

# 1        **Structure and dynamics of SARS-CoV-2 proofreading exoribonuclease ExoN**

2

3    Nicholas H. Moeller<sup>1,2,3†</sup>, Ke Shi<sup>1,2,3†</sup>, Özlem Demir<sup>4†</sup>, Surajit Banerjee<sup>5</sup>, Lulu Yin<sup>1,2,3</sup>, Christopher  
4    Belica<sup>1,2,3</sup>, Cameron Durfee<sup>1,2,3</sup>, Rommie E. Amaro<sup>4</sup>, Hideki Aihara<sup>1,2,3\*</sup>

5

6    <sup>1</sup>Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota,

7    Minneapolis, Minnesota, 55455, USA

8    <sup>2</sup>Institute for Molecular Virology, University of Minnesota, Minneapolis, Minnesota, 55455, USA

9    <sup>3</sup>Masonic Cancer Center, University of Minnesota, Minneapolis, Minnesota, 55455, USA

10   <sup>4</sup>Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093,

11   USA

12   <sup>5</sup>Northeastern Collaborative Access Team, Cornell University, Advanced Photon Source, Lemont, IL

13   60439, USA

14   † Co-first authors

15   \*Correspondence: [aihar001@umn.edu](mailto:aihar001@umn.edu)

16

## 17    **Abstract**

18    High-fidelity replication of the large RNA genome of coronaviruses (CoVs) is mediated by a 3'-to-5'

19    exoribonuclease (ExoN) in non-structural protein 14 (nsp14), which excises nucleotides including

20    antiviral drugs mis-incorporated by the low-fidelity viral RNA-dependent RNA polymerase (RdRp) and

21    has also been implicated in viral RNA recombination and resistance to innate immunity. Here we

22    determined a 1.6-Å resolution crystal structure of SARS-CoV-2 ExoN in complex with its essential co-

23    factor, nsp10. The structure shows a highly basic and concave surface flanking the active site,

24    comprising several Lys residues of nsp14 and the N-terminal amino group of nsp10. Modeling suggests

25    that this basic patch binds to the template strand of double-stranded RNA substrates to position the 3'

26 end of the nascent strand in the ExoN active site, which is corroborated by mutational and computational  
27 analyses. Molecular dynamics simulations further show remarkable flexibility of multi-domain nsp14  
28 and suggest that nsp10 stabilizes ExoN for substrate RNA-binding to support its exoribonuclease  
29 activity. Our high-resolution structure of the SARS-CoV-2 ExoN-nsp10 complex serves as a platform  
30 for future development of anti-coronaviral drugs or strategies to attenuate the viral virulence.

31

## 32 **Introduction**

33 The 29.9 kb single-stranded RNA genome of SARS-CoV-2, the causative agent of the global COVID-19  
34 pandemic, is replicated and transcribed by the viral RNA-dependent RNA polymerase (RdRp, nsp12)  
35 (1-3). Unlike the high-fidelity cellular replicative DNA polymerases, viral RdRp enzymes including the  
36 coronavirus (CoV) RdRp do not contain a proofreading exonuclease domain to ensure high fidelity. The  
37 resulting higher mutation rate ( $10^{-4}$  to  $10^{-6}$  substitutions/nucleotide/round of replication) is generally  
38 thought to promote rapid viral adaptation in response to selective pressure (4-6). However, the lack of  
39 proofreading activity in RdRp poses a particular challenge for the replication of coronaviruses, which  
40 feature the largest known RNA virus genomes (27 ~ 32 kb, up to twice the length as the next-largest  
41 non-segmented RNA viral genomes) (7, 8). It has been reported that SARS-CoV nsp12 is the fastest  
42 viral RdRp known but with an error rate more than one order of magnitude higher than the generally  
43 admitted error rate of viral RdRps (9), clearly necessitating a unique proofreading mechanism.

44 To mitigate the low fidelity of RdRp, all coronaviruses encode a 3'-to-5' exoribonuclease (ExoN)  
45 in nsp14 (10-12). Mutations of SARS-CoV-2 nsp14 exhibit strong association with increased genome-  
46 wide mutation load (13, 14), and genetic inactivation of ExoN in engineered SARS-CoV and murine  
47 hepatitis virus (MHV) leads to 15 to 20-fold increases in mutation rates (7, 15, 16). Furthermore, in a  
48 mouse model, SARS-CoV with inactivated ExoN shows a mutator phenotype with decreased fitness and  
49 lower virulence over serial passage, suggesting a potential strategy for generating a live, impaired-  
50 fidelity coronavirus vaccine (17). Alternatively, recent studies show that ExoN inactivation is lethal for

51 SARS-CoV-2 and Middle East Respiratory Syndrome (MERS)-CoV (18), hinting at additional functions  
52 for ExoN in viral replication. Indeed, the ExoN activity has been reported to mediate extensive viral  
53 RNA recombination required for subgenomic mRNA synthesis during normal replication of CoVs  
54 including SARS-CoV-2 (19), and it was shown to be required for resistance to the antiviral innate  
55 immune response for MHV (20). ExoN inactivation also significantly increases the sensitivity of CoVs  
56 to nucleoside analogs that target RdRp, which is consistent with the biochemical activity of ExoN to  
57 excise mutagenic or chain-terminating nucleotides mis-incorporated by RdRp (21-23). These  
58 observations combine to suggest that chemical inhibition of ExoN could be an effective antiviral  
59 strategy against CoVs. In this study, we determined a high-resolution crystal structure of the SARS-  
60 CoV-2 ExoN-nsp10 complex and studied its biochemical activities. Furthermore, we used molecular  
61 dynamics (MD) simulations to better understand the dynamics of nsp14, nsp10, and their interaction  
62 with RNA.

63

## 64 **Results**

65 The multifunctional SARS-CoV-2 nsp14 consists of the N-terminal ExoN domain involved in  
66 proofreading and the C-terminal guanine N7 methyl transferase (N7-MTase) domain that functions in  
67 mRNA capping. We co-expressed in bacteria the full-length 527-residue SARS-CoV-2 nsp14 or its N-  
68 terminal fragment (residues 1 to 289) containing only the ExoN domain, with full-length 139-residue  
69 nsp10 in both cases and purified the heterodimeric complexes. The nsp14-nsp10 and ExoN-nsp10  
70 complexes both showed the expected 3'-to-5' exonuclease activity on a 5'-fluorescently labeled 20-  
71 nucleotide (nt) RNA (LS2U: 5'-GUCAUUCUCCUAAGAAGCUU; similar to 'LS2' used previously in  
72 SARS-CoV ExoN studies (21)) (**Fig. 1A, B**). Although LS2U RNA by itself served as a substrate, more  
73 extensive degradation was observed when it was annealed to an unlabeled 40-nt template strand  
74 (LS15A\_RNA; **Table 1**) to generate a double-stranded (ds) RNA with a 20-nt 5'-overhang. Introducing  
75 a base-mismatch at the 3' end of the degradable strand by using an alternative bottom strand

76 (LS15\_RNA; **Table 1**) had no discernable effect on the processing by either complex (**Fig. 1A, B**).  
77 When DNA was used as the template strand (LS15\_DNA; **Table 1**) to generate an RNA/DNA  
78 heteroduplex substrate that is expected to take the A-form conformation similarly to dsRNA, the activity  
79 was observed but weaker than for dsRNA. No nuclease activity was observed on a 5'-fluorescently  
80 labeled 20-nt DNA (LS2\_DNA; **Table 1**), whether the template strand was RNA (LS15\_RNA), DNA  
81 (LS15\_DNA; **Table 1**), or absent. A 20-nt poly-U RNA (U20\_RNA; **Table 1**), which is less likely to  
82 adopt secondary structures than LS2U, did not serve as a substrate by itself but was degraded  
83 extensively when supplemented with a complementary 30-nt poly-A RNA (A30\_RNA; **Table 1**) (**Fig.**  
84 **1C**). Collectively, these results show that the N-terminal ExoN domain of SARS-CoV-2 nsp14 is  
85 sufficient for binding to nsp10 to form an active exoribonuclease complex that preferentially degrades  
86 dsRNA. For comparison, we also generated a corresponding SARS-CoV ExoN-nsp10 complex, which  
87 showed similar activities to SARS-CoV-2 ExoN-nsp10 (**Fig. 1C, Supplementary Fig. 1**).

88 Previous X-ray crystallographic studies have provided the structure of SARS-CoV nsp14-nsp10  
89 complex at resolutions ranging from 3.2 to 3.4 Å (21, 24). To obtain higher resolution view of a CoV  
90 exoribonuclease complex and to reveal possible structural difference between SARS-CoV and SARS-  
91 CoV-2 ExoN, we have crystallized the SARS-CoV-2 ExoN-nsp10 complex. An ExoN variant with a  
92 nuclease-inactivating mutation (E191Q) (**Fig. 1C, Supplementary Fig. 1**) was used in our  
93 crystallographic studies as it was expressed more robustly and generated a more stable complex with  
94 nsp10 than wild-type ExoN. We obtained crystals under two different conditions, one containing  
95 ammonium tartrate and the other containing magnesium chloride (MgCl<sub>2</sub>), albeit in the same crystal  
96 form. The structures were determined by molecular replacement phasing and refined to 1.64 and 2.10-Å  
97 resolution for the tartrate and magnesium-bound crystals, respectively (**Fig. 2A, Table 2**). The final  
98 models consist of nsp14 residues Asn3 to Arg289 (Val287 for the lower resolution structure) and nsp10  
99 residues Ala1 to Cys130, with two zinc ions bound to each polypeptide chain. As expected from high  
100 sequence conservations, SARS-CoV-2 ExoN-nsp10 complex shows high structural similarity to its

101 counterpart from SARS-CoV (root-mean-square deviation of 0.95 Å for all main chain atoms against  
102 5C8T (24)), whose shape was previously described to resemble ‘hand (ExoN) over a fist (nsp10)’ (21)  
103 (**Fig. 2B**). A superposition between the SARS-CoV and SARS-CoV-2 ExoN-nsp10 structures shows  
104 only relatively small (3.0 Å or less) deviations in several regions of the complex, including the tip of the  
105 ‘fingers’ region of ExoN comprising nsp14 residues 40 ~ 50, and surface-exposed loops of nsp10  
106 (**Supplementary Fig. 2**).

107 While our structures of SARS-CoV-2 ExoN-nsp10 obtained in the two different crystallization  
108 conditions are highly similar to each other, they show notable differences in the exonuclease active site  
109 located around the ‘knuckles’ of ExoN. In the crystal grown in the presence of MgCl<sub>2</sub>, we observed a  
110 magnesium ion octahedrally coordinated by Asp90, Glu92, Asp273, and three water molecules (**Fig. 2C**,  
111 **Supplementary Fig. 3d**). Another magnesium ion required for the conserved two-metal ion mechanism  
112 of 3'-5' editing exonucleases (25, 26) was not observed. The previously reported SARS-CoV nsp14-  
113 nsp10 structures also showed only one metal ion, bound at an alternative site between Asp90 and  
114 Glu191 (21, 24). This site is unoccupied in our structure presumably due to the E191Q mutation. In  
115 contrast, the higher resolution tartrate-bound structure shows a unique configuration of metal-free active  
116 site (**Fig. 2D**, **Supplementary Fig. 3c**). Without the magnesium ion, Asp90 takes two distinct  
117 conformers with its carboxylate group in orthogonal orientations. Glu92 is pointed away from  
118 Asp90/Asp273 and hydrogen-bonded to Gln108 side chain, whereas His268 in turn is flipped away from  
119 Glu92. A comparison between the Mg<sup>2+</sup>-bound and free structures shows a significant rearrangement for  
120 residues Gly265 to Val269 including the main chain atoms, accompanying an inward movement of  
121 His268 upon Mg<sup>2+</sup>-binding (**Fig. 2E**). These observations demonstrate high flexibility of the ExoN  
122 active site in the absence of divalent metal co-factors.

123 To obtain an idea about how SARS-CoV-2 ExoN-nsp10 complex engages RNA substrates, we  
124 modeled an RNA-bound ExoN-nsp10 structure based on the double-stranded (ds) RNA-bound structures  
125 of Lassa virus nucleoprotein (NP) exonuclease domain, which is another DEDDh-family 3'-to-5'

126 exoribonuclease. A superposition of the Lassa NP-RNA complex (27, 28) on ExoN-nsp10 based on their  
127 conserved catalytic residues (Lassa NP: D389/E391/D466/H528/D533 according to the numbering in  
128 4FVU (27), vs. SARS-CoV-2 ExoN: D90/E92/E191/H268/D273) places the A-form dsRNA in a  
129 shallow groove on ExoN surface adjacent to the active site, with remarkable shape complementarity  
130 (**Fig. 3 B, C**). In this model, the sugar-phosphate backbone of the non-degradable (template) RNA  
131 strand tracks a positively charged patch on the ExoN surface including Lys9 and Lys61, whereas the 3'  
132 end of its complementary (degradable) strand is presented to the active site. The extensive protein  
133 contacts made by the non-degradable strand in a dsRNA substrate is consistent with the preference for  
134 dsRNA substrates by SARS-CoV-2 ExoN as shown above (**Fig. 1**) and by SARS-CoV ExoN reported  
135 earlier (29). Notably, we observed ordered tartrate ions from the crystallization condition bound to this  
136 basic patch in our crystal structure, potentially mimicking RNA backbone phosphate interactions  
137 (**Supplementary Figs. 3a, 3b, and 4**).

138 Our hypothetical model described above suggests that the basic patch of ExoN helps position the  
139 substrate RNA for exonucleolytic degradation. Lys9 and Lys61 are involved in the RNA backbone  
140 interaction in our model. In addition, Lys139 is located farther down along the basic patch toward the  
141 direction of the 5'-overhang of the template strand (**Fig. 3A**). Thus, we tested the activities of SARS-  
142 CoV-2 ExoN with single amino acid substitutions, K9A, K61A, and K139A. These ExoN mutants were  
143 co-expressed with nsp10 and purified as heterodimeric complexes. In the exoribonuclease assay using  
144 the RNA substrates described above, all three lysine-to-alanine mutants showed lower activity than  
145 wild-type ExoN (**Fig. 4**). In particular, the K9A and K61A substitutions caused severer defect than  
146 K139A, consistent with our dsRNA-binding model (**Fig. 3 B, C**). While the precise conformation of  
147 LS2U RNA in the absence of a complementary strand is unknown, its binding to ExoN must also  
148 depend on these Lys residues, underscoring the importance of electrostatic interactions with RNA by the  
149 mutated lysine residues in the ExoN activity.

150 Previous studies showed that the exoribonuclease activity of nsp14 is strongly stimulated by  
151 nsp10 for both SARS-CoV and SARS-CoV-2 (29-32). In our crystal structure, the N-terminal residues  
152 of ExoN and those of nsp10 are wrapped around each other in a ‘criss-cross’ arrangement and forming  
153 several hydrogen-bond contacts, including one between nsp14 Lys9 and nsp10 Ala1 (**Supplementary**  
154 **Fig. 3a**). In addition, the first  $\alpha$ -helix of nsp10 interacts with the ExoN loop harboring nsp14 Lys61,  
155 where the main chain amide group of Lys61 is hydrogen-bonded to the side chain of nsp10 Ser15 (**Fig.**  
156 **3A**). In the absence of nsp10 supporting the RNA-binding groove from the back (**Fig. 3C**,  
157 **Supplementary Figs. 5, 6**), the N-terminal residues of ExoN including nsp14 Lys9 and those around  
158 Lys61 are likely to be more flexible. Moreover, the terminal amino group of nsp10 Ala1 is part of the  
159 basic patch and involved in direct RNA backbone contact in our protein-RNA docking model (**Fig. 3B**,  
160 **Supplementary Fig. 5**). These observations may together explain the strong stimulation of ExoN  
161 activity by nsp10.

162 To obtain further insights into the role of nsp10 and to support our RNA-binding model, we  
163 performed explicitly solvated, all-atom molecular dynamics (MD) simulations of full-length SARS-  
164 CoV-2 nsp14, constructed from our ExoN-nsp10 co-crystal structure and a homology model of the C-  
165 terminal N7-MTase domain. Three independent copies of MD simulations totaling 2.6- $\mu$ s were  
166 performed for each of nsp14 alone, nsp14-nsp10 complex, and the nsp14-nsp10-RNA complex based on  
167 our docking model described above. In addition, three independent copies of Gaussian-accelerated MD  
168 simulations (GAMD) totaling 0.6- $\mu$ s were performed for each system to enhance conformational  
169 sampling. Comparing trajectories of these simulations for the 3 systems, the most noticeable difference  
170 is an extreme flexibility of the ‘fingers’ region of ExoN primarily comprising its N-terminal residues  
171 (nsp14 residues 1-60), which showed large deviations from the starting model and eventually became  
172 highly disordered in the absence of nsp10. A principal component analysis for the 3 systems show that  
173 the conformational space sampled by nsp14 is significantly larger in the absence of nsp10 (**Fig. 5A**,  
174 **Supplementary Fig. 7**).

175 The first principal component (PC1), which is broadly sampled by all 3 systems, corresponds to  
176 a large hinge motion of the N7-MTase domain (~50 Å translocation at the distal end, **Supplementary**  
177 **Fig. 8, Supplementary animation 1**). In the conformation with minimal PC1 (**Fig. 5B, left**), the  
178 substrates (S-adenosyl methionine [SAM] and GpppA)-binding cleft of the N7-MTase domain abuts  
179 against the ExoN domain, leading to occlusion of the substrates. On the other extreme with maximal  
180 PC1, the cleft is more open to the solvent (**Fig. 5B, right**). The second principal component (PC2)  
181 corresponds to an ordered-to-disordered transition of the ‘fingers’ region of ExoN, which shows a large  
182 population of disordered conformations only for the nsp14-alone system as mentioned above (**Fig. 5C,**  
183 **Supplementary Fig. 9, Supplementary animations 2, 3**). Although folding of the core of the ExoN  
184 domain does not depend on nsp10, residues Lys9 and Lys61 important for RNA-binding and the  
185 surrounding residues show increased flexibility in the absence of nsp10, confirming our prediction  
186 above (**Table 3, Fig. 5 D, E, Supplementary Fig. 9, Supplementary animation 3**). The dsRNA  
187 molecule in the nsp14-nsp10-RNA complex was stable throughout the simulation with direct RNA  
188 phosphate contacts by nsp14 Lys9, Lys61, and the terminal amino group of nsp10 maintained, providing  
189 further support for our model for dsRNA-binding (**Fig. 5 F, G, Supplementary Fig. 10**). An ionic  
190 interaction between Ala1 of nsp10 and RNA backbone phosphate was particularly persistent and  
191 observed for 97 % of the time during the simulations (3.2 Å distance cutoff), which led to a significant  
192 stabilization of this residue in the presence of RNA (**Table 3**). Lastly, it is also worth noting that an  
193 analysis of the internal dynamics of the N7-MTase domain indicates several highly mobile regions,  
194 including loops (residues 289-300, 355-362) flanking the substrates-binding cleft and a loop (residues  
195 454-470) adjacent to the third zinc finger motif of nsp14 distal to the N7-MTase active site  
196 (**Supplementary Fig. 11**).

197

198 **Discussion**

199 Our X-ray crystallographic, biochemical, and computational analyses shed light on the substrate  
200 preference, structure, and dynamics of the SARS-CoV-2 ExoN-nsp10 exoribonuclease complex and  
201 further identified important roles of nsp10 in RNA substrate binding. It is particularly notable that the  
202 ExoN-nsp10 complex preferentially degrades dsRNA substrates. This is in contrast to the proofreading  
203 exonuclease domain of high-fidelity DNA polymerases, whose active site engages the single-stranded  
204 DNA 3' end in partially melted double-stranded substrates (25, 33), and suggests a unique mechanism of  
205 proofreading. The extensive ExoN/nsp10 interface buries a total of 2203 Å<sup>2</sup> of surfaces from both  
206 proteins, spanning both the ‘fingers’ and ‘palm’ regions of ExoN. Folding of the fingers region depends  
207 on its interaction with nsp10, which involves several critical residues including nsp10 Tyr96 (31) (**Fig.**  
208 **2A, Supplementary Fig. 6**). On the other hand, an interesting feature for the interaction in the palm  
209 region includes the insertion of Phe16 and Phe19 from the first α-helix of nsp10 into a deep  
210 hydrophobic pocket of ExoN, which is essential for the stable complex formation (31). Notably, this  
211 hydrophobic pocket is located on the backside from the ExoN active site, where nsp10 Phe19 side chain  
212 makes van der Waals contacts with the main chain of an ExoN α-helix harboring one of the catalytic  
213 residues Glu191 (**Supplementary Fig. 6**). Thus, targeting said pocket of ExoN by small molecules to  
214 block its interaction with nsp10 or potentially to allosterically modulate its catalytic activity could be a  
215 possible strategy of inhibition.

216 MD simulations revealed remarkable flexibility in full-length nsp14 (**Supplementary Figs. 7, 8,**  
217 **and Supplementary animations 1, 2**), which affects solvent accessibility of the SAM/GpppA-binding  
218 cleft and may play an important role in the catalytic cycle of N7-MTase (**Fig. 5B**). Similar  
219 conformational variation, albeit with a much smaller magnitude, was previously observed between two  
220 SARS-CoV nsp14 molecules in the asymmetric unit of a crystal (**Supplementary Fig. 8**) (21). Although  
221 this hinge motion was observed for all 3 systems (nsp14-alone, nsp14-nsp10, and nsp14-nsp10-RNA) in  
222 our simulations, they showed different distributions of the PC1 value (**Supplementary Fig. 7**). In  
223 addition, conformational sampling in the nsp14-alone system shows several clusters with distinct

224 combinations of PC1 and PC2 values (**Fig. 5A, left**), suggesting that there may be a long-range  
225 interaction between the N-terminal fingers region of ExoN and the C-terminal N7-MTase domain. These  
226 observations are consistent with earlier studies showing that single amino acid substitutions R84A and  
227 W86A within the ExoN domain completely abolished, while a deletion of the N-terminal 61 residues  
228 significantly enhanced, the N7-MTase activity of SARS-CoV nsp14 (34). These mutations in ExoN may  
229 have modulated the PC1 motion of nsp14 to affect its N7-MTase activity. Conversely, although we  
230 showed in this study that the N7-MTase domain is not essential for the exoribonuclease activity of nsp14  
231 *in vitro*, SARS-CoV nsp14 N7-MTase residues Tyr498 and His487 were shown to be required for  
232 RdRp/nsp12 binding (21), which is presumably important in proofreading. Thus, it is likely that the  
233 ExoN and N7-MTase domains are functionally dependent on each other, where proper dynamics may be  
234 key to support their respective activities and possible coordination. We hope that our structural and  
235 functional studies will help future development of ExoN inhibitors to impede the replication of SARS-  
236 CoV-2 and related coronaviruses.

237

## 238 **Methods**

### 239 *Protein expression and purification*

240 SARS-CoV-2 (GenBank: MN908947.3) nsp14 and nsp14(1-289) were co-expressed with nsp10 in *E.*  
241 *coli* strain BL21(DE3) under the control of T7 promoters. To facilitate purification, a 6xHis tag was  
242 added to the N-terminus of nsp14 and nsp14(1-289) with a human rhinovirus (HRV) 3C protease  
243 cleavage site. A methionine residue was added to nsp10 to enable translation. Transformed bacteria were  
244 cultured in LB medium at 37 °C to the mid-log phase, induced with 0.5 mM and 50 μM (final  
245 concentrations) of Isopropyl β-D-1-thiogalactopyranoside and zinc chloride, respectively, and further  
246 incubated at 18 °C overnight before being pelleted by centrifugation. Collected bacteria were disrupted  
247 by the addition of hen egg white lysozyme and sonication in 20 mM Tris-HCl, pH 7.4, 0.5 M NaCl, 5  
248 mM β-mercaptoethanol, and 5 mM imidazole. The lysate was cleared by centrifugation at 63,000 x g for

249 1 hour at 4 °C, after which the protein complex in the supernatant was captured by nickel-affinity  
250 chromatography and eluted by a linear gradient of imidazole. Eluted proteins were digested with HRV  
251 3C protease overnight at 4 °C, concentrated by ultrafiltration, and passed through a Superdex75 size-  
252 exclusion column operating with the same buffer as above except not containing imidazole. The nsp14-  
253 nsp10 complexes eluted as a heterodimer were concentrated by ultrafiltration and frozen in small  
254 volume aliquots in liquid nitrogen for storage at -80 °C. The ExoN mutant derivatives were generated by  
255 site-directed mutagenesis and purified using the same procedure. The protein concentrations were  
256 determined based on UV absorbance at 280 nm measured on a Nanodrop8000 spectrophotometer and  
257 theoretical extinction coefficients calculated from the protein amino acid sequences.

258

### 259 *Crystallization and structure determination*

260 Purified nsp14(1-289, E191Q)-nsp10 complex (**Supplementary Fig. 12**) at 17 mg ml<sup>-1</sup> was crystallized  
261 using the hanging drop vapor diffusion method, by mixing the protein solution with an equal volume of  
262 reservoir solution including either 0.2 M di-ammonium tartrate, pH 7.0, 20 % polyethylene glycol (PEG)  
263 3,350 (condition 1), or 0.1 M MgCl<sub>2</sub>, 0.1 M Tris-HCl pH 8.5, 20 % PEG 4,000 (condition 2). Both  
264 conditions produced thin needles crystals. The crystals were cryo-protected with ethylene glycol and  
265 flash-cooled by plunging in liquid nitrogen. X-ray diffraction data were collected at the Northeastern  
266 Collaborative Access Team (NE-CAT) beamlines of the Advanced Photon Source (Lemont, IL) and  
267 processed using XDS (35). The structure of the SARS-CoV-2 ExoN-nsp10 complex was determined by  
268 molecular replacement phasing by PHASER (36), using the crystal structures of SARS-CoV nsp14-  
269 nsp10 complex (PDB ID: 5C8T) (24) as the search model. Iterative model building and refinement were  
270 performed using COOT (37) and PHENIX (38), respectively. A summary of data collection and model  
271 refinement statistics is shown in **Table 2**. Structure images were generated using PyMOL  
272 (<https://pymol.org/>).

273

274 *Exonuclease activity assays*

275 The 5'-fluorescein labeled oligonucleotides (**Table 1**) at 750 nM, in the presence or absence of  
276 equimolar complementary unlabeled strands, were incubated with 50 nM nsp14 (or its ExoN domain  
277 alone)-nsp10 complexes in 42 mM Tris-HCl, pH 8.0, 0.94 mM MgCl<sub>2</sub>, 0.94 mM dithiothreitol, and  
278 0.009 % Tween-20. After incubation at 37 °C for 10 min, the reactions were stopped by the addition of  
279 formamide to 67 % and heating to 95°C for 10 min. The reaction products were separated on a 15 %  
280 TBE-Urea gel, which was scanned on a Typhoon FLA 9500 imager.

281

282 *Molecular dynamics simulations*

283 A homology model of full-length SARS-CoV-2 nsp14 was generated for sequence of YP\_009725309.1  
284 and taking SARS-CoV nsp14 crystal structure (PDB ID: 5NFY) (21) as a template in Schrödinger Prime  
285 module (39). The nsp14 ExoN domain of the homology model was then replaced with the crystal  
286 structure of SARS-CoV-2 nsp14 ExoN in complex with nsp10 obtained in this study. E191Q mutation  
287 in the crystal structure was reverted computationally to the wild type. For a nsp14-nsp10-RNA model,  
288 RNA was modeled based on Lassa NP-RNA complex (PDB ID: 4FVU) (27) and the second Mg ion at  
289 the active site was modeled based on a Mn<sup>2+</sup> ion found in Lassa NP-RNA complex (PDB ID: 4GV9)  
290 (28). Three systems were prepared from this model: 1. Full-length nsp14 alone, 2. Full-length nsp14-  
291 nsp10 complex, 3. Full-length nsp14-nsp10-RNA complex. Protonation states of titratable amino acids  
292 were determined using PropKa analysis (40). Each of these systems were explicitly solvated in TIP3P  
293 water box and ions were added to achieve 0.2 M salt concentration. Amber ff14SB (41) and RNA.OL3  
294 force fields are used for protein and RNA, respectively. For zinc ions and zinc-coordinating residues,  
295 Cationic Dummy Atom (CADA) parameters were used (42). Conventional MD simulations (cMD) were  
296 performed with NAMD2.14 program (43), while Gaussian-accelerated MD simulations (GAMD) were  
297 performed with Amber20 program (44). First, each system was minimized in 4 consecutive steps by  
298 gradually decreasing restraints. Subsequently, each system was heated from 0 to 310 K slowly, and then

299 equilibrated for about 1 ns by gradually decreasing restraints in 3 consecutive steps. For cMD, three  
300 independent copies (2x 1  $\mu$ s and 1x 0.6  $\mu$ s) of simulation were run for each system. For GAMD, three  
301 independent copies of 0.2  $\mu$ s of simulation were run for each system using dual boost method following  
302 a 20-ns MD run to calculate parameters for GAMD production runs. All cMD and GAMD simulations  
303 were performed at 310 K and 1 atm and with a 2 fs timestep. For each system, 32,000 data points with  
304 0.1 ns intervals were collected from simulations and analyzed. Stability of MD simulations are shown  
305 with RMSD plots of nsp14 ExoN domain (**Supplementary Fig. 13**). MDTraj (45) was used for some of  
306 the MD trajectory analysis.

307

### 308 **Acknowledgements**

309 We thank Daniel Harki and Reuben Harris for thoughtful comments. This work was supported by grants  
310 from the US National Institutes of Health (NIGMS R35-GM118047 to H.A., R01-GM132826 to R.E.A.,  
311 and NCI P01-CA234228 to R.E.A. and H.A.), NSF RAPID MCB-2032054, an award from the RCSA  
312 Research Corp., and a UC San Diego Moores Cancer Center 2020 SARS-COV-2 seed grant to R.E.A.  
313 This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines,  
314 which are funded by the US National Institutes of Health (NIGMS P30 GM124165). The Pilatus 6M  
315 detector on 24-ID-C beamline is funded by a NIH-ORIP HEI grant (S10 RR029205). This research used  
316 resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User  
317 Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No.  
318 DE-AC02-06CH11357 and those of the Minnesota Supercomputing Institute. We are grateful for the  
319 efforts of the Texas Advanced Computing Center (TACC) Frontera team and for the compute time made  
320 available through a Director's Discretionary Allocation (made possible by the National Science  
321 Foundation award OAC-1818253).

322

### 323 **Data Availability**

324 The atomic coordinates and structure factors for the SARS-CoV-2 ExoN-nsp10 complex structures have  
325 been deposited in the RCSB Protein Data Bank, with the accession codes 7MC5 and 7MC6.

326

### 327 **Competing Interests**

328 The authors have no competing interests to declare.

329

### 330 **Reference:**

- 331 1. Hillen HS, *et al.* (2020) Structure of replicating SARS-CoV-2 polymerase. *Nature* 584(7819):154-  
332 156.
- 333 2. Posthuma CC, Te Velhuis AJW, & Snijder EJ (2017) Nidovirus RNA polymerases: Complex  
334 enzymes handling exceptional RNA genomes. *Virus Res* 234:58-73.
- 335 3. Sola I, Almazan F, Zuniga S, & Enjuanes L (2015) Continuous and Discontinuous RNA Synthesis in  
336 Coronaviruses. *Annu Rev Virol* 2(1):265-288.
- 337 4. Drake JW & Holland JJ (1999) Mutation rates among RNA viruses. *Proc Natl Acad Sci U S A*  
338 96(24):13910-13913.
- 339 5. Jenkins GM, Rambaut A, Pybus OG, & Holmes EC (2002) Rates of molecular evolution in RNA  
340 viruses: a quantitative phylogenetic analysis. *J Mol Evol* 54(2):156-165.
- 341 6. Sanjuan R, Nebot MR, Chirico N, Mansky LM, & Belshaw R (2010) Viral mutation rates. *J Virol*  
342 84(19):9733-9748.
- 343 7. Denison MR, Graham RL, Donaldson EF, Eckerle LD, & Baric RS (2011) Coronaviruses: an RNA  
344 proofreading machine regulates replication fidelity and diversity. *RNA Biol* 8(2):270-279.
- 345 8. Gorbalenya AE, Enjuanes L, Ziebuhr J, & Snijder EJ (2006) Nidovirales: evolving the largest RNA  
346 virus genome. *Virus Res* 117(1):17-37.
- 347 9. Shannon A, *et al.* (2020) Rapid incorporation of Favipiravir by the fast and permissive viral RNA  
348 polymerase complex results in SARS-CoV-2 lethal mutagenesis. *Nat Commun* 11(1):4682.
- 349 10. Minskaia E, *et al.* (2006) Discovery of an RNA virus 3'->5' exoribonuclease that is critically  
350 involved in coronavirus RNA synthesis. *Proc Natl Acad Sci U S A* 103(13):5108-5113.
- 351 11. Robson F, *et al.* (2020) Coronavirus RNA Proofreading: Molecular Basis and Therapeutic  
352 Targeting. *Mol Cell* 79(5):710-727.
- 353 12. Smith EC & Denison MR (2013) Coronaviruses as DNA wannabes: a new model for the  
354 regulation of RNA virus replication fidelity. *PLoS Pathog* 9(12):e1003760.
- 355 13. Eskier D, Suner A, Oktay Y, & Karakulah G (2020) Mutations of SARS-CoV-2 nsp14 exhibit strong  
356 association with increased genome-wide mutation load. *PeerJ* 8:e10181.
- 357 14. Takada K, Takahashi Ueda M, Watanabe T, & Nakagawa S (2020) Genomic diversity of SARS-  
358 CoV-2 can be accelerated by a mutation in the nsp14 gene. *bioRxiv* 2020.12.23.424231; doi:  
359 <https://doi.org/10.1101/2020.12.23.424231>
- 360 15. Eckerle LD, *et al.* (2010) Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is  
361 revealed by complete genome sequencing. *PLoS Pathog* 6(5):e1000896.
- 362 16. Eckerle LD, Lu X, Sperry SM, Choi L, & Denison MR (2007) High fidelity of murine hepatitis virus  
363 replication is decreased in nsp14 exoribonuclease mutants. *J Virol* 81(22):12135-12144.

- 364 17. Graham RL, *et al.* (2012) A live, impaired-fidelity coronavirus vaccine protects in an aged,  
365 immunocompromised mouse model of lethal disease. *Nat Med* 18(12):1820-1826.
- 366 18. Ogando NS, *et al.* (2020) The Enzymatic Activity of the nsp14 Exoribonuclease Is Critical for  
367 Replication of MERS-CoV and SARS-CoV-2. *J Virol* 94(23).
- 368 19. Gribble J, *et al.* (2021) The coronavirus proofreading exoribonuclease mediates extensive viral  
369 recombination. *PLoS Pathog* 17(1):e1009226.
- 370 20. Case JB, *et al.* (2018) Murine Hepatitis Virus nsp14 Exoribonuclease Activity Is Required for  
371 Resistance to Innate Immunity. *J Virol* 92(1).
- 372 21. Ferron F, *et al.* (2018) Structural and molecular basis of mismatch correction and ribavirin  
373 excision from coronavirus RNA. *Proc Natl Acad Sci U S A* 115(2):E162-E171.
- 374 22. Agostini ML, *et al.* (2018) Coronavirus Susceptibility to the Antiviral Remdesivir (GS-5734) Is  
375 Mediated by the Viral Polymerase and the Proofreading Exoribonuclease. *mBio* 9(2).
- 376 23. Smith EC, Blanc H, Surdel MC, Vignuzzi M, & Denison MR (2013) Coronaviruses lacking  
377 exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and  
378 potential therapeutics. *PLoS Pathog* 9(8):e1003565.
- 379 24. Ma Y, *et al.* (2015) Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10  
380 complex. *Proc Natl Acad Sci U S A* 112(30):9436-9441.
- 381 25. Beese LS & Steitz TA (1991) Structural basis for the 3'-5' exonuclease activity of Escherichia coli  
382 DNA polymerase I: a two metal ion mechanism. *EMBO J* 10(1):25-33.
- 383 26. Chen P, *et al.* (2007) Biochemical characterization of exoribonuclease encoded by SARS  
384 coronavirus. *J Biochem Mol Biol* 40(5):649-655.
- 385 27. Hastie KM, King LB, Zandonatti MA, & Saphire EO (2012) Structural basis for the dsRNA  
386 specificity of the Lassa virus NP exonuclease. *PLoS One* 7(8):e44211.
- 387 28. Jiang X, *et al.* (2013) Structures of arenaviral nucleoproteins with triphosphate dsRNA reveal a  
388 unique mechanism of immune suppression. *J Biol Chem* 288(23):16949-16959.
- 389 29. Bouvet M, *et al.* (2012) RNA 3'-end mismatch excision by the severe acute respiratory  
390 syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. *Proc Natl*  
391 *Acad Sci U S A* 109(24):9372-9377.
- 392 30. Bouvet M, *et al.* (2014) Coronavirus Nsp10, a critical co-factor for activation of multiple  
393 replicative enzymes. *J Biol Chem* 289(37):25783-25796.
- 394 31. Saramago M, *et al.* (2021) New targets for drug design: Importance of nsp14/nsp10 complex  
395 formation for the 3'-5' exoribonucleolytic activity on SARS-CoV-2. *bioRxiv* 2021.01.07.425745;  
396 doi: <https://doi.org/10.1101/2021.01.07.425745>
- 397 32. Baddock HT, *et al.* (2020) Characterisation of the SARS-CoV-2 ExoN (nsp14ExoN-nsp10)  
398 complex: implications for its role in viral genome stability and inhibitor identification. *bioRxiv*  
399 2020.08.13.248211; doi: <https://doi.org/10.1101/2020.08.13.248211>.
- 400 33. Beese LS, Derbyshire V, & Steitz TA (1993) Structure of DNA polymerase I Klenow fragment  
401 bound to duplex DNA. *Science* 260(5106):352-355.
- 402 34. Chen Y, *et al.* (2013) Structure-function analysis of severe acute respiratory syndrome  
403 coronavirus RNA cap guanine-N7-methyltransferase. *J Virol* 87(11):6296-6305.
- 404 35. Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):125-132.
- 405 36. McCoy AJ, *et al.* (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658-674.
- 406 37. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta*  
407 *Crystallogr D Biol Crystallogr* 66(Pt 4):486-501.
- 408 38. Liebschner D, *et al.* (2019) Macromolecular structure determination using X-rays, neutrons and  
409 electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* 75(Pt 10):861-877.

- 410 39. Schrödinger (2021) Schrödinger Release 2021-1: Prime, Schrödinger, LLC, New York, NY.  
411 40. Dolinsky TJ, Nielsen JE, McCammon JA, & Baker NA (2004) PDB2PQR: an automated pipeline for  
412 the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* 32(Web Server  
413 issue):W665-667.  
414 41. Maier JA, *et al.* (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone  
415 Parameters from ff99SB. *J Chem Theory Comput* 11(8):3696-3713.  
416 42. Pang YP (1999) Novel Zinc Protein Molecular Dynamics Simulations: Steps Toward  
417 Antiangiogenesis for Cancer Treatment. *J Mol Model* 5:196–202.  
418 43. Phillips JC, *et al.* (2020) Scalable molecular dynamics on CPU and GPU architectures with NAMD.  
419 *J Chem Phys* 153(4):044130.  
420 44. Case DA, *et al.* (2020) AMBER 2020, University of California, San Francisco.  
421 45. McGibbon RT, *et al.* (2015) MDTraj: A Modern Open Library for the Analysis of Molecular  
422 Dynamics Trajectories. *Biophys J* 109(8):1528-1532.  
423 46. Baker NA, Sept D, Joseph S, Holst MJ, & McCammon JA (2001) Electrostatics of nanosystems:  
424 application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* 98(18):10037-10041.  
425

426 **Table 1: Oligonucleotides used in biochemical assays**

427	<b>LS2U</b>
428	/56-FAM/rGrUrCrArUrUrCrUrCrCrUrArArGrArArGrCrUrU
429	<b>LS15A_RNA</b>
430	rCrUrArUrCrCrCrCrArUrGrUrGrArUrUrUrUrArCrArArGrCrUrUrCrUrUrArGrGrArGrArArUrGrArC
431	<b>LS15_RNA</b>
432	rCrUrArUrCrCrCrCrArUrGrUrGrArUrUrUrUrArCrUrArGrCrUrUrCrUrUrArGrGrArGrArArUrGrArC
433	<b>LS2_DNA</b>
434	/56-FAM/GTCATTCTCCTAAGAAGCTA
435	<b>LS15_DNA</b>
436	CTATCCCCATGTGATTTTACTAGCTTCTTAGGAGAATGAC
437	<b>U20_RNA</b>
438	/56-FAM/rU
439	<b>A30_RNA</b>
440	rA
441	'56-FAM' denotes 5' 6-Fluorescein. 'r' denotes ribonucleotide.

442 **Table 2: Summary of X-ray data collection and model refinement statistics**

443

	ExoN-nsp10 (7MC5)	ExoN-nsp10-Mg <sup>2+</sup> (7MC6)
<b>Data collection</b>		
Wavelength (Å)	0.979	0.979
Resolution range (Å)	57.7 - 1.64 (1.70 - 1.64)	42.6 - 2.10 (2.18 - 2.10)
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Unit cell ( <i>a,b,c</i> in Å)	63.74 67.48 111.25	61.67 70.32 108.54
Total reflections	258196 (22096)	105896 (10815)
Unique reflections	58702 (5273)	27756 (2767)
Multiplicity	4.4 (4.2)	3.8 (3.9)
Completeness (%)	98.81 (90.43)	98.25 (99.43)
$\langle I/\sigma(I) \rangle$	12.57 (1.48)	10.70 (1.96)
R <sub>merge</sub>	0.148 (1.22)	0.078 (0.928)
R <sub>meas</sub>	0.166 (1.40)	0.091 (1.082)
R <sub>p.i.m.</sub>	0.076 (0.660)	0.045 (0.543)
CC <sub>1/2</sub>	0.995 (0.394)	0.997 (0.524)
<b>Refinement</b>		
Reflections, working set	58626 (5273)	27755 (2768)
Reflections, test set	2826 (251)	1364 (132)
R <sub>work</sub>	0.166 (0.354)	0.197 (0.306)
R <sub>free</sub>	0.197 (0.371)	0.219 (0.346)
No. of non-H atoms	3890	3447
Macromolecules	3264	3221
Ligands	117	42
Solvent	509	184
Protein residues	417	415
R.m.s. deviations		
Bond length (Å)	0.011	0.001
Bond angles (°)	1.10	0.41
Ramachandran plot		
Favored (%)	96.85	96.84
Allowed (%)	2.91	2.92
Outliers (%)	0.24	0.24
Average B factor (Å <sup>2</sup> )		
Macromolecules	26.61	44.43
Ligands	24.60	44.11
Solvent	37.61	54.76
	36.94	47.72

444 Statistics for the highest-resolution shell are shown in parentheses.

445

446 **Table 3: Root-mean-square fluctuations (RMSF, in Å) of the catalytic residues and RNA-binding**  
447 **residues in the 3 simulated systems.**

448 RMSF of C $\alpha$  atoms were calculated after aligning trajectories to the initial model with respect to C $\alpha$   
449 atoms of residues 71-289 (core of the ExoN domain). RMSF of all atoms for each residue is presented in  
450 parenthesis. Catalytic residues of ExoN are underlined.

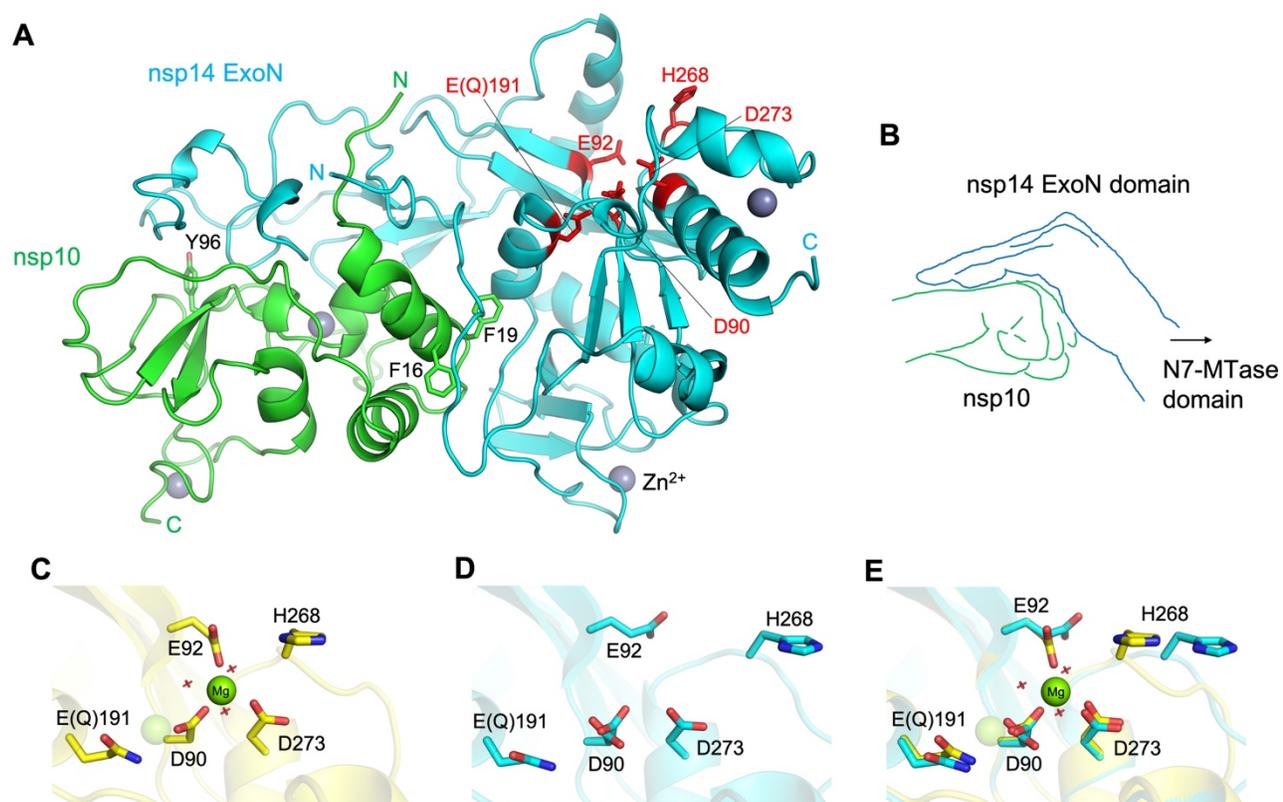
451

	<b>nsp14</b>	<b>nsp14-nsp10</b>	<b>nsp14-nsp10-RNA</b>
<b><u>D90</u></b> (nsp14)	0.44 (0.56)	0.40 (0.52)	0.35 (0.39)
<b><u>E92</u></b> (nsp14)	0.61 (1.18)	0.63 (1.16)	0.39 (0.76)
<b><u>E191</u></b> (nsp14)	0.58 (0.76)	0.56 (0.75)	0.38 (0.61)
<b><u>H268</u></b> (nsp14)	1.75 (2.33)	1.66 (2.21)	1.34 (2.05)
<b><u>D273</u></b> (nsp14)	0.58 (0.97)	0.60 (0.92)	0.39 (0.45)
<b>K9</b> (nsp14)	1.80 (2.59)	1.16 (1.62)	0.55 (0.67)
<b>K61</b> (nsp14)	2.81 (3.49)	1.60 (2.26)	0.73 (1.25)
<b>K139</b> (nsp14)	0.95 (1.56)	0.86 (1.52)	0.62 (1.18)
<b>A1</b> (nsp10)		4.04 (4.12)	0.75 (0.82)

452

453

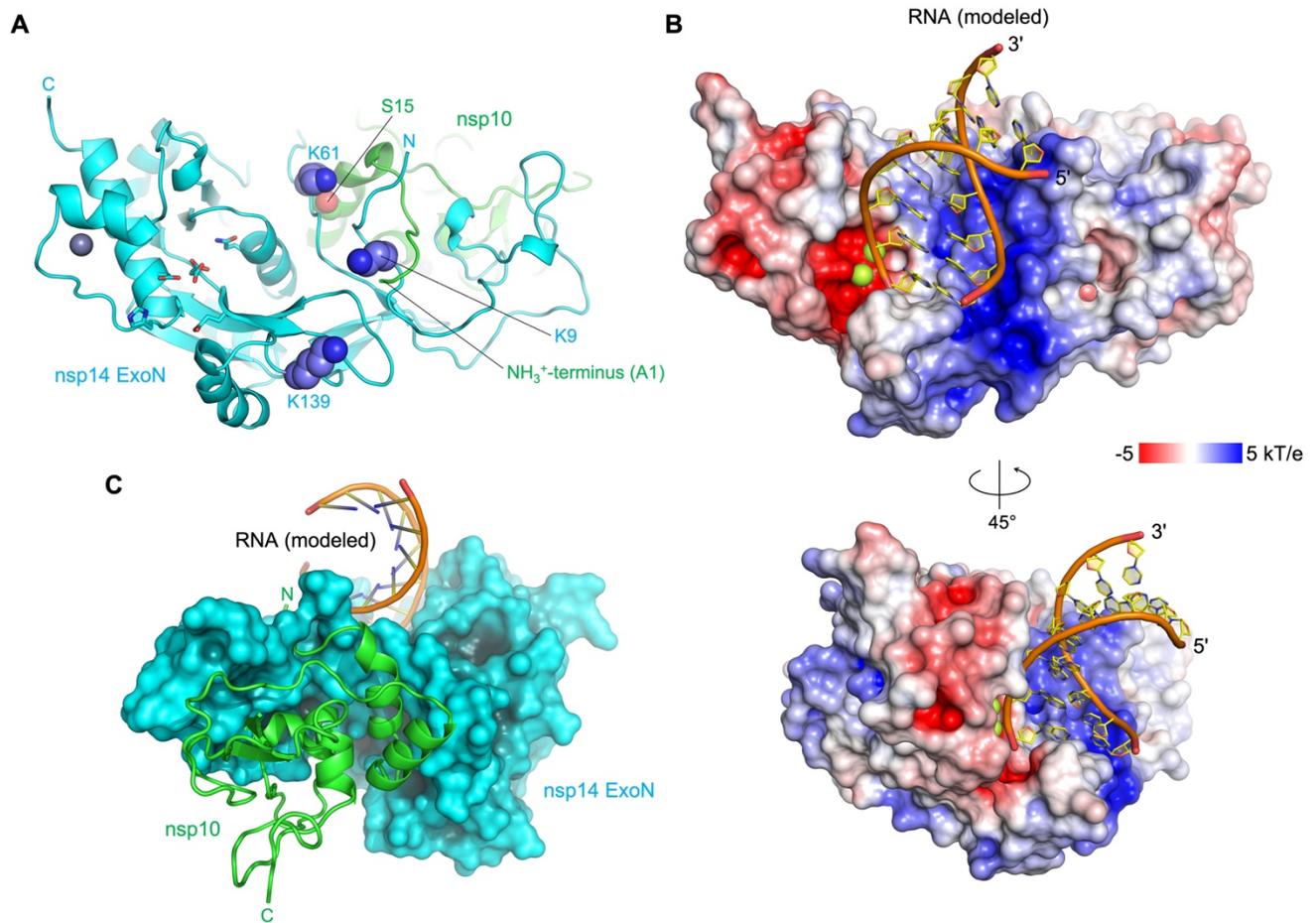




463

464 **Fig. 2. SARS-CoV-2 ExoN-nsp10 structure and its active site flexibility**

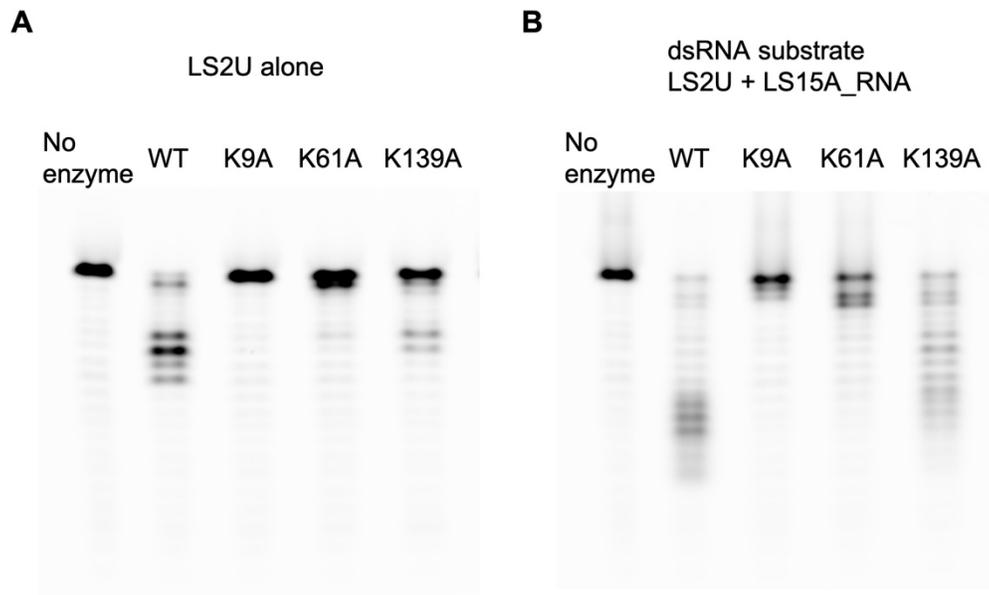
465 **A**, Overall structure of nsp14(1-289)-nsp10 complex. The N-terminal ExoN domain of nsp14 is shown  
466 in cyan and nsp10 in green. The ExoN active site residues are highlighted as red sticks. Key aromatic  
467 residues of nsp10 in the protein-protein interface are also shown as sticks. Gray spheres represent zinc  
468 ions. **B**, A schematic illustration of hand (ExoN) over a fist (nsp10). **C**, ExoN active site in the presence  
469 of Mg<sup>2+</sup>. The magnesium ion is shown as a solid sphere scaled at half the van der Waals radius. The  
470 second Mg<sup>2+</sup>-binding site, indicated by a transparent sphere, is unoccupied in our structure presumably  
471 due to the E191Q mutation. **D**, Mg<sup>2+</sup>-free active site as observed in the tartrate-bound crystal. Asp90  
472 side chain shows a dual conformation. **E**, Superposition of **C** and **D**, highlighting the conformational  
473 changes upon Mg<sup>2+</sup>-binding.



474

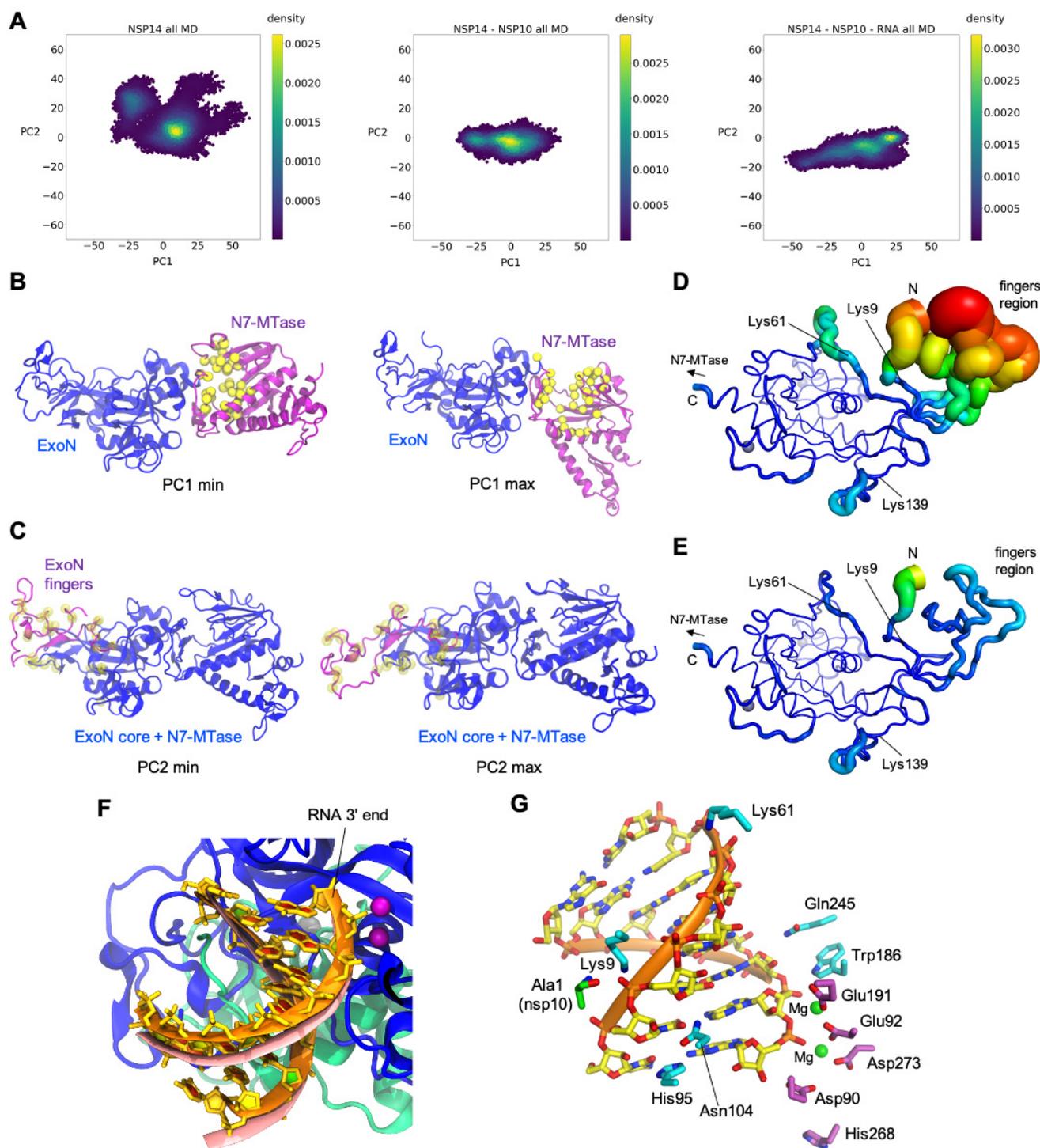
475 **Fig. 3. Location of the basic patch and an ExoN-nsp10-RNA complex model**

476 **A**, Locations of the ExoN lysine residues forming the basic patch. Note that Lys9 and Lys61 interact  
477 with the N-terminus (Ala1) and Ser15 of nsp10, respectively. **B**, A hypothetical model of ExoN-nsp10-  
478 dsRNA complex, viewed from two different orientations. The protein surface is colored according to the  
479 electrostatic potential calculated using APBS (46). **C**, Backside of the ExoN-nsp10-dsRNA model,  
480 viewed from the ExoN-nsp10 interface. Nsp10 is shown as green ribbon.



481 **Fig. 4. Activities of ExoN lysine mutants**

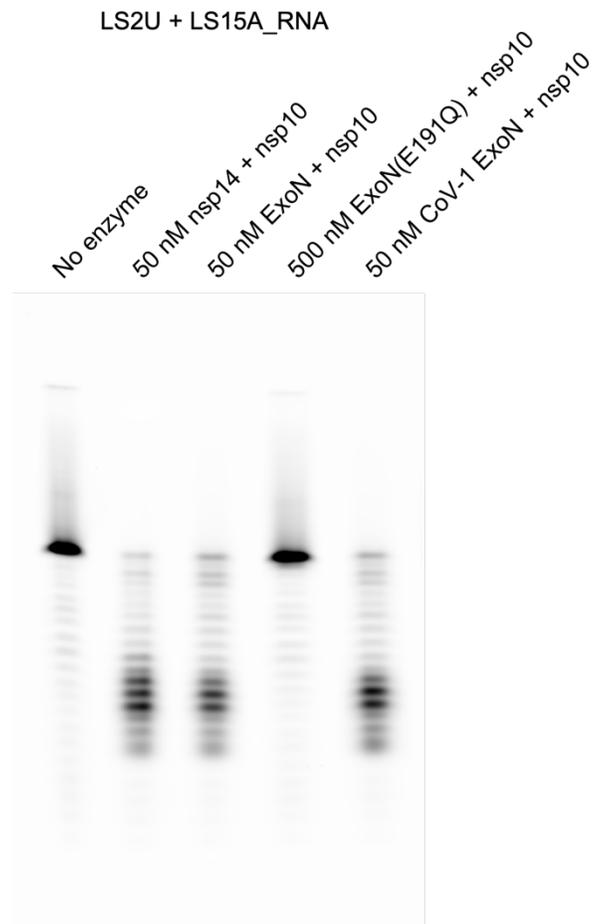
482 Exoribonuclease activities of SARS-CoV-2 ExoN-nsp10 complex and its lysine-to-alanine point mutant  
483 derivatives. **A**, Processing of LS2U RNA without a complementary strand. **B**, LS2U RNA annealed with  
484 the fully complementary LS15A RNA (dsRNA substrate). Please see **Table 1** for the substrate  
485 sequences.



486 **Fig. 5. MD simulations.**

487 **A**, Principal component analysis depicting differential conformational sampling for the 3 systems in MD  
 488 simulations. **B**, Structures that correspond to PC1 minimum and maximum values for the nsp14-alone  
 489 system. N7-MTase and ExoN domains of nsp14 are depicted in purple and blue ribbons, respectively.  
 490 Yellow spheres represent the C $\alpha$  atoms of residues that constitute the binding site of SAM and GpppA

491 substrates of N7-MTase based on homology to SARS-CoV nsp14 N7-MTase crystal structures (PDB  
492 ID: 5C8S and 5C8T)(24). **C**, Structures that correspond to PC2 minimum and maximum values for the  
493 nsp14-alone system. N-terminal region (residues 1-71) of nsp14 is depicted in purple ribbons while the  
494 rest of nsp14 is depicted in blue ribbons. Transparent yellow spheres represent the C $\alpha$  atoms of nsp14  
495 residues that constitute nsp10 binding site. **D**, ExoN domain in nsp14-alone system with root-mean-  
496 square fluctuations (RMSF) of C $\alpha$  atoms depicted on the structure with varying tube thickness and color  
497 (low in blue to high in red). The view is similar to that in **Fig. 3A**. **E**, ExoN domain of nsp14-nsp10  
498 system with C $\alpha$  RMSF depicted on the structure with varying tube thickness and color. **F**, RNA after 1  
499  $\mu$ s MD simulation (in orange ribbons) of nsp14-nsp10-RNA system superimposed onto RNA of the  
500 starting model (salmon). Nsp14 and nsp10 are depicted as blue and green ribbons, respectively. Dark  
501 purple spheres represent two Mg ions in the active site. **G**, RNA after 1  $\mu$ s MD simulation of the nsp14-  
502 nsp10-RNA system, with nsp14 ExoN domain (cyan) or nsp10 (green) residues making persistent  
503 hydrogen-bond or salt bridge interactions with RNA in MD simulations shown as sticks. The active site  
504 residues of ExoN are also shown (purple sticks) with two Mg<sup>2+</sup> ions as green spheres.



505

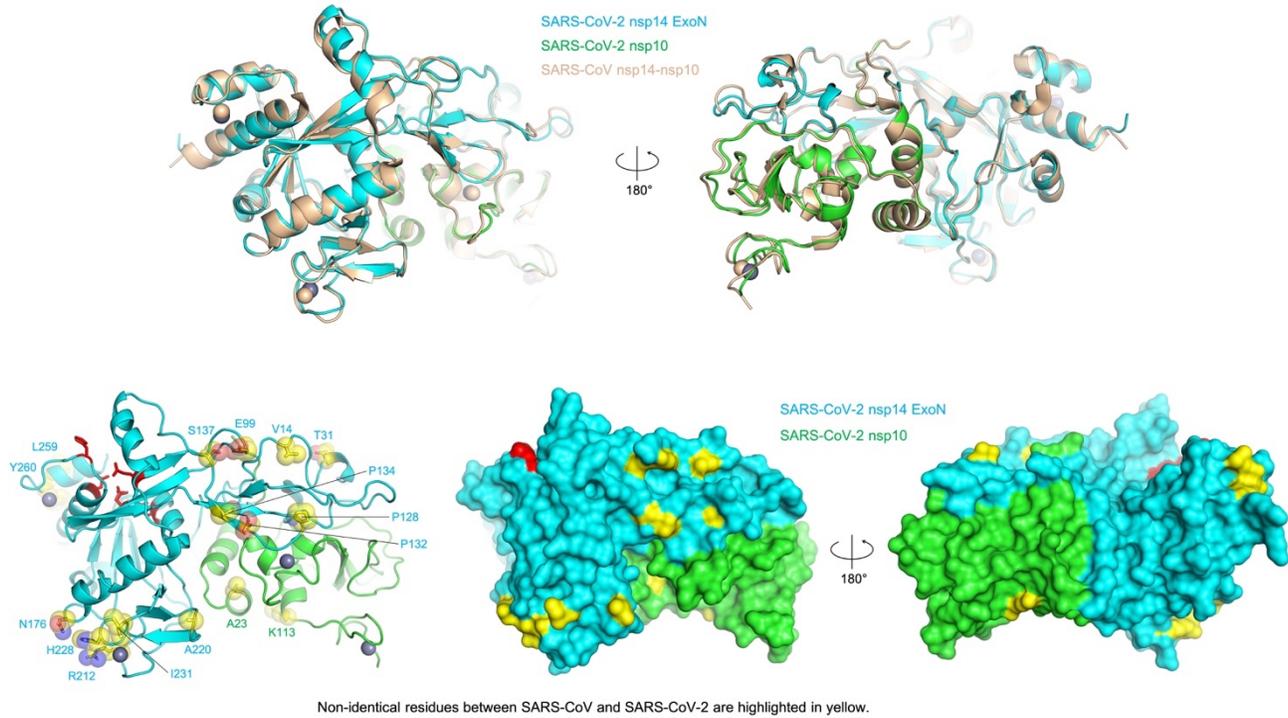
506 **Supplementary Fig. 1 | Comparison of exoribonuclease activities**

507 Exoribonuclease activities of SARS-CoV-2 nsp14-nsp10, ExoN-nsp10, and SARS-CoV ExoN-nsp10

508 complexes on a double-stranded RNA substrate. The inactive E191Q mutant enzyme was tested at a 10

509 times higher protein concentration.

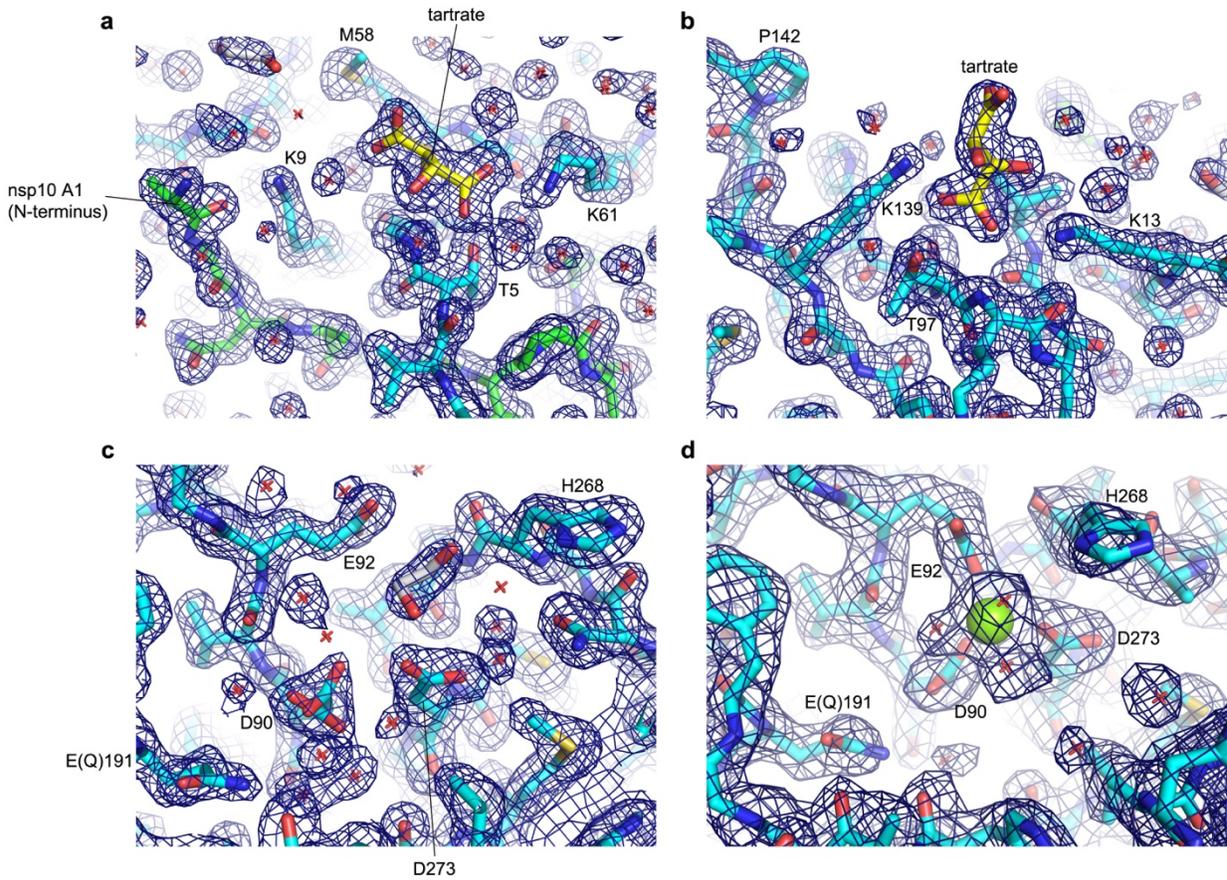
510



511

512 **Supplementary Fig. 2 | SARS-CoV vs. SARS-CoV-2 ExoN-nsp10 structure comparison**

513 **Top**, A superposition between SARS-CoV (PDB ID: 5C8T) (24) and SARS-CoV-2 (this study) ExoN-  
514 nsp10 structures. **Bottom**, Difference in the amino acid sequence between SARS-CoV and SARS-CoV-  
515 2 mapped on the ExoN-nsp10 structure and highlighted in yellow. The active site residues are shown in  
516 red.



517 **Supplementary Fig. 3 | Electron density maps**

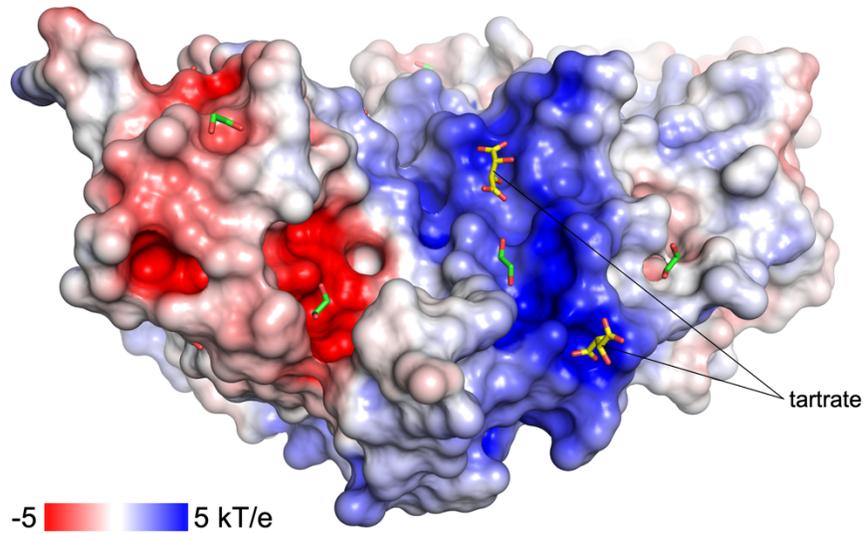
518 2mFo-DFc map contoured at 1.0  $\sigma$  is shown for the higher resolution (1.64 Å) tartrate-bound structure in  
519 **a-c**, and for the lower resolution (2.10 Å) Mg<sup>2+</sup>-bound structure in **d**.

520 **a**, Region including Lys9 and Lys61 of nsp14/ExoN and the N-terminus of nsp10 (The crystallized  
521 protein has additional methionine residue on the N-terminus, which is likely to be disordered) with a  
522 bound tartrate molecule.

523 **b**, Region including Lys139 and Lys13 with a tartrate molecule bound between the two lysine side  
524 chains.

525 **c**, Mg<sup>2+</sup>-free active site. An ethylene glycol molecule used as the cryo-protectant was observed.

526 **d**, Mg<sup>2+</sup>-bound active site.

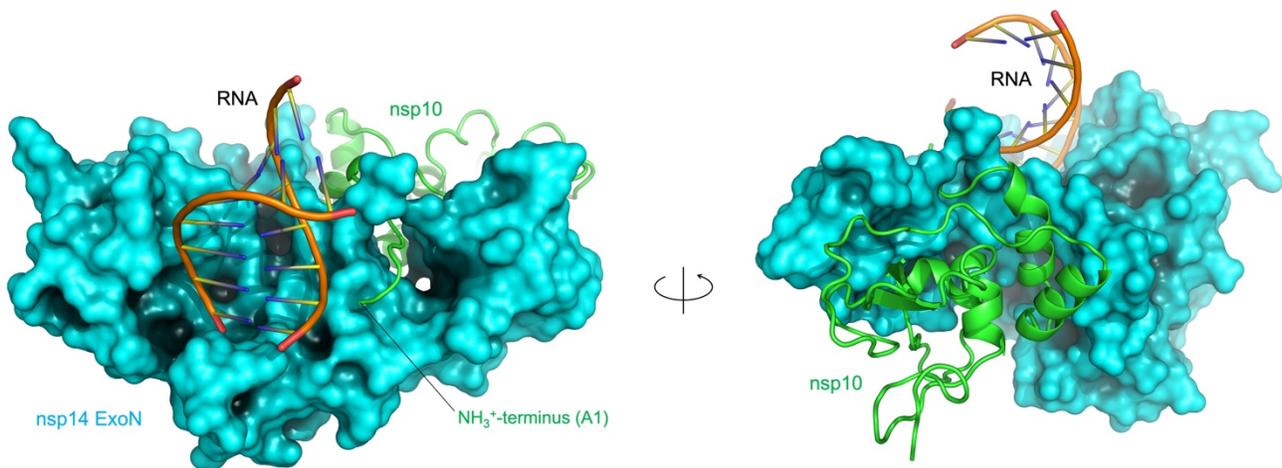


527 **Supplementary Fig. 4 | Tartrate ions bound on the basic patch of ExoN-nsp10 complex**

528 Electrostatic surface potential of ExoN-nsp10 with tartrate or ethylene glycol bound on the protein  
529 surface.

530

531



532

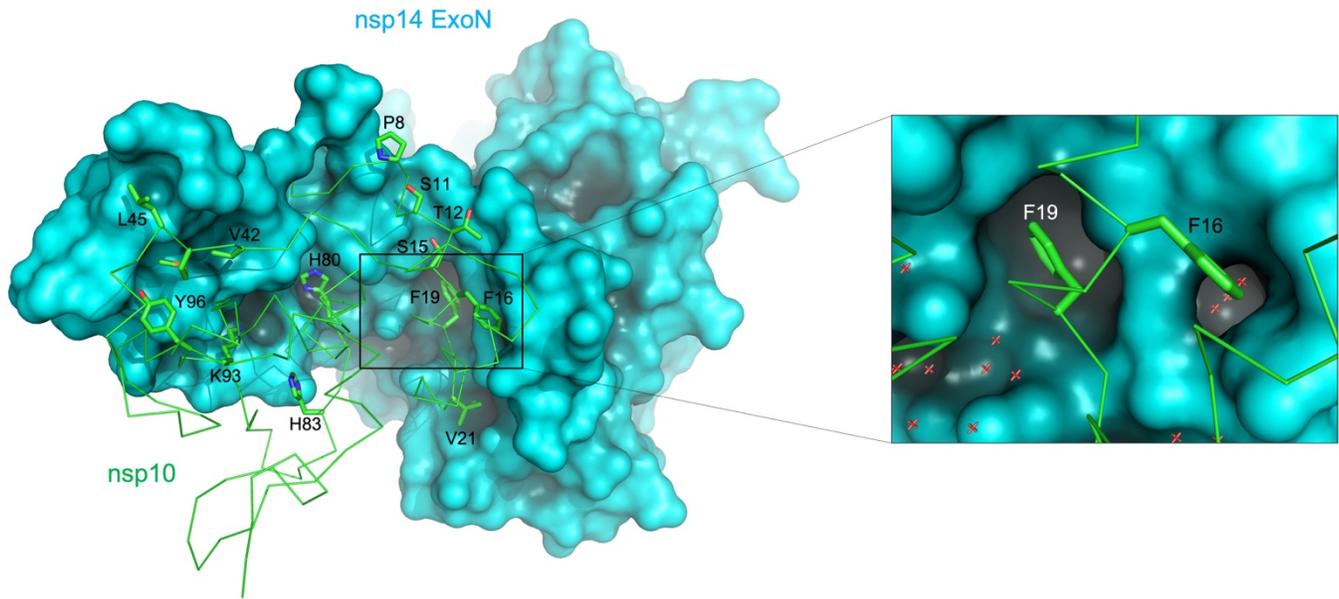
533 **Supplementary Fig. 5 | ExoN-nsp10-RNA complex model (an additional view)**

534 The image on the right is same as **Fig. 3C**.

535

536

537



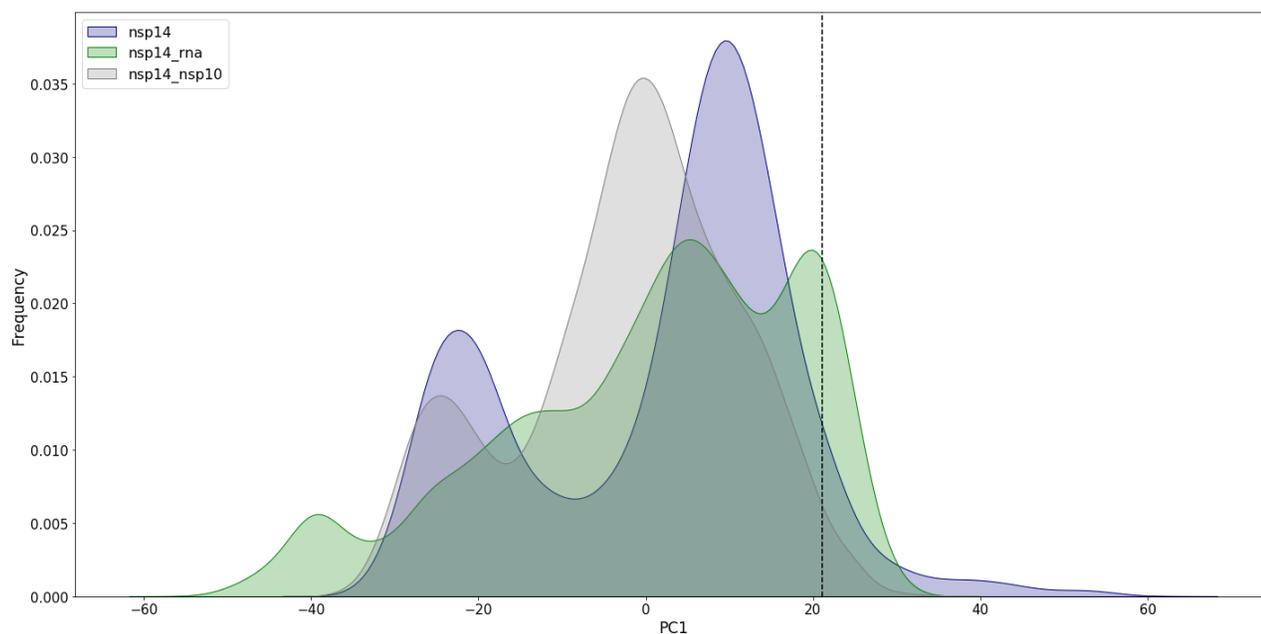
538

539 **Supplementary Fig. 6 | ExoN-nsp10 interface**

540 ExoN is shown in solid surface and nsp10 in wire-frame representations, respectively. Some (not all) of  
541 the nsp10 side chains involved in the protein-protein interaction are shown as sticks. A zoomed view of  
542 the hydrophobic pocket that accommodates Phe16 and Phe19 of nsp10 is shown on the right. Red  
543 crosshairs represent water molecules.

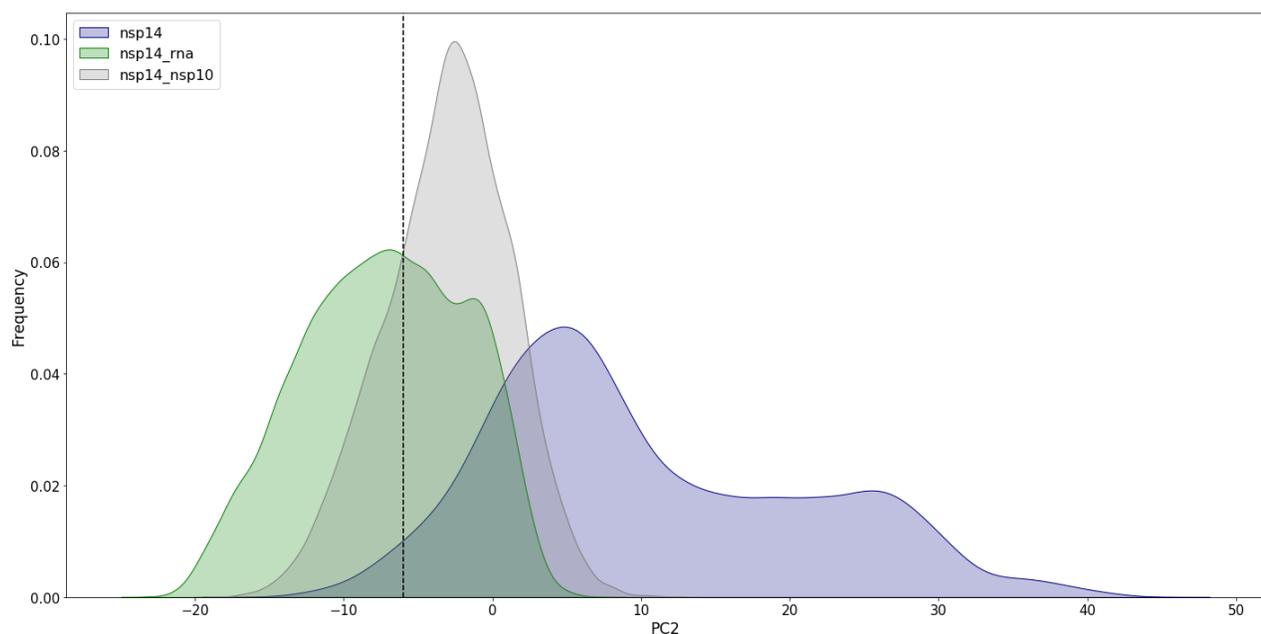
544

545



546

547



548

549 **Supplementary Fig. 7 | Distribution of principal components 1 and 2 (PC1 and PC2) observed in**  
550 **MD simulations.** The dashed line in each plot indicates the value calculated for the starting structure.

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

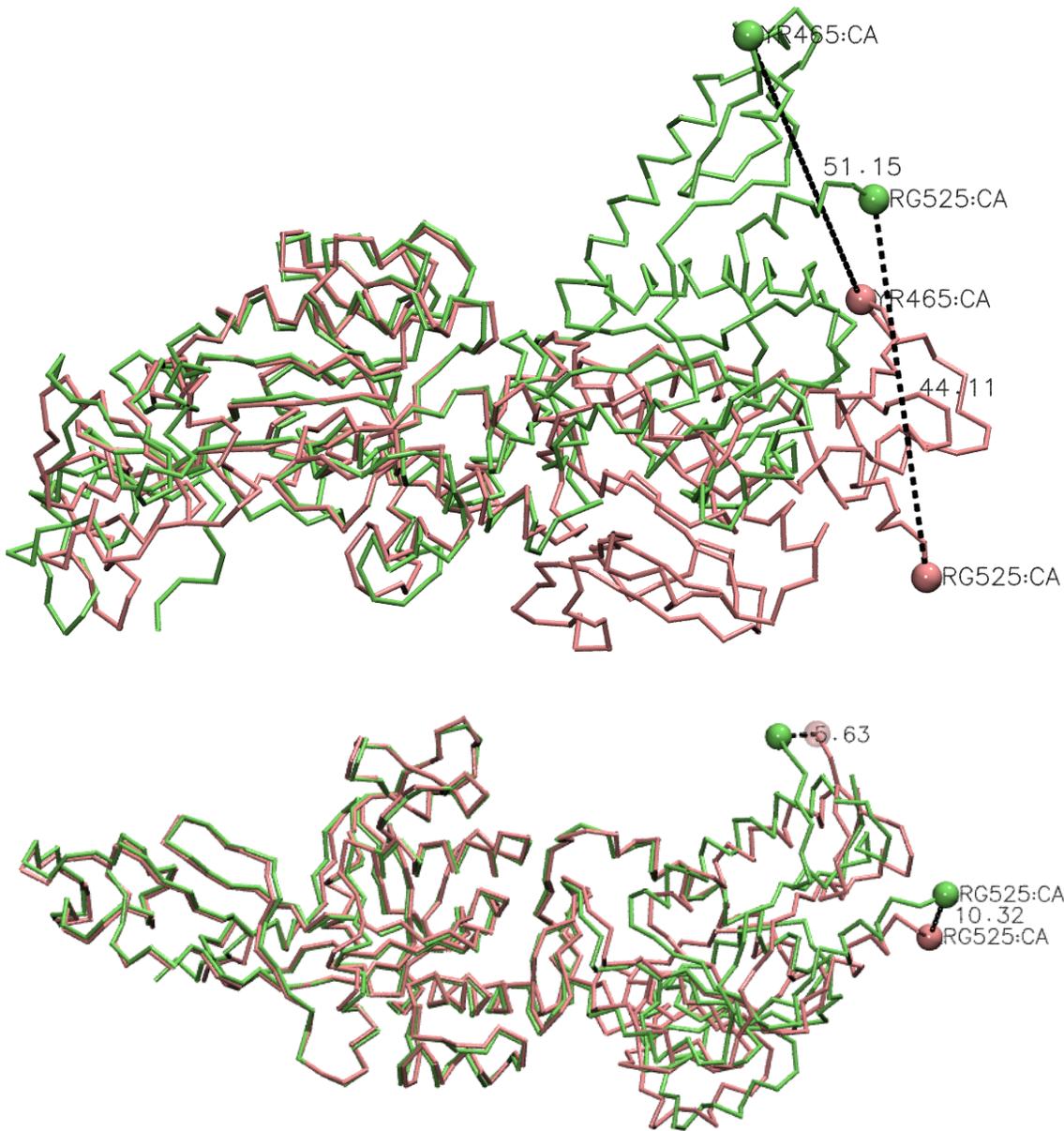
566

567

568

569

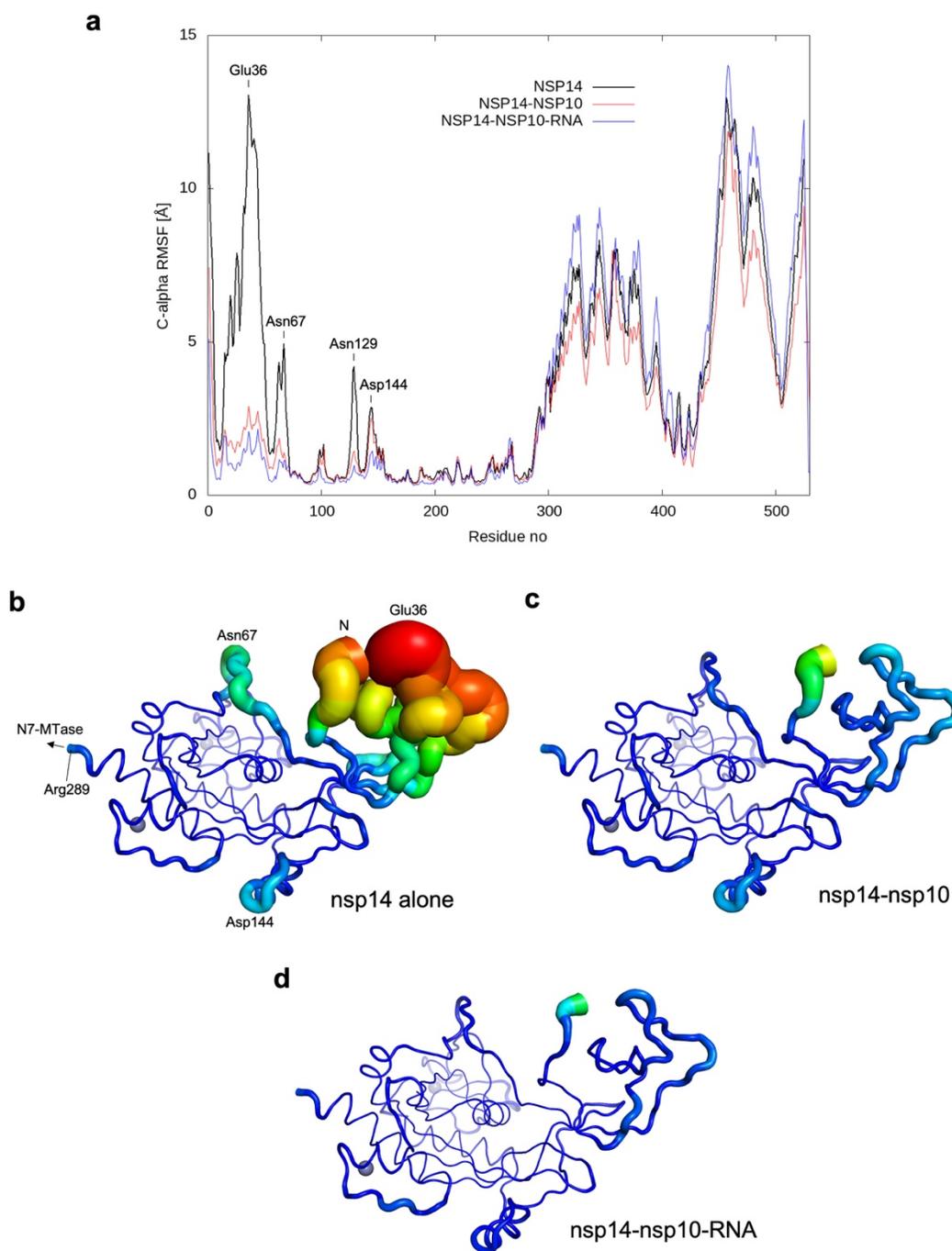
570



571 **Supplementary Fig. 8 | Conformational flexibility of full-length nsp14.**

572 **a**, Superposition of the structures that correspond to PC1 minimum and PC1 maximum values in MD  
573 simulations based on the ExoN domain, showing large displacement of the N7-MTase domain (This  
574 conformational change is also shown in **Supplementary animation 1**). **b**, Similar conformational  
575 variability, albeit with a smaller magnitude, observed between chains A and B of SARS-CoV nsp14  
576 crystal structure (PDB ID: 5NFY) (21).

577



578 **Supplementary Fig. 9 | Internal dynamics of ExoN domain observed in MD simulations.**

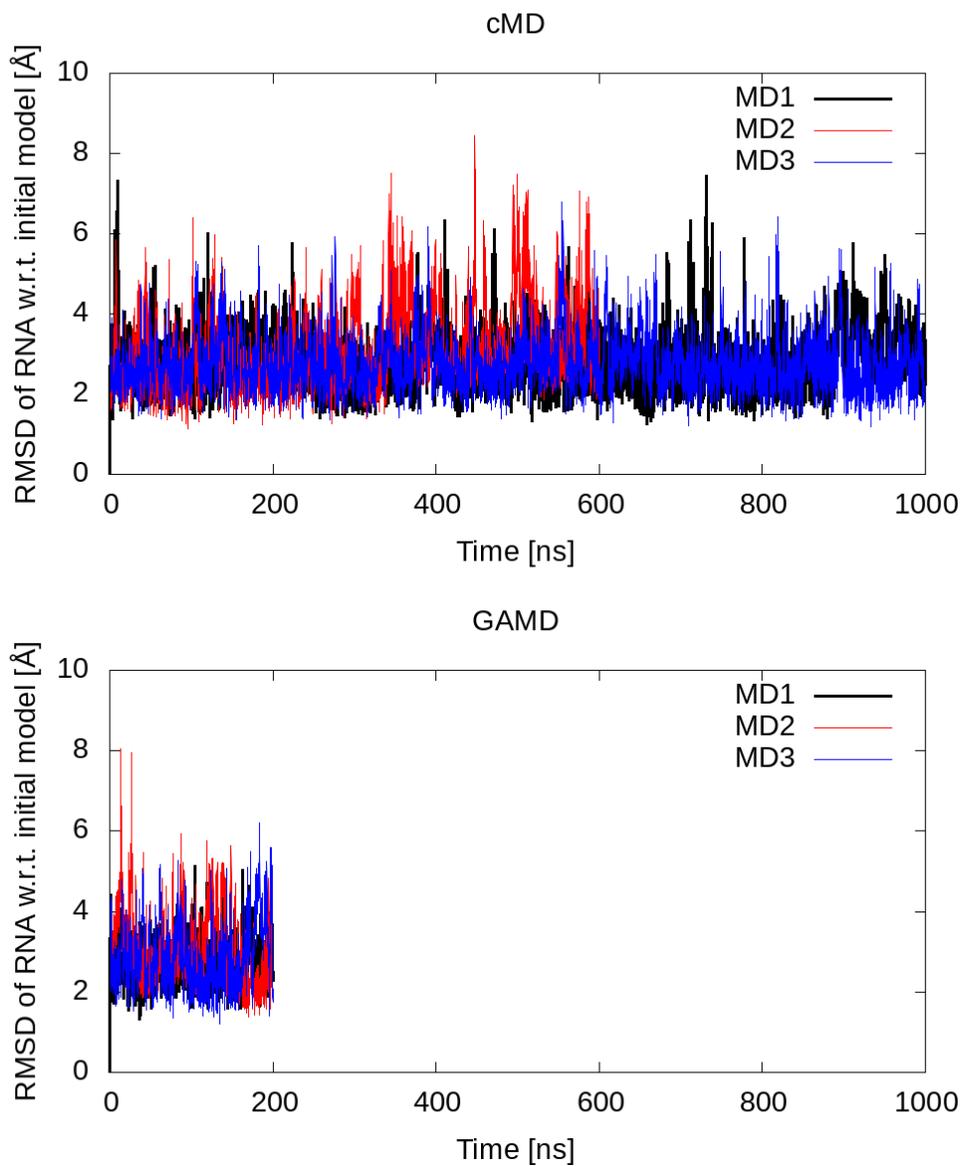
579 **a**, Root-mean-square fluctuations (RMSF) for nsp14 C $\alpha$  atoms in MD simulations of the three systems  
580 after aligning their trajectories to the starting structure with respect to C $\alpha$  atoms of nsp14 residues 71-

581 289. **b-d**, RMSF for nsp14 alone (**b**), nsp14-nsp10 (**c**), and nsp14-nsp10-RNA (**d**), depicted by tube

582 thickness and color. Panels **b** and **c** are same as **Fig. 5 D** and **E**, respectively, and panels **b-d** correspond

583 to the 3 frames in **Supplementary animation 3**.

584



585

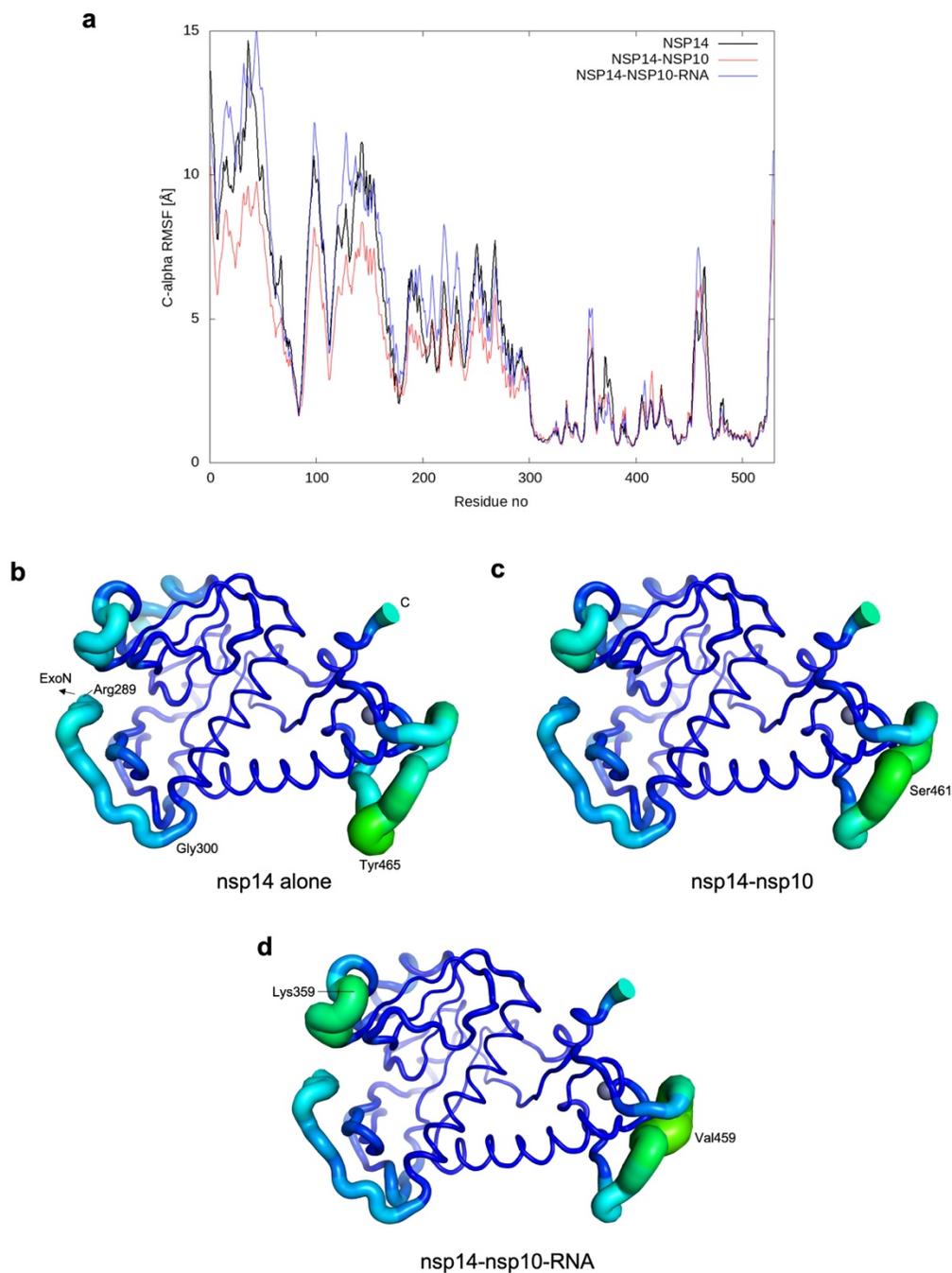
586 **Supplementary Fig. 10 | Stability of RNA in MD simulations.**

587 Root-mean-square deviation (RMSD) of RNA atoms calculated for conventional and Gaussian-

588 accelerated MD (cMD and GAMD) simulations after aligning the trajectories with respect to C $\alpha$  atoms

589 of nsp14 residues 71-289.

590



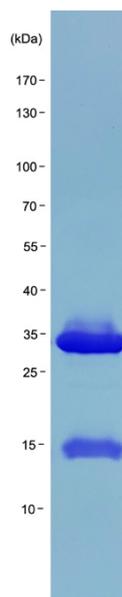
591 **Supplementary Fig. 11 | Internal dynamics of N7-MTase domain observed in MD simulations.**

592 **a**, Root-mean-square fluctuations (RMSF) for nsp14 C $\alpha$  atoms in MD simulations of the three systems  
593 after aligning their trajectories to the starting structure with respect to C $\alpha$  atoms of nsp14 residues 300-  
594 525. **b-d**, RMSF for nsp14 alone (b), nsp14-nsp10 (c), and nsp14-nsp10-RNA (d), depicted by tube  
595 thickness and color.

596

597

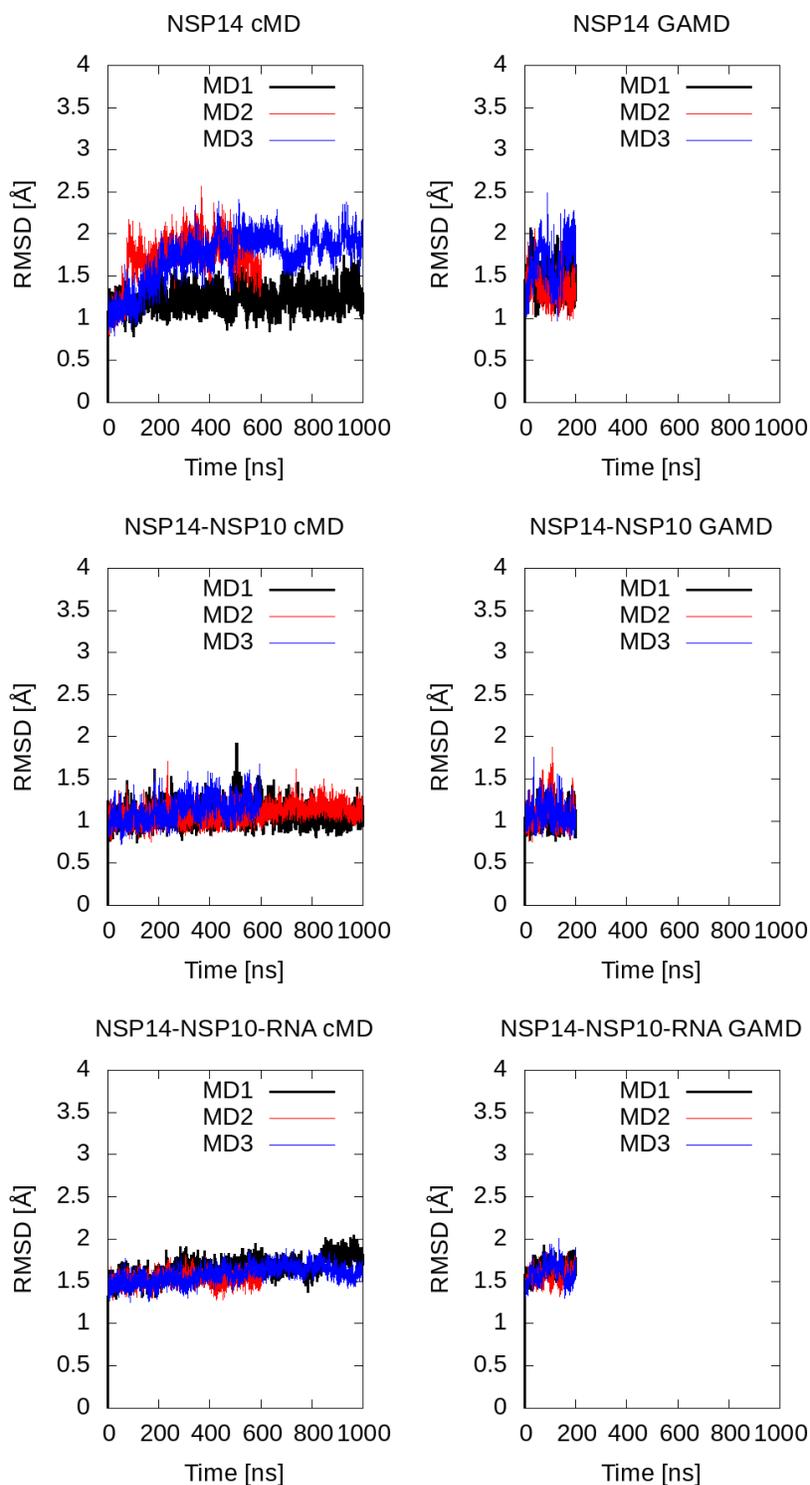
598



599

600 **Supplementary Fig. 12 | SDS-PAGE of purified SARS-CoV-2 ExoN(E191Q)-nsp10 complex.**

601 This protein complex was used in the crystallographic studies.



602

603 **Supplementary Fig. 13 | Stability of MD simulations.**

604 Root-mean-square deviation (RMSD) of the nsp14 ExoN domain C $\alpha$  atoms (residues 71-289) with  
605 respect to the initial model throughout MD simulations.