigma,
632 A5757). Protein expression was induced at O.D 1.2 by adding 0.4 mM IPTG and overnight
633 incubation (16 hours) at 18°C. Cultures were centrifuged at 3,000 x g for 10 min and the bacterial
634 pellet was resuspended in lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, 10% glycerol 17.4
635 µg/ml PMSF, 15 mM imidazole pH 8.0 and 5 mM β-ME) and lysed by sonication. Lysates were
636 centrifuged for 30 min at 30,000 x g and the supernatants incubated in Ni-NTA resin for 1-2 h at
637 4°C. The Ni-NTA resin was washed with 50 mM Tris pH 8.0, 300 mM NaCl, 10% glycerol, 30
638 mM imidazole pH 8.0, 1 mM DTT and each protein was eluted in 50 mM Tris pH 8.0, 300 mM
639 NaCl, 10% glycerol, 300 mM imidazole pH 8.0 and 1 mM DTT. The eluted protein was incubated
640 with 5 µg/ml Ulp1 overnight at 4°C. Proteins were diluted with 50 mM Tris pH 8.0, 10% glycerol,
641 1 mM DTT buffer to lower the NaCl concentration to 30 mM and subsequently ran through a Hi-
642 Trap CptoQ column where the flowthrough contained purified nsp9. NaCl was added to each
643 protein to a final concentration of 150 mM, concentrated in an Amicon Ultra-15 with a 10k
644 MWCO, aliquoted, and stored at -80°C.

645 *Nsp12*

646 8xHis or 10xHis-Sumo-nsp12 and respective mutant plasmids (with native N-termini following
647 the diGly motif in Sumo) were transformed into LOBSTR-BL21(DE3)-RIL *E. coli* cells under 50
648 µg/ml kanamycin exposure. 5 ml LB Miller growth medium starter cultures containing 50 µg/ml
649 kanamycin and 34 µg/ml chloramphenicol were grown for 2-4 h at 37°C and then transferred to

650 1L growth medium containing the same antibiotics. Protein expression was induced at O.D. 0.8-
651 1.2 by adding 0.4 mM IPTG and overnight incubation (16 hours) at 18°C. Cultures were
652 centrifuged at 3,000-3.500 x g for 10 min and the bacterial pellet was resuspended in lysis buffer
653 (50 mM Tris pH 8.0, 300 mM NaCl, 10% glycerol, 17.4 µg/ml PMSF, 15 mM imidazole pH 8.0
654 and 5 mM β-ME) and lysed by sonication. Lysates were centrifuged for 30 min at 30,000-35,000
655 x g and the supernatants incubated in Ni-NTA resin for 1-2 h at 4°C. The Ni-NTA resin was
656 washed with high salt buffer: 50 mM Tris pH 8.0, 1 M NaCl, 10% glycerol, 30 mM Imidazole and
657 5 mM β-ME, followed by a high imidazole wash: 50 mM Tris pH 8.0, 300 mM NaCl, 10%
658 glycerol, 75 mM Imidazole and 5 mM β-ME, and the protein was eluted in 50 mM Tris pH 8.0,
659 300 mM NaCl, 10% glycerol, 300 mM imidazole pH 8.0 and 1 mM DTT. The eluted proteins were
660 incubated with 5 µg/ml Ulp1 overnight at 4°C. Proteins were further purified by size exclusion
661 chromatography using a Superdex 200 10/300 increase column, or Superdex 200 16/600 in 50 mM
662 Tris pH 8.0, 150-300 mM NaCl, 1 mM DTT. Fractions containing nsp12 were pooled,
663 concentrated in an Amicon Ultra-15 with a 30-50k MWCO centrifugal filter, aliquoted, and stored
664 at -80°C.

665 *Nsp13*

666 6xHis-Sumo-nsp13 (used in **Extended Data Fig. 12**), or 10xHis-Sumo-nsp13 (used for [α^{32} P]-
667 GTP conversion into GDP) and respective mutant plasmids (with native N-termini following the
668 diGly motif in Sumo) were transformed into Rosetta (DE3) *E. coli* cells under 50 µg/ml kanamycin
669 exposure. 5 ml LB Miller growth medium starter cultures containing 50 µg/ml kanamycin and 34
670 µg/ml chloramphenicol were grown for 2-4 h at 37°C and then transferred to 1 L of growth medium
671 containing the same antibiotics. Protein expression was induced at O.D. 1.0 by adding 0.4 mM
672 IPTG and overnight incubation (16 hours) at 18°C. Cultures were centrifuged at 3,000-3.500 x g
673 for 10 min and the bacterial pellet was resuspended in lysis buffer (50 mM Tris pH 8.0, 300 mM
674 NaCl, 10% glycerol, 17.4 µg/ml PMSF, 15 mM imidazole pH 8.0 and 5 mM β-ME) and lysed by
675 sonication. Lysates were centrifuged for 30 min at 30,000-35,000 x g and the supernatants
676 incubated in Ni-NTA resin for 1-2 h at 4°C. Ni-NTA resin for 6xHis-Sumo-nsp13 was washed
677 with 50 mM Tris pH 8.0, 300 mM NaCl, 10% glycerol, 30 mM imidazole pH 8.0 and 1 mM DTT.
678 10xHis-Sumo-nsp13 with high salt buffer: 50 mM Tris pH 8.0, 1 M NaCl, 10% glycerol, 30 mM
679 Imidazole and 5 mM β-ME, followed by a high imidazole wash: 50 mM Tris pH 8.0, 300 mM

680 NaCl, 10% glycerol, 75 mM imidazole and 5 mM β -ME, and the protein was eluted in 50 mM Tris
681 pH 8.0, 300 mM NaCl, 10% glycerol, 300 mM imidazole pH 8.0 and 1 mM DTT. The eluted
682 protein was incubated with 5 μ g/ml Ulp1 overnight at 4°C. Proteins were buffer-exchanged or
683 dialysed into a buffer containing 50 mM Bis-Tris pH 6.0, 30 mM NaCl, 10% glycerol and 1 mM
684 DTT followed by ion-exchange chromatography in a 5/50 MonoS column. Fractions containing
685 nsp13 were pooled and further purified by size exclusion chromatography using a Superdex 200
686 10/300 increase or Superdex 200 16/600 column in 50 mM Tris pH 8.0, 150 mM NaCl, 10%
687 glycerol (only in 6xHis), 1 mM DTT. Fractions containing nsp13 were pooled, concentrated in an
688 Amicon Ultra-15 with a 30-50k MWCO, aliquoted, and stored at -80°C.

689 *eIF4E*

690 For production of GST-eIF4E K119A, LOBSTR-BL21(DE3)-RIL cells were transformed with
691 pGEX-2T-GST-eIF4E K119A³⁶ and were grown in LB supplemented with 100 μ g/L Ampicillin,
692 34 μ g/L chloramphenicol. Protein expression was induced at O.D. 1.0 by adding 0.4 mM IPTG
693 and overnight incubation (16 hours) at 18 °C. Cultures were centrifuged at 3,000-3.500 x g for 10
694 min and the bacterial pellet was resuspended in lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl,
695 17.4 μ g/ml PMSF, and 5 mM β -ME) and lysed by sonication. Lysates were centrifuged for 30 min
696 at 35,000 x g and the supernatants incubated with Pierce Glutathione resin for 1-2 h at 4 °C. The
697 resin was washed with lysis buffer, and the GST-eIF4E K119A eluted with 50 mM Tris-HCl pH
698 8.0, 300 mM NaCl, 50 mM glutathione, 1 mM DTT. The protein was purified over Size Exclusion
699 Chromatography on Superdex 200 16/600 in 50mM Tris-HCl pH 8.0, 300 mM NaCl, 1 mM DTT,
700 concentrated, and stored as above.

701 *Ipp1*

702 For production of yeast inorganic pyrophosphatase (*ipp1*), The *S. cerevisiae* *ipp1* CDS was cloned
703 into pProEx2 containing a N-terminal 6xHis-TEV linker and was transformed into Rosetta *E. coli*
704 cells under 100 μ g/ml ampicillin exposure. 5 ml LB Miller growth medium starter cultures
705 containing 100 μ g/ml ampicillin were grown for 2-4 h at 37°C and then transferred to 1L growth
706 medium containing the same antibiotics. Protein expression was induced at O.D. 0.8-1.2 by adding
707 0.4 mM IPTG and overnight incubation (16 hours) at 18°C. Cultures were centrifuged at 3,000-
708 3.500 x g for 10 min and the bacterial pellet was resuspended in lysis buffer (50 mM Tris pH 8.0,
709 300 mM NaCl, 17.4 μ g/ml PMSF, 15 mM imidazole pH 8.0 and 5 mM β -ME) and lysed by

710 sonication. Lysates were centrifuged for 30 min at 35,000 x g and the supernatants incubated in
711 Ni-NTA resin for 1 h at 4°C. Ni-NTA resin was washed with high salt buffer: 50 mM Tris pH 8.0,
712 1 M NaCl, 30 mM imidazole and 5 mM β-ME, and was eluted in 50 mM Tris pH 8.0, 50 mM
713 NaCl, 300 mM imidazole pH 8.0 and 1 mM DTT. The eluted protein was loaded onto a Capto
714 HiRes Q 5/50 column (Cytiva) equilibrated in 50 mM Tris 8.0, 50 mM NaCl, 1 mM DTT, eluted
715 with 0-50% gradient of buffer containing 1 M NaCl. Protein was further purified by size exclusion
716 chromatography using a Superdex 200 16/600 in 25 mM Tris pH 7.5, 50 mM NaCl, 2 mM DTT.
717 Fractions containing YIPP were pooled, concentrated, and stored as above.

718 *Nsp8-nsp9 fusion*

719 The 6xHis-Sumo-nsp8-nsp9 plasmid and mutants (N1A, N1D and N2A) were transformed into
720 Rosetta (DE3) *E. coli*. Cells were grown in Terrific broth media in the presence of 50 µg/ml
721 kanamycin and 25 µg/ml chloramphenicol to OD 1.0 and induced with 0.4 mM IPTG for 16 hours
722 at 18 °C. Cultures were centrifuged at 3500 x g for 15 minutes, and the pellets resuspended in
723 lysis buffer (50 mM Tris, pH 8.0; 500 mM NaCl; 25 mM imidazole; 10% glycerol) in the presence
724 of 1 mM PMSF. Cells were lysed by sonication and lysates cleared by centrifugation at 25000 x g
725 for 1 hour. The lysate was passed over Ni-NTA beads, which were washed with lysis buffer.
726 Protein samples were eluted with elution buffer (50 mM Tris, pH 8.0; 300 mM NaCl; 300 mM
727 imidazole; 5% glycerol) and cleaved overnight at 4 °C with Ulp Sumo protease. Protein samples
728 were further purified into cleavage assay buffer (50 mM Tris, pH 7.4; 150 mM NaCl; 5% glycerol)
729 by size exclusion chromatography using a Superdex 75 Increase 10/300 GL column.

730 Guanylyltransferase (GTase) activity assays

731 GTase activity assays were performed as described in Yan et al. ²¹. Reactions were assembled in
732 20 µL containing 50 mM Tris pH 8.0, 5 mM KCl, 1 mM DTT, 0.005 U/ml inorganic
733 pyrophosphatase, 10 uM 5'-pppACCCCCCCCCCCCCCCCCC-3' (5'-pppRNA^{A19C}), 1.25 mM
734 RNase inhibitor, and where indicated, 0.5 µM nsp12, nsp12^{D218A}, nsp13, nsp13^{K288A} or 1 U/ml of
735 vaccinia capping enzyme (VCE). Reactions were started with 1 mM MgCl₂, 100 µM [α-³²P] GTP,
736 (specific radioactivity = 1000 cpm/pmol) and incubated for 1 hr at 37°C. Half of the reaction was
737 stopped by the addition of 0.8 U/ml proteinase K and incubated for 30 min at 37°C prior to the
738 addition of 2X RNA loading dye (Novex) and incubated for 3 mins at 95°C. Reaction products

739 were resolved in a 15% TBE-Urea PAGE gel. The gel was then stained with toluidine blue O and
740 the ^{32}P signal detected via autoradiography.

741 The other half of the GTase reactions were treated with 10 U/ml P1 nuclease for 1 h at 37°C.
742 Reactions were then split in half again with one half treated with 1 U/ml Quick CIP for 30 min at
743 37°C. Reactions were spotted on a PEI cellulose thin-layer chromatography (TLC) plate and
744 developed in a 0.4 M ammonium sulfate $(\text{NH}_4)_2\text{SO}_4$ solvent system. The plate was dried and the
745 ^{32}P signal was detected via autoradiography.

746 NMPylation assays

747 NMPylation reactions were carried out in 20 μL containing 50 mM Tris (pH 7.5), 5 mM KCl, 1
748 mM DTT, 16 μM nsp7, nsp8 or nsp9 (and mutants) and 4.8 nM nsp12 (and mutants). Reactions
749 were started with 1 mM MgCl_2 or MnCl_2 , 200 μM [α - ^{32}P] ATP, [α - ^{32}P] UTP, [α - ^{32}P] GTP, or [α -
750 ^{32}P] CTP (specific radioactivity = 1000 cpm/pmol). The reactions were incubated at 37°C for 5
751 minutes and stopped by adding 2 μL of 500 mM EDTA, followed by addition of 5X SDS-PAGE
752 sample buffer with 10% β -ME and incubated for 3 minutes at 95°C. Reaction products were
753 resolved by SDS-PAGE on a 4-20% gradient gel and visualized by staining with Coomassie
754 Brilliant Blue. The ^{32}P signal was detected via autoradiography and scintillation counting.

755 Nsp9 NMPylation kinetics

756 NMPylation reactions were carried out in a 20 μL reaction containing 50 mM Tris (pH 7.5), 5 mM
757 KCl, 1 mM DTT, 16 μM nsp9, and 4.8 nM nsp12. Reactions were started by adding MnCl_2 and
758 [α - ^{32}P] ATP, CTP, GTP, or UTP as indicated. The final concentration in the reaction was 0.5 to
759 200 μM (specific radioactivity = \sim 5000 cpm/pmol) of the indicated nucleotide triphosphate and 1
760 mM MnCl_2 . The reactions were incubated at 37°C for 5 minutes and stopped by adding 2 μL of
761 500 mM EDTA, followed by addition of 5X SDS-PAGE sample buffer + β -ME and boiling for 2-
762 5 minutes. Reaction products were resolved by SDS-PAGE on a 4-20% gradient gel and visualized
763 by staining with Coomassie Brilliant Blue. Incorporation of ^{32}P was quantified by excising the
764 nsp9 bands from the gel and scintillation counting. Background radioactivity was subtracted from
765 each measurement. Rate measurements were fit to Michaelis-Menten kinetic models and K_m and
766 V_{max} for substrates were calculated by nonlinear regression using Prism 9.3.0 for macOS
767 (GraphPad Software, San Diego, California USA, www.graphpad.com).

768 NMR

769 For NMR studies, non-isotopically enriched AMPylated nsp9 was dissolved in 50mM Tris buffer
770 at pH 7.5, 150mM NaCl, 1mM DTT and 10% D₂O for spectrometer locking. The final protein
771 concentration of this solution was 0.5mM. A total volume of 500uL was then used with a 5mm
772 NMR tube to record all the spectra.

773 All NMR experiments were run on a Bruker Avance III spectrometer operating at 600MHz (1H)
774 and equipped with a 5mm proton-optimized quadruple resonance cryogenic probe. The
775 temperature of the sample was regulated at 308K throughout data collection.

776 A one-dimensional (1D) 31P spectrum was recorded with 8192 scans and a repetition delay of
777 1.5sec for a total collection time of 3.5 hours. The 31P spectral window and offset were set to
778 17ppm and 2.6ppm, respectively. Waltz16 decoupling was used on 1H during 31P acquisition.

779 To observe contacts between 31P and the nearest 1H nuclei, a two-dimensional (2D) 1H,31P-
780 HSQC spectrum was recorded with 1024 and 22 complex points in the direct 1H and indirect 31P
781 dimensions, respectively. The spectral window and offset were set to 16.7ppm and 4.7ppm for the
782 1H dimension and 3.4ppm and 2.6ppm for the 31P dimension, respectively. Each FID was
783 accumulated with 1536 scans with a repetition delay of 1sec for a total recording time of
784 approximately 21 hours.

785 A 2D 1H,31P-HSQC-TOCSY spectrum was recorded using similar spectral window and offset
786 parameters for the 1H and 31P dimensions as the HSQC spectrum described above. A 60ms long
787 1H-1H TOCSY pulse train using a DIPSI-2 sequence and a field strength of 10KHz was tagged at
788 the end of the HSQC sequence to observe signals from 1H nuclei that are further away from 31P.
789 Given the lower sensitivity of this experiment, each FID was accumulated with 4096 scans and a
790 repetition delay of 1sec was used for a total recording time of 2 days and 14 hours.

791 Using a similar pulse sequence, a 2D 1H,1H-HSQC-TOCSY spectrum was also recorded by
792 evolving the indirect 1H dimension instead of 31P. The spectral window for the 1H indirect
793 dimension was set to 4.2ppm, while the offset was maintained at 4.7ppm as for the direct 1H
794 dimension. 40 complex points were recorded for then indirect 1H dimension, using 2048
795 accumulations for each FID and a repetition delay of 1sec for a total recording time of 2 days and
796 14 hours.

797 All 2D spectra were processed using NMRPipe³⁷ and analysed with NMRFAM-SPARKY³⁸.

798 Intact mass analysis

799 Protein samples were analysed by LC/MS, using a Sciex X500B Q-TOF mass spectrometer
800 coupled to an Agilent 1290 Infinity II HPLC. Samples were injected onto a POROS R1 reverse-
801 phase column (2.1 x 30 mm, 20 µm particle size, 4000 Å pore size) and desalted. The mobile phase
802 flow rate was 300 µL/min and the gradient was as follows: 0-3 min: 0% B, 3-4 min: 0-15% B, 4-
803 16 min: 15-55% B, 16-16.1 min: 55-80% B, 16.1-18 min: 80% B. The column was then re-
804 equilibrated at initial conditions prior to the subsequent injection. Buffer A contained 0.1% formic
805 acid in water and buffer B contained 0.1% formic acid in acetonitrile.

806 The mass spectrometer was controlled by Sciex OS v.1.6.1 using the following settings: Ion source
807 gas 1 30 psi, ion source gas 2 30 psi, curtain gas 35, CAD gas 7, temperature 300 °C, spray voltage
808 5500 V, declustering potential 80 V, collision energy 10 V. Data was acquired from 400-2000 Da
809 with a 0.5 s accumulation time and 4 time bins summed. The acquired mass spectra for the proteins
810 of interest were deconvoluted using BioPharmaView v. 3.0.1 software (Sciex) in order to obtain
811 the molecular weights. The peak threshold was set to $\geq 5\%$, reconstruction processing was set to
812 20 iterations with a signal-to-noise threshold of ≥ 20 and a resolution of 2500.

813 RNAylation assays

814 RNAylation reactions were typically carried out in a 10 µL volume containing 50 mM Tris (pH
815 7.5), 5 mM KCl, 1 mM DTT, 0.4 µg yeast inorganic pyrophosphatase, 20 µM nsp9, and 2 µM
816 nsp12. Reactions were started by adding MnCl₂ and 5'-pppRNA^{LS10} to a final concentration of 1
817 mM and 100 µM, respectively. Reactions were incubated at 37°C for the indicated time points and
818 stopped by addition of 5X SDS-PAGE sample buffer + β-ME and boiling the samples for 5
819 minutes. Reaction products were resolved by SDS-PAGE on a 4-20% gradient gel and visualized
820 by Coomassie staining.

821 For the time course comparing the RNAylation of nsp9 N1A and N2A mutants (i.e. **Fig. 2d**),
822 reactions were performed as above, except with 2.4 µM nsp12. At each indicated time point,
823 reactions were stopped by addition of 5X SDS-PAGE sample buffer + β-ME and boiling the
824 samples for 5 minutes. Reaction products were resolved by SDS-PAGE on a 4-20% gradient gel
825 and visualized with Coomassie staining.

826 For reactions testing RNA length specificity (i.e. **Fig. 2e**), 7.5 μ L of a reaction master mix
827 containing nsp9, nsp12, and yeast inorganic pyrophosphatase was added to 2.5 μ L of start mix
828 consisting of $MnCl_2$ and the indicated 5'-pppRNA. The final reaction conditions were as follows:
829 50 mM Tris pH 7.5, 5 mM KCl, 1 mM DTT, 0.4 μ g yeast inorganic pyrophosphatase, 20 μ M nsp9,
830 2 μ M nsp12, 1 mM $MnCl_2$, and 100 μ M of the indicated RNA. Reactions were incubated for 30
831 minutes at 37°C, then stopped by addition of 5X SDS-PAGE sample buffer + β -ME and boiling
832 the samples for 5 minutes. Reaction products were resolved by SDS-PAGE on a 4-20% gradient
833 gel and visualized by Coomassie staining.

834 For RNAylation reactions comparing different RNA sequences (i.e. **Fig. 2f**) or nsp9 mutants (i.e.
835 **Fig. 5h**), reactions were performed as above, except using 1 μ M nsp12. Reactions were incubated
836 for 5 minutes and stopped by addition of 5X SDS-PAGE sample buffer + β -ME and boiling the
837 samples for 5 minutes. Reaction products were resolved by SDS-PAGE on a 4-20% gradient gel
838 and visualized by Coomassie staining.

839 Purification of nsp9-pRNA^{LS10} species

840 Purified native nsp9 (0.8 mg/mL, 65 μ M) was incubated at room temperature overnight with 130
841 μ M of 5'-pppRNA^{LS10} and \sim 0.8 μ M of nsp12 in presence of 0.05 mg/ml yeast inorganic
842 pyrophosphatase and 1 mM $MnCl_2$, in the reaction buffer (50 mM Tris 7.5, 5 mM KCl, 1 mM
843 DTT). The samples were clarified by centrifugation to remove any precipitate and applied directly
844 onto a Cpto HiRes Q 5/50 column (Cytiva) equilibrated in 50 mM Tris 8.0, 50 mM NaCl, 1 mM
845 DTT. An elution gradient of 0-50% with 1 M NaCl was applied over 30 column volumes. Under
846 these conditions, RNA and nsp9-pRNA^{LS10} bound the column and unmodified nsp9 did not. nsp9-
847 pRNA^{LS10} and unreacted RNA^{LS10} eluted as a peak doublet around 70 mS/cm. Fractions were
848 pooled, and further purified over Superdex 75 increase 10/300 GL (50 mM Tris 8.0, 300 mM NaCl,
849 1 mM DTT), separating nsp9-pRNA^{LS10} from unreacted RNA^{LS10}, and nsp12. nsp9-pRNA^{LS10} was
850 quantified by spectrophotometry with an estimated extinction coefficient of $\epsilon_{260}=130,650 M^{-1}cm^{-1}$.
851 We also generated nsp9-pRNA^{LS10} in the absence of inorganic pyrophosphatase. This nsp9-
852 pRNA^{LS10} was used in control reactions to test for PP_i hydrolysis, necessary to ensure that PP_i
853 mediated deRNAylation reactions shown in **Fig. 3e** do not suffer from pyrophosphate hydrolysis.
854 The results were like those presented in the Figure, thus confirming that there was no
855 contaminating inorganic pyrophosphatase in the assays.

856 DeRNAylation of nsp9-pRNA^{LS10}

857 DeRNAylation reactions were typically performed in a 10 μ L reaction volume consisting of 50
858 mM Tris pH 7.5, 5 mM KCl, 1 mM DTT, 20 μ M nsp9-pRNA^{LS10}, 1 μ M nsp12, 1 mM MgCl₂, and
859 500 μ M GDP. Reactions were started by adding MgCl₂/GDP and incubated at 37°C for 5-60
860 minutes as indicated. Reactions were stopped by addition of 5X SDS-PAGE sample buffer + β -
861 ME and boiling the samples for 5 minutes. Reaction products were resolved by SDS-PAGE on a
862 4-20% gradient gel and visualized with Coomassie staining.

863 For deRNAylation reactions comparing various nucleotide triphosphates (NTP) and nucleotide
864 diphosphates (NDP), reactions were performed in 10 μ L volume consisting of 50 mM Tris pH 7.5,
865 5 mM KCl, 1 mM DTT, 20 μ M nsp9-pRNA^{LS10}, 500 nM nsp12, 1 mM MgCl₂, and 500 μ M of the
866 indicated NTP or NDP. Reactions were incubated for 5 minutes at 37°C and stopped by addition
867 of 5X SDS-PAGE sample buffer with β -ME and boiling for 5 minutes. Reaction products were
868 resolved by SDS-PAGE on a 4-20% gradient gel and visualized with Coomassie staining.

869 Generation of [α -³²P]-GDP using nsp13

870 To generate [α -³²P]-GDP from [α -³²P]-GTP, 0.3-1 mM of [α -³²P]-GTP (specific activity ~2000
871 cpm/pmol) was incubated with 0.5-1 mg/mL nsp13 (and in some cases yeast cet1 NTPase) in 20
872 μ L reaction buffer (depending on amount needed) consisting of 50 mM Tris (pH 7.5), 5 mM KCl,
873 1 mM DTT, 2 mM MgCl₂. Reactions were started by addition of enzyme and allowed to proceed
874 for 30 minutes at 37°C. Following the 30-minute incubation, reactions were boiled at 95°C for 5
875 minutes to inactivate nsp13 or cet1.

876 Generation of radiolabelled GpppA-RNA^{LS10} from nsp9-pRNA^{LS10} and [α -³²P]-GDP

877 Reactions were performed in a 10 μ L volume containing 50 mM Tris pH 7.5, 5 mM KCl, 1 mM
878 DTT, 15 μ M nsp9-pRNA^{LS10}, 377 nM nsp12, 1 mM MgCl₂, and 500 μ M [α -³²P]-GDP (specific
879 radioactivity = ~2000 cpm/pmol). Reactions were started by addition of [α -³²P]-GDP/MgCl₂
880 mixture (generated as described above) and incubated for 30 minutes at 37°C. As a control, VCE
881 was used but with [α -³²P]-GTP. VCE assays were generally performed as described in the NEB
882 Capping Protocol (M2080) with the following modifications: the reaction contained 20 μ M of 5'-
883 pppRNA^{LS10}, 500 μ M [α -³²P]-GTP (specific activity ~2000cpm/pmol), and did not contain SAM.

884 Reactions were stopped by the addition of 2X TBE-Urea sample buffer, boiled for 5 minutes, and
885 resolved by UREA-PAGE (20%).

886 GDP inhibition of RNAylation (one pot capping assays)

887 Reactions were performed in 50 mM Tris pH 7.5, 5 mM KCl, 1 mM DTT, 1 mM MgCl₂, 1 mM
888 MnCl₂, and contained 20 μM nsp9, 2 μM nsp12, and 100 μM 5'-pppRNA^{LS10} in 20 μL volume.
889 The [α -³²P]-GDP was prepared as described earlier (using 400 μM GTP, [α -³²P]-GTP at specific
890 activity ~1,500 cpm/pmol) and diluted to final reaction concentrations in the range of 6.25-100
891 μM. Reactions were started by the addition of nsp12. GDP was added either before the addition of
892 nsp12 (t=0), or after 30 minutes of preincubation. After an additional 30 minutes, the reactions
893 were split in half and stopped by the addition of 5x SDS-PAGE or 2x Formamide loading dyes
894 and the products were analysed by 4-20% gradient SDS-PAGE gel, or 15% 19:1 TBE UREA-
895 PAGE gel, respectively.

896 LC-MS/MS analysis of GpppA

897 De-RNAylation reactions (in triplicate) were performed in 20 μL of buffer solution containing 50
898 mM Tris pH 7.5, 5 mM KCl, 20 μM nsp9-pRNA^{LS10}, 2 μM wild-type nsp12 or the D218A mutant,
899 1 mM MgCl₂, and 100 μM GDP. Reactions were started by adding MgCl₂ and GDP and allowed
900 to proceed for 1 hour at 37°C. After the 1-hour incubation, the reactions were supplemented with
901 2 μL of 10X P1 buffer and 1 μL of Nuclease P1 enzyme and allowed to proceed for an additional
902 30 minutes at 37°C. Reactions were stopped by boiling for 5 minutes and submitted for LC-MS/MS
903 analysis.

904 For standards, 20 μL of blank reaction buffer (50 mM Tris pH 7.5, 5 mM KCl, 1 mM DTT) plus
905 40 μL blank reaction buffer containing 0.33 μM (final) m7GpppA (m7G(5')ppp(5')A RNA Cap
906 Structure Analog, New England Biolabs, S1405S) as an internal standard (IS) was spiked with
907 varying concentrations of GpppA (New England Biolabs, S1406L). Reaction samples (20 μL)
908 were diluted with blank reaction buffer, containing 0.33 μM (final) m7GpppA IS, at 1:2 to a total
909 volume of 60 μL. Standards and samples were mixed with 60 μL of 100% methanol, vortexed,
910 and then spun for 5 min at 16,100 x g. Supernatant was removed and analysed by LC-MS/MS
911 using a Sciex (Framingham, MA) QTRAP® 6500+ mass spectrometer coupled to a Shimadzu
912 (Columbia, MD) Nexera X2 LC. GpppA was detected with the mass spectrometer in positive
913 MRM (multiple reaction monitoring) mode by following the precursor to fragment ion transition

914 772.9 → 604.0. A Thermo Scientific BioBasic AX column (2.1 x 50 mm, 5 micron packing) was
915 used for chromatography with the following conditions: Buffer A: 8:2 dH₂O:Acetonitrile + 10 mM
916 ammonium acetate, pH 6, Buffer B: 7:3 dH₂O:Acetonitrile + 1 mM ammonium acetate, pH 10.5,
917 0.5 mL/min flow rate, 0-1 min 0%B, 1-2.5 min gradient to 35%B, 2.5-5 min 35%B, 5-7 min
918 gradient to 65%B, 7-10 min 65%B, 10-10.5 min gradient to 100%B, 10.5-15 min 100%B, 15-15.5
919 min gradient to 0%B, 15.5-20.5 min 0%B. m7GpppA (transition 787.1 → 508.0) was used as an
920 internal standard. Peak areas were determined and data were further analysed using the Sciex
921 Analyst 1.7.2 software package. Back-calculation of standard curve samples were accurate to
922 within 15% for 100% of these samples at concentrations ranging from 0.001 μM to 10 μM. A limit
923 of detection (LOD) was defined as a level three times that observed in blank reaction buffer and
924 the limit of quantitation (LOQ) as the lowest point on the standard curve that gave an analyte signal
925 above the LOD and within 20% of nominal upon back-calculation. The LOQ for GpppA was 0.005
926 μM.

927 Methyltransferase assays

928 In 10 μL reactions, 40 μM nsp9-pRNA^{LS10} was incubated with 2 μM nsp12 in presence of 1 mM
929 MgCl₂ and 100 μM [³²P]GDP (generated as above) for 60 min at 37 °C. Reactions were filled to
930 15 μL with nsp14 and SAM (final 0.05 mg/mL and 100 uM respectively), and incubated for
931 another 30 minutes. Treatment with nsp10/16 can be done concurrently with nsp14, however
932 nsp10/14 complex partially processes RNA^{LS10}, resulting in a mobility shift ³⁹. Thus, prior to
933 addition of nsp10 and nsp16, nsp14 exonuclease activity was removed by heat inactivation (5
934 minutes at 95 °C). For 2'-O methylation, reactions were supplemented with nsp10/16 and fresh
935 SAM to final concentrations of 0.05 mg/mL nsp10, 0.05 mg/mL nsp16, 100 μM SAM in final 20
936 μL volume. Vaccinia reactions were conducted as per manufacturer's instructions. Reactions were
937 stopped by adding 2x formamide loading dye and were separated on 20% TBE-UREA
938 polyacrylamide gels (19:1). Radioactivity was visualized by autoradiography and RNA by
939 toluidine staining. For TLC analysis, bands with detectible ³²P signal were excised, fragmented,
940 and incubated overnight at 55°C in elution buffer (1 M Ammonium acetate, 0.2% SDS, 20 mM
941 EDTA), rotating top-over-bottom. Solutions were filtered using 0.22 μm centrifugal filters,
942 supplemented with 23 ug Glyco Blue co-precipitant (Invitrogen) and precipitated for 1 hr at -20°C
943 by addition of isopropyl alcohol to a final concentration of 60%. The pellets were washed once
944 with 70% EtOH and reconstituted in 10 μL of P1 buffer with P1 nuclease (NEB). After 30 minutes

945 at 37°C, reactions were supplemented with Quick CIP (NEB) and rCut Smart buffer, to a final
946 volume of 12 uL. After 30 min of further incubation, reactions were spotted onto PEI-Cellulose F
947 TLC plates and resolved in 0.4 M Ammonium Sulfate mobile phase. Beforehand, TLC plates were
948 prepared by development in water, removing yellow discoloration. The ³²P signal was detected
949 by autoradiography, and compared with cold standards of GTP, GDP, GpppA and ^{m7}GpppA
950 detected by absorption of plate fluorescence, excitable with a UV lamp λ=265 nm. The position of
951 ^{m7}GpppA_{2'-OMe} was determined from the Vaccinia capping enzyme and Vaccinia 2'-O-
952 Methyltransferase control reaction.

953 Reactions with ¹⁴C-labelled SAM were conducted as above, with two differences: cold GDP was
954 used at 100 μM (Millipore Sigma, G7127), and 55 μM [¹⁴C]SAM (Perkin Elmer), SAM was used
955 at the supplied radioactivity of 52.6 mCi/mmol (~117 cpm/pmol), with no further dilution using
956 cold SAM.

957 GST-eIF4E pulldown of ⁷MeGpppA-RNA

958 Capping reactions were set up in 20 μL and contained 2 μM nsp12, 0.04 mg/ml nsp14 WT or
959 D331A, 30 μM nsp9-pRNA^{LS10}, 100 μM [^{α-32}P]GDP (specific radioactivity = 1,000 cpm/pmol),
960 100 μM SAM, 2 mM MgCl₂. The reaction buffer was 50 mM Tris 8.0, 5 mM KCl, 1 mM DTT.
961 Vaccinia capping enzyme controls were performed according to manufacturer's instructions, with
962 the same nucleotide and SAM concentrations as nsp12 reactions. After 90 minutes of incubation
963 at 37 °C, 15.6 ug of GST-eIF4E K119A was added, along with 15 μL of Glutathione resin (Pierce,
964 Thermo Scientific). Reactions were filled to 700 μL with 50 mM Tris 8.0, 150 mM NaCl, 1 mM
965 DTT, and nutated for 1 hour. The resin was washed 3x with 500 μL of 50 mM Tris 8.0, 150 mM
966 NaCl, 1 mM DTT, and radioactive signal was quantified by scintillation counting.

967 Cryo-EM grid preparation

968 To form nsp12/7/8 core complex (RTC), native nsp12, nsp7 and nsp8 were incubated in 1:2:4
969 molar ratio and run over Superdex 200 increase 10/300 GL to separate unassociated monomers.
970 The purified complex was concentrated using spin concentrators (Amicon 10k MWCO, Sigma-
971 Millipore), and quantified by spectrophotometry. A 3x molar excess of nsp9 over RTC was added,
972 followed by 0.05 mM final DDM detergent immediately prior to freezing. Final concentration of
973 nsp12/7/8 was 2 mg/mL. Buffer contained 50 mM Tris 7.5, 150 mM NaCl, 1 mM DTT, 2 mM

974 MnCl₂, 1 mM UMP-NPP. Copper Quantifoil 1.2/1.3 mesh 300 grids were used to freeze 3.5 μL of
975 sample at 100% relative humidity using Vitrobot mk. IV (Thermofisher).

976 Cryo-EM data collection

977 Prior to data collection, sample grids were screened on a Talos Artica microscope at the Cryo
978 Electron Microscopy Facility (CEMF) at UT Southwestern. Cryo-EM data of NSP12/7/8/9
979 complex were collected on a Titan Krios microscope at Cryo-Electron Microscopy Facility
980 (CEMF) at UT Southwestern Medical Center, with the post-column energy filter (Gatan) and a K3
981 direct detection camera (Gatan), using SerialEM⁴⁰. 4,770 movies were acquired at a pixel size of
982 0.55 Å in super-resolution counting mode, with an accumulated total dose of 54 e⁻/Å² over 50
983 frames. The defocus range of the images was set to be -1.0 to -2.5 μm.

984 Image processing and 3D reconstruction

985 Unless described otherwise, all datasets were processed with Relion⁴¹. Movies were aligned and
986 summed using MotionCor2⁴², with a downsampled pixel size of 1.09 Å. The CTF parameters
987 were calculated using Gctf⁴³, and images with estimated CTF max resolution better than 5 Å^o were
988 selected for further processing. 4,196,086 particles were picked using crYOLO⁴⁴ from 4,757
989 images, and extracted with a re-scaled pixel size of 2.19 Å. 663,999 particles were selected and
990 re-extracted after multiple rounds of 2D and 3D classifications in Relion with the original pixel
991 size of 1.09 Å. An additional round of 3D classification was carried out, followed by particle
992 reduction with a homemade script to remove particles from dominant orientations. The remaining
993 89,945 particles were subjected to 3D refinement, CTF refinement and particle polishing
994 sequentially. A final round of 3D classification with a reference mask led to 39,985 particles, which
995 were then imported into cryoSPARC⁴⁵ for one round of non-uniform refinement. The map
996 resolution was reported at 3.18 Å from cryoSPARC with the gold standard FSC method.

997 Nsp5 cleavage reactions

998 Concentrated protein samples were diluted in cleavage buffer (50 mM Tris, pH 7.4; 150 mM NaCl;
999 5% glycerol) and each reaction was performed in a total volume of 10 μL. Initial experiments
1000 measured the nsp5 concentration dependence of the nsp8-nsp9 cleavage reaction. To measure time
1001 dependence, nsp5 (2.5 μM final concentration) was added to the nsp8-nsp9 fusion protein (12.5
1002 μM final concentration). Reactions were incubated at 37 °C for varying amounts of time (0 to 80

1003 minutes) and terminated by boiling the samples for five minutes in the presence of SDS-PAGE
1004 loading buffer. Reaction products were resolved on a 4-20% gradient tris-glycine gel and products
1005 were visualized by Coomassie staining.

1006 Model building and refinement

1007 Model was build using PDB 7CYQ as a template ²¹. Model was manually rebuilt into the map
1008 using Coot ⁴⁶, and refined using Phenix real space refinement ⁴⁷. Model validation was performed
1009 using MolProbity software ⁴⁸.

1010 Bioinformatics

1011 A representative subset of NiRAN domain sequences, provided as a multiple alignment in the
1012 original NiRAN publication ⁵, was supplemented by additional sequences used in this study
1013 (SARS-CoV-2, OC43, 229E strains). The human SELO sequence was added according to a
1014 FATCAT structural alignment ⁴⁹ between SARS-CoV-2 nsp12 and bacterial SelO (PDB identifiers
1015 7cyq and 6eac). The alignment was visualized using the ESPrict server ⁵⁰.

1016 SARS-CoV-2 infection experiments

1017 *Plasmid Construction*

1018 To generate recombinant SARS-CoV-2 expressing ZsGreen mutants, the infectious clone pCC1-
1019 4K-SARS-CoV-2-Wuhan-Hu-1-ZsGreen was used as the parental backbone ⁸. To generate nsp9
1020 (N1A, N1D, N2A) and nsp12 (K73A) mutants, mutations were introduced by overlap extension
1021 PCR. Briefly, 2 fragments for each mutant were created by PCR using Ex Taq DNA Polymerase
1022 (Takara). These fragments shared homology at the 3' end of Fragment 1 and 5' end of Fragment
1023 2. The resulting 2 fragments were then used as a template for a third fragment to PCR a full-length
1024 amplicon containing the mutation flanked by PacI and MluI sites on the ends. Mutations were
1025 confirmed by DNA sequencing. To generate nsp12 mutants (D218A, D760A), a SARS-CoV-2
1026 shuttle vector (ps1180.SARS-CoV-2-shuttle) was created using ps1180.delXhoISacII plasmid as
1027 the backbone. Using Gibson cloning, a restriction enzyme linker that contained unique restriction
1028 sites specific to the SARS-CoV-2 genome was inserted. Smaller fragments of the SARS-CoV-2
1029 genome were digested from the pCC1-4K-SARS-CoV-2-Wuhan-Hu-1-ZsGreen plasmid and
1030 ligated into ps1180.SARS-Cov-2-shuttle plasmid to create 3 new plasmids (MluI/SacI fragment
1031 for D218A region; SacI/Bsu36I fragment for D760A region). gBlocks containing mutations were

1032 synthesized by IDT and introduced into the SARS-CoV-2 shuttle vector by Gibson Assembly
1033 following standard protocols. To reassemble the full length parental pCC1-4K-SARS-CoV-2-
1034 Wuhan-Hu-1-ZsGreen containing new mutants, pCC1-4K-SARS-CoV-2-Wuhan-Hu-1-ZsGreen
1035 was digested with PacI/MluI, MluI/SacI, and SacI/Bsu36I restriction enzymes. A roughly 28-35
1036 kb fragment was purified for each digest. PCR Amplicons (nsp9^{N1A}, nsp9^{N1D}, nsp9^{N2A}, nsp12^{K73A})
1037 were digested with PacI/MluI and 5.3 kb fragment was purified. The SARS-CoV-2 shuttle
1038 plasmids were digested as follows: nsp12^{D218A} (MluI/SacI releasing 1.4 kb fragment); nsp12^{D760A}
1039 (SacI/Bsu36I/PvuI releasing 3 kb fragments). All fragments were purified using QIAexII Gel
1040 Purification Kit following standard protocol (Qiagen). Fragments were ligated together at a 3:1
1041 ratio overnight. Ligated DNA was precipitated using 7.5 M Ammonium acetate, Glycogen, and
1042 Isopropanol, followed by an ethanol wash. The DNA was electroporated into TransforMAX
1043 EPI300 Electro competent *E. coli* (Lucigen). An overlapping 8-fragment PCR strategy was used
1044 to verify individual colonies by colony PCR. Confirmed colonies were grown in 10 ml Tryptic
1045 Soy Broth (TSB; Sigma) containing 12.5 µg/ml Chloramphenicol for 6-8 hours, shaking at 37°C.
1046 A 10 ml culture was inoculated into 100 ml TSB/Chloramphenicol culture and incubated
1047 overnight, shaking at 37°C. Overnight culture was diluted 1:5 into fresh TSB/Chloramphenicol
1048 containing 0.1% Arabinose and incubated an additional 5 hrs. Bacteria was pelleted and DNA was
1049 isolated using a homemade midi prep protocol followed by Machery-Nagel NucleoBond Xtra Midi
1050 Kit (Fisher). Full-length infectious clone plasmid was confirmed by restriction digestion and 8-
1051 fragment PCR. Oligonucleotides used are shown in **Table S3**.

1052 *Virus production*

1053 To generate virus from DNA-based infectious clones, 3 µg of plasmid were transfected into 2
1054 individual 6 wells of 400,000 BHK-21J cells using X-treme Gene9 Transfection Reagent (Sigma).
1055 Three days post transfection, the supernatant from 2 individual wells was combined and 3 ml was
1056 transferred to a T25 flask containing 1 x 10⁶ VeroE6-C1008-TMPRSS2 cells and 2 ml serum-free
1057 MEM. After 4 days, 250 µl supernatant was added to 750 µl TriReagent for RNA extraction and
1058 RT-qPCR. T25 flasks were fixed with 4% paraformaldehyde, imaged on Nikon Eclipse Ti and
1059 processed with ImageJ.

1060 *RNA Extraction/RT-qPCR*

1061 RNA was isolated using the Direct-zol RNA mini prep kit following manufacturer's instructions
1062 (ZymoResearch). A 20 µl reaction contained 5 µl RNA, 5 µl TaqMan Fast Virus 1-Step Master
1063 Mix, and 1.8 µl SARS-CoV-2 primer/probe set containing 6.7 µM each primer/1.7 µM probe (final
1064 concentration of primer/probe were 600 nM/150 nM probe). SARS-CoV-2 primers and probe were
1065 designed as recommended by the Center for Disease Control
1066 (<https://www.cdc.gov/coronavirus/2019-ncov/lab/rt-pcr-panel-primer-probes.html>). All
1067 oligonucleotides were synthesized by LGC Biosearch Technologies. RT was performed at 50°C
1068 for 5 minutes, followed by inactivation at 95°C for 2 minutes, and 40 cycles of PCR (95°C for 3
1069 seconds, 60°C for 30 seconds) on a QuantStudio 3 (Applied Biosystems).

1070 *Cells*

1071 BHK-21J cells (a generous gift from C. Rice) were grown in MEM (Gibco) supplemented with
1072 10% FBS and 1X NEAA. VeroE6-C1008 cells (ATCC) were transduced with lentiviral vector
1073 SCRBBL-TMPRSS2, selected and maintained in MEM supplemented with 10% FBS, 1X NEAA,
1074 and 8 µg/ml Blasticidin.

1075

1076 **References**

- 1077
- 1078 1 Chen, Y. & Guo, D. Molecular mechanisms of coronavirus RNA capping and methylation.
1079 *Virology* **31**, 3-11, doi:10.1007/s12250-016-3726-4 (2016).
- 1080 2 Malone, B., Urakova, N., Snijder, E. J. & Campbell, E. A. Structures and functions of
1081 coronavirus replication-transcription complexes and their relevance for SARS-CoV-2 drug
1082 design. *Nat Rev Mol Cell Biol*, doi:10.1038/s41580-021-00432-z (2021).
- 1083 3 Walker, A. P. *et al.* The SARS-CoV-2 RNA polymerase is a viral RNA capping enzyme.
1084 *Nucleic Acids Res* **49**, 13019-13030, doi:10.1093/nar/gkab1160 (2021).
- 1085 4 van Vliet, A. L., Smits, S. L., Rottier, P. J. & de Groot, R. J. Discontinuous and non-
1086 discontinuous subgenomic RNA transcription in a nidovirus. *EMBO J* **21**, 6571-6580,
1087 doi:10.1093/emboj/cdf635 (2002).
- 1088 5 Lehmann, K. C. *et al.* Discovery of an essential nucleotidylating activity associated with a
1089 newly delineated conserved domain in the RNA polymerase-containing protein of all
1090 nidoviruses. *Nucleic Acids Res* **43**, 8416-8434, doi:10.1093/nar/gkv838 (2015).
- 1091 6 Chen, Y. *et al.* Functional screen reveals SARS coronavirus nonstructural protein nsp14 as
1092 a novel cap N7 methyltransferase. *Proc Natl Acad Sci U S A* **106**, 3484-3489,
1093 doi:10.1073/pnas.0808790106 (2009).
- 1094 7 Bouvet, M. *et al.* In vitro reconstitution of SARS-coronavirus mRNA cap methylation.
1095 *PLoS Pathog* **6**, e1000863, doi:10.1371/journal.ppat.1000863 (2010).
- 1096 8 Rihn, S. J. *et al.* A plasmid DNA-launched SARS-CoV-2 reverse genetics system and
1097 coronavirus toolkit for COVID-19 research. *PLoS Biol* **19**, e3001091,
1098 doi:10.1371/journal.pbio.3001091 (2021).
- 1099 9 V'Kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology and
1100 replication: implications for SARS-CoV-2. *Nat Rev Microbiol* **19**, 155-170,
1101 doi:10.1038/s41579-020-00468-6 (2021).
- 1102 10 Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat
1103 origin. *Nature*, doi:10.1038/s41586-020-2012-7 (2020).
- 1104 11 Beigel, J. H. *et al.* Remdesivir for the Treatment of Covid-19 - Final Report. *N Engl J Med*
1105 **383**, 1813-1826, doi:10.1056/NEJMoa2007764 (2020).
- 1106 12 Fischer, W. *et al.* Molnupiravir, an Oral Antiviral Treatment for COVID-19. *medRxiv*,
1107 doi:10.1101/2021.06.17.21258639 (2021).
- 1108 13 Sreelatha, A. *et al.* Protein AMPylation by an Evolutionarily Conserved Pseudokinase. *Cell*
1109 **175**, 809-821 e819, doi:10.1016/j.cell.2018.08.046 (2018).
- 1110 14 Kirchdoerfer, R. N. & Ward, A. B. Structure of the SARS-CoV nsp12 polymerase bound
1111 to nsp7 and nsp8 co-factors. *Nat Commun* **10**, 2342, doi:10.1038/s41467-019-10280-3
1112 (2019).
- 1113 15 Dudkiewicz, M., Szczepinska, T., Grynberg, M. & Pawlowski, K. A novel protein kinase-
1114 like domain in a selenoprotein, widespread in the tree of life. *PLoS One* **7**, e32138,
1115 doi:10.1371/journal.pone.0032138 (2012).
- 1116 16 Slanina, H. *et al.* Coronavirus replication-transcription complex: Vital and selective
1117 NMPylation of a conserved site in nsp9 by the NiRAN-RdRp subunit. *Proc Natl Acad Sci*
1118 *U S A* **118**, doi:10.1073/pnas.2022310118 (2021).
- 1119 17 Kikkert, M. Innate Immune Evasion by Human Respiratory RNA Viruses. *J Innate Immun*
1120 **12**, 4-20, doi:10.1159/000503030 (2020).

- 1121 18 Ramanathan, A., Robb, G. B. & Chan, S. H. mRNA capping: biological functions and
1122 applications. *Nucleic Acids Res* **44**, 7511-7526, doi:10.1093/nar/gkw551 (2016).
- 1123 19 Daffis, S. *et al.* 2'-O methylation of the viral mRNA cap evades host restriction by IFIT
1124 family members. *Nature* **468**, 452-456, doi:10.1038/nature09489 (2010).
- 1125 20 Ivanov, K. A. *et al.* Multiple enzymatic activities associated with severe acute respiratory
1126 syndrome coronavirus helicase. *J Virol* **78**, 5619-5632, doi:10.1128/JVI.78.11.5619-
1127 5632.2004 (2004).
- 1128 21 Yan, L. *et al.* Cryo-EM Structure of an Extended SARS-CoV-2 Replication and
1129 Transcription Complex Reveals an Intermediate State in Cap Synthesis. *Cell* **184**, 184-193
1130 e110, doi:10.1016/j.cell.2020.11.016 (2021).
- 1131 22 Conti, B. J., Leicht, A. S., Kirchdoerfer, R. N. & Sussman, M. R. Mass spectrometric based
1132 detection of protein nucleotidylation in the RNA polymerase of SARS-CoV-2. *Commun*
1133 *Chem* **4**, doi:10.1038/s42004-021-00476-4 (2021).
- 1134 23 Shannon, A. *et al.* Protein-primed RNA synthesis in SARS-CoVs and structural basis for
1135 inhibition by AT-527. *bioRxiv*, 2021.2003.2023.436564, doi:10.1101/2021.03.23.436564
1136 (2021).
- 1137 24 Ogino, T. & Banerjee, A. K. Unconventional mechanism of mRNA capping by the RNA-
1138 dependent RNA polymerase of vesicular stomatitis virus. *Mol Cell* **25**, 85-97,
1139 doi:10.1016/j.molcel.2006.11.013 (2007).
- 1140 25 Ogino, T. & Green, T. J. RNA Synthesis and Capping by Non-segmented Negative Strand
1141 RNA Viral Polymerases: Lessons From a Prototypic Virus. *Front Microbiol* **10**, 1490,
1142 doi:10.3389/fmicb.2019.01490 (2019).
- 1143 26 Sonenberg, N. & Gingras, A. C. The mRNA 5' cap-binding protein eIF4E and control of
1144 cell growth. *Curr Opin Cell Biol* **10**, 268-275, doi:10.1016/s0955-0674(98)80150-6
1145 (1998).
- 1146 27 Wang, B., Svetlov, D. & Artsimovitch, I. NMPylation and de-NMPylation of SARS-CoV-
1147 2 nsp9 by the NiRAN domain. *Nucleic Acids Res* **49**, 8822-8835, doi:10.1093/nar/gkab677
1148 (2021).
- 1149 28 Taylor, S. S. & Kornev, A. P. Protein kinases: evolution of dynamic regulatory proteins.
1150 *Trends Biochem Sci* **36**, 65-77, doi:10.1016/j.tibs.2010.09.006 (2011).
- 1151 29 Malone, B., Campbell, E. A. & Darst, S. A. CoV-er all the bases: Structural perspectives
1152 of SARS-CoV-2 RNA synthesis. *Enzymes* **49**, 1-37, doi:10.1016/bs.enz.2021.06.004
1153 (2021).
- 1154 30 Chen, J. *et al.* Structural Basis for Helicase-Polymerase Coupling in the SARS-CoV-2
1155 Replication-Transcription Complex. *Cell*, doi:10.1016/j.cell.2020.07.033 (2020).
- 1156 31 Yan, L. *et al.* Coupling of N7-methyltransferase and 3'-5' exoribonuclease with SARS-
1157 CoV-2 polymerase reveals mechanisms for capping and proofreading. *Cell* **184**, 3474-3485
1158 e3411, doi:10.1016/j.cell.2021.05.033 (2021).
- 1159 32 Perry, J. K. *et al.* An atomistic model of the coronavirus replication-transcription complex
1160 as a hexamer assembled around nsp15. *J Biol Chem* **297**, 101218,
1161 doi:10.1016/j.jbc.2021.101218 (2021).
- 1162 33 Ogino, T., Yadav, S. P. & Banerjee, A. K. Histidine-mediated RNA transfer to GDP for
1163 unique mRNA capping by vesicular stomatitis virus RNA polymerase. *Proc Natl Acad Sci*
1164 *U S A* **107**, 3463-3468, doi:10.1073/pnas.0913083107 (2010).
- 1165 34 Liang, B. *et al.* Structure of the L Protein of Vesicular Stomatitis Virus from Electron
1166 Cryomicroscopy. *Cell* **162**, 314-327, doi:10.1016/j.cell.2015.06.018 (2015).

1167 35 Gordon, D. E. *et al.* A SARS-CoV-2 protein interaction map reveals targets for drug
1168 repurposing. *Nature* **583**, 459-468, doi:10.1038/s41586-020-2286-9 (2020).

1169 36 Trotman, J. B., Gilmier, A. J., Mukherjee, C. & Schoenberg, D. R. RNA guanine-7
1170 methyltransferase catalyzes the methylation of cytoplasmically recapped RNAs. *Nucleic
1171 Acids Res* **45**, 10726-10739, doi:10.1093/nar/gkx801 (2017).

1172 37 Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on
1173 UNIX pipes. *J Biomol NMR* **6**, 277-293, doi:10.1007/BF00197809 (1995).

1174 38 Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for
1175 biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325-1327,
1176 doi:10.1093/bioinformatics/btu830 (2015).

1177 39 Ogando, N. S. *et al.* The Enzymatic Activity of the nsp14 Exoribonuclease Is Critical for
1178 Replication of MERS-CoV and SARS-CoV-2. *J Virol* **94**, doi:10.1128/JVI.01246-20
1179 (2020).

1180 40 Mastronarde, D. N. Automated electron microscope tomography using robust prediction
1181 of specimen movements. *J Struct Biol* **152**, 36-51, doi:10.1016/j.jsb.2005.07.007 (2005).

1182 41 Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure
1183 determination. *J Struct Biol* **180**, 519-530, doi:10.1016/j.jsb.2012.09.006 (2012).

1184 42 Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for
1185 improved cryo-electron microscopy. *Nat Methods* **14**, 331-332, doi:10.1038/nmeth.4193
1186 (2017).

1187 43 Zhang, K. Gctf: Real-time CTF determination and correction. *J Struct Biol* **193**, 1-12,
1188 doi:10.1016/j.jsb.2015.11.003 (2016).

1189 44 Wagner, T. *et al.* SPHIRE-crYOLO is a fast and accurate fully automated particle picker
1190 for cryo-EM. *Commun Biol* **2**, 218, doi:10.1038/s42003-019-0437-z (2019).

1191 45 Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for
1192 rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296,
1193 doi:10.1038/nmeth.4169 (2017).

1194 46 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot.
1195 *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493
1196 (2010).

1197 47 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular
1198 structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221,
1199 doi:10.1107/S0907444909052925 (2010).

1200 48 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular
1201 crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21,
1202 doi:10.1107/S0907444909042073 (2010).

1203 49 Li, Z., Jaroszewski, L., Iyer, M., Sedova, M. & Godzik, A. FATCAT 2.0: towards a better
1204 understanding of the structural diversity of proteins. *Nucleic Acids Res* **48**, W60-W64,
1205 doi:10.1093/nar/gkaa443 (2020).

1206 50 Robert, X. & Gouet, P. Deciphering key features in protein structures with the new
1207 ENDscript server. *Nucleic Acids Res* **42**, W320-324, doi:10.1093/nar/gku316 (2014).

1208