1   Despite the odds: formation of the SARS-CoV-2 methylation complex

2

3   Alex Matsuda[1], Jacek Plewka[1], Yuliya Chykunova[1,2], Alisha N. Jones[3,4], Magdalena Pachota[1,2], Michał Rawski[1], André

4   Mourão[3], Abdulkarim Karim[1,5,6], Leanid Kresik[1], Kinga Lis[1,7], Igor Minia[8], Kinga Hartman[9], Ravi Sonani[10], Grzegorz

5   Dubin[10], Michael Sattler[3,4], Piotr Suder[9], Paweł Mak[11], Grzegorz M. Popowicz[3,4,*], Krzysztof Pyrć[1,*], Anna Czarna[1,*]

6

7

8   1. Virogenetics Laboratory of Virology, Malopolska Centre of Biotechnology, Jagiellonian University,
9   Gronostajowa 7a, 30-387 Krakow, Poland.
10   2. Microbiology Department, Faculty of Biochemistry, Biophysics and Biotechnology, Jagiellonian University,
11   Gronostajowa 7, 30-387 Krakow, Poland.
12   3. Helmholtz Zentrum München, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany
13   4. Bavarian NMR Center, Department of Chemistry, Technical University of Munich, Lichtenbergstr. 4, 85748
14   Garching, Germany.
15   5. Department of Biology, College of Science, Salahaddin University-Erbil, Kirkuk Road, 44002 Erbil, Kurdistan
16   Region, Iraq.
17   6. Department of Community Health, College of Health Technology, Cihan University-Erbil, 100 street, 44001
18   Erbil, Kurdistan Region, Iraq.
19   7. Faculty of Chemical Engineering and Technology, Cracow University of Technology, Warszawska 24, 31-155,
20   Poland.
21   8. Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Berlin Institute for Medical Systems
22   Biology, Berlin 13125, Germany.
23   9. Department of Analytical Chemistry and Biochemistry, Faculty of Materials Science and Ceramics, AGH
24   University of Science and Technology, Mickiewicza 30, 30-059 Krakow, Poland.
25   10. Protein Crystallography Research Group, Malopolska Centre of Biotechnology, Jagiellonian University,
26   Gronostajowa 7a, 30-387 Krakow, Poland.
27   11. Department of Analytical Biochemistry, Faculty of Biochemistry, Biophysics and Biotechnology, Jagiellonian
28   University, Gronostajowa 7, 30-387 Krakow, Poland

29

30

31   * co-senior and corresponding authors.

32

33

34   **Abstract**

35   Coronaviruses protect their single-stranded RNA genome with a methylated cap during replication. The capping

36   process is initiated by several nonstructural proteins (nsp) encoded in the viral genome. The methylation is

37   performed by two methyltransferases, nsp14 and nsp16 where nsp10 acts as a co-factor to both. Aditionally,

38   nsp14 carries an exonuclease domain, which operates in the proofreading system during RNA replication of the

39   viral genome. Both nsp14 and nsp16 were reported to independently bind nsp10, but the available structural

40   information suggests that the concomitant interaction between these three proteins should be impossible due to

41   steric clashes. Here, we show that nsp14, nsp10, and nsp16 can form a heterotrimer complex. This interaction is

42   expected to encourage formation of mature capped viral mRNA, modulating the nsp14's exonuclease activity, and

43   protecting the viral RNA. Our findings show that nsp14 is amenable to allosteric regulation and may serve as a

44   novel target for therapeutic approaches.

Introduction

The coronaviral genome is of positive polarity and serves as a substrate for the translational machinery of the cell after its release to the cytoplasm. The first and only product of the genomic mRNA translation is a large and non-functional 1a/1ab polyprotein. The polyprotein matures by autoproteolytic processing carried out by two viral SARS-CoV-2 proteases – main protease, M^{pro} and papin-like protease, PL^{pro}. This leads to the generation of a set of non-structural proteins (nsp), which are responsible for the viral replication process and remodeling of the intracellular environment. Once nsps reshape the cell to form a viral factory, the genomic RNA is copied and a set of subgenomic mRNAs is produced in a peculiar, discontinuous transcription process. These subgenomic mRNAs are monocistronic and serve as templates for the production of structural and accessory proteins required for the formation, assembly and release of progeny viruses[1]. The activity of particular nsps has been previously described, showing the complex network of interactions of multi-functional components. However, their coordinated action has not been fully understood.

Works by Gao *et al.*[2], Yan *et al.*[3], Wang *et al.*[4], and Kabinger *et al.*[5] shed light on the scaffold of the replication complex formed by nsp12 (polymerase) and two co-factors nsp7 and nsp8, that create functional machinery able to replicate the viral RNA. Next, an extended elongation complex was described, where nsp12/7/8 is accompanied by nsp13 helicase. This complex is suggested to serve as the basic replication module[6].

Coronaviruses are known for their large genomes, requiring high-fidelity replication to maintain their integrity. While the SARS-CoV-2 nsp12 polymerase is highly processive, it is error-prone and does not provide sufficient fidelity. It has been previously demonstrated that a proofreading system is encoded in coronaviral genomes[7-9]. Nsp14 carries an N-terminal exonuclease (ExoN) domain that serves in this role. ExoN is a member of the DEDDh exonuclease superfamily and exhibits 3′-5′ exonuclease activity, removing incorrectly incorporated nucleotides from the 3′ terminus of the newly formed RNA. ExoN has additionally been proposed to play a role during discontinuous replication of coronaviruses. Nsp14 has been shown to associate with the replicatory complex, with nsp10 as a co-factor, modulating and enhancing nsp14 exoribonuclease activity[7-9].

Apart from the supporting role in replication, nsp14 has a second important function. It takes part in cap formation after genome copying is finalized[7]. Capping of viral mRNAs is essential for their function and integrity. It enables translation initiation and protects viral mRNA from recognition as foreign by cellular sensors, thereby preventing the induction of innate immune responses[10]. Cap formation is a tightly regulated process consisting of four consecutive enzymatic reactions. First, nsp13 triphosphatase removes the γ-phosphate of the 5′-triphosphate end (pppA)[11,12]; next, nsp12 guanylyltransferase (the nidovirus RdRp-associated nucleotidyltransferase, NiRAN domain)[3] transfers GMP to the 5′ phosphate to form the core structure of the cap (GpppA); the GpppA is methylated at the N7 position by the nsp14 N7-methyltransferase domain (^{7Me}GpppA); subsequently, ribose in the first ribonucleotide is methylated at the 2′-O-position by the nsp16 2′-O-methyltransferase[13-15]. Capping is regulated by nsp9, which binds nsp12 near the NiRAN active site[16]. The described process results in a functional cap (^{7Me}GpppA$_{2'OMe}$), completing the genome replication. While there is a good structural understanding of the interactions between nsp12, nsp13, and nsp14 during replication, nsp16 remains an orphan, and its methylation process is not fully understood.

82   Here, we studied the interaction between the two methyltransferases encoded in the genome of the

83   coronavirus–nsp16 and nsp14. Both protteins bind to nsp10 and both are required for complete methylation.

84   Therefore, their spatial proximity mediated by nsp10 would appear as beneficial for the methylation efficacy. Prior

85   *in silico* analysis of structural information on nsp10/14 and nsp10/16 complexes suggested that simultaneous

86   binding of both exonucleases to nsp10 should be impossible due to steric hindrance[17].Interestingly, both nsp14

87   and nsp16 interact with nsp10 by typical, well-defined protein-protein interfaces containing deep-reaching

88   lipophilic residues and solvent-protected hydrogen bonds. The element that causes steric overlap between nsp14

89   and nsp16 is a peculiar N-terminal "lid" domain of nsp14. This lid is mostly devoid of secondary structure and lacks

90   characteristic protein-protein interaction complementarity. Interestingly, recent structures of nsp14 without

91   nsp10 show massive rearrangement of the lid region confirming its structural flexibility[18,19]. This prompted us to

92   postulate that a local structural rearrangement of the nsp14 N-terminal lid region is therefore required and

93   possible, thus facilitating the heterotrimer complex formation.

94

95   ## Results

96   ### *Biochemical evidence of heterotrimer formation*

97   Nsp10, nsp14, and nsp16 were co-expressed in *E. coli* and purified. A single peak was obtained in size exclusion

98   chromatography (SEC), indicating that all three proteins co-migrate (Fig 1a). Because no prior crosslinking was

99   used, this result suggests the formation of a heterotrimer complex. Sodium dodecyl sulphate-polyacrylamide gel

100   electrophoresis (SDS-PAGE) analysis of the peak containing the putative complex, together with mass

101   spectrometry identification of components, indicates the presence of all three proteins (Fig. 1b). We further show

102   that the putative complex migrates as a single major band in native PAGE (Extended Data Fig. 1a). When the band

103   was excised from the gel and analyzed by SDS-PAGE, three bands were identified, corresponding in molecular

104   weight to nsp14, nsp16, and nsp10 (Extended Data Fig. 1b). MS analysis of proteins contained in the major band

105   derived from the native PAGE resulted in the identification of all three components of interest (Extended Data

106   Table 1), further suggesting heterotrimer complex formation. LC-MS additionally allowed the assessment of the

107   stoichiometry of the complex. By analyzing the signals at 254 and 280 nm, with further MS-based identification of

108   the proteins under each UV chromatographic peak, stoichiometry was consistently established at 1.2:1:1

109   (nsp10:nsp14:nsp16) (Extended Data Fig. 2 and Extended Data Table 2).

110   The kinetics of the heterotrimer complex formation were assessed by MicroScale Thermophoresis (MST).

111   Nsp14 and nsp16 were separately expressed as histidine-tagged constructs and labeled with a high-affinity His-

112   tag specific fluorophore dye. When labeled nsp14 was titrated with unlabeled nsp10, a dose-dependent increase

113   in thermophoretic signal was observed, indicating an interaction with a K_d of 2.4 ± 0.2 µM, which is in the

114   agreement with the previously reported affinity with a K_d of 1.1 ± 0.9 µM[20] (Fig. 1c). A comparable effect[21] was

115   observed when labeled nsp16 was titrated with unlabeled nsp10 with a K_d of 0.24 ± 0.01 µM. Labeled nsp14 did

116   not directly interact with unlabeled nsp16, but when labeled nsp14 was titrated with unlabeled nsp10/16 complex,

117   a dose-dependent increase in thermophoretic signal was observed (Fig. 1c), which was interpreted as the

118   heterotrimer complex formation. Fitting the experimental data allowed the determination of the K_d characterizing

119   the interaction at 0.28 ± 0.01 µM.

120   To further characterize the heterotrimer complex and its components, we analyzed thermal denaturation

121   profiles using nanoDSF. Each of the individual system components (nsp10, nsp14, and nsp16) was characterized

122   by a single characteristic denaturation temperature (Fig. 1d), indicating that structures of the functional domains

123   (if any) within particular components collapse in a coordinated manner upon temperature increase. Nsp16 was

124   least temperature stable (T_m=45°C), while nsp10 (T_m=50°C) was most stable. However, the differences in melting

125   temperatures were small, with all components demonstrating overall similar stability. Nsp14 was stabilized by the

126   complex formation with nsp10 (nsp10/14 complex: T_m=55°C), while nsp10 binding did not significantly affect the

127   thermal stability of nsp16 (nsp10/16 complex: T_m=45.5°C). The heterotrimer is characterized by a single sharp

128   thermal denaturation peak with a characteristic melting temperature of 51.7°C. This temperature, distinctly

129   different from that of any of the individual or binary-complexed components, gives a further evidence for the

130   heterotrimer complex formation.

131
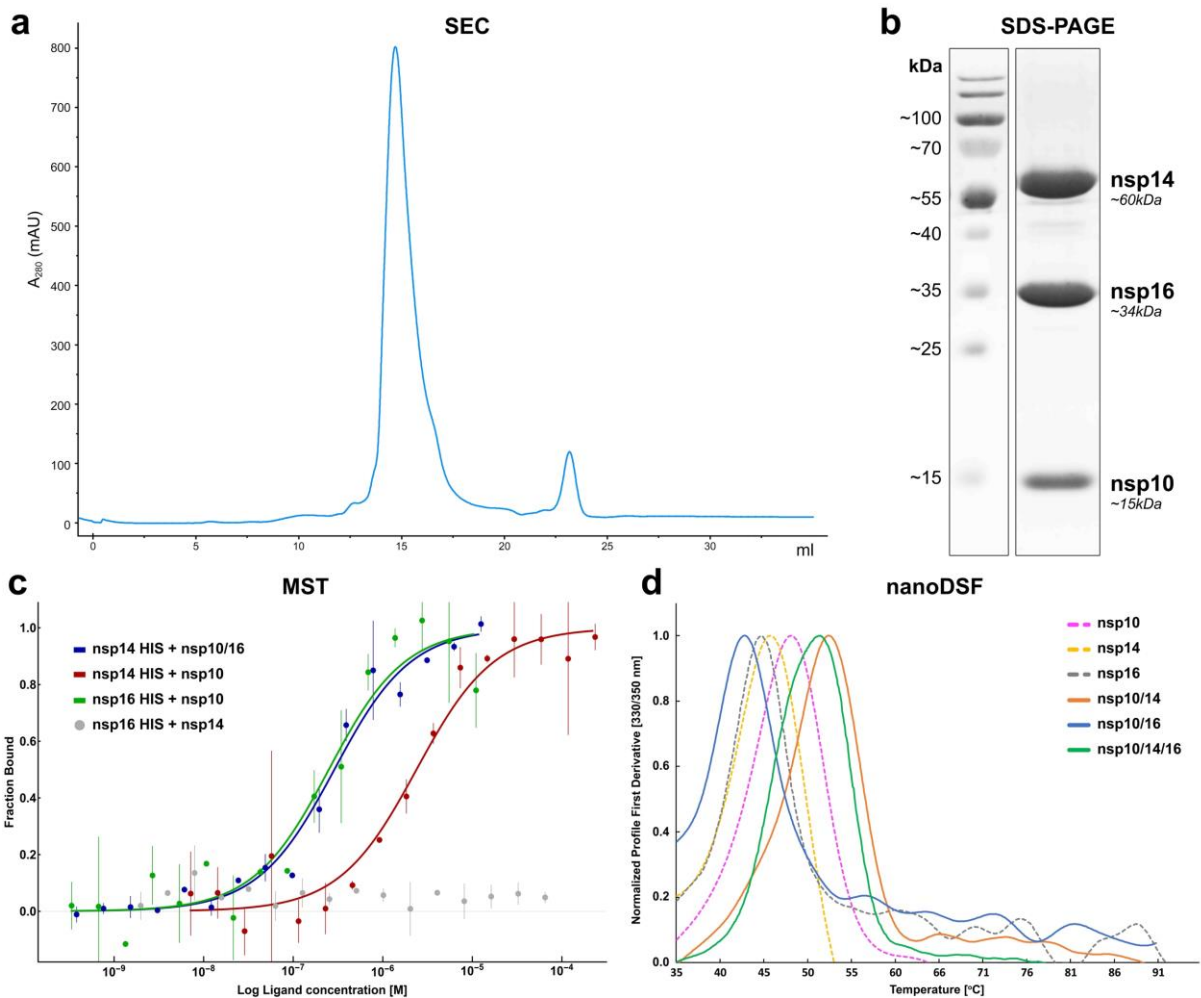
Fig. 1 | In vitro formation of the nsp10/14/16 heterotrimer. a, SEC chromatogram showing co-migration of nsp10, nsp14, and nsp16 proteins in a single peak at ~15 mL. **b,** SDS-PAGE analysis of the major peak from panel **a**. Three major proteins corresponding in size to nsp10, nsp14, and nsp16 (Mw of 15, 60, and 33.5 kDa, respectively) are present in roughly equal quantities. **c,** MST analysis of the intra-complex affinities. Experimental data are represented as mean (dots) with error bars. Binding model fit is represented as a solid line. **d,** The melting profiles of heterotrimer complex and complex components, determined using nanoDSF.

Functional consequences of nsp10/14/16 heterotrimer complex formation

139 GpppA is methylated at position N7 by nsp14 N7-methyltransferase domain, yielding $N^7MeGpppA^{22}$; SAM is used

140 as a donor of the methyl group (Fig. 2a). Here, we tested whether the heterotrimer complex formation affects

141 nsp14 catalyzed methylation. To follow the N7-methyltransferase activity of nsp14, we used an indirect assay to

142 monitor changes in the level of a second reaction product, SAH, by HTRF. In the absence of one of the substrates

143 (SAM or RNA) or the tested enzyme (negative controls), there was no activity validating the experimental setup

144 (Extended Data Fig. 3a). Nsp14 showed no preference for the nascent nucleotide methylating both GpppG and

145 GpppA, as demonstrated by the production of equal levels of SAH when using either nucleotide as a substrate (Fig.

146 2b). The binary complex of nsp10/14 and the ternary complex N7-methyltransferase activities were comparable

147 to that of nsp14 alone, indicating that binding to nsp10 or the heterotrimer complex formation had no significant

148 influence on the N7-methyltransferase activity of nsp14. N7-methyltrasnferase activity was also not reported for

149 N7-methylated substrate (O4 oligo - 5' (N7-MeGppp) ACA UUU GCU UCU GAC 3'). Nsp16, on the other hand,

150 presenting 2'-O-methyltransferase activity in the complex with its obligatory partner nsp10 does not show a signal

151 on non-methylated substrates and modetare activity on the O4-oligo substrate. Similarly to N7-methyltransferase

152 activity, also 2'-O-methyltransferase activity is not affected by the heterotrimer complex formation. As expected,

153 the pan-methyltransferase inhibitor sinefungin[23] halted methyltransferase activities of all repoted here proteins.

154 We have also tested if the nsp14 exoribonuclease (ExoN) activity affects methyltransferase activity by mutating

155 the ExoN binding site. However, no difference in N7-methyltrasnferase activity was observed, decoupling those

156 two activities from each other (Extended Data Fig. 3b). Alongside N7-methyltransferase activity, nsp14 harbors an

157 ExoN domain characterized by exoribonuclease activity; an activity essential for proofreading during virus

158 replication. However, *in vitro,* nsp14 is characterized by the high processivity and non-specifically degrades nucleic

159 acids[17,24,25]. As the nuclease activity must be tightly controlled, we assessed whether the complex formation

160 regulates this process. We evaluated the binding and nuclease activity of nsp10/14 and nsp10/16 compared to

161 nsp10/14/16 with two synthetic RNAs of CoV-RNA1-G 5'-GGGGGGGGGGGCGCGUAGUUUUCUACGCG-3' and CoV-

162 RNA1-A 5'-AAAAAAAAAAACGCGUAGUUUUCUACGCG-3'.

163 Addition of nsp10/14 and nsp10/16 (equimolar at 1.6μM) results in the retardation of the RNA relative to an

164 RNA only control, thus indicating a binding interaction between the RNA and protein heterodimers (Fig. 2c native

165 gel, Extended Data Fig. 4). Interestingly, upon addition of nsp10/14/16 (equimolar at 1.6 μM), we observed

166 increased retardation of the RNA-heterotrimer protein complex relative to the nsp10/14 and nsp10/16

167 heterodimer complexes (Fig. 2c, native gel). The observed broadened band suggests that nsp10/14/16 forms a

168 tetrameric complex with each of the RNA substrates.

169 As expected, addition of nsp10/14 results in the degradation of the RNA substrates (as indicated by the

170 decreased intensity of the RNA band in the presence of protein relative to the RNA only control, observed both by

171 native and nondenaturing gel analysis) (Fig. 2c, denaturing gel). Interestingly, we observed that the nuclease

172 activity of nsp10/14 is reduced for the nsp10/14/16 heterotrimer, suggesting that nsp14 nuclease activity is

173 modulated by nsp16 (Fig. 2c, denaturing gel). To assess this, we added increasing amounts of nsp16 to preformed

174 nsp10/14 CoV-RNA1-A RNA complex and monitored the degradation of the RNA. Upon increasing the

175 concentration of nsp16, the degradation activity of nsp14 is reduced, starting at a near 1:1 ratio (Fig. 2d). This

176 effect is unlikely to be caused by the protective effect of RNA sequestering by the nsp16 as the RNA-nsp10/16

177 affinity is reported to be in high micromolar regime[26] (~100 µM) and yet protective functions of nsp16 appear at

178 much lower protein concentrations equal to the nsp10/14 complex concentration. Recent crystallographic data

179 suggest that nsp14 exonuclease domain is controlled by lid rearrangement caused by nsp10 binding[18]. In

180 combination with our binding shift and degradation assays, this data suggests that binding of nsp10/16 to nsp14

181 causes additional allosteric change that inhibits unwanted exonuclease activity in favor of the methyltransferase

182 one.
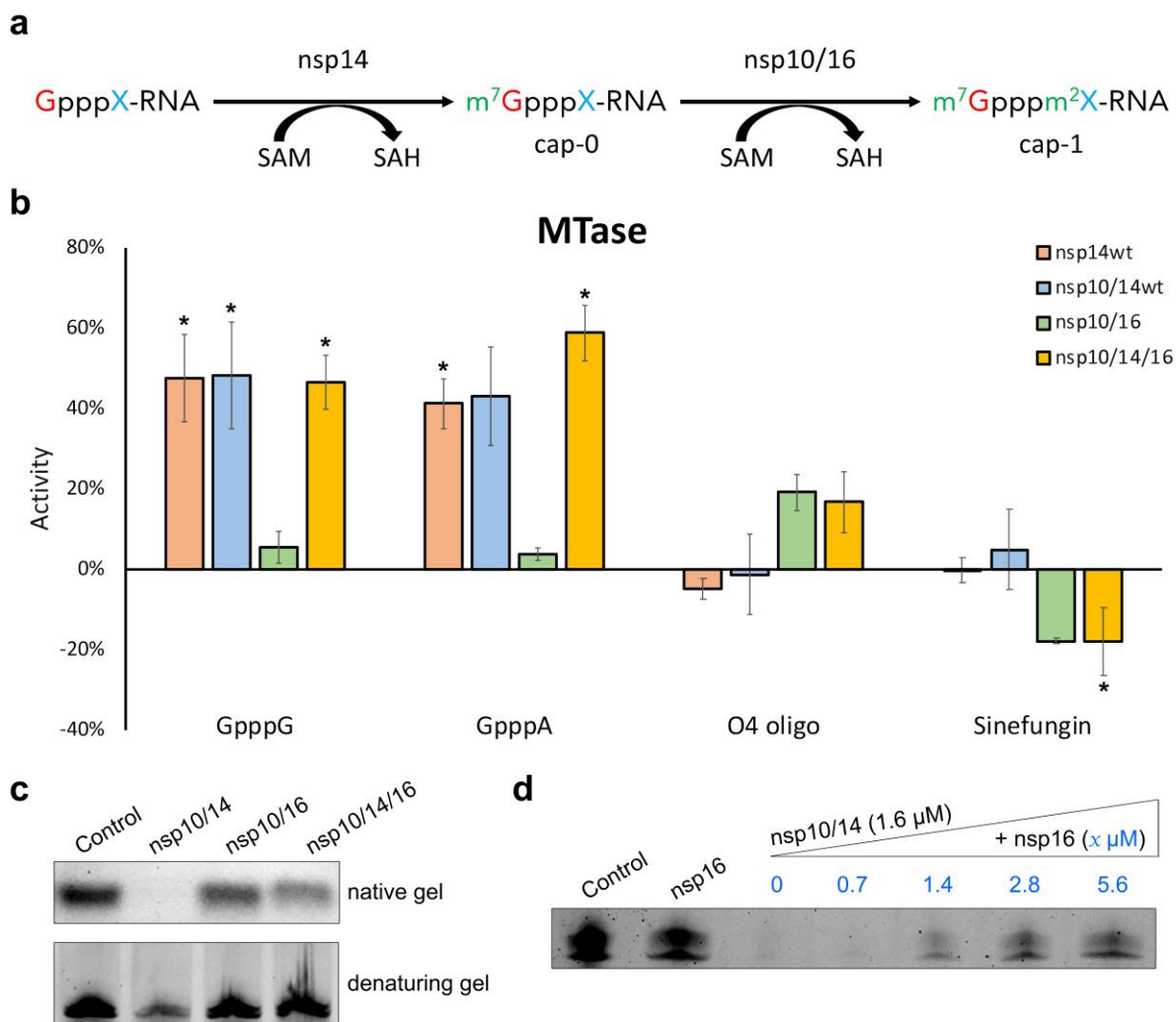
183

184 *Fig. 2| The formation of nsp10/14/16 heterotrimer modulates the ribonuclease, but not methyltransferase processivity. a, Schematic view of mRNA methylation. X*

185 *represents a nascent nucleotide that could be adenine (A) or guanine (G). m[7] represents methylation of the first guanine at position N7 by nsp14. m[2] indicates 2'-O*

186 *methylation of the nascent mRNA nucleotide. b, Modulation of the methyltransferase activity by the protein partners. The results were normalized using the SAH*

187 *calibration curve and denoted µM of methylated product. All experiments were performed in duplicate. Average values with error bars (SD) are shown; * p < 0.05 in*

188 *comparison to the condition with no activity. c, Analysis of RNA binding (native gel) and degradation (denaturing gel) potential of indicated nsp complexes. d,*

189 *Determination of stoichiometry of the heterotrimer complex limiting the exonuclease activity of nsp14. Titration of nsp10/14 (1.6µM) with indicated amounts of*

190 *nsp16 (blue). (c, d) CoV-RNA1-A was used as a substrate. Similar results are obtained for CoV-RNA1-G (Extended Data Fig. 4).*

191

192 *Lid hypothesis explains the nsp10/14/16 heterotrimer complex formation*

193 A number of high-resolution crystal structures are available for the nsp10/14 and nsp10/16 complexes. Overlay of

194 the structures by a common component (nsp10) demonstrates that fifty N-terminal amino acids of nsp14 overlap

195 with nsp16 at the surface of nsp10. The significant steric clash produced by this overlap precludes the heterotrimer

196   complex assembly, which is mediated by the concomitant interaction of nsp14 and nsp16 with nsp10 (Fig. 3a-c).

197   As such, the formation of the heterotrimer complex would require a significant structural rearrangement within

198   either nsp14 or nsp16. Analysis of the interactions between nsp10, nsp14 and nsp16 suggests that the former

199   binding surface contains a weakly interacting component – the N-terminal region of nsp14 (for rigorous analysis,

200   see Discussion). We hypothesize that within the nsp10/14/16 heterotrimer complex, nsp16 displaces the N-

201   terminal "lid" of nsp14 at the interface with nsp10.
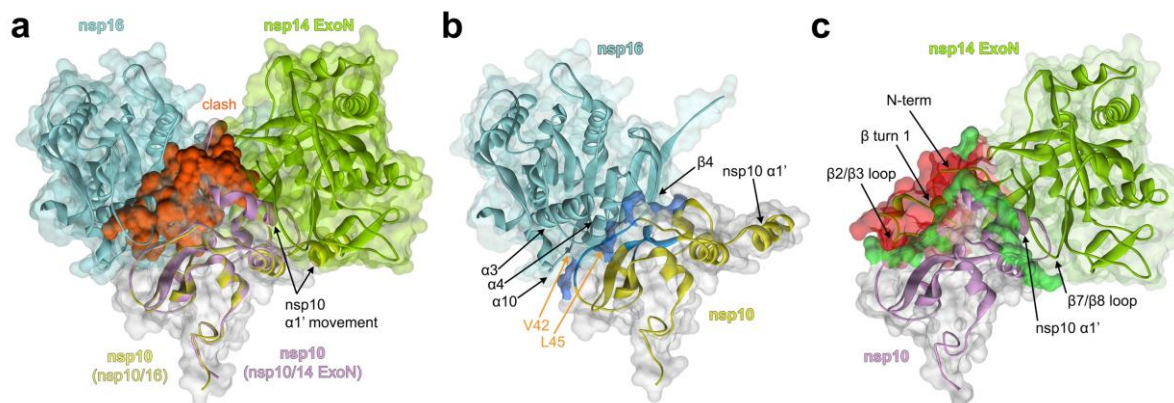
202
203   **Fig. 3 | In silico modelling of the nsp10/14/16 complex. a**, Nsp10-centered alignment of nsp10/16 (PDB ID: 6WVN) and nsp10/14 ExoN domain (PDB ID: 7DIY). Nsp16
204   is shown in cyan, and the associated nsp10 is shown in yellow. nsp14 ExoN domain is shown in green and the associated nsp10 in purple. The structural clash
205   between the nsp14 ExoN and nsp10/16 surfaces is shown in red. **b**, The interface between nsp10 and nsp16. The strong hydrophobic interaction between nsp10
206   and nsp16 is shown in blue. The nsp10 residues V42 and L45 are shown as sticks. **c**, The interface between nsp10 and the nsp14 ExoN. The hydrogen bonds between
207   nsp10 and nsp14 are shown in green. All panels: Different orientation of nsp10 α1'-helices in respective complexes with nsp16 and nsp14 is shown. Arrows indicate
208   major structural features constituting the interface.
209

210   To evaluate the above hypothesis, we created a "lid"-truncated mutant of nsp14 (nsp14Δ) missing the initial

211   50 amino acids and evaluated its interaction with nsp10. Nsp14Δ still formed a complex with nsp10, characterized

212   by a K_d of 1.5 μM. This value did not differ significantly from that characterizing the complex involving full length

213   nsp14 (K_d=2.4 μM). The above data indicate that the interactions of the N-terminal region of nsp14 with nsp10 do

214   not contribute significantly to the affinity of either component, indirectly supporting the "lid" rearrangement

215   hypothesis in the heterotrimer complex formation.

216

217   *Structural characterization of the heterotrimer complex*

218   Nsp14, nsp10/14, nsp10/16 and nsp10/14/16 were characterized by SEC-SAXS[15]. The retention times of tested

219   biomolecules correlate with theoretical molecular weights, assuming 1:1(:1) molar ratios[16] (Extended Data Fig. 5).

220   Guinier analysis of scattering profiles established that the largest radii of gyration (R_g) characterized the

221   heterotrimer complex (40.3±1.0 Å). Nsp14 and the nsp10/14 complexes were characterized by significantly

222   smaller R_gs (28.0±0.5 and 30.1±1.8 Å, respectively); the nsp10/16 complex was characterized by the smallest (R_g =

223   21.0±1.5 Å) (Extended Data Table 3). This data corresponds well with the expected molecular weights of tested

224   complexes at 1:1:1 stoichiometry, further supporting the stoichiometry of the heterotrimer complex.

225   The estimated molecular weights of a scatterer differ from the expected values, perhaps as the result of the

226   flexibility of the tested system (SAXS signal averages all conformations). Nonetheless, the relative values follow

227   the expected pattern, with nsp10/16 characterized by the lowest and the nsp10/14/16 heterotrimer complex by

228   the highest molecular weight, as determined by SAXS.

229   When reconstructed in real space using the indirect Fourier transform using software GNOM, the scattering

230   profiles present roughly Gaussian shapes with significant tailing for nsp10/14/16 and nsp10/14, indicating an

231   elongated globular nature for the protein complexes, with peaks overlapping with radii of gyration obtained using

232   Guinier analysis. Calculated maximal distances within scatterers support the trend established above, with

233   nsp10/16 constituting the smallest complex at 80.0 Å, nsp14 at 95.8 Å, nsp10/14 at 122.0 Å, and the nsp10/14/16

234   at 140.0 Å (longest axis).

235   Molecular envelopes which best represent the scattering profiles were calculated using DAMMIF software[27]

236   (Fig. 4). Crystal structures of nsp10/16, nsp14 and nsp10/14 and a model of the nsp10/14/16 heterotrimer

237   complex, created assuming the "lid" hypothesis, were fitted into the envelopes using the SUPCOMB software. The

238   crystal structures of nsp14 and nsp10/14 fit the molecular envelopes poorly, suggesting that these two remain

239   flexible in solution (Fig. 4a and Fig. 4b). The structure of nsp10/16 fills the envelope tightly, suggesting complex

240   rigidity in solution. The initial model of the heterotrimer complex already filled the envelope relatively well and

241   was further optimized via normal mode analysis using the SREFLEX software[28]. SREFLEX rotates and translates rigid

242   body domains of the input model within the constraints of flexible loops, optimizing the fit to the experimental

243   scattering curve. The resulting heterotrimer model was characterized by a value of 1.08 for goodness-of-fit to the

244   experimental SAXS data, suggesting a very likely solution. The model fits the envelope tightly, suggesting the

245   rigidification of nsp14 structure upon the heterotrimer complex formation (comparing to nsp10/14 envelope fit).

246   The decomposition of the nsp10/14/16 scattering profile into volume fractions calculated from the binary

247   complexes nsp10/14, nsp10/16 crystal structures or extracted individual protein in the software Oligomer suggest

248   that the signal can be devided 1:1 into nsp10/16 and nsp14 (Extended Data Table 3). This furter implies that

249   nsp10/16 interface within the nsp10/14/16 is retained, while it is nsp14 that undergoes structural rearrangements

250   upon the heterotrimer complex formation.

251
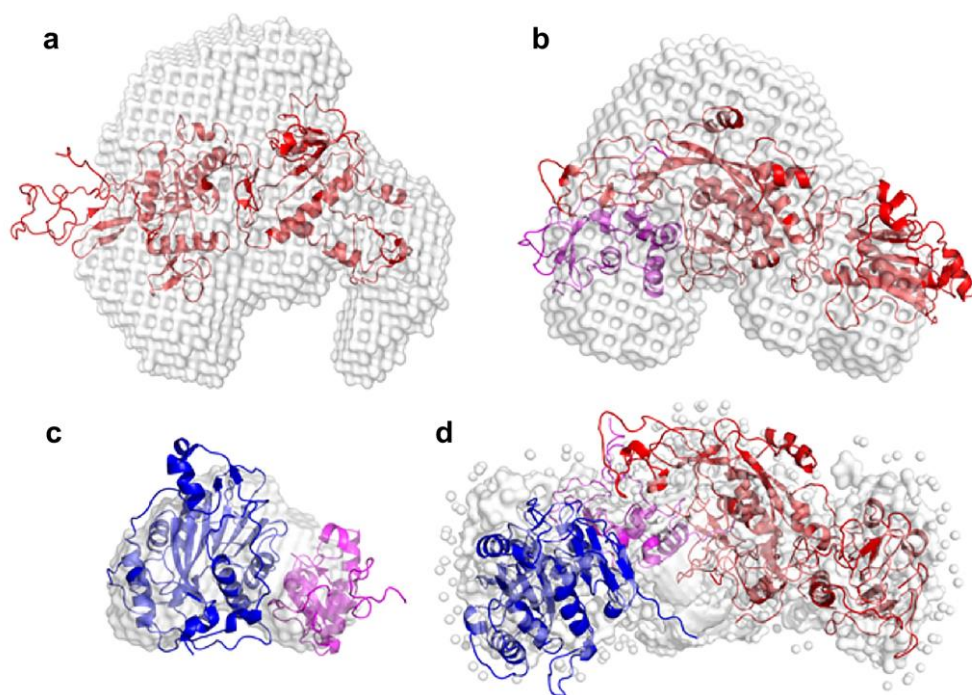
252
253   **Fig. 4 | Molecular envelopes representing the experimental SAXS scattering profiles** of nsp14 (**a**), nsp10/14 (**b**), nsp10/16 (**c**), nsp10/14/16 (**d**). Overlaid are best fits of crystallographic / theoretical models of relevant complexes. Color coding: nsp10 in magenta, nsp14 in red, nsp16 in blue, molecular envelopes in grey.

254   The heterotrimer complex formed by nsp10, nsp14 and nsp16 was further characterized by transmission

255   electron microscopy. Negatively stained samples of a heterotrimer complex formed from full-length components

256   yielded a non-homogenous particle distribution, which precluded structural analysis (Extended Data Fig. 6).

257   However, when nsp14 methyltransferase domain was truncated out of the structure leaving nsp14 ExoN, the

258   particle distribution became more homogenous, allowing the convergent classification (Fig. 5a) and structural

259   analysis. The reconstruction obtained from the negative-stained transmission electron micrographs at 20 Å

260   resolution shows elongated particles with approximate dimensions of ~ 10 × 5 × 4.5 nm (Fig. 5b). Necking is evident

261   in the center of the particles, indicating that two larger structural components are connected by a component of

262   a lower molecular weight (nsp10). The SAXS-derived structural model of the heterotrimer complex fits the

263   experimental NS-TEM map reasonably well with CCmask/CCbox of 0.59/0.80. Nsp14 ExoN and nsp16 fit the two

264   globular regions of the map, connected by a region with a density corresponding to nsp10. The central cavity

265   suggested by SAXS data is visible in the electron microscope-derived map, indicating an overall match between

266   the 3D reconstitutions obtained using each method.
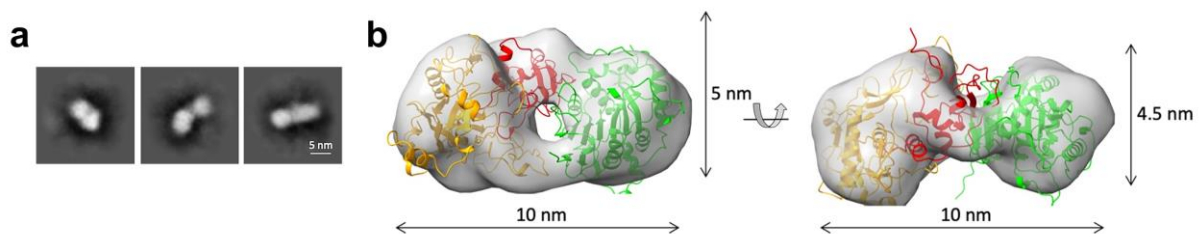
267
268 **Fig. 5 | Transmission Electon Microscopy characterization of nsp10/14Δ/16 complex. a,** Representative 2D-classes obtained by template-free 2D classification of
269 particles picked from NS-TEM micrographs of nsp10/14/16 complex. **b,** Rigid body fit of SAXS-derived structure of the heterotrimer complex into NS-TEM-derived
270 3D reconstitution map. Nsp14, nsp10, and nsp16 are represented as orange, red, and green ribbon models, respectively. The 3D reconstitution map is shown as a
271 transparent grey surface.
272

273

274   Discussion

275   It was demonstrated earlier that nsp10 serves as a protein co-factor of both nsp14 and nsp16. One might logically

276   expect that nsp10 could bring nsp14 and nsp16 together into a heterotrimer complex and that the kinetics and

277   processivity of the process would be positively influenced by the spatial proximity of enzymes catalysing

278   consecutive reactions in the capping pathway. The crystal structures available to date, however, suggest

279   otherwise. The binding interfaces of nsp14 and nsp16 overlap at the surface of nsp10, suggesting that the

280   heterotrimer complex is not feasible without a major structural rearrangement of either nsp14 or nsp16.

281   Careful analysis of existing data demonstrates that the nsp10/16 interface relies on a solid network of

282   hydrophobic interactions mediated by a rigid central antiparallel β1-sheet of nsp10; the helices α2, α3 and α4; a

283   coiled-coil region connecting helix α1 and the sheet β1 and α helices 3,4 and 10; as well as β-sheet 4 (Fig. 3b). In

284   particular, Val42 and Leu45 of nsp10 are embedded into hydrophobic pockets formed by helices α3, α4, and α10

285   of nsp16.

286   Despite the fact that the interface between nsp10 and nsp14 buries a larger surface area compared to the

287   nsp10/16 interface, the affinity characterizing the components of the former complex is almost an order of

288   magnitude weaker than that of the latter. This prompted us to speculate that a significant part of the nsp10/14

289   interaction may not significantly contribute to affinity. The nsp10/14 interaction surface contains two major

290   regions. Interactions within the N-terminal region of nsp14 involve primarily loops and other poorly structured

291   regions, such as the N-terminal coil-coiled region that interacts with the α1' helix of nsp10, β-turn 1, and loops

292   between β-sheet 2/3 and β-sheet 7/8 (Fig. 3c). Many of those interactions are present within the region that would

293   overlap with nsp16 if the complexes were aligned. Indeed, nsp14Δ retains nsp10 binding properties and is

294   characterized by an affinity comparable to wild type nsp14, supporting the claim that the N-terminal region does

295   not significantly contribute to interaction. This allowed us to hypothesize that the heterotrimer complex is formed

296   when nsp16 displaces the N-terminal region ("lid") of nsp14 at the surface of nsp10 (Fig. 6). Indeed, both nsp14

297   and nsp14Δ readily form a heterotrimer complex with nsp16 and nsp10.
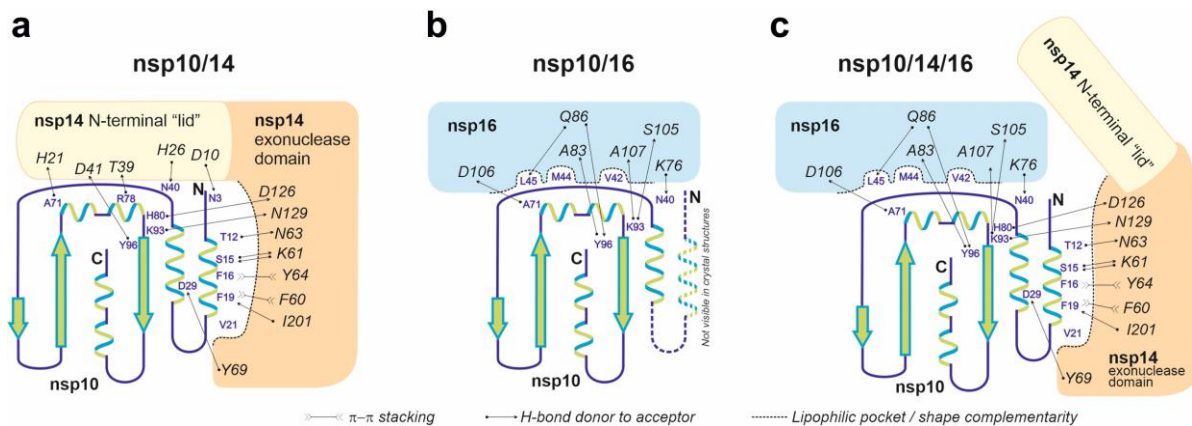
298
299   Fig. 6 | Schematic diagram of interactions guiding the affinities of nsp10, nsp14 and nsp16. a, The α1 helix of nsp10 provides several deeply buried lipophilic and π-
300   stacking interactions with the exonuclease domain of nsp14 and the interaction may be described in terms of shape complementarity. In turn, the interaction of
301   the N-terminal "lid" region (amino acids 1-50) of nsp14 with nsp10 shows poor shape complementarity and is characterized only by a low number of solvent-
302   exposed hydrogen bonds. b, Nsp16 binds nsp10 at the site overlapping that involved in binding of the "lid" of nsp14, but nsp10/16 interaction is characterized by
303   a well-developed interface involving deep lipophilic pockets and solvent-shielded hydrogen bonds. The α1 helix of nsp10 is not defined in the crystal structures of
304   the nsp10/16 complexes (6W4H, 6YZ1), indicating it is flexible and not involved in binding. c, The formation of the nsp10/14/16 heterotrimer complex is
305   accompanied by "lid" displacement and stabilization of the α1 helix.
306

307   The "lid" hypothesis is supported by low-resolution structural data provided in this work. The gyration radii

308   and molecular weight of the heterotrimer complex derived from SEC-SAXS experiments are higher than the

309   gyration radii and molecular weights of any of the components or binary complexes (Extended Data Fig. 5a). The

310   molecular envelope derived *ab inito* from the scattering profile of the heterotrimer complex perfectly fits the

311   model suggested by the "lid" hypothesis. Additionally, SAXS demonstrates that the heterotrimer complex

312   formation stabilizes conformational flexibility of the nsp10/14 complex.

313   NS-TEM reconstruction of 3D volume characterizing the nsp10/14Δ/16 heterotrimer complex accommodates

314   the SAXS-derived model with high confidence, further supporting the lid hypothesis. We were unsuccessful in our

315   effort to use cryoEM to determine a high-resolution structure of the heterotrimer complex composed of full-

316   length components due to low particle homogeneity. Low resolution reconstruction was nonetheless possible and

317   again accommodated the SAXS derived model well (Extended Data Fig. 7), further supporting the structural

318   arrangement of the components within the heterotrimer complex.

319   Overall, in this study we provide evidence that nsp14, nsp10 and nsp16 form a heterotrimer complex

320   characterized by 1:1:1 stoichiometry, built around nsp10. The architecture of the complex follows the general

321   arrangement previously observed in nsp10/16 and nsp10/14 complexes, but nsp16 displaces the "lid" (N-terminal)

322   of nsp14 at the nsp10 surface (Fig. 6). The heterotrimer complex brings together two consecutive activities

323   required for RNA cap formation (nsp14 associated N7-methyltransferase and nsp16 2'-O-methyltransferase), likely

324   contributing to the processivity of the capping process. The heterotrimer complex formation does not, however,

325   influence the methyltransferase activities. Further, the heterotrimer complex formation mitigates the 3'→5'

326   exonuclease activity of nsp14, preventing the excessive degradation of viral nucleic acid and allowing the complex

327   to switch from the proofreading mode to the methylation mode.

328

329

Materials and Methods

Protein expression and complex purification

332   Constructs of nsp10, comprising amino acids $4254 - 4392$, nsp14, comprising amino acids $5926 - 6452$, and nsp16,

333   comprising amino acids $6799 - 7096$, of SARS-CoV-2 polyprotein 1ab optimized for expression in *E. coli* were

334   ordered from GeneArt and subcloned into expression vector pETDuet-1. For the nsp14 catalytic mutant (np14-

335   cat), D90 and E92 were replaced with alanines[29]. Plasmids were then co-transformed into *E. coli* strain BL21.

336   Transformed BL21 (DE3) *E. coli* cells were grown in Terrific Broth medium (TB; Bioshop), supplemented with

337   100 µg/ml of ampicillin (Sigma), at 37°C overnight and used as a starter culture for the large-scale expression in

338   TB. After the culture reached $OD_{600}=1.2 - 1.4$, it was induced with 0.5mM isopropyl-D-1-thiogalactopyranoside

339   (IPTG; Sigma), and protein expression was carried out at 18 °C for 16 hours. Bacterial pellets were collected by

340   centrifugation at 6'000 rpm for 10 min at 4°C, resuspended in lysis buffer (50 mM Tris-HCl pH 8.5, 300 mM NaCl,

341   5 mM $MgCl_2$, 5% v/v glycerol, 5 mM β-mercaptoethanol) and disrupted by sonication at 80% amplitude for 15 min

342   at 10°C (3s, 3s pulse). The lysed sample was clarified by centrifuging for 1 h at 25'000 rpm at 4°C. The supernatant

343   was collected and incubated with 2 ml of Ni-NTA Agarose (Jena Bioscience) pre-equilibrated with the lysis buffer

344   for 2 hours at 4°C. Purification was carried out in a gravity-flow column. Resin was washed with 50 BV (bed volume)

345   of buffer A (50 mM Tris-HCl pH 8.5, 300 mM NaCl, 5 mM $MgCl_2$, 5 mM β-mercaptoethanol, 10 mM imidazole) and

346   20 BV of buffer B (50 mM Tris-HCl pH 8.5, 300 mM NaCl, 5 mM $MgCl_2$, 5 mM β-mercaptoethanol, 20 mM

347   imidazole). The protein complex was eluted with buffers C, D and E (50 mM Tris-HCl pH 8.5, 150 mM NaCl, 5 mM

348    $MgCl_2$, 5 mM β-mercaptoethanol, 100 mM (C) / 250 mM (D) / 350 mM imidazole (E), 5 × 2 BV each). The eluted

349   fractions were concentrated to 5 ml and loaded onto a HiLoad 26/600 Superdex 200 prep grade (GE Healthcare)

350   equilibrated with the SEC buffer (50 mM Tris-HCl pH 8.5, 150 mM NaCl, 5 mM $MgCl_2$, 2 mM β-mercaptoethanol).

351   Confirmed by SDS-PAGE, the protein complex fractions were collected. TEV protease was added in a combination

352   of a final concentration of 10 mM β-mercaptoethanol and incubated with gentle rocking at 4°C for approximately

353   12h. When the His-tag had been successfully cleaved from the protein complex, the sample was concentrated to

354   5 ml and loaded over the HiLoad 26/600 Superdex 200 prep grade (GE Healthcare). In order to clear the impurities

355   which had been eluted with the complex, a reverse binding was performed by incubating for 15 minutes, rolling

356   at 4°C with 100 µl of Ni-NTA Agarose (GE Healthcare) pre-equilibrated with the SEC buffer. The sample was

357   separated with a gravity-flow column, and the Ni-NTA agarose was washed three times with the SEC buffer with

358   150 mM NaCl and finally with SEC buffer supplemented with another 150 mM NaCl (totaling 300 mM NaCl). The

359   flow-through samples contained the clean protein complex, as verified by SDS-PAGE. The protein complex was

360   collected, concentrated, and loaded onto the Superdex 200 Increase 10/300 GL (GE Healthcare) equilibrated with

361   the final SEC buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 2 mM β-mercaptoethanol). The protein

362   complex was collected and concentrated up to 5 mg/ml for further applications such as crystallization, inhibitor

363   and activity screening. After every stage of purification, the protein content and purity were evaluated with SDS-

364   PAGE and InstantBlue (Sigma) gel staining (Extended Data Fig. 1c).

365

366   *Protein identification from gel bands - LC-MS/MS analysis*

367   Protein identification was performed at the Proteomics and Mass Spectrometry Core Facility, Malopolska Centre

368   of Biotechnology, Jagiellonian University, Krakow. Samples were prepared, measured and analyzed as described

369   in Pabis *et al.*[30] with minor changes. Briefly, gel bands were destained by alternating washing with 25% and 50%

370   acetonitrile (ACN) in 25 mM ammonium bicarbonate (ABC). Then, protein reduction was performed with 50 mM

371   DTT in 25 mM ABC (45 min of incubation at 37°C) followed by alkylation with 55 mM iodoacetamide (1 h of

372   incubation at room temperature in the dark). In the next steps, gel bands were washed with 50% ACN in 25 mM

373   ABC, dehydrated in 100% ACN, dried and rehydrated in 20 µl of trypsin solution (10 ng/µl in 25 mM ABC). After

374   rehydration, 40 µl of 25 mM ABC was added and samples were left for overnight incubation at 37°C. Protein

375   digestion was stopped by adding trifluoroacetic acid (TFA) to the concentration of about 0.5%. Peptides present

376   in the solution were collected and additionally extracted from the gel by dehydration with 100% ACN. The obtained

377   peptide mixtures were dried and suspended in a loading buffer (2% ACN with 0.05% TFA) for LC-MS/MS analysis,

378   carried out with a nanoHPLC (UltiMate 3000 RSLCnano System, Thermo Fisher Scientific) coupled to a Q Exactive

379   mass spectrometer (Thermo Fisher Scientific). Peptides were loaded onto a trap column (Acclaim PepMap 100

380   C18, 75 µm × 20 mm, 3 µm particle, 100 Å pore size, Thermo Fisher Scientific) at a flow rate of 5 µl/min and

381   separated on an analytical column (Acclaim PepMap RSLC C18, 75 µm × 500 mm, 2 µm particle, 100 Å pore size,

382   Thermo Fisher Scientific) at 50°C with a 60 min gradient of ACN, from 2% to 40%, in the presence of 0.05% formic

383   acid at a flow rate of 250 nl/min. The eluting peptides were ionized in a Digital PicoView 550 nanospray source

384   (New Objective) and measured with Q Exactive operated in a data-dependent mode. A Top8 method was used

385   with 35 s of dynamic exclusion. MS and MS/MS spectra were acquired with a resolution of 70'000 and 35'000,

386   respectively. The ion accumulation times were adjusted to ensure parallel filling and detection. The acquired LC-

387   MS/MS data were processed with the use of Proteome Discoverer platform (v.1.4; Thermo Scientific) and searched

388   with an in-house MASCOT server (v.2.5.1; Matrix Science, London, UK) against the database of common protein

389   contaminants (cRAP database) with manually added sequences for the proteins of interest. The following

390   parameters were applied for the database search: enzyme: trypsin; missed cleavages: up to 1; fixed modifications:

391   carbamidomethyl (C); variable modifications: oxidation (M); peptide mass tolerance: 10 ppm; fragment mass

392   tolerance: 20 mmu. Additionally, the SwissProt database, restricted to *E. coli* taxonomy, was searched to assess

393   contamination with host proteins.

394

395   *Stoichiometry determination*

396   For protein quantitation, sample separation was carried out following a simple protocol using the Prominence

397   HPLC system (2×LC-20AD pumps, SPD-M20A diode array detector, DGU-20 degasser, all from Shimadzu Corp.,

398   Kyoto, Japan). For gradient separation, a Kinetex 2.6 µm/100A C18 100 mm/2.1 mm ID HPLC column was used

399   (00D-4462AN, Phenomenex, Torrance, CA, USA). Solvents used for separation: A = water + formic acid (99.9:0.1,

400   v/v), B = acetonitrile + formic acid (99.9:0.1, v/v). All solvents were supplied by Merck local distributor (Merck

401   KgaA, Darmstadt, Germany). Gradient was set as follows: t (time)=0 min, 25% B; t=20 min, 75% B; t=20.5 min, 90%

402   B; t=25 min, 90% B; t=25.5 min, 25% B; t=35 min, 25% B (end). Flow rate was set to 0.3 ml/min. Diode array

403   detector settings: wavelength range: 200-350 nm (deuterium lamp only), sampling frequency: 5 Hz. Data

404   acquisition and data processing were controlled by LCsolution software, ver. 1.25 (Shimadzu Corp., Kyoto, Japan).

405   To confirm protein content under every chromatographic peak taken for protein quantitation, mass spectrometry-

406   based identification was used. The protocol for protein identification, applied with minor changes, is available

407   elsewhere[31]. Briefly, fractions acquired during protein separation were freeze-dried using CentriVap system

408   (Labconco, Kansas City, MO, USA) and redissolved in 70 μl 50 mM ABC (pH=7.8). Next, reduction and alkylation of

409   cysteine residues were done using DTT and following iodoacetamide 50 mM ABC solutions (both reagents: 5 mM

410   final concentrations). In both cases, 10 min incubation in 80°C with shaking was applied. After cooling down,

411   trypsin (Gold-MS grade, Promega, Madison, WI, USA) was added in a final concentration of 2 pmol per sample.

412   Samples were incubated overnight at 37°C with shaking, then freeze-dried again and redissolved in 30 μl of 4%

413   acetonitrile/water solution acidified by 0.1% formic acid (v/v/v). NanoLC-MS/MS analyses were performed on an

414   Ultimate 3000 system (Thermo, Waltham, MA, USA) connected on-line to AmaZon SL, equipped with nanoFlowESI

415   ion source (Bruker-Daltonics, Bremen, Germany). Parameters of nanoLC system: column Acclaim PepMap100,

416   C18, l=10 cm/75 μm I.D., precolumn PepMap100, C18, l=1 cm/1 mm I.D., gradient settings: solvent A = water with

417   0.1% formic acid (v/v), solvent B = acetonitrile with 0.1% formic acid (purity: MS-grade; Merck KGaA, Darmstadt,

418   Germany), t (time)=0 min, 6% B; t=5 min, 6% B; t=50 min, 55% B; t=50.1 min, 80% B; t=53 min, 80% B; t=54 min,

419   6% B; t=58 min, 6% B (end). Sample injection volume was usually in the range 3-5 μl. Flow rate: 300 nl/min, flow

420   rate for sample introduction on precolumn: 30 μl/min. Mass spectrometer settings were as follows: capillary

421   voltage: 4'200V; heated capillary temperature: 150°C; MS scan range: 375-1600 m/z; MS/MS scan range: 200-

422   2000 m/z; resolution: enhanced; scanning frequency: ca. 1 Hz; ICC (Ion Charge Control): 250'000 ions/trap cycle;

423   fragmentation ions selection range: 450-1'600 m/z (with minor exclusions); minimal ion intensity selected for

424   fragmentation: at least 1×10^6 units. Both instruments were run under the HyStar ver. 4.1 SR1 (Bruker-Daltonics,

425   Bremen, Germany). Data analysis was performed in Bruker's Compass DataAnalysis 4.4 SR1. Acquired data were

426   converted into mgf files using built-in scripts, introduced into Mascot search engine (ver. 2.6, Matrixscience,

427   London, UK), and searched against SwissProt and in-house created database. Mascot settings: enzyme: trypsin;

428   missing cleavages: 1; taxonomy: all; fixed modifications: carbamidomethylation; variable modifications: oxidation-

429   Met; peptide tolerance: 1.2 Da; #13C: 1; MS/MS tolerance: 0.6 Da; peptide charge: +1,+2,+3; instrument: ion-trap.

430

431   *MicroScale Thermophoresis*

432   His-tag proteins were labeled with Monolith His-Tag Labeling Kit RED-tris-NTA 2nd Generation according to

433   manufacturer's guidelines. Labeled proteins were diluted in PBS containing 0.05% Tween-20 up to 80 nM

434   concentration and mixed with tested ligands. Samples were allowed 30 min incubation at RT prior measurements.

435   Measurements were performed on Monolith NT.115 in duplicates using Excitation Power: 80% and MST Power:

436   high in Monolith NT.115 Capillaries.

437

NanoDSF

NanoDSF was performed in standard capillaries using Tycho equipment. Proteins and their respective complexes were measured at 1 mg/mL using default ramp temperature. The resulting melting temperatures were reported as the first derivative of the fluorescence ratio.

Methyltransferase activity

The methyltransferase activity of wild type heterotrimer nsp10/14/16 or the heterotrimer with mutated nsp14 protein (ExoN mutant and catalytic mutant – D90A and E92A) was measured using the EPIgeneous Methyltransferase kit from Cisbio as previously described[32]. Individual kit reagents were reconstituted according to the manufacturer's instructions. Briefly, the methyltransferase reaction was incubated for 20 minutes at room temperature in 8 μl reaction volume with 100 nM nsp10/14/16 wild type or mutated heterotrimer, 1 μM Ultrapure SAM (Cisbio), 187.5 μM RNA cap analogue (GpppA or GpppG, New England Biolabs) or 18.75 μM Cap 0 RNA oligo (TriLink) in reaction buffer consisting of 20 mM Tris-HCl pH 7.4, 150 mM NaCl, and 0.5 mM DTT. The reaction was quenched by the addition of 2 μl of 5M followed by the addition of 2 μl Detection Buffer 1 (Cisbio) to the reaction mixture. After 10 min, 4 μl of 16× SAH-d2 conjugate solution (Cisbio) was added. After 5 min, 4 μl of 1× α-SAH Tb Cryptate antibody solution was added to the reaction mixture and incubated for 1 hour at room temperature. Homogenous Time-Resolved Fluorescence (HTRF) measurements were performed on a SpectraMax iD5 plate reader (Molecular Devices) according to the manufacturer's guideline (excitation wavelength 340 nm, emission wavelengths 665 and 620 nm, top mode, 100 flashes, optimal gain, z position calculated from the negative control (no enzyme), lag time of 60 μs and the integration time of 500 μs). The resulting data were background-subtracted and normalized as follows. The ratio of 665 to 620 nm wavelength was calculated. The data was background-corrected on the averaged signal for the buffer control. Next, the data was normalized for each series individually on the wells not containing the enzyme.

Binding and exonuclease activity assays

Substrate for binding and exonuclease assays had a primary Sequence (5' to 3') of CoV-RNA1-X (X = G, or A) XXXXXXXXXXXCGCGUAGUUUUCUACGCG. The CoV-RNA1-A and G RNAs were ordered from IDT as a PAGE-purified and desalted oligos. Protein complexes (equimolar at 1.6 μM) were incubated with 100 ng of RNA in a buffer containing 20 mM HEPES pH 7.5, 100 mM NaCl, 5% v/v glycerol, 10 mM $MgCl_2$, and 5 mM β-mercaptoethanol, followed by analysis via a 20% urea-denaturing PAGE or a native 1% TBE agarose gel.

SAXS

Samples were measured in SEC-SAXS mode at BM29/ESRF, Grenoble France on 12.12.2020. (session ID MX2341). Samples (100 μL) were measured on Agilent AdvanceBio SEC 300 with 50mM Tris-HCl pH 8.5, 150mM NaCl, 5mM $MgCl_2$, and 2 mM β-mercaptoethanol running phase at 0.16 ml/min flowrate. The measurements were performed at 0.99 Å wavelength. The sample to detector distance was set at 2.83 m with Pilatus2M detector for data acquisition.

476 *CryoEM*

477 Out of 100 randomly chosen micrographs, 1'000 particles were manually picked without any structural knowledge

478 about the complex to minimize bias and assigned with 2D classes that were used in the *ab initio* model built. Out

479 of generated 3D classes, one was manually picked and used for the training of TOPAZ neural networks, which in

480 turn picked the next interaction of particles that were used to retrain Topaz. With this approach, approximately

481 0.5 M particles were selected to generate 50 2D classes, out of which 19 were manually picked.

482

483 *Crosslinking and Negative-Staining Transmission Electron Microscopy*

484 Following SEC in HEPES buffer, the purified protein complex was crosslinked with bis(sulfosuccinimidyl)suberate

485 (BS³, Thermo Scientific), an amine-to-amine crosslinker. The protein sample was incubated with 0.5 mM BS^3 (from

486 50 mM stock) for 30 min at RT. The crosslinking reaction was quenched with 1M Tris pH 7.5 to a final 50 mM Tris,

487 and incubated 15 min at RT. The excess BS^3 crosslinker was cleared by sequential dilution and concentration with

488 Tris buffer. Concentrated samples were kept at 4°C or stored at -80°C for further measurements. Negative-stain

489 transmission electron microscopy (NS-TEM) measurements were done in Formvar/carbon-supported 400 mesh

490 copper grids, suspended in air with a negative lock tweezer. The purified protein complex (0.03 mg/mL) was

491 applied on glow-discharged Formvar/carbon-supported 400 mesh copper grids and negatively stained with 1%

492 neutralized uranyl-acetate. Grids were imaged using the JEOL JEM 2100HT electron microscope (Jeol Ltd, Tokyo,

493 Japan) at accelerating voltage 200 kV. Images were taken by using a 4kx4k camera (TVIPS) equipped with EMMENU

494 software ver. 4.0.9.87.

495

496 *Image processing and 3-D reconstitution*

497 Collected micrographs were processed using cryoSPARC 3.1.1. Initially, 9'350 particles were picked from

498 micrographs using Blob Picker. Picked particles were subjected to a template-free 2D classification, from which

499 1'216 particles were selected and subjected to 3D reconstitution using the ab-initio reconstitution job. The

500 nsp10/14/16 complex map derived from SAXS data was used for a rigid-body fit in to 3D-reconstitution map using

501 Dock in map.

502

Acknowledgements

516   ## Declaration of interests

517   None.

518

519   REFERENCES

520   1   Pyrc, K., Jebbink, M. F., Berkhout, B. & van der Hoek, L. Genome structure and transcriptional
521   regulation of human coronavirus NL63. *Virol J* **1**, 7, doi:10.1186/1743-422X-1-7 (2004).
522   2   Gao, Y. *et al.* Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science*
523   **368**, 779-782, doi:10.1126/science.abb7498 (2020).
524   3   Yan, L. *et al.* Cryo-EM Structure of an Extended SARS-CoV-2 Replication and Transcription
525   Complex Reveals an Intermediate State in Cap Synthesis. *Cell* **184**, 184-193 e110,
526   doi:10.1016/j.cell.2020.11.016 (2021).
527   4   Wang, Q. *et al.* Structural Basis for RNA Replication by the SARS-CoV-2 Polymerase. *Cell* **182**,
528   417-428 e413, doi:10.1016/j.cell.2020.05.034 (2020).
529   5   Kabinger, F. *et al.* Mechanism of molnupiravir-induced SARS-CoV-2 mutagenesis. *Nat Struct*
530   *Mol Biol* **28**, 740-746, doi:10.1038/s41594-021-00651-0 (2021).
531   6   Yan, L. *et al.* Coupling of N7-methyltransferase and 3'-5' exoribonuclease with SARS-CoV-2
532   polymerase reveals mechanisms for capping and proofreading. *Cell* **184**, 3474-3485 e3411,
533   doi:10.1016/j.cell.2021.05.033 (2021).
534   7   Sevajol, M., Subissi, L., Decroly, E., Canard, B. & Imbert, I. Insights into RNA synthesis, capping,
535   and   proofreading   mechanisms   of   SARS-coronavirus. *Virus   Res* **194**,   90-99,
536   doi:10.1016/j.virusres.2014.10.008 (2014).
537   8   Bouvet, M. *et al.* RNA 3'-end mismatch excision by the severe acute respiratory syndrome
538   coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. *Proc Natl Acad Sci*
539   *U S A* **109**, 9372-9377, doi:10.1073/pnas.1201130109 (2012).
540   9   Eckerle, L. D., Lu, X., Sperry, S. M., Choi, L. & Denison, M. R. High fidelity of murine hepatitis
541   virus replication is decreased in nsp14 exoribonuclease mutants. *J Virol* **81**, 12135-12144,
542   doi:10.1128/jvi.01296-07 (2007).
543   10   Daffis, S. *et al.* 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family
544   members. *Nature* **468**, 452-456, doi:10.1038/nature09489 (2010).
545   11   Ivanov, K. A. *et al.* Multiple enzymatic activities associated with severe acute respiratory
546   syndrome coronavirus helicase. *J Virol* **78**, 5619-5632, doi:10.1128/JVI.78.11.5619-5632.2004
547   (2004).
548   12   Ivanov, K. A. & Ziebuhr, J. Human coronavirus 229E nonstructural protein 13: characterization
549   of duplex-unwinding, nucleoside triphosphatase, and RNA 5'-triphosphatase activities. *J Virol*
550   **78**, 7833-7838, doi:10.1128/JVI.78.14.7833-7838.2004 (2004).
551   13   Decroly, E. *et al.* Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing
552   (nucleoside-2'O)-methyltransferase activity. *J Virol* **82**, 8071-8084, doi:10.1128/JVI.00407-08
553   (2008).
554   14   Chen, Y. *et al.* Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a
555   novel   cap   N7   methyltransferase. *Proc   Natl   Acad   Sci   U   S   A* **106**,   3484-3489,
556   doi:10.1073/pnas.0808790106 (2009).
557   15   Bouvet, M. *et al.* In vitro reconstitution of SARS-coronavirus mRNA cap methylation. *PLoS*
558   *Pathog* **6**, e1000863, doi:10.1371/journal.ppat.1000863 (2010).

559 16   Slanina, H. *et al.* Coronavirus replication-transcription complex: Vital and selective NMPylation
560   of a conserved site in nsp9 by the NiRAN-RdRp subunit. *Proc Natl Acad Sci U S A* **118**,
561   doi:10.1073/pnas.2022310118 (2021).
562 17   Bouvet, M. *et al.* Coronavirus Nsp10, a critical co-factor for activation of multiple replicative
563   enzymes. *J Biol Chem* **289**, 25783-25796, doi:10.1074/jbc.M114.577353 (2014).
564 18   Czarna, A. *et al.* Refolding of lid subdomain of SARS-CoV-2 nsp14 upon nsp10 interaction
565   releases exonuclease activity. Preprint at *bioRxiv*, 2022.2002.2017.480845,
566   doi:10.1101/2022.02.17.480845 (2022).
567 19   Newman, J. A., Imprachim, N., Yosaatmadja, Y. & Gileadi, O.   (RCSB Protein Data Bank, 2022).
568 20   Kozielski, F. *et al.* Identification of fragments binding to SARS-CoV-2 nsp10 reveals ligand-
569   binding sites in conserved interfaces between nsp10 and nsp14/nsp16. *RSC Chem Biol* **3**, 44-
570   55, doi:10.1039/d1cb00135c (2022).
571 21   Aouadi, W. *et al.* Binding of the Methyl Donor S-Adenosyl-l-Methionine to Middle East
572   Respiratory Syndrome Coronavirus 2'-O-Methyltransferase nsp16 Promotes Recruitment of
573   the Allosteric Activator nsp10. *J Virol* **91**, doi:10.1128/JVI.02217-16 (2017).
574 22   Schlee, M. & Hartmann, G. Discriminating self from non-self in nucleic acid sensing. *Nature*
575   *Reviews Immunology* **16**, 566-580, doi:10.1038/nri.2016.78 (2016).
576 23   Ahmed-Belkacem, R. *et al.* Synthesis of adenine dinucleosides SAM analogs as specific
577   inhibitors of SARS-CoV nsp14 RNA cap guanine-N7-methyltransferase. *Eur J Med Chem* **201**,
578   112557, doi:10.1016/j.ejmech.2020.112557 (2020).
579 24   Riccio, A. A., Sullivan, E. D. & Copeland, W. C. Activation of the SARS-CoV-2 NSP14 3'-5'
580   exoribonuclease by NSP10 and response to antiviral inhibitors. *J Biol Chem* **298**, 101518,
581   doi:10.1016/j.jbc.2021.101518 (2021).
582 25   Krafcikova, P., Silhan, J., Nencka, R. & Boura, E. Structural analysis of the SARS-CoV-2
583   methyltransferase complex involved in RNA cap creation bound to sinefungin. *Nat Commun*
584   **11**, 3717, doi:10.1038/s41467-020-17495-9 (2020).
585 26   Viswanathan, T. *et al.* Structural basis of RNA cap modification by SARS-CoV-2. *Nat Commun*
586   **11**, 3718, doi:10.1038/s41467-020-17496-8 (2020).
587 27   Franke, D. & Svergun, D. I. DAMMIF, a program for rapid ab-initio shape determination in
588   small-angle scattering. *J Appl Crystallogr* **42**, 342-346, doi:10.1107/S0021889809000338
589   (2009).
590 28   Panjkovich, A. & Svergun, D. I. Deciphering conformational transitions of proteins by small
591   angle X-ray scattering and normal mode analysis. *Phys Chem Chem Phys* **18**, 5707-5719,
592   doi:10.1039/c5cp04540a (2016).
593 29   Ma, Y. *et al.* Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10
594   complex. *Proc Natl Acad Sci U S A* **112**, 9436-9441, doi:10.1073/pnas.1508686112 (2015).
595 30   Pabis, M. *et al.* Molecular basis for the bifunctional Uba4-Urm1 sulfur-relay system in tRNA
596   thiolation and ubiquitin-like conjugation. *EMBO J* **39**, e105087,
597   doi:10.15252/embj.2020105087 (2020).
598 31   Bodzon-Kulakowska, A. *et al.* Proteomic analysis of striatal neuronal cell cultures after
599   morphine administration. *J Sep Sci* **32**, 1200-1210, doi:10.1002/jssc.200800464 (2009).
600 32   Basu, S. *et al.* Identifying SARS-CoV-2 antiviral compounds by screening for small molecule
601   inhibitors of Nsp14 RNA cap methyltransferase. *Biochem J* **478**, 2481-2497,
602   doi:10.1042/bcj20210219 (2021).

603

604
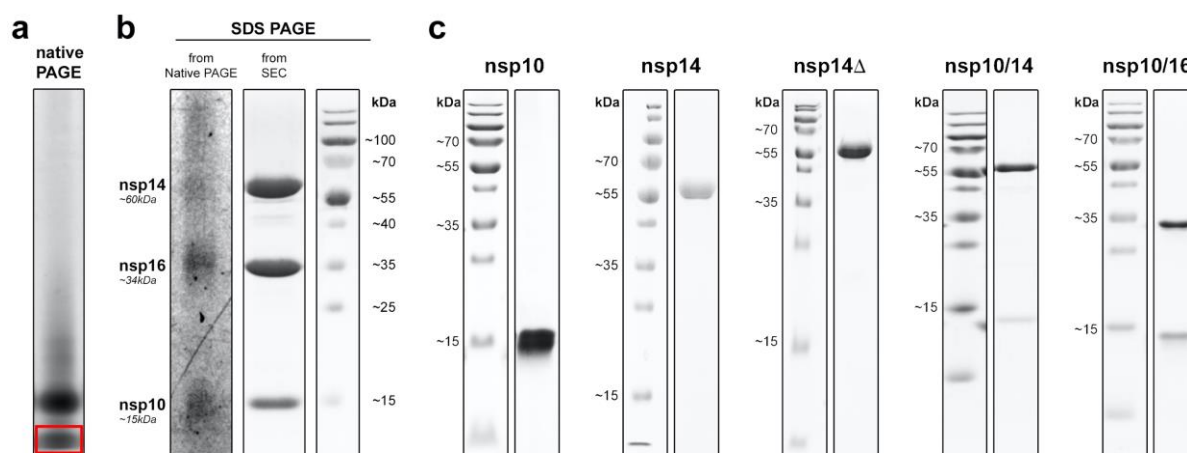
Extended Data

Extended Data Fig. 1 | a, Native-PAGE of heterotrimer. **b**, SDS-PAGE of the band excited from the native-PAGE (red box showing disintegration of the heterotrimer into nsp14, nsp16, and nsp10. **c**, SDS-PAGE illustrating the purity of nsp10, nsp14, nsp14Δ, nsp10/14, and nsp10/16 proteins.

Extended Table 1. Qualitative MS analysis of the native PAGE gel band indicated in red box, in Extended Data Fig. 1a.

Description	Score	Coverage
>sp|Nsp14tag_PROTEIN|	10'464.37	85.61
>sp|Nsp14_PROTEIN|	10'156.82	85.28
>sp|Nsp16tag_PROTEIN	9'601.49	86.83
>sp|Nsp10tag_PROTEIN|	7'572.82	100.00
>sp|Nsp10_PROTEIN|	7'408.26	100.00
>sp|K1C10_HUMAN|	3'205.43	56.83
>sp|K2C1_HUMAN|	2'357.77	47.90
>sp|K1C9_HUMAN|	2'156.15	48.96
>sp|TRYP_PIG|	1'814.17	35.50
>sp|K22E_HUMAN|	1'653.50	43.26
>sp|3264P0DTD13569_PROTEIN|	398.47	22.55
>sp|K1C15_SHEEP|	215.66	9.27
>sp|CAS1_BOVIN|	169.83	22.43
>sp|ALBU_HUMAN|	155.74	8.05
>sp|ALBU_BOVIN|	146.29	8.40
>sp|CASB_BOVIN|	59.51	8.48
>sp|MINDY3mod_HUMAN|	21.06	1.34
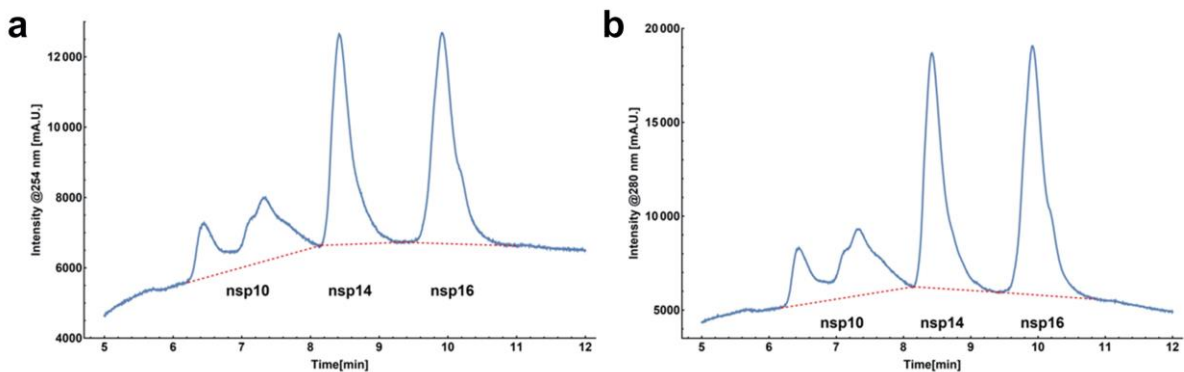
613

614 Extended Data Fig. 2 | Mass spectroscopy of the heterotrimer at 254 and 280 nm.

615

616 Extended Data Table 2 | Summary of the peak volumes from Extended Data Fig. 2 presenting 1.2:1:1 ratio of nsp10/14/16.

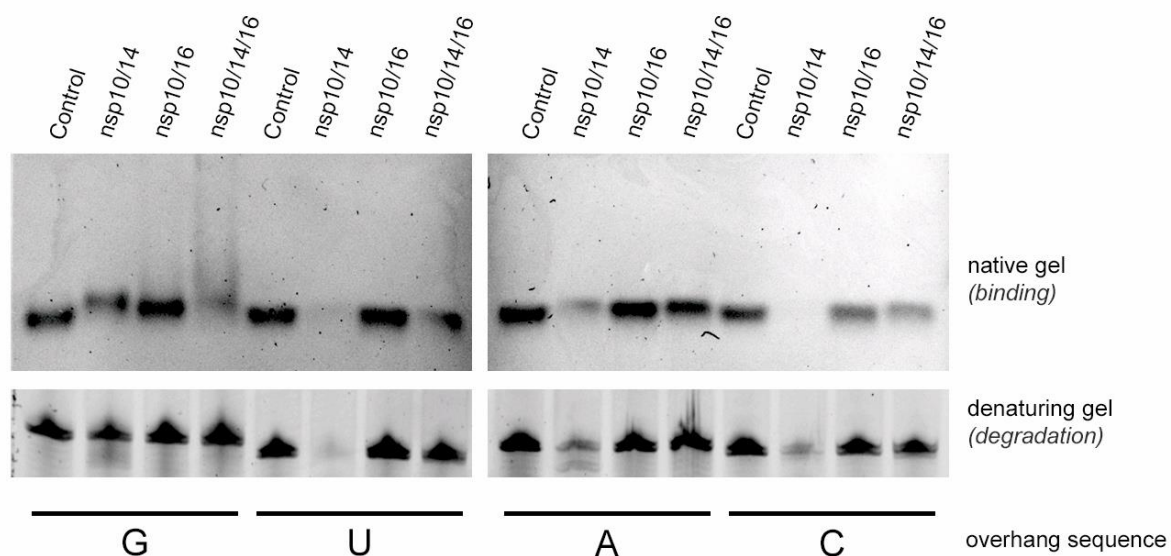|	Peak 1 (fr 6.3-8.0)	Peak 2 (fr. 8.2-9.3)	Peak 3 (fr. 9.6-10.7)	Average surface area	Total	Percent	Ratio
nsp10	64.5% (100)	22% (19.6)	13.5% (10.9)	145'756	205'562	38%	1.2
nsp14	0% (0.0)	90% (80.4)	10% (8.2)	187'293	167'747	31%	1.0
nsp16	0% (0.0)	0% (0.0)	100% (80.9)	211'899	171'638	31%	1.0

617
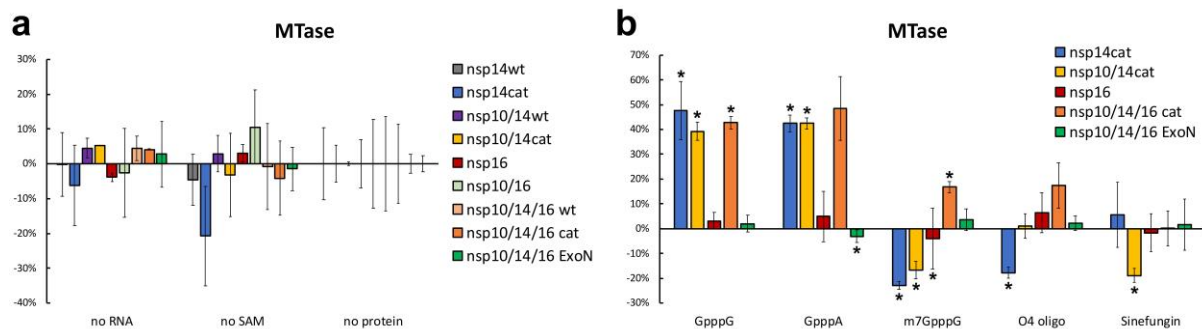
618
619 Extended Data Fig. 3 | Modulation of the methyltransferase activity by the protein partners.

620

621 Extended Data Fig. 4 | Uncropped RNA gells presented on Fig. 2.

622

623 **Extended Data Table 3.** Summary of the structural parameters derived from scattering profiles.

a. Sample details

	nsp14	nsp10/14	nsp10/16	nsp10/14/16
Source organism	*Severe acute respiratory syndrome coronavirus 2*			
Source	*E. coli* BL21			
UniProt sequence ID (residues in construct)	P0DTD1 (5926-6452)	P0DTD1 (5926-6452)	-	P0DTD1 (5926-6452)
	-	P0DTC1 (4254-4392)	P0DTC1 (4254-4392)	P0DTC1 (4254-4392)
	-	-	P0DTD1 (6799-7096)	P0DTD1 (6799-7096)
Molecular weight from chemical composition (kDa)	60	75	48.5	108.5
SEC–SAXS column	AdvanceBio Bio SEC 300 column			
Loading concentration (mg/mL)	4	3.5	2.6	2.8
Injection volume (µL)	100			
Flow rate (mg/mL)	0.16			
Running phase composition	50mM Tris, 150mM NaCl, 5mM $MgCl_2$, βME 2mM pH8.5			

b. SAXS data-collection parameters

Instrument	BM29, ESRF, Grenoble France
Detector	Pilatus2M
Wavelength	0.99 Å
Sample to detector distance	2.83 m

c. Structural parameters

	nsp14	nsp10/14	nsp10/16	nsp10/14/16
Guinier analysis				
$I(0)$ (cm^{-1})	8.43 ± 0.13	3.22 ± 0.13	1.73 ± 0.09	5.26 ± 0.11
R_g (Å)	28.0 ± 0.5	30.1 ± 1.8	21.0 ± 1.5	40.3 ± 1.0
qR_g ($Å^{-1}$)	0.59 − 1.51	0.59 − 1.32	0.52 − 1.39	0.76 − 1.4
P(r) analysis				
R_g(Å)/ $I(0)$ (cm^{-1})	29.4/8.61	32.5/3.28	21.1/1.64	41.2/5.23
Guinier R_g(Å)/ $I(0)$ (cm^{-1})	29.4/8.61	32.7/3.28	21.2/1.64	41.5/5.23
r_{max}(Å)	95.8	122.0	80.0	140.0
Total quality estimate	0.94	0.78	0.75	0.77
Molecular weight estimate/predicted (kDa)	33.1/60	28.9/75	20.6/48.5	83.2/108.5
Oligomerization state	monomeric			

d. Shape model-fitting results

DAMMIF (10 runs)	nsp14	nsp10/14	nsp10/16	nsp10/14/16
q_{max} range for fitting($Å^{-1}$)	0.26	0.26	0.21	0.17
Symmetry, anisotropy assumptions	P1, none			
NSD (standard deviation)	1.42 (0.07)	1.10 (0.07)	1.29 (0.14)	0.80 (0.08)
Chi-squared	1.16	1.09	1.07	1.08
Resolution (from SASRES) (Å)	43 ± 3	39.3 ± 3	33.3 ± 3	38.3 ± 3
SASDBD IDs	SASDKT6	SASDKU6	SASDKV6	SASDKW6

e. Oligomer volume fractions

| volume fractions Chi^2 1.06				
nsp10/16 (6YZ1)	nsp16	nsp14	nsp10	nsp10/14 (5CU8U)
53%	0%	47%	0%	0%

624
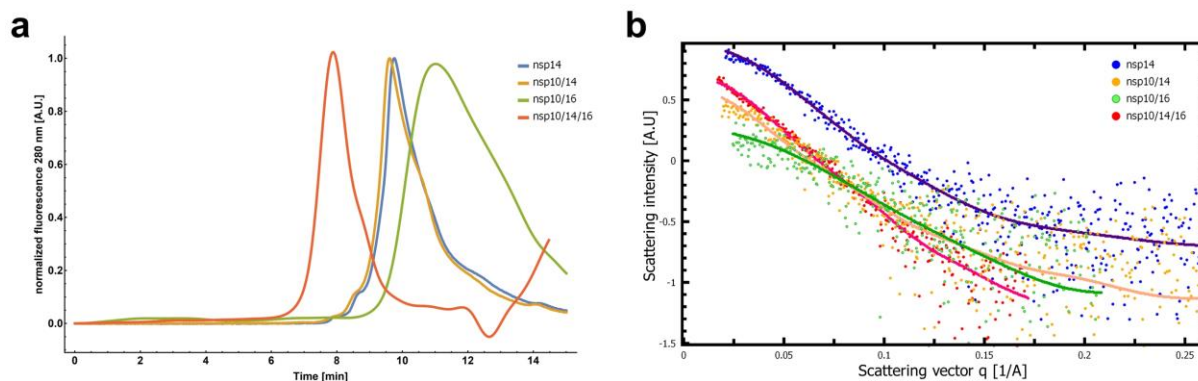
Extended Data Fig. 5 | SEC-SAXS. a, The SEC profiles for nsp10/14/16 (red), nsp10/14 (yellow), nsp14 (blue) and nsp10/16 (green) directly before flow cell for SAXS measurement. The 280 nm fluorescence intensity was normalized for clarity. **b,** SAXS scattering profiles resulting from merging the signal form SEC-SAXS experiments for nsp10/14/16 (red), nsp10/14 (yellow), nsp14 (blue) and nsp10/16 (green). The solid lines represents the fit of the experimental data to the real space models. In both cases the position closer to the left (lower elution volumes or the curvature change at smaller scattering vector values q) indicate larger objects that were analyzed. Therefore, nsp10/14/16 represents the largest from the analyzed protein complexes, while nsp10/16 the smallest, which is in the agreement with their theoretical masses calculated from the amino acid sequences.
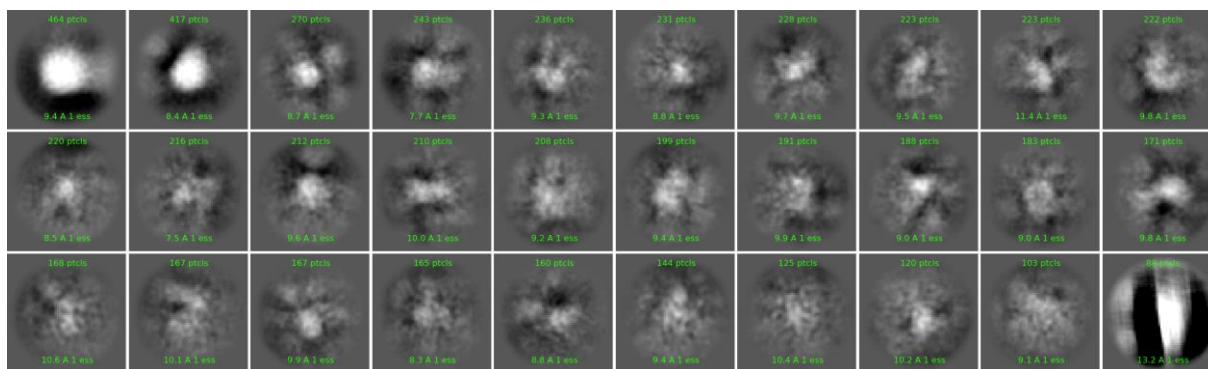
Extended Data Fig. 6 | 2D classes picked in cryoSparc for full length heterotrimer.

CryoEM

We attempted to solve a high-resolution heterotrimer structure using cryoEM techniques. Highly purified heterotrimer solution was vitrified on grids using Vitrobot under various conditions. Resulting grids were measured using Titan Krios G3i at Solaris, Poland. Data analysis was performed using cryoSPARC software. We managed to train the TOPAZ neural networks to pick approximately 0.5 M particles used to generate 50 2D classes, out of which 20 were manually picked. Generated classes, though noisy, exhibit croissant-like shape with four distinct domains that correspond to nsp16, nsp10 and two domains of nsp14 (highlighted with arrows on the class that shows side-on view). Unfortunately, we were not able to reconstruct a high-resolution 3D structure from collected data. Our best attempt shown below in cyan has ca. 9 Å resolution. The fitted heterotrimer hybrid model based on SAXS data (gold) shows overall good fit with the extra volume that may arise from the flexible nature of the heterotrimer.
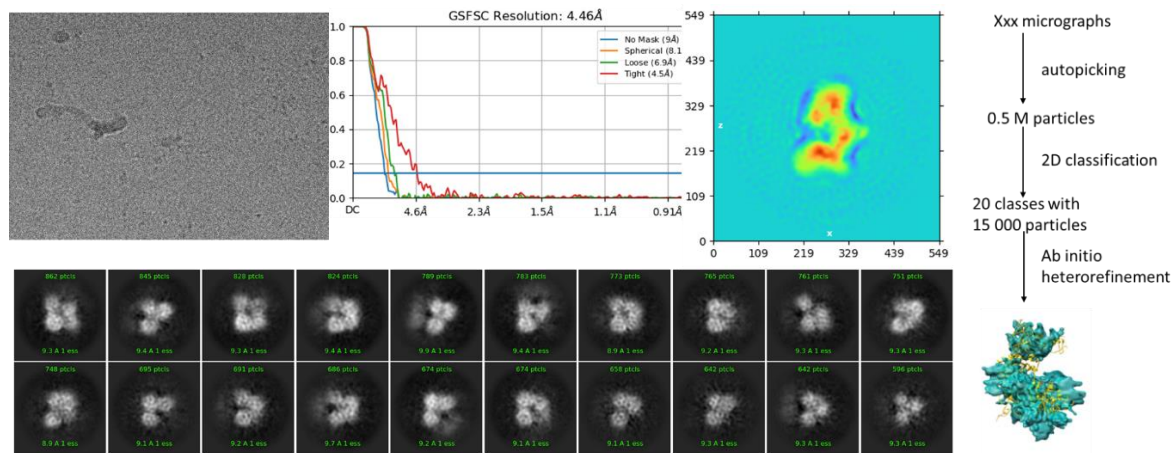
647
648 **Extended Data Fig. 7 | An overview of cryoEM data analysis in cryoSPARC.** The overlay of generated 3D model from cryoEM (cyan) and SAXS model (gold).
649