bioRxiv preprint doi: https://doi.org/10.1101/2020.09.30.320762. this version posted October 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

# 1 Hallmarks of Alpha- and Betacoronavirus non-

# 2 structural protein 7+8 complexes

Boris Krichel<sup>1</sup>, Ganesh Bylapudi<sup>2</sup>, Christina Schmidt<sup>3</sup>, Clement Blanchet<sup>4</sup>, Robin
 Schubert<sup>3</sup>, Lea Brings<sup>3</sup>, Martin Koehler<sup>5</sup>, Renato Zenobi<sup>5</sup>, Dmitri Svergun<sup>4</sup>, Kristina
 Lorenzen<sup>3</sup>, Ramakanth Madhugiri<sup>2</sup>, John Ziebuhr<sup>2</sup>, Charlotte Uetrecht<sup>1,3\*</sup>

- <sup>6</sup> <sup>1</sup>Heinrich Pette Institute, Leibniz Institute for Experimental Virology, Hamburg, Germany;
- <sup>7</sup> <sup>2</sup> Institute of Medical Virology, Justus Liebig University Giessen, Giessen, Germany;
- 8 <sup>3</sup>European XFEL GmbH, Schenefeld, Germany;
- 9 <sup>4</sup>EMBL Hamburg c/o DESY, Notkestraße 85, 22607 Hamburg, Germany
- <sup>5</sup>ETH Zurich D-CHAB Lab of Organic Chemistry, Zürich, Switzerland;
- 11

12 \*correspondence to <u>charlotte.uetrecht@xfel.eu</u>

13

#### 14 Abstract

15 Coronaviruses infect many different species including humans. The last two decades have 16 seen three zoonotic coronaviruses with SARS-CoV-2 causing a pandemic in 2020. 17 Coronaviral non-structural proteins (nsp) built up the replication-transcription complex 18 (RTC). Nsp7 and nsp8 interact with and regulate the RNA-dependent RNA-polymerase and 19 other enzymes in the RTC. However, the structural plasticity of nsp7+8 complex has been 20 under debate. Here, we present the framework of nsp7+8 complex stoichiometry and 21 topology based on a native mass spectrometry and complementary biophysical techniques of 22 nsp7+8 complexes from seven coronaviruses in the genera Alpha- and Betacoronavirus 23 including SARS-CoV-2. Their complexes cluster into three groups, which systematically form 24 either heterotrimers or heterotetramers or both, exhibiting distinct topologies. Moreover, 25 even at high protein concentrations mainly heterotetramers are observed for SARS-CoV-2 nsp7+8. From these results, the different assembly paths can be pinpointed to specific
residues and an assembly model is proposed.

#### 28 Introduction

29 Seven coronaviruses (CoV) from six coronavirus species are known to cause infections in 30 humans. While four of these viruses (HCoV-229E, HCoV-NL63, HCoV-OC43, HCoV-HKU1) 31 predominantly cause seasonal outbreaks of (upper) respiratory tract infections with mild 32 disease symptoms in most cases, three other coronaviruses (SARS-CoV, MERS-CoV and 33 SARS-CoV-2) of recent zoonotic origin are associated with lower respiratory tract disease 34 including acute respiratory distress syndrome (ARDS) [1-3]. SARS-CoV-2 is the etiologic 35 agent of COVID-19, a respiratory disease with a wide spectrum of clinical presentations and 36 outcomes. First detected in December 2019, it quickly became pandemic with numbers still 37 growing (>30 million confirmed cases, >1,000,000 deaths, by end September 2020) [4-6]. 38 COVID-19 caused major perturbations of historical dimensions in politics, economics and 39 healthcare. Moreover, coronaviruses are important, widespread animal pathogens as 40 illustrated by feline intestine peritonitis virus (FIPV) causing a severe and often fatal disease 41 in cats [7] or porcine coronaviruses [8], such as transmissible-gastroenteritis virus (TGEV) or porcine epidemic diarrhea virus (PEDV), the latter causing massive outbreaks and economic 42 43 losses in swine industry.

44 The viral replication machinery is largely conserved across the different coronavirus species 45 from the four currently recognized genera Alpha-, Beta-, Gamma- and Deltacoronavirus 46 (subfamily Orthocoronavirinae, family Coronaviridae) [9]. The key components are generally 47 referred to as nonstructural proteins (nsp) and encoded by the viral replicase genes (ORFs 1a 48 and 1b) and translated as parts of the replicase polyproteins pp1a (nsp1-11) or pp1ab (nsp1-49 16). Translation of the ORF1b-encoded C-terminal part of pp1ab requires a ribosomal (-1)-50 frameshift immediately upstream of the ORF1a stop codon. Two proteases called PLpro (one 51 or two protease domains in nsp3) and Mpro (also called 3CLpro or nsp5) facilitate polyprotein 52 processing into 16 (sometimes 15) mature nsps. The majority of these nsps form a membrane-53 anchored, highly dynamic protein-RNA machinery, the replication-transcription complex 54 (RTC), which mediates replication of the ~30 kb single-strand (+)-sense RNA genome and 55 production of subgenomic mRNAs [9, 10].

The main CoV-RTC building block is the fastest known RNA-dependent RNA-polymerase (RdRp) residing in the nsp12 C-terminal domain [11]. For RdRp activity, nsp12 requires binding to its cofactors nsp7 and nsp8 [12]. Recently, high-resolution structures illuminated two binding sites at nsp12, the first for an nsp7+8 (1:1) heterodimer and the second for a single nsp8 [13-16]. For *in vitro* RdRp activity assays, different methods were used to assemble the polymerase complex [11, 17, 18]. So far, the highest processivity *in vitro* was obtained by mixing nsp12 with a flexibly linked nsp7L8 fusion protein.

63 Recently, we reported that SARS-CoV nsp7 and nsp8 form a heterotetramer (2:2) in solution, in which nsp7 subunits have no self-interaction and rather sandwich an nsp8 scaffold with 64 putative head-to-tail interactions [19]. Current knowledge of full-length nsp7+8 complexes is 65 mainly based on two X-ray crystal structures, each of which displays a different quaternary 66 conformation. First, a SARS-CoV nsp7+8 (8:8) hexadecamer is assembled from four (2:2) 67 68 heterotetramers with similar topologies but two distinct conformations, T1 and T2, which are 69 both consistent with our in solution structure [19] [20]. Second, in a feline coronavirus (FIPV) 70 nsp7+8 (2:1) heterotrimer, nsp8 is associated to two nsp7 molecules that self-interact [21]. 71 Moreover, structures of SARS-CoV and SARS-CoV-2 with N-terminally truncated forms of 72 nsp8, thus lacking the self-interaction domain, revealed heterotetrameric nsp7+8 complexes 73 around an nsp7 scaffold [22, 23].

74 Current knowledge of coronavirus nsp7+8 complexes suggests a remarkable architectural 75 plasticity but is unsupportive of deducing common principles of complex formation. 76 Moreover, it is unknown if the quaternary structure of nsp7+8 is conserved within a given 77 coronavirus species or between genera. To fill these knowledge gaps, we analyzed nsp7+8 78 complexes derived from seven viruses of the Alpha- and Betacoronavirus genera, including a 79 range of human coronaviruses, namely SARS-CoV, SARS-CoV-2, MERS-CoV and HCoV-80 229E (Table S 2). We used native mass spectrometry (MS) to illustrate the landscape of 81 nsp7+8 complexes in vacuo, collision induced dissociation tandem MS (CID-MS/MS) to 82 reconstruct complex topology and complementary methods to verify the results [24, 25]. Our 83 findings reveal distinct sets of nsp7+8 complexes for the different CoV species. The results 84 hint at the properties that lead to complex heterogeneity and suggest common principles of 85 complex formation based on two conserved binding sites.

# 86 Experimental Results

88

# 87 Native MS illustrates the landscape of nsp7+8 complexes.



89 Figure 1: Three complexation groups formed by CoV nsp7+8. Representative mass spectra showing distinct 90 nsp7+8 complexation patterns that were classified into the three groups A, B and AB. Complex formation 91 triggered by MPro (M) mediated cleavage of 15 µM CoV nsp7-8-His6or MERS-CoV nsp7-11-His6 precursors in 92 300 mM ammonium acetate (AmAc), 1 mM DTT, pH 8.0. (A) SARS-CoV-2 representing group A forms 93 heterotetramers (nsp7+8 (2:2), red), (B) FIPV from group B forms heterotrimers (nsp7+8 (2:1), blue) and (C) HCoV-94 229E from group AB forms both complex stoichiometries (2:2 and 2:1). (D) In case of MERS-CoV, also group AB, 95 cleavage of the longer nsp7-11-His6 precursor resulted in additional processing intermediates (labeled A, B, C). In 96 all three complexation groups, nsp7+8 (1:1) heterodimeric intermediates are observed (green). For spectra of all 97 seven CoV nsp7+8 complexes see Figure S 2.

98 To ensure authentic nsp7 and nsp8 N- and C-termini, which allow for optimal nsp7+nsp8 99 complex assembly, the proteins are expressed as nsp7-8-His<sub>6</sub> polyprotein precursors and 100 cleaved by their cognate protease M<sup>pro</sup> (Figure S 1). Native MS provides an overview of mass 101 species in solution, while CID-MS/MS confirms the stoichiometry of protein complexes. 102 Distinct oligomerization patterns of nsp7+8 (1:1) heterodimers, (2:1) heterotrimers and (2:2) heterotetramers in the different CoV allowed us to categorize their nsp7+8 complexes into 103 104 three groups (Figure 1, Figure S 2). SARS-CoV and SARS-CoV-2 (species Severe acute 105 respiratory syndrome-related coronavirus, genus Betacoronavirus) represent nsp7+8 group A 106 complex formation pattern (Figure 1 A). Consistent with our previous work, SARS-CoV 107 nsp7+8 complexes exist primarily as a heterotetramer comprising two copies of each nsp7 108 and nsp8 (2:2) [19]. Expectedly, SARS-CoV-2 nsp7+8 form identical (2:2) complexes given the 109 high sequence identity of 97.5 % in the nsp7-8 region (Table S 2). Next, relative peak 110 intensities in native MS of nsp7+8 complexes are converted in a semi-quantitative analysis 111 into abundances of complex species [26]. The heterodimer (2-4%) is much less abundant 112 than the heterotetramer (96-98 %) suggesting high affinity and hence efficient conversion of 113 heterodimeric intermediates into heterotetramers. Hence, group A only forms two types of 114 nsp7+8 complexes, heterodimers (1:1) and -tetramers (2:2), with the latter clearly being 115 predominant.

In FIPV and TGEV from the species *Alphacoronavirus 1*, genus *Alphacoronavirus*, nsp7 and nsp8 proteins share a sequence identity of 93.9% (Table S 2). Their nsp7+8 complexes are assigned to group B forming predominantly nsp7+8 (2:1) heterotrimers (83 %) and to a lesser extent heterodimers (1:1) (~17 %) (Figure 1 B). An nsp7+8 (2:1) heterotrimeric structure has previously been reported for FIPV but not for TGEV or any other CoV. The association of a single nsp8 with two nsp7 indicates that group B nsp7+8 complexes lack the ability to form tetramers around an nsp8 scaffold.

123 The third oligomerization pattern is observed for nsp7+8 of HCoV-229E and PEDV, which 124 represent different species in the genus Alphacoronavirus. They share only 70.9 % sequence 125 identity in the nsp7-8 region and even less (42-62 %) with the other CoV species examined 126 (Table S 2). PEDV and HCoV-229E nsp7+8 form three major types of oligomers with slightly 127 different efficiencies: heterodimers (1:1), heterotrimers (2:1) and heterotetramers (2:2) 128 (HCoV-229E: 20 %, 12 % and 69 %; PEDV: 52 %, 6 % and 42 %, respectively) (Figure 1 C). By 129 forming both, heterotrimers and -tetramers, these complexes combine properties described 130 above for groups A and B, and are hence categorized into a separate group named 131 accordingly AB. This begs the question whether assembly pathways and structures of 132 heterotetramers in group A and AB are similar. Either, two heterodimers form a heterotetramer around an nsp8 scaffold as in group A [19] or alternatively the heterotrimer recruits another nsp8 subunit to the complex, thus employing an nsp7 core [21]. The latter pathway has recently been reported for SARS-CoV-2 nsp7+8 heterotetramers containing *N*terminally truncated nsp8 [23].

137 Additionally, nsp7+8 complexation after Mpro mediated cleavage of a MERS-CoV nsp7-11-138 His6 precursor is compared (Figure 1 D). This larger precursor, comprising nsp7, nsp8, nsp9 139 nsp10 and nsp11, behaves similar to nsp7-9-His6 (Figure S1) and is used because initial 140 attempts to cleave nsp7-8 only constructs failed. Proteolytic processing of this polyprotein precursor leads to cleavage intermediates (Figure 1 D). Such processing intermediates have 141 142 been proposed to occur intracellularly and to function distinctly from the individual nsps in e.g. regulation of RTC assembly and viral RNA synthesis [27]. Here, signal intensities of 143 144 these intermediates provide insights into the processing sequence. Surprisingly, the 145 dominant intermediate is nsp10-11-His6, despite the small size of nsp11 and a hence expected 146 high accessibility of the nsp10/11 cleavage site. Therefore, slow cleavage and prolonged 147 presence of an nsp10-11 intermediate may have functional implications warranting further 148 studies. Notably, in many CoV polyproteins the nsp10/11 and/or nsp10/12 cleavage sites 149 contain replacements (Pro in MERS-CoV) of the canonical P2 Leu residue conserved 150 throughout most M<sup>pro</sup> cleavage sites, suggesting that slow or incomplete cleavage is 151 beneficial for these particular sites. Moreover, this cleavage site has different C-terminal 152 contexts in the two CoV replicase polyproteins, nsp10-11 in pp1a and nsp10-12 in pp1ab. 153 While the structure of the small nsp11 (~1.5 kDa) is unknown, nsp12 is a large folded protein 154 (~105 kDa), which potentially improves the accessibility of the nsp10/12 site for M<sup>pro</sup>. Similar 155 effects have been observed for the nsp8/9 cleavage site, which is efficiently cleaved in the 156 protein but not in peptide substrates [19, 28]. The question remains if unprocessed nsp10-11 157 and/or nsp10-12 intermediates exist in virus-infected cells for prolonged times to fulfill specific functions. Other detected intermediates are nsp7-8-9 and nsp9-10 lacking nsp11-His6. 158 159 Particularly, the nsp9-10 intermediate has not been identified in our analysis of SARS-CoV 160 nsp7-10 processing, suggesting differences in the in vitro processing order between SARS-161 CoV and MERS-CoV.

MERS-CoV nsp7+8 forms heterodimers (1:1), -trimers (2:1) and -tetramers (2:2) (73 %, 8 % and 19 %, respectively), thus demonstrating a group AB complexation pattern. However, we cannot confirm the heterotrimer (2:1) formation by CID-MS/MS due to spectral complexity. Moreover, because of incomplete cleavage as is evident from the cleavage intermediates, signals assigned to the nsp7+8 heterodimer likely overlap with signals of unprocessed nsp7-8. Thus, complete cleavage of nsp7-8 could shift the peak fractions from heterodimer to heterotrimer or -tetramer.

#### 169 Homodimerization of subunits and precursors.

170 In the mass spectra of nsp7+8 complexes, monomers and homodimers of nsp7 and nsp8 are 171 also observed. While nsp7 homodimers are identified for all seven CoV species tested, nsp8 172 homodimers are only detected for SARS-CoV and SARS-CoV-2, which belong to group B 173 forming exclusively nsp7+8 heterotetramers around a dimeric nsp8 scaffold (Figure S 3). 174 Moreover, the oligomeric states of the different uncleaved nsp7-8 precursors are probed. 175 Notably, precursors from group B CoVs are mostly monomeric, whereas precursors from 176 group AB and A CoVs are in varying equilibria between monomers and dimers (Figure S 4). 177 The different oligomerization propensities of precursors suggest that molecular interactions 178 driving dimerization of nsp7-8 precursors could critically affect subsequent nsp7+8 179 oligomerization. For SARS-CoV and SARS-CoV-2, the nsp7-8 dimer affinity is low as two-180 fold dilution to 9  $\mu$ M shifted the equilibrium towards a monomeric state (Figure S 5). This is 181 in line with our previous findings [19], in which C-terminally extended SARS-CoV nsp7-9-182 His6 and His6-nsp7-10 polyprotein constructs were mainly monomeric, suggesting that the 183 presence of the extra C-terminal sequence further destabilizes an already weak dimerization.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.30.320762. this version posted October 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.



## 184 Collision-induced dissociation reveals complex topology



186 Figure 2: Gas-phase dissociation reveals complex topology. CID-MS/MS product ion spectra (A and C), 187 dissociation pathways and topology maps (B and D) for HCoV-229E nsp7+8 (2:1) heterotrimers and (2:2) 188 heterotetramers are shown. With increasing collisional voltage protein complexes are successively stripped from 189 their subunits revealing alternative dissociation pathways. The remaining dimeric species expose direct subunit 190 interactions in the nsp7+8 complexes (grey boxes). Charge states are labeled. (A and B) The heterotetramers (2:2) 191 undergo two consecutive losses resulting in dimeric product ions of nsp7+8 (1:1) and nsp82. These products 192 indicate that nsp7:nsp8 and nsp8:nsp8 have direct interfaces in heterotetramers. (C and D) HCoV-229E 193 heterotrimers dissociate into the dimeric products nsp7+8 (1:1) and nsp72 indicating direct interfaces between 194 nsp7:nsp8 and nsp7:nsp7 in heterotrimers. All CoV heterotrimers follow similar dissociation pathways, also all 195 CoV heterotetramers follow a common dissociation route, which allow a topological reconstruction of two 196 distinct complex architectures (Figure S 6-Figure S 10).

197 To deduce the complex topology in the different groups of nsp7+8 interaction patterns, we

198 applied CID-MS/MS using successive subunit dissociations to dissect conserved interactions.

199 CID-MS/MS of the HCoV-229E nsp7+8 heterotetramer (2:2) reveals two dissociation

200 pathways, in which first one nsp7 subunit is ejected from the complex followed by another

201 nsp7 or an nsp8 subunit. After two consecutive losses, the product ions are nsp7+8 (1:1) and

nsp8<sup>2</sup> dimers, providing evidence for specific subunit interfaces in the complex (Figure 2 A).
From these results, the complex topology is deduced as a heterotetramer based on an nsp8<sup>2</sup>
dimer scaffold, in which each nsp8 binds only one nsp7 subunit. Strikingly, this is identical
to our previously reported SARS-CoV nsp7+8 heterotetramer (2:2) architecture [19]. In fact,
all nsp7+8 (2:2) heterotetramers of groups A and AB (SARS-CoV-2, SARS-CoV, PEDV,
HCoV-229E and MERS-CoV) resulted in similar dissociation pathways, subunit interfaces
and topology maps, suggesting that these structures are similar across these diverse CoVs.

209 Next, the dissociation pathway of the HCoV-229E nsp7+8 (2:1) heterotrimers is monitored in 210 CID-MS/MS (Figure 2 B). After ejection of one nsp7 or nsp8 subunit, product dimers of 211 nsp7+8 (1:1) and nsp7<sub>2</sub> are detected, indicating specific subunit interfaces between nsp7:nsp8 212 and nsp7:nsp7. Again, similar dissociation pathways and subunit interfaces are found for 213 group B and AB heterotrimers (FIPV, TGEV, HCoV-229E and PEDV). Topological 214 reconstructions reveal a heterotrimer forming a tripartite interaction between one nsp8 and 215 two nsp7 subunits. These results agree with the reported X-ray structure of FIPV nsp7+8 [21] 216 and indicate that heterotrimers of these CoVs species have similar arrangements. In turn, this 217 implies that heterotrimers and heterotetramers follow distinct assembly paths.

#### 218 Chemical cross-linking confirms the formation of specific complexes.

219 To further support the native MS results, which relies on spraying from volatile salt solutions (e.g. ammonium acetate, AmAc), complementary methods compatible with conventional 220 221 buffers supplemented with sodium chloride are applied. To provide additional evidence for 222 specific nsp7+8 complex formation, the FIPV and HCoV-229E nsp7+8 complexes are stabilized via crosslinking with glutaraldehyde and subjected to XL-MALDI MS (Cross-223 linking Matrix-assisted laser ionization MS) (Figure S 11). Peak areas in the MALDI mass 224 225 spectra are assigned to FIPV nsp7+8 heterodimer, heterotrimer and -tetramer (8.5 %, 9.5 % and 4.2 %, respectively), and HCoV-229E nsp7+8 heterodimer, heterotrimer and -tetramer 226 227 (6.7%, 5.3% and 6.0%, respectively). The results suggest a higher abundance of nsp7+8 228 heterodimer and -trimer complexes in FIPV than in HCoV-229E, while HCoV-229E contains more heterotetramers. This largely agrees with the results from native MS. However, the 229 230 MALDI mass spectra show high background of virtually all possible nsp7+8 stoichiometries 231 (<200,000 m/z), probably due to over-crosslinking with the rather unspecific glutaraldehyde.

To refine these results, nsp7+8 complexes are stabilized with the amine specific crosslinker BS<sup>3</sup> and analyzed by SDS-PAGE (Figure S 12). Multiple stoichiometries are identified with a few prominent bands highlighting the main complexes generated. These bands are assigned to SARS-CoV nsp7+8 heterodimers and -tetramers, FIPV nsp7+8 heterodimers and -trimers and HCoV-22E nsp7+8 heterodimers, heterotrimers and -tetramers of HCoV-229E, providing additional support for the classification of nsp7+8 complexes into groups A (SARS-CoV), B (FIPV) and AB (HCoV-229E).

# Light scattering provides insights into complexation at high protein concentrations





242 Figure 3: DLS and SAXS reveal oligomeric state of SARS-CoV-2 nsp7+8 at higher protein concentrations. At 243 increasing protein concentrations, the hydrodynamic radius ( $R_0$ ) remains stable but becomes more homogenous 244 in DLS. (A) shows two exemplary DLS plots (1 mg/ml; 15 mg/ml) and (B) how the radius Ro develops with 245 increasing protein concentration. Theoretical hydrodynamic radii (R<sub>0,theo</sub>) of heterotetramer and hexadecamer 246 candidate structures are indicated (red dashed lines). (C) displays SAXS curves collected at different solute 247 concentrations and (D) the fit of the curve computed from T1 tetramer (red line) to the SAXS data collected at 248 1.2 mg/mL (blue dots with error bars). (E) Radius of gyration ( $R_g$ ) and (F) molecular weight estimated from the 249 SAXS data both stabilize with increasing concentration like in DLS on values that are in agreement with the R0,theo 250 of the T1 nsp7+8 (2:2) heterotetramer.

To test the stoichiometry at higher protein concentrations in solution, dynamic light scattering (DLS) of SARS-CoV-2 nsp7+8 from 1 to 15 mg/mL is performed (Figure 4 A). No significant increase of the hydrodynamic radius (*R*<sub>0</sub>) occurrs with increasing concentration. At the same time, the measured radii become more stable and fluctuate less, which suggests a shift towards higher complex homogeneity and a reduced fraction of free nsp7 and nsp8.

256 For SARS-CoV-2, no complex structure is available for full-length nsp7+8 proteins but, 257 previously, a SARS-CoV nsp7+8 (8:8) has been reported using X-ray crystallography [20], 258 where high protein concentrations are deployed. In order to relate the average experimental 259 hydrodynamic radius ( $R_{0,exp} = 4.25 \pm 0.61$  nm) to candidate structures, the theoretical hydrodynamic radius is calculated for the SARS-CoV nsp7+8 (8:8) hexadecamer 260 261  $(R_{0,\text{theo}} = 5.80 \pm 0.29 \text{ nm})$  and a subcomplex thereof, a putative nsp7+8 heterotetramer (2:2) in 262 T1 conformation ( $R_{0,\text{theo}} = 4.52 \pm 0.27 \text{ nm}$ ) (Figure S 13). This is the only model with full-length 263 nsp8 that agrees with the stoichiometry and topology determined by native MS. At 264 physiologically relevant concentrations from 1 to 10 mg/mL, the average experimental 265 hydrodynamic radius agrees well with the theoretical hydrodynamic radius of the 266 heterotetramer T1. Hence a heterotetramer is likely the prevailing species in solution (Figure 267 4B).

268 To underpin the DLS results, SAXS data are collected on solutions of nsp7+8 at 269 concentrations ranging from 1.2 to 47.7 mg/ml. The normalized SAXS intensities increase at 270 low angles with increasing concentration (Figure 4 B and Table S 3), suggesting a change in 271 the oligomeric equilibrium and a formation of larger oligomers. This trend is well illustrated 272 by the evolution of the apparent radius of gyration and molecular weight of the solute 273 determined from the SAXS data (Figure 4 E and F). The increase in the effective molecular 274 weight, from about 50 to 80 kDa suggests that the change in oligomeric state is limited and 275 that the tetrameric state (MW<sub>theo</sub>: 62 kDa) remains predominant in solution.

The SAXS data at low concentrations (<4 mg/mL) fit well the computed scattering from heterotetramer T1 but misfits appear at higher concentration (Figure 4 D, structure of T1 shown in Figure S 13and the discrepancy  $\chi^2$  reported in Table S 4). Mixtures of heterotetramers and hexadecamers cannot successfully fitted to the higher concentration data either (Figure S 14). To further explore the oligomeric states of nsp7+8, a dimer of T1 is used to simultaneously fit the curves collected at different concentrations by a mixture of heterotetramers and -octamers. Reasonable fits to all SAXS data are obtained with volume fractions of heterooctamers growing from 0 to 0.52 with increasing concentration (Figure S 14). Based on the flexibility of the molecule and the multiple possible binding sites between nsp7 and nsp8, it is not surprising that larger assemblies are observed at very high solute concentrations. The SAXS and DLS results provide evidence that the nsp7+8 (2:2) heterotetramer is the prevailing stoichiometry in solution at physiological concentrations (1-10 mg/mL with volume fractions between 1 and 0.7).

# Potential implications of sequence conservation on heterotrimer and -tetramer formation.

291 To extend this analysis, we select candidate structures in agreement with the stoichiometry 292 and topology observed (Figure S 10). For the heterotetramer, two conformers of nsp7+8 (2:2) 293 subcomplexes, T1 and T2, of correct architecture can be extracted from the larger SARS-CoV 294 nsp7+8 hexadecamer [20] (pdb 2ahm), (Figure S 13 A). Both conformers constitute a head to tail interaction of two nsp7+8 heterodimers mediated by an nsp8-nsp8 interface. Notably, 295 296 nsp8 in T1 is more extended, revealing an almost full-length amino acid sequence (2-193), 297 while in T2 the nsp8 N-terminal 35 to 55 residues are unresolved. For the heterotrimeric 298 complexes, the only deposited structure is the FIPV nsp7+8 (2:1) heterotrimer [21] (pdb 299 3ub0), which agrees well with our experimental topology (Figure S 13 B).

300 In order to identify molecular determinants for heterotrimer or -tetramer formation, the candidate structures are examined for molecular contacts (van der Waals radius -0.4 Å). The 301 302 conservation of contact residues is evaluated in a sequence alignment to identify possible 303 determinants of different stoichiometries (Figure S 15). Notably, most amino acids lining 304 subunit interfaces in heterodimers, -trimers and -tetramers are conserved. The interfaces in 305 the candidate structures occupy two common structural portions of the nsp8 subunit (Figure 306 4 A). The first binding site (BS I) is located between the nsp8 head and shaft domain, 307 responsible for binding of nsp7 (I) in heterodimer formation, as seen in all available high 308 resolution structures of nsp7+8 [20-23] and the polymerase complex [13-16]. The second 309 binding site (BS II) appears highly variable in terms of its binding partner and lies at the nsp8 310 elongated N-terminus. In fact, one largely conserved motif (res60-70) is responsible for the 311 main contacts in the entire candidate complexes selected based on our data: nsp7+8 (2:2) T1 312 and T2 for the heterotetramer and nsp7+8 (2:1) for the heterotrimer. The respective 313 sidechains take positions on one side of the nsp8  $\alpha$ -helix and have the ability to form interactions with either mainly nsp7 (partly nsp8) in the SARS-CoV nsp7+8 (2:2) heterotetramer T1, mainly nsp8 (partly nsp7) in the SARS-CoV nsp7+8 heterotetramer T2 or only nsp7 in the FIPV nsp7+8 (2:1) heterotrimer (Figure S 16 A-C). Due to its sequence conservation, it is unlikely that alone this motif at BS II has a decisive impact for heterotrimer or -tetramer formation.



Figure 4: Candidate structures and sequence conservation of nsp7+8 heterotrimers and -tetramers. Candidate
 structures for nsp7+8 heterotetramer and -trimer are chosen based on experimental stoichiometry and topology

322 in solution. (A) shows two conformers of SARS-CoV nsp7+8 (2:2) heterotetrameric subcomplexes, T1 (left) and T2 323 (middle), from the larger (8:8) heterohexadecamer (pdb 2ahm) and the FIPV nsp7+8 (2:1) heterotrimer (right, pdb 324 3ub0). Complexes exhibit two similar binding sites in nsp8, BSI (orange) and BSII (red). For simplification 325 binding sites are only shown for one nsp8 subunit (green). BS II is additionally labelled with nsp8 residue (res) 326 number forming the main interaction patch (BS II contact residues see Figure S 15). (B) Sequence alignment of 327 nsp8 (green, res55-92) and nsp7 (yellow, res74-78) is displayed for the seven CoVs. Two heterotrimer or -tetramer 328 specific contact sites (red) exhibit sequence conservation well in line with the complexation groups determined by 329 native MS. (C) In SARS-CoV T2, nsp8 Glu77 comes into contact with nsp8 Glu77, a unique heterotetramer 330 interaction, (D) which in a homology model of HCoV-229E based on T2 is replaced by nsp8 Asn78. (E) The FIPV 331 heterotrimer structure shows nsp7 Phe76, binding to nsp8 Met87 and Leu90, a contact unique for the heterotrimer 332 forming species. Insets show magnification of contact sites. Closest distances (3.0-3.8 Å) for relevant residues in 333 contact (red) are given.

334 Therefore, unique interactions could exist, which explain the shift in complex stoichiometry 335 from heterotrimer to -tetramer observed in the different CoVs categorized into group A, AB 336 and B (Figure 4 B). Here, we identify a possibly heterotetramer stabilizing contact site in T2, 337 where nsp8 Glu77 self-interacts with nsp8II Glu77, which gives the complex density and compactness (Figure 4 C). This residue is only present in nsp8 of SARS-CoV and SARS-CoV-338 339 2 from group A and MERS-CoV of group AB. However, homology models suggest that in 340 the other tetramer forming complexes of group AB, HCoV-229E and PEDV, nsp8 Asn78 341 could partially replace this interaction (Figure 4 D). This is different in group B viruses, 342 forming only heterotrimers, where residues at these positions are nsp8 Val77 and Asp78, 343 with the Asp78 possibly being solvent exposed and hence unable to replace this interaction. 344 Furthermore, we also identify a contact site possibly stabilizing the heterotrimer in the 345 crystal structure of the FIPV nsp7+8 (2:1), which reveals that a second subunit of nsp7 346 (nsp7II) is locked via Phe76 to nsp8 (Figure 4 E). Importantly, this residue is uniquely 347 conserved among trimer forming complexes of group B and group AB, but replaced by nsp7 348 Leu76 in the strictly heterotetramer forming group A.

349 These findings are compared to the recently released structure of the polymerase complex 350 (pdb 6xez, Figure S 16 D and E), comprising nsp7+8+12+13(1:2:1:2) [29]. The residues 351 potentially responsible for a shift in quaternary structure, nsp8 Glu77 or Asp78 and nsp7 352 Phe76, are distant from any protein-protein or protein-RNA interaction and thus are not 353 expected to play a role in polymerase complex formation. Surprisingly, the identical set of 354 residues in BS II supports all interactions (Glu60, Met62, Ala63, Met67 and Met70) between 355 nsp8b and nsp12/nsp13.1 and between nsp8a and nsp13b (Figure S 16 D and E). Notably, within the polymerase complex amino acids involved in RNA binding point in the opposite 356 357 direction of the protein interfaces and have little or no role in nsp7+8 complex formation.

#### 359 Discussion

360 Our findings reveal the nsp7+8 quaternary composition of seven CoVs representing five coronavirus species of the genera Alpha- and Betacoronavirus. Viruses of the same species 361 362 (SARS-CoV/SARS-CoV-2 and TGEV/FIPV, respectively) produce the same type of nsp7+8 363 complexes. Next to a conserved nsp7+8 heterodimer (1:1), the inherent specificity of nsp7+8 364 complex formation categorizes them into three groups: group A forming only 365 heterotetramers (2:2), group B forming only heterotrimers (2:1) and group AB forming both 366 heterotetramers (2:2) and heterotrimers (2:1). Complexes of the same stoichiometry exhibit a 367 conserved topology, consisting of an nsp8 homodimeric scaffold for the heterotetramers and 368 an nsp7 homodimeric core for the trimers. Candidate structures based on our results 369 highlight Alpha- and Betacoronavirus-wide conserved binding sites on nsp8, named BS I and 370 BS II, which provide the modular framework for a variety of complexes. Furthermore, 371 unique molecular contacts for the complex groups have the potential to determine the ability 372 and preference for heterotrimer and/or heterotetramer formation.

373 We provide evidence that, even at high concentrations, the SARS-CoV-2 nsp7+8 374 heterotetramer (2:2) represents the predominant species. In order to relate our results to in 375 vivo conditions, we consider the following aspects: According to maximum molecular 376 crowding [30], polyproteins pp1a and pp1ab can reach a maximum of 125-450 µM, which 377 translates to 3.9-11.7 mg/mL nsp7+8. This range is covered by our DLS and SAXS analysis. In 378 absence of other interaction partners, we expect that, in vivo, the nsp7+8 (2:2) heterotetramer 379 represents the predominant nsp7+8 complex of SARS-CoV-2 and other heterotetramer 380 forming CoVs of complexation groups A and AB.

The heterotetramer candidate structures and models presented here are based on the conformers T1 or T2 of the SARS-CoV heterohexadecamer structure, which contains fulllength nsp8 [20]. Although, our results cannot clarify, if one of these conformers is the biologically relevant structure existing in solution, the combined evidence provided here strongly suggests structural similarity to T1/T2. Considering the crystallographic origin of T1/T2 and the overlap of binding sites, the heterotetramer could well be a flexible and dynamic structure in solution.

In contrast to our findings, a SARS-CoV nsp7+8 hexadecamer structure has been reported
[20]. However, this structure has been derived from X-ray crystallography, hence showing a

390 static, frozen state, where the crystal lattice formation favors stabilized arrangements that 391 could differ from the solution state of the protein complexes. In the case of nsp8, the flexible 392 *N*-terminus could inhibit crystal formation, and has been removed in some studies [22, 23]. 393 Alternatively, it may stabilize specific interactions, thereby promoting crystal formation by 394 binding to one of the multiple interfaces presented between nsp7 and nsp8, resulting in a 395 physiologically irrelevant larger oligomeric structure. The SAXS data presented here 396 partially supports this scenario at high protein concentrations but also confirms a 397 predominantly heterotetrameric assembly in solution. Thus, a potential shift of quaternary 398 structure from heterotetramer towards a higher-order complex, such as a heterohexadecamer 399 appears unlikely unless triggered e.g. by binding to nucleic acids as has been repeatedly 400 described for nsp7+8 complexes [20].

All seven CoV nsp7 and nsp8 proteins shown here also form heterodimers (1:1). Such 401 402 heterodimeric subcomplexes with nsp7 bound to nsp8 BSI have been observed in all 403 deposited complex structures containing nsp7+8 [20-23] or nsp12 [13-16]. Therefore, the 404 heterodimer represents the most basic form of nsp7+8 complexes and likely serves as 405 universal substructure building block in the coordinated assembly of functional RTCs of 406 CoVs from the genera Alpha- and Betacoronavirus. Moreover, heterotrimer and -tetramer 407 formation are based on a second canonical binding site at the nsp8 N-terminal domain, BS II. 408 This site appears to have a high propensity to form complexes with various binding partners 409 (e.g. nsp7+8, nsp12 or nsp13). Accordingly, our analysis suggests that the nsp8 BS II strives for occupation. The nsp7+8 quaternary composition, topology and analysis of binding sites 410 411 presented here allow us to reconstruct and propose a model of the complex formation 412 pathway (Figure 5).

413 The preference for heterotrimer and -tetramer can probably be pinpointed to just a few 414 amino acids within nsp8 BS II or interacting with it. Here, we identify two contacts that could have unique discriminatory potential for promoting heterotrimeric (nsp7 Phe76) or 415 416 heterotetrameric (nsp8 Glu77 and Asn78) quaternary structures. Notably in presence of nsp7 417 Phe76 and nsp8 Asn78, as observed for group AB, the heterotetramer is always more 418 abundant than the heterotrimer. However, compared to the entire BS II, theses contacts only 419 represent a small share of the binding interface and contribute little interaction energy 420 through van der Waals forces. Nevertheless, the unique position of their contacts could 421 critically determine the types of interactions with one or another binding partner.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.30.320762. this version posted October 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.



423 Figure 5: Proposed model for nsp7+8 complex formation. (A) For complexes of group A, heterodimers form via 424 nsp8 BS I, which quickly dimerize via BS II into a heterotetramer. A theoretic route via a preformed nsp8 scaffold 425 is unlikely to play a role in heterotetramer formation since no nsp7+8 (1:2) intermediates are observed for 426 complexes of SARS-CoV or SARS-CoV-2. Moreover,, nsp7 and nsp8 occupy neighboring positions in the replicase 427 polyproteins, thus favoring their interaction (in cis) at early stages in the infection cycle (when intracellular viral 428 polyprotein concentrations are low) over intermolecular interactions between different replicase polyprotein 429 molecules as is also evident from the low dimerization ability of the precursors. (B) For group B complexes, we 430 propose the formation of a heterodimer intermediate via nsp8 BS I or BS II and subsequent recruitment of a 431 second nsp7, resulting an nsp7+8 (2:1) heterotrimer. This is also supported by the relatively high peak fractions of 432 heterodimers detected. Group AB complexes can use both complexation pathways. In line with this, the proteins 433 also produce a relatively high heterodimer signal but, ultimately, prefer to form heterotetramers rather than 434 heterotrimers.

435 Since the critical residues required for nsp7+8 complex formation have no overlap with 436 nsp12 interaction sites, direct docking of preformed heterotrimers and -tetramers to nsp12 437 can be expected. Furthermore, heterotrimeric and -tetrameric structures are compatible with 438 accommodation of specific RNA structures similar to what has been suggested for 439 heterohexadecameric nsp7+8 by Rao et al. [14]. Notably, if heterotrimeric or -tetrameric nsp7+8 structures were associated with nsp12, the binding site for nsp13 would be blocked, 440 441 which may have regulatory implications for CoV replication. Altogether, these conserved 442 binding mechanisms and overlapping binding sites confirm the proposed role of nsp8 as a 443 major interaction hub within the CoV RTC [31], and indicate critical regulatory functions by specific nsp7+8 complexes. 444

Finally, we can only speculate about possible reasons for the existence of different nsp7+8 complexes: (1) Similar kinetic stability due to occupation of both binding sites (both structures exist because they are equally efficient in occupying BS I and BS II), (2) unknown functional relevance in CoV replication (e.g. specificity to RNA structures channeled to the nsp12 RdRp) or (3) adaption to host factors and possible regulatory functions.

In summary, our work shows, and provides a framework to understand, the characteristic distribution and structures of nsp7+8 (1:1) heterodimer, (2:1) heterotrimers and (2:2) heterotetramers in representative alpha- and betacoronaviruses. The nsp7+8 structure in solution can be used to investigate its independent functional role in the formation of active polymerase complexes and, possibly, regulation and coordination of polymerase and other RTC activities, for example in the context of antiviral drug development targeting different subunits of CoV polymerase complexes reconstituted *in vitro*.

#### 457 Material and Methods

#### 458 **Cloning and gene constructs**

459 The codon optimized sequence for the SARS-CoV-2 nsp7-8 region (NC 045512.2) was 460 synthesized by Eurofins scientific SE with overhangs suitable for insertion into pASK-IBA33plus plasmid DNA (IBA Life Sciences). A golden gate assembly approach using Eco31I 461 462 (BsaI) (Thermo Fisher Scientific) was used to shuttle the gene into the plasmid. Linker and 463 tag of the expression construct SARS-CoV-2 nsp7-8-His6 contained the C-terminal amino acids -SGSGSARGS-His6 (SGSG residues as P1'-P4' of Mpro cleavage site and SARGS-His6 464465 residues as the default linker of pASK vectors). The SARS-CoV nsp7-8 pASK33+ plasmid generated previously [21, 32] was used for the expression of SARS-CoV nsp7-8-His6 466 containing the C-terminal amino acids -SARGS-His6. The expression plasmid for SARS-CoV 467 Mpro was generated as described by Xue et al.[33]. To produce nsp7-8-SGSGSARGS-His6 468469 precursor proteins in Escherichia coli, the nsp7-8 coding sequences of HCoV-229E (HCoV-229E; GenBank accession number AF304460), FIPV (FIPV, strain 79/1146; DQ010921), SARS-470 471 CoV (strain Frankfurt-1; AY291315), PEDV (PEDV, strain CV777, NC\_003436) and TGEV 472 (TGEV, strain Purdue; NC\_038861) were amplified by reverse transcription-polymerase 473 chain reaction (RT-PCR) from viral RNA isolated from cells infected with the respective 474 viruses and inserted into pASK3-Ub-CHis6 using restriction- and ligation-free cloning 475 methods as described before (Tvarogová et al., 2019). Similarly, the nsp7-9 or nsp7-11 coding 476 region of MERS-CoV (strain HCoV-EMC; NC\_019843), was amplified by RT-PCR from 477 infected cells and inserted into pASK3-Ub-CHis6. The HCoV-229E and FIPV nsp5 coding 478 sequences were cloned into pMAL-c2 plasmid DNA (New England Biolabs) for expression 479 as MBP fusion proteins containing a *C*-terminal His6-tag. Primers used for cloning and 480 mutagenesis are available upon request.

#### 481 **Expression and purification**

482 SARS-CoV M<sup>pro</sup> was produced with authentic ends as described in earlier work [33]. To 483 produce the precursors, SARS-CoV and SARS-CoV-2 nsp7-8-His<sub>6</sub>, BL21 Rosetta2 (Merck 484 Millipore) were transformed, grown in culture flasks to  $OD_{600} = 0.4-0.6$ , then induced with 50 µM anhydrotetracycline and continued to grow at 20 °C for 16 h. For pelleting, cultures 485 486 were centrifuged ( $6000 \times g$  for 20 min) and cells were frozen at -20 °C. Cell pellets were lysed 487 in 1:5 (v/v) buffer B1 (40 mM phosphate buffer, 300 mM NaCl, pH 8.0) with one freeze-thaw cycle, sonicated (micro tip, 70 % power, 6 times on 10 s, off 60 s; Branson digital sonifier SFX 488 489 150) and then centrifuged (20,000  $\times$  g for 45 min). Proteins were isolated with Ni<sup>2+</sup>-NTA beads 490 (Thermo Fisher Scientific) in gravity flow columns (BioRad). Proteins were bound to beads 491 equilibrated with 20 column volumes (CV) B1 + 20 mM imidazole, then washed with 20 CV 492 B1 + 20 mM imidazole followed by 10 CV of B1 + 50 mM imidazole. The proteins were eluted 493 in eight fractions of 0.5 CV B1 + 300 mM imidazole. Immediately after elution, fractions were 494 supplemented with 4 mM DTT. Before analysis with native MS, Ni<sup>2+</sup>-NTA eluted fractions 495 containing the polyprotein were concentrated to 10 mg/mL and further purified over a 496 10/300 Superdex 200 column (GE healthcare) in 20 mM phosphate buffer, 150 mM NaCl, 497 4 mM DTT, pH 8.0. The main elution peaks contained nsp7-8. For quality analysis, SDS-498 PAGE was performed to assess the sample purity.

499 To obtain a pre-purified SARS-CoV-2 nsp7+8 complex for DLS and SAXS, eluate fractions 500 from the Ni2+-NTA column containing the nsp7-8-His6 were concentrated and the buffer 501 exchanged with a PD-10 desalting column (GE Healthcare) equilibrated with 50 mM Tris, 502 pH 8.0, 100 mM NaCl, 4 mM DTT and 4 mM MgCl<sub>2</sub> (SEC-buffer). Then nsp7-8-His<sub>6</sub> was 503 eluted with 3.5 mL SEC-buffer and subsequently cleaved with MPro-His6 (1:5, MPro-His6: nsp7-504 8-His6) for 16 h at RT. Mpro-His6 was removed with Ni-NTA agarose and the cleaved nsp7+8 505 complex was subjected to a HiLoad 16/600 Superdex 75 pg size exclusion column 506 equilibrated with SEC-buffer.

507 The HCoV-229E, PEDV, FIPV and TGEV nsp7-8-His<sup>6</sup> and MERS-CoV nsp7-11-His<sup>6</sup> precursor 508 proteins were produced and purified as described before (Tvarogová et al., 2019) with a 509 slightly modified storage buffer. Anion exchange chromatography fractions of the peak 510 containing the desired protein were identified by SDS-PAGE, pooled and dialyzed against 511 storage buffer (50 mM Tris-Cl, pH 8.0, 200 mM NaCl and 2 mM DTT).

512 MBP-nsp5-His6 fusion proteins were purified using Ni2+-IMAC as described before 513 (Tvarogová et al., 2019). To produce HCoV-229E and FIPV MBP-nsp5-His<sub>6</sub>, E. coli TB1 cells 514 were transformed with the appropriate pMAL-c2-MBP-nsp5-His6 construct and grown at 515 37 °C in LB medium containing ampicillin (100 µg/mL). When an OD<sub>600</sub> of 0.6 was reached, 516 protein production was induced with 0.3 mM isopropyl β-D-thiogalactopyranoside (IPTG) 517 and cells were grown for another 16 h at 18 °C. Thereafter, the cultures were centrifuged 518  $(6000 \times g \text{ for } 20 \text{ min})$  and the cell pellet was suspended in lysis buffer (20 mM Tris-Cl, pH 8.0, 519 300 mM NaCl, 5% glycerol, 0.05% Tween-20, 10 mM imidazole and 10 mM β-520 mercaptoethanol) and further incubated with lysozyme at 4 °C (0.1 mg/mL) for 30 min. 521 Subsequently, cells were lysed by sonication and cell debris was removed by centrifugation 522 for 30 min at 40,000  $\times$  g and 4°C. The cell-free extract was bound to pre-equilibrated Ni<sup>2+</sup>-NTA 523 (Qiagen) matrix for 3 h. Ni<sup>2+</sup>-IMAC elution fractions were dialyzed against buffer comprised 524 of 20 mM Tris-Cl, pH 7.4, 200 mM NaCl, 5 mM CaCl2 and 2 mM DTT and cleaved with 525 factor Xa to release nsp5-His6. Then, nsp5-His6 was passed through an amylose column and 526 subsequently bound to Ni2+-NTA matrix to remove any remaining MBP. Following elution 527 from the Ni<sup>2+</sup>-NTA column, nsp5-His6 was dialyzed against storage buffer (20 mM Tris-Cl, 528 pH 7.4, 200 mM NaCl and 2 mM DTT) and stored at -80 °C until further use.

#### 529 Native mass spectrometry

To prepare samples for native MS measurements, M<sup>pro</sup> was buffer exchanged into 300 mM AmAc, 1 mM DTT, pH 8.0 by two cycles of centrifugal gel filtration (Biospin mini columns, 6,000 MWCO, BioRad) and the precursors were transferred into 300 mM AmAc, 1 mM DTT, pH 8.0 by five rounds of dilution and concentration in centrifugal filter units (Amicon, 10,000 MWCO, Merck Millipore). Cleavage and complex formation was started by mixing nsp7-8-His<sub>6</sub> and protease M<sup>pro</sup> with final concentrations of 15 µM and 3 µM, respectively. Three independent reactions were started in parallel and incubated at 4 °C overnight. Tips for nano-electrospray ionization (nanoESI) were pulled in-house from borosilicate capillaries (1.2 mm outer diameter, 0.68 mm inner diameter, with filament, World Precision Instruments) with a micropipette puller (P-1000, Sutter Instruments) using a squared box filament ( $2.5 \times 2.5$  mm, Sutter Instruments) in a two-step program. Subsequently, tips were gold-coated using a sputter coater (Q150R, Quorum Technologies) with 40 mA, 200 s, tooling factor 2.3 and end bleed vacuum of  $8 \times 10^{-2}$  mbar argon.

543 Native MS was performed at a nanoESI quadrupole time-of-flight (Q-TOF) instrument (Q-544 TOF2, Micromass/Waters, MS Vision) modified for higher masses [34]. Samples were ionized 545 in positive ion mode with voltages applied at the capillary of 1300-1500 V and at the cone of 130-135 V. The pressure in the source region was kept at 10 mbar throughout all native MS 546 547 experiments. For desolvation and dissociation, the pressure in the collision cell was 548  $1.5 \times 10^{-2}$  mbar argon. For native MS, accelerating voltages were 10 - 30 V and quadrupole profile 1,000 - 10,000 m/z. For CID-MS/MS, acceleration voltages were 30 - 200 V. Raw data 549 550 were calibrated with CsI (25 mg/mL) and analyzed using MassLynx 4.1 (Waters). Peak 551 deconvolution and determination of relative intensity was performed using UniDec [35]. All 552 determined masses are provided (Table S 1).

#### 553 XL-MALDI

554 Pre-purified FIPV nsp7+8 and HCoV-229E nsp7+8 at 20 µM were cross-linked with 0.15 % 555 glutaraldehyde (Sigma-Aldrich) at 4 °C for 25 min before diluting them to 1 µM in MALDI 556 matrix solution (sinapinic acid 10 mg/mL in acetonitrile/water/TFA, 49.95/49.95/0.1, v/v/v) 557 and spotting (1 µL) onto a stainless steel MALDI target plate. The MALDI-TOF/TOF mass 558 spectrometer (ABI 4800, AB Sciex) equipped with a high-mass detector (HM2, CovalX) was 559 used in linear mode. For acquiring mass spectra (1,000 to 1,000,000 m/z) spots were ionized 560 with a Nd:YAG laser (355 nm) and 500 shots per spectrum were accumulated. Obtained raw 561 data were smoothed and analyzed using mMass (v5.5.0, by Martin Strohalm [36]).

562 **DLS** 

To check the monodispersity of the samples and to study the stoichiometry of the nsp7+8 complexes, DLS measurements were performed with the Spectro Light 600 (Xtal Concepts). The complex was concentrated to various concentrations and samples were spun down for 10 min at 12,000 rpm and 4 °C. A Douglas Vapour batch plate (Douglas instruments) was 567 filled with paraffin oil and 2  $\mu$ L of each sample was pipetted under oil. DLS measurements 568 for each sample were performed at 20 °C with 20 measurements for 20 s each, respectively.

#### 569 SAXS

SAXS data were collected on the P12 beamline of EMBL at the PETRA III storage ring (DESY, 570 Hamburg). X-ray wavelength of 1.24 Å (10 keV) was used for the measurements, scattered 571 572 photons were collected on a Pilatus 6M detector (Dectris), with a sample to detector distance 573 of 3 m. Data were collected on 22 concentrations ranging from 1.2 to 48 mg/mL nsp7+8 in 574 50 mM Tris, pH 8.0, 100 mM NaCl, 4 mM DTT and 4 mM MgCl<sub>2</sub>, pure buffer was measured 575 between samples. For each data collection, 20 frames of 100 ms were collected. 2D scattering images were radially averaged and normalized to the beam intensity. The frames were 576 577 compared using the program Cormap [37, 38], and only similar frames were averaged and used for further analysis to avoid possible beam-induced effects. Scattering collected on the 578 579 pure buffer was subtracted from that of the sample and the resulting curves were normalized 580 to the protein concentration to obtain the scattering of nsp7+8 complexes.

581 The data processing pipeline SASflow was used for data reduction and calculation of the 582 overall SAXS parameters [38]. Molecular weights were inferred from different molecular 583 calculation methods using a Bayesian assessment [39]. The program CRYSOL was used to 584 compute the theoretical curves from the atomic structures [40]. Volume fractions of the 585 components of the oligomeric mixtures were computed and fitted to the data using the 586 program OLIGOMER [41]. The dimer of T1 was built by the program SASREFMX [42], 587 which builds a dimeric model that fits best, in mixture with the monomeric T1, multiple 588 scattering curves collected at different concentrations.

#### 589 Visualization

590 Molecular graphics and analyses were performed with UCSF ChimeraX, developed by the 591 Resource for Biocomputing, Visualization, and Informatics at the University of California, 592 San Francisco, with support from National Institutes of Health R01-GM129325 and the Office 593 of Cyber Infrastructure and Computational Biology, National Institute of Allergy and 594 Infectious Diseases.

#### 595 Acknowledgements

596 We would like to thank the XBI User consortium for providing the instrumentation and597 laboratories at the European XFEL.

## 598 Author contributions

- 599 Conceptualization and methodology B.K., C.S., C.U. G.B., R.M. and J.Z.; plasmid
- 600 construction and protein production B.K., C.S., G.B. , L.B.; providing research materials
- 601 C.U., K.L. and J.Z.; XL-MALDI-MS B.K., M.K. and R.Z.; SAXS: C.B., D.S. Investigation -
- 602 B.K., G.B., C.S., R.S. Discussion of results B.K., C.S., C.U., K.L., R.M., G.B., J.Z.; Formal
- 603 analysis and visualization B.K.; Original Draft B.K. and C.U.; Writing, review and editing -
- 604 B.K. and C.U. with help from all authors.

#### 605 Funding

- 606 The Heinrich Pette Institute, Leibniz Institute for Experimental Virology is supported by the
- 607 Free and Hanseatic City of Hamburg and the Federal Ministry of Health. C.U. and B.K. are
- supported by EU Horizon 2020 ERC StG-2017 759661. C.U. and B.K. also received funding
- 609 through Leibniz Association SAW-2014-HPI-4 grant. B.K. acknowledges funding from COST
- 610 BM1403.The work of J.Z., G.B. and R.M. was supported by grants (to J.Z.) from the Deutsche
- 611 Forschungsgemeinschaft (SFB 1021, A01; IRTG 1384) and the LOEWE program of the state of
- 612 Hessen (DRUID, B02).

#### 614 Supplement:

Table S 1: Mass table. Data from at least three representative native mass spectra were analyzed with MassLynx
4.1 and UniDec. Mtheo is calculated based on amino acid sequence. Full width half-maximum (FWHM) values

617 derived from main peaks, charge state given in parenthesis. Three data points are used for calculation of average

618 (Avg) and standard deviation (SD) (only two for PEDV complexes).

	Oligomer	M <sub>theo</sub> (Da)	Avg M <sub>exp</sub> ± SD (Da)	Avg FWHM (Da) and charge state	nsp7+8 relative intensity fraction (%)
SARS-CoV-	nsp7+8 (1:1)	31121	31140 ± 10	62 (11+)	4 ± 2
2	nsp7+8 (2:1)	-	-	-	-
	nsp7+8 (2:2)	62242	62300 ± 20	155 (15+)	96 ± 2
SARS-CoV	nsp7+8 (1:1)	31129	31123 ± 2	20 (11+)	2.3 ± 0.2
	nsp7+8 (2:1)	-	-	-	-
	nsp7+8 (2:2)	62258	62267 ± 2	65 (15+)	97.6 ± 0.2
MERS-CoV	nsp7+8 (1:1)	30931	30954.0 ± 0.0	20 (10+)	73 ± 11
	nsp7+8 (2:1)	40012	40025 ± 2	22 (11+)	8 ± 1
	nsp7+8 (2:2)	61898	61920 ± 10	75 (14+)	19 ± 11
PEDV	nsp7+8 (1:1)	30895	30927 ± 1	20 (10+)	52 ± 4
	nsp7+8 (2:1)	40102	40238 ± 3	27 (11+)	6 ± 1
	nsp7+8 (2:2)	61723	61863 ± 1	45 (15+)	42 ± 2
HCoV-229E	nsp7+8 (1:1)	30901	30920 ± 10	20 (10+)	20 ± 7
	nsp7+8 (2:1)	40188	40145 ± 1	29 (11+)	12 ± 1
	nsp7+8 (2:2)	61793	61839 ± 1	45 (15+)	69 ± 6
TGEV	nsp7+8 (1:1)	30930	30962 ± 6	22 (11+)	17 ± 3
	nsp7+8 (2:1)	40401	40470 ± 20	68 (12+)	83 ± 3
	nsp7+8 (2:2)	-	-	-	-
FIPV	nsp7+8 (1:1)	308334	30850 ± 5	26 (11+)	17 ± 2
	nsp7+8 (2:1)	40234	40400 ± 100	208 (12+)	83 ± 2
	nsp7+8 (2:2)	-	-	-	-

619

620 **Table S 2: Sequence identity matrix of nsp7-8 species.** Values for pairwise sequence identity are given in 621 percent. For the multiple sequence alignment with identity matrix output the SIAS Sequence identity and

622 similarity tool has been used provided by Secretaria general de sciencia, technologica e innovacion of Spain

623 (<u>http://imed.med.ucm.es/Tools/sias.html</u>). As input parameter, length of the smallest sequence was selected.

SARS- CoV-2	100						
SARS- CoV	98	100					
HCoV- 229E	43	44	100				
PEDV	42	42	71	100			
FIPV	42	43	63	62	100		
TGEV	42	42	64	61	94	100	
MERS	28	28	21	21	20	20	100
%	SARS- CoV-2	SARS- CoV	HCoV- 229E	PEDV	FIPV	TGEV	MERS

#### 626 Table S 3: SAXS experimental parameters and analysis methods.

EMBL P12 (PETRAIII, DESY, Hamburg)
0.2 × 0.12
0.12
0.03 – 5
2 (20 × 0.1)
293
SASFLOW
PRIMUS
CRYOL
OLIGOMER
SASREFMX

## 627 Table S 4: Overall parameters computed from the SAXS curves.

<i>c</i> (mg/mL)	<i>R</i> g (nm)	MW Bayes (kDa)	Fit to tetramer T1	Mixture of tetramer and dimer of tetramer		Mixture of tetramer and hexadecamer		
			X <sup>2</sup>	X <sup>2</sup>	Tetramer volume fraction	X <sup>2</sup>	Tetramer volume fraction	
1.2	3.25	46	1.35	1.34	1	1.34	1	
1.2	3.41	58	1.27	1.27	1	1.27	1	
2.1	3.37	56	1.31	1.41	0.97	1.47	1	
3.9	3.46	60	2.57	2.54	0.9	3.7	0.98	
4.4	3.41	62	1.97	2.2	0.9	3.34	0.97	
7.7	3.56	62	5.13	2.3	0.8	4.84	0.95	
9.1	3.67	67	14.32	2.45	0.69	9.54	0.91	
10.4	3.79	62	22.95	4.4	0.73	15.02	0.93	
10.6	3.79	74	21.79	2.86	0.66	12.75	0.91	
11.5	3.66	62	4.94	2.14	0.79	4.77	0.94	
12.8	3.73	74	62.98	5.13	0.63	30.59	0.9	
13.1	3.53	62	19.63	5.29	0.77	15.77	0.94	
16.6	3.82	74	17.89	2.27	0.64	9.96	0.9	
20	3.67	68	40.19	5.59	0.68	20.25	0.91	
21.1	3.82	74	67.56	4.62	0.61	29.52	0.89	
25.8	3.8	74	73	4.69	0.57	27.08	0.88	
26.6	3.86	74	58.68	6.49	0.62	24.02	0.89	
31.1	3.91	83	148.76	7.82	0.52	45.67	0.86	
32.1	3.82	76	111.04	10.92	0.61	41.36	0.89	
47.7	3.95	86	86.73	6.57	0.47	21.69	0.84	

## 629 Table S 5: Results table, overview of main experimental results for nsp7+8 complexes of seven CoV species.

630 Results from left to right: Categorization of complexes, quantitation of complexes by native MS (values are given 631 in percent), identified homodimers nsp72 and nsp82, and dimerization of precursor (clearly assigned species 632 indicated with "yes" and if no assignment possible with minus(-)), nsp8 and nsp7 possible mutations that shift 633 heterotrimer to heterotetramer formation.

	Group	Complex nsp7+8 (1:1)	Complex nsp7+8 (2:1)	Complex nsp7+8 (2:2)	Dimer nsp7 (2)	Dimer nsp8 (2)	Dimer nsp7-8	nsp8 Glu77/ Asn78	nsp7 Phe76
SARS- CoV-2	А	4	-	96	yes	yes	yes	Glu77	-
SARS-CoV	А	2	-	98	yes	yes	yes	Glu77	-
MERS- CoV	AB	73	8	19	yes	-	*different precursor	Glu77	Phe76
HCoV- 229E	AB	20	12	69	yes	-	yes	Asn78	Phe76
PEDV	AB	52	6	42	yes	-	yes	Asn78	Phe76
FIPV	В	17	83	-	yes	-	-	-	Phe76
TGEV	В	17	83	-	yes	-	-	-	Phe76

634

#### 636 Table S 6: Amino acid sequences of precursor constructs.

SKMSDV ASEFSSL AMQTML	>SARS-CoV-2_nsp7-8-His <sub>6</sub> KCTSVVLLSVLQQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMVSLLSVLLSMQGAVDINKLCEEMLDNRATLQAI .PSYAAFATAQEAYEQAVANGDSEVVLKKLKKSLNVAKSEFDRDAAMQRKLEKMADQAMTQMYKQARSEDKRAKVTS FTMLRKLDNDALNNIINNARDGCVPLNIIPLTTAAKLMVVIPDYNTYKNTCDGTTFTYASALWEIQQVVDADSKIVQLSEIS MDNSPNLAWPLIVTALRANSAVKLQ <b>SGSGSARGSHHHHHH</b>
SKMSDV ASEFSSL AMQTML	SARS-CoV_nsp7-8-His <sub>6</sub> KCTSVVLLSVLQQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMVSLLSVLLSMQGAVDINRLCEEMLDNRATLQAI PSYAAYATAQEAYEQAVANGDSEVVLKKLKKSLNVAKSEFDRDAAMQRKLEKMADQAMTQMYKQARSEDKRAKVTS FTMLRKLDNDALNNIINNARDGCVPLNIIPLTTAAKLMVVVPDYGTYKNTCDGNTFTYASALWEIQQVVDADSKIVQLSEI NMDNSPNLAWPLIVTALRANSAVKLQ <b>SARGSHHHHHH</b>
SKLTDLK SSFVGM MHSLLF0	>HCoV-229E_nsp7-8-His <sub>6</sub> CCTNVVLMGILSNMNIASNSKEWAYCVEMHNKINLCDDPETAQELLLALLAFFLSKHSDFGLGDLVDSYFENDSILQSVA IPSFVAYETARQEYENAVANGSSPQIIKQLKKAMNVAKAEFDRESSVQKKINRMAEQAAAAMYKEARAVNRKSKVVSA GMLRRLDMSSVDTILNMARNGVVPLSVIPATSAARLVVVVPDHDSFVKMMVDGFVHYAGVVWTLQEVKDNDGKNVHL KDVTKENQEILVWPLILTCERVVKLQ <b>SGSGHHHHH</b>
SKLTDIK ASTYVG SAMHSLI	>PEDV_nsp7-8-His <sub>6</sub> CSNVVLLGCLSSMNVSANSTEWAYCVDLHNKINLCNDPEKAQEMLLALLAFFLSKNSAFGLDDLLESYFNDNSMLQSV LPSYVIYENARQQYEDAVNNGSPPQLVKQLRHAMNVAKSEFDREASTQRKLDRMAEQAAAQMYKEARAVNRKSKVV .FGMLRRLDMSSVDTILNLAKDGVVPLSVIPAVSATKLNIVTSDIDSYNRIQREGCVHYAGTIWNIIDIKDNDGKVVHVKEV TAQNAESLSWPLVLGCERIVKLQ <b>SGSGHHHHHH</b>
SKLTEMI AYAALPS LLFGMLI	>FIPV_nsp7-8-His <sub>6</sub> KCTNVVLLGLLSKMHVESNSKEWNYCVGLHNEINLCDDPDAVLEKLLALIAFFLSKHNTCDLSDLIESYFENTTILQSVAS SWIAYEKARADLEEAKKNDVSPQLLKQLTKACNIAKSEFEREASVQKKLDKMAEQAAASMYKEARAVDRKSKIVSAMHS KKLDMSSVNTIIEQARNGVLPLSIIPAASATRLIVVTPNLEVLSKVRQENNVHYAGAIWSIVEVKDANGAQVHLKEVTAAN ELNITWPLSITCERTTKLQ <b>SGSGHHHHHH</b>
SKLTEM YAALPSV FGMLKKI	>TGEV_nsp7-8-His <sub>6</sub> KCTNVVLLGLLSKMHVESNSKEWNYCVGLHNEINLCDDPEIVLEKLLALIAFFLSKHNTCDLSELIESYFENTTILQSVASA VIALEKARADLEEAKKNDVSPQILKQLTKAFNIAKSDFEREASVQKKLDKMAEQAAASMYKEARAVDRKSKIVSAMHSLL LDMSSVNTIIDQARNGVLPLSIIPAASATRLVVITPSLEVFSKIRQENNVHYAGAIWTIVEVKDANGSHVHLKEVTAANELN LTWPLSITCERTTKLQ <b>SGSGHHHHHH</b>
SKLTDL SEFSHLA MQTMLF DSNENL DGFVSVE GAPLTNO	>MERS-CoV_nsp7-11-His6 KCTSVVLLSVLQQLHLEANSRAWAFCVKCHNDILAATDPSEAFEKFVSLFATLMTFSGNVDLDALASDIFDTPSVLQATL ATFAELEAAQKAYQEAMDSGDTSPQVLKALQKAVNIAKNAYEKDKAVARKLERMADQAMTSMYKQARAEDKKAKIVSA GMIKKLDNDVLNGIISNARNGCIPLSVIPLCASNKLRVVIPDFTVWNQVVTYPSLNYAGALWDITVINNVDNEIVKSSDVV TWPLVLECTRASTSAVKLQNNEIKPSGLKTMVVSAGQEQTNCNTSSLAYYEPVQGRKMLMALLSDNAYLKWARVEGK ELQPPCKFLIAGPKGPEIRYLYFVKNLNNLHRGQVLGHIAATVRLQAGSNTEFASNSSVLSLVNFTVDPQKAYLDFVNAG CVKMLTPKTGTGIAISVKPESTADQETYGGASVCLYCRAHIEHPDVSGVCKYKGKFVQIPAQCVRDPVGFCLSNTPCNV CQYWIGYGCNCDSLRQAALPQ <b>SKDSNFLNHHHHH</b>

637



640 Figure S 1: Mpro mediated processing of precursor protein constructs. SDS-PAGE analysis of Mpro (nsp5)-641 mediated processing and generation of coronavirus nsp7+8 complexes with authentic N- and C-termini from 642 polyprotein precursors nsp7-8, nsp7-8-9 and nsp7-11. The nsp7+8 complexes of HCoV-229E, PEDV, FIPV, TGEV, 643 and SARS-CoV were produced from the respective nsp7-8-His6 precursors and MERS-CoV was produced from an 644 nsp7-9-His6 precursor. Precursor proteins were purified by Ni<sup>2+</sup>-IMAC and ion-exchange chromatography. Then, 645 15 µg protein was cleaved with M<sup>pro</sup> (nsp5-His<sub>6</sub>, 5 µg) for 48 h at 4 °C. Subsequently, His<sub>6</sub>-tag containing cleavage 646 products were removed by passing the material through a Ni2+-IMAC column and nsp7+8 complexes were 647 enriched by ion-exchange chromatography. (A) SDS-PAGE showing the purified M<sup>pro</sup> (nsp5-His<sub>6</sub>) – lanes 1, 5, 9, 648 13, 17; nsp7-8- His6 - lanes 2, 6, 10, 14, 18; M<sup>pro</sup>-mediated cleavage reaction - lanes 3, 7, 11, 15, 19; enriched nsp7+8 649 complexes - lanes 4, 8, 12, 16, 20. (B) SDS-PAGE showing the purified Mpro (nsp5-His6) - lane 1; nsp7-8-9-His6 -650 lane 2; M<sup>pro</sup>-mediated cleavage reaction – lane 3; enriched nsp7+8 complex – lane 4. (C) SDS-PAGE showing the 651 purified Mpro (nsp5-His6) - lane 1; nsp7-8-9-10-11-His6 - lane 2; Mpro mediated cleavage reaction - lane 3. Lane M, 652 marker proteins with molecular masses in kD indicated to the left. Black arrows on the right indicate the identities 653 of proteins generated from precursor proteins by M<sup>pro</sup>-mediated cleavage. Gray arrowheads indicate aberrant in 654 vitro cleavage products of nsp8 as observed previously for SARS-CoV [22]. nsp - nonstructural protein. +/-655 indicate the presence or absence of the respective proteins.





657 Figure S 2: Native MS of nsp7+8 complexes of seven CoVs representing five different CoV species. 658 Representative mass spectra showing distinct nsp7+8 complexation patterns that were classified into the three 659 groups A, B and AB. Complex formation triggered by Mpro (M) mediated cleavage of 15 µM nsp7-8-His6 or MERS-660 CoV nsp7-11-His6 precursors in 300 mM AmAc, 1 mM DTT, pH 8.0. (A) SARS-CoV and (B) SARS-CoV-2 from group A forming nsp7+8 (2:2) heterotetramers (red), (C) FIPV and (D) TGEV from group B forming nsp7+8 (2:1) 661 662 heterotrimers (blue) as well as (E) HCoV-229E and (F) PEDV from group AB forming both complex 663 stoichiometries. (G) MERS-CoV, also from group AB, produced from an nsp7-11 precursor, additionally results in 664 several processing intermediates that allow for an estimation of relative cleavage efficiencies at different cleavage

sites. All groups form nsp7+8 (1:1) heterodimers as intermediate state (green).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.30.320762. this version posted October 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.



Figure S 3: Zoom in on nsp7 and nsp8 monomers and homodimers from spectra in Figure S 2. Representative
native mass spectra showing region 1500-3600 *m/z* (for full spectra see Figure S 2) of cleaved nsp7-8 or nsp7-11
precursors showing that homodimers of nsp8 (green, nsp82) were detected for SARS-CoV (A) and SARS-CoV-2
(B), but not for (C) FIPV, (D) TGEV, (E) HCoV-229E, (F) PEDV and (G) MERS-CoV. Homodimers of nsp7 (yellow, nsp72) were detected for all seven CoVs.



**Figure S 4: Mass and oligomeric state of nsp7-8 precursors.** Native MS of nsp7-8 precursors (**A-G**) sprayed at 18  $\mu$ M from 300 mM AmAc, pH 8.0 and 1 mM DTT. Dominant charge envelope is highlighted (blue box). Labelled are charge states and molecular mass. (**C**) Inset shows mass heterogeneity in FIPV nsp7-8. The experimental molecular weight  $M_{exp}$  of the precursors agrees with the sequence-derived theoretical  $M_{theo}$ (Table S 1). Only FIPV nsp7-8 contained two mass species separated by ~110 Da. This heterogeneity was attributed to the precursor's central nsp8 domain following M<sup>pro</sup> processing. Assignment to an amino acid





Figure S 5: Monomer-dimer equilibrium in SARS-CoV and SARS-CoV nsp7-8 precursors. Native mass spectra
of nsp7-8 precursors of SARS-CoV (A) 18 μM and (B) 9 μM and SARS-CoV-2 (C) 18 μM and (D) 9 μM. Proteins at
18 μM were diluted to 9 μM, incubated for 10 min and then sprayed from 300 mM AmAc, pH 8.0 and 1 mM DTT.
Charge states are labelled.



Figure S 6: Topological reconstruction from CID-MS/MS of TGEV and FIPV nsp7+8 heterotrimers. (A) Three
 product ion spectra of TGEV 11+ nsp7+8 (2:1) heterotrimers are shown at different collisional voltages (CV).
 Increased collisional activation allows assignment of additional charge states and product species. (B) Product ion
 spectra of the FIPV 11+ nsp7+8 (2:1) show a similar dissociation pattern as its homologue from TGEV. (C)
 Schematic pathway of dissociation allows for a topological reconstruction of the TGEV and FIPV heterotrimers.



694

Figure S 7: Topological reconstruction from CID-MS/MS of PEDV and HCoV-229E nsp7+8 heterotrimers. (A)
Product ion spectra of PEDV 11+ nsp7+8 (2:1) heterotrimers shows nsp7 homodimeric species, revealing the core
interaction of this protein complex. (B) Product ion spectra of HCoV-229E 11+ nsp7+8 (2:1) heterotrimers are
shown. Compared to the dissociation pattern of PEDV homologue, the HCoV-229E nsp7 and nsp8 high charge
products have distinct intensities and charge state distributions but the detected product species are similar. (C)
Schematic pathway of dissociation allows for a topological reconstruction of the PEDV and HCoV-229E
heterotrimer.



703 Figure S 8: Topological reconstruction from CID-MS/MS of PEDV and HCoV-229E nsp7+8 heterotetramers. 704 (A) Product ion spectra of PEDV 16+ nsp7+8 (2:2) heterotetramers are shown at two different CV. At 80 V CV, the 705 dissociation of one nsp7 or alternatively one nsp8 occurs, as indicated by the high and low charge products. 706 Notably, the high charge state of the dissociated nsp8 suggests an extended shape of the ion. At 140 V CV, nsp8 707 dissociates at lower charge states suggesting alternative dissociation pathways. (B) Product ion spectra of HCoV-708 229E 16+ nsp7+8 (2:2) heterotetramers at 140 V are similar to the PEDV heterotetramers. (C and D) The zoom 709 (2000-9000 m/z), excluding the nsp8 high charge states, reveals additional product species (e.g. the nsp8<sub>2</sub> 710 homodimer). (E) Schematic pathway of dissociation allows for a topological reconstruction of the PEDV and 711 HCoV-229E heterotetramer.



712

Figure S 9: Topological reconstruction from CID-MS/MS of MERS-CoV nsp7+8 heterotetramers. (A) Product
ion spectra of MERS-CoV 15+ nsp7+8 (2:2) heterotetramers are shown at two different CV. At 100 V CV, the
dissociation of one nsp7 or, alternatively, one nsp8 occurs as indicated by the high and low charge products. At
140 V CV, other dissociation products increase in intensity (e.g. nsp82 homodimer and nsp7+8(1:1) heterodimer).
(B) Schematic pathway of dissociation allows topological reconstruction of the MERS-CoV heterotetramers.



Figure S 10: Topological reconstruction from CID-MS/MS of SARS-CoV and SARS-CoV-2 nsp7+8
 heterotetramers. Product ion spectra of (A) SARS-CoV-2 and (B) SARS-CoV 17+ nsp7+8 (2:2) heterotetramers are
 shown. High charge state precursor allows for efficient dissociation already at relatively low activation (CV 80 V).
 Product ion species can be clearly assigned from the overview spectra. (C) Schematic pathway of dissociation
 allows topological reconstruction of the SARS-CoV-2 and SARS-CoV heterotetramers.



726 Figure S 11: MALDI-MS of nsp7+8 complexes from FIPV and HCoV-229E stabilized with crosslinker. (A) 727 MALDI mass spectrum of FIPV nsp7+8 crosslinked with 0.15 % glutaraldehyde for 25 min at 4 °C. Inset shows the 728 most abundant nsp7+8 complexes: (1:1) heterodimer (32.4 kDa) and (2:1) heterotrimer (42.5 kDa). Abundance is 729 determined from relative peak areas as indicated (red). (B) MALDI mass spectrum of HCoV-229E nsp7+8 730 crosslinked with 0.15 % glutaraldehyde for 25 min at 4 °C. Inset shows most abundant nsp7+8 complexes: (1:1) 731 heterodimer (32.2 kDa) and (2:2) heterotetramer (64.6 kDa). Abundance is determined from relative peak areas as 732 indicated (blue). Symbols depict stoichiometry of mass species of nsp7 (yellow) and nsp8 (green) with highest 733 signal strength. Masses are higher in crosslinked samples due to the additional glutaraldehyde molecules. Mass 734 spectra were not calibrated. Each spectrum shown was generated from three MALDI spots. Most signals, except 735 for the HCoV-229E heterotetramer, above 50,000 m/z are low abundant and likely due to over-crosslinking.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.30.320762. this version posted October 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.



737 Figure S 12: SDS-PAGE analysis of chemically cross-linked HCoV-229E, FIPV, and SARS-CoV nsp7+8 738 complexes. The 5  $\mu$ g protein of nsp7+8 complexes were crosslinked with 10  $\mu$ M BS<sup>3</sup> (ThermoFisher) in reaction 739 buffer (20 mM HEPES-KOH, pH 8.0, 30 mM KCl, and 2 mM β-mercaptoethanol). Crosslinking was carried out at 740 37 °C for 30 min and quenched with 50 mM AmAc for another 30 min at 37 °C. After terminating the crosslinking 741 reaction, the samples were mixed with an excess of Laemmli sample buffer (50 mM Tris-HCl, pH 6.8, 2.5 % (w/v) 742 SDS, 10 % (v/v) glycerol, and 0.01 % (w/v) bromophenol blue) and analyzed on 12 % SDS-PAGE. Lanes 1, 3, 5 -743 nsp7+8 complexes not treated with BS3 (-) and lanes 2, 4, 6 - nsp7+8 complexes treated with BS3. Lane M, marker 744 proteins; molecular masses in kD are indicated to the left. Black arrows on the right indicate the different 745 oligomeric states of the nsp7+8 complexes obtained by crosslinking.



Figure S 13: Candidate structures in agreement with the observed stoichiometries and topologies observed. (A)
For the full-length heterotetramer, an isolated structure does not exist. However from the larger SARS-CoV
nsp7+8 hexadecamer [20] (pdb 2ahm), two conformer subcomplexes of nsp7+8 (2:2), T1 and T2, can be extracted.
Both conformers constitute a head-to-tail interaction of two heterodimers by an nsp8:nsp8 interface. Notably,
nsp8 in T1 is more extended, containing an almost full-length amino acid sequence (2-193), while in T2 the nsp8
N-terminal 35 to 55 residues are unresolved. (B) For the trimeric complexes, the only deposited structure is
FIPV nsp7+8 (2:1) trimer [21] (pdb 3ub0), which agrees well with our experimental topology.



Figure S 14: Fits to all SAXS data. Experimental data (blue with experimental errors) for SARS-CoV-2 nsp7+8
 complexes fitted with a mixture of T1 and hexadecamer (red), or T1 and dimer of T1 (green).

1       10       20	1       10       20       20       30       40       50       63         SLITEMAGE INVERSEMENTER MENDEN IN COLLEMENTING COLLEMENTER INTO DE AMILENTIALE FLANTIALE FL		nsp7									
PPV SELTE MAGEN VLICELSKIM V SNREE NY COLLENSKIM VSNREE NY COLLENS HILCO DE MUEELLAN HILF HI PEDU SELTE DIK TN VLICELSKIM VSN BENNY JANE KENNY COLLENS HILCO DE MUEELLAN HILF HI PEDU SELTE DIK SNV HILCO SNNV JANE SNNV JANE KENNY COLLENS HILCO DE MUEELLAN HILL HI SERSOW SELTE DIK SNV HILCO SNNV JANE SNNV JANE KENNY COLLENS HILLONDE KARELLAN HILF HI PEDU SELTE DIK SNV HILCO SNNV JANE SNNV JANE KENNY COLLENS HILLONDE KARELLAN HILF HILF SARSOW SELTE DIK SNV HILCO SNNV JANE SNNV JANE KENNY COLLENS HILLONDE KARELLAN HILF HILF SARSOW SELTE DIK SNV HILF SVY HILF SNNS SNV HILF SNV HILF SNNS SNV HILF	PPV TGEV T		1 •	10 •	20 •	30	40 •	50 •	60 •			
Inc. 22926       ST. TD ILKG TINGUT GOTES NAM ILLES NE ALE ST.	Inc. V22920       STATE TOLKG THE VIEW OLD SAME TARE AND VERY AND THE VIEW OLD PERSON OF TALLAR FLUE PERSON       Int. TOLKG THE VIEW OLD PERSON OF TALLAR FLUE VIEW OLD PERSON OF TALE AND VIEW OLD PERSON O	FIPV TGEV	SKLTEMP	KCHNVVLIGI KCHNVVLIGI	LSKMHVESNS ISKMHVESNS	KEWNYCVGL KEWNYCVGL	HNEINLCDDP HNEINLCDDP	DAVLEKLLALI EtVLEKLLALI	AFFL			
PEOV SARS-COV SARS-COV SARS-COV SARS-COV SARS-COV SARS-COV SARS-COV SARS-COV SARS-COV SARS-COV SARS-COV SARS-COV SARS-COV SARS-COV SARS-COV SARS-COV SKINTCDLSD TO TO TO TO TO TO TO TO TO TO	PEOU PEOU PEOU PEOU PEOU SARS-COV PEOU PE	HCoV-229E	SKLTDL	KCTNVVLMGI	L SNMN I A SN	KEWAYCVEMI	HNKINLCDDP	ETAQELLLALL	AFFL			
SARS-CoV2 SIMS DVICE TS VVPLUS V COLRVESSENT DA COLLETED LIAARD TTPALE RAMY SIL SVIL SARS-CoV2 SIMS DVICE TS VVPLUS V COLRVESSENT DA CV LEND LIAARD TTPALE RAMY SIL SVIL INSPE TGEV SKINT CD LG D TIES TENTT ILG SVAR ANAALPS WIAYER ARADLE PARKING'S LOVI SKINT CD LG D TIES TENTT ILG SVAR ANAALPS WIAYER ARADLE PARKING'S LOVI PEV SKINT CD LG D TIES TENTT ILG SVAR ANAALPS WIAYER ARADLE PARKING'S LOVI PEV SKINT CD LG D TIES TENTT ILG SVAR ANAALPS WIAYER ARADLE PARKING'S LOVI PEV SKINS DF CLOP TIES TENTT ILG SVAR ANAALPS WIAYER ARADLE PARKING'S LOVI SKINS ARCOV SKINT CD LG D TIES TENTT ILG SVAR ANAALPS WIAYER ARADLE PARKING'S LOVI SKINS ARCOV SKINT CD LG TUES TENTT ILG SVAR ANAALPS WIAYER ARADLE PARKING'S LOVING S LOVIN SKINS ARCOV SKINT CD LG TUES TENTT ILG SVAR ANAALPS WIAYER ARADLE PARKING'S LOVING S LOVIN SKINS ARCOV SKINT TENDET SVGAR ANAALPS WIAYER ARADLE PARKING'S LOVING S LOVIN SARS-COV2 SG GANDLIN HOUSE MUDINAL TG AILS DT S SUBSTAN ARADLE PARKING'S LOVING S LOVIN SARS-COV2 SG GANDLIN HOUSE MUDINAL TG AILS DT S SUBSTAN ARADLE STAN AND THE SUBSTAN ARADLE STORE AND	SARS-Cov2 SARS-Cov2 STMSDVICTSVVILSVICQURVESSITERACCOLLERD LLAR DITEATER MVS.LdviL nsp8 70 80 10 10 10 10 10 10 10 10 10 1	PEDV MERS-CoV	SKLTDI SKLTDI	KCSNVVLLGC KCTSVVLLSV	LSSMNVSANS LOOLHLEANS	TEWAYCVDL Rawafcvkc	HNKINLCNDP HNDILAATDP	EKAQEMLLALL SEAFEKFVSLF	AFFL ATLM			
SARG-002         INFER DYNET I DYNET DY	SARG-0042         Insp8         (1)         (1)         TOP         SK HENT COLSPAN THE INCOMA ANALD SWITALD SWITALDEN RADLEE (RADLES TENDING) POILS         HOUVER SK HENT COLSPAN THE INCOMA ANALD SWITALDEN RADLEE (RADLES TENDING) POILS         HOUVER SK HENT COLSPAN THE INCOMA ANALD SWITALDEN RADLEE (RADLES TENDING) POILS         HOUVER SK HENT COLSPAN THE INCOMA ANALD SWITALDEN RADLEE (RADLES TENDING) POILS         HOUVER SK HENT COLSPAN THE INCOMA ANALD SWITA RADLES TENDING (NORS) POILS         MERSCOV         SK HEST CLOPE UD SV TENDIS HOUVER STUDE SY VALLE AT PALE ELSON ON NORS (S. POILS POILS SK VILLES TENDING) POILS (S. POILS POILS SK VILLES TENDING)         SK HEST CLOPE UD ANA DE POILS SUBALLES TENDING (S. POILS POILS POILS POILS SK VILLES TENDING (S. POILS PO	SARS-CoV	SKMSDV	KCTSVVLLSV	LÕÕLRVESSS	KLWAQCVQLI	HNDILLAKDT	TEAFEKMVSLL	SVLL			
Insp8         TO       Q       Q       Q       Q       Q         FIFY       SK H HTCDLDS D LIES YE ENT IL IG SVAA SVAAL PS WILLYEK A RADLE PARKIND'S POLLY         KK H HTCDLS D LIES YE ENT IL IG SVAA SVAAL PS WILLYEK A RADLE PARKIND'S POLLY         WERS-CCV         SK HATCDLS D LIES YF END IS MCGVAA TWO DP SVIJET RADOVE PARVING SE POLY         SK HATCDLS D LIES YF END IS MCGVAA TWO DP SVIJET RADOVE PARVING SE POLY         SK HATCDLS D LIES YF END IS MCGVAA TWO DP SVIJET RADOVE PARVING SE POLY         SK HATCDLS D LIES YF END IS MCGVAA TWO DP SVIJET RADOVE PARVING SE POLY         SARS-COV         SK HATCDLS D LIES YF END IS MCGVAA TWO DP SVIJET RADOVE PARVING SE POLY         SARS-COV         ST SCAVDING TWO DE SVIJET RADOVE PARVING SE POLY         SARS-COV         ST SCAVDING TWO DE SVIJET RADOVE PARVING SE POLY         SARS-COV         ST SCAVDING TWO PARVING SE POLY	Insp8         FIPU         SKHNTCDLSE FIRSTYENTETING VAR AVAALES WIXALES WIXALE ARADLE BARKIND VAR POLICK         FIPU         SKHNTCDLSE FIRSTYENTETING VAR AVAALES WIXALES WIXE ARADLE BARKIND VAR POLICK         PEOV         SKHNTCDLSE FIRSTYENTETING VAR AVAALES WIXE ARADLE BARKIND VAR POLICK         PEOVER SKINSTYLE ARADLE BARKIND VAR EVER SVITTER ARADLE BARKIND VAR EVER SUPERIOR VAR EVER SVITTER ARADLE BARKIND VAR ARADLE BARKIND VAR ARADLE BARKIND VAR EVER SVITTER ARADLE BARKIND VAR ARADLE	SARS-Cov-2	SKMSDV		DOGTRAFSE		INDILLAKDI	IEAFDEMVSEL	<u> </u>			
nsp8         Y       Y       Y         Y <th cols<="" th=""><th>INPRESENT 1100 VARAMYAALP SWIALESKA ADULEEAKKINDUS POLIK         YEY         <th colspan="2" th="" y<=""><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th></th></th></th>	<th>INPRESENT 1100 VARAMYAALP SWIALESKA ADULEEAKKINDUS POLIK         YEY         <th colspan="2" th="" y<=""><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th></th></th>	INPRESENT 1100 VARAMYAALP SWIALESKA ADULEEAKKINDUS POLIK         YEY         YEY <th colspan="2" th="" y<=""><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th></th>	<th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>									
Insp8         70       0       7       7         FIPU       SK HINT CDL'S DI TES YFENNET TO SVAAS AVAALP SWITAYEKA RADLER KINDVS. POILLE         PEOV       SK HINT CDL'S DI TES YFENNET TO SVAAS AVAALP SWITAYEKA RADLER KINDVS. POILLE         PEOV       SK HINT CDL'S DI TES YFENNET TO SVAAS AVAALP SWITAYEKA RADLER KINDVS. POILLE         MERS-COV         SK HINT CDL'S DI TES YFENNET TO COVAS AVAAS AVAALP SWITAYEKA RADLER KINDVS. POILLE         SK HINT CDL'S DI TES YFENNET TO COVAS AVAAS AVAALP SWITAYEKA RADLER AVANACISS. POILLE         SK HINT CDL'S DI TES YFENNET TO COVAS AVAAS AVAALP SWITAYEKA RADLER AVAANCISS. POILLE         SK HINT CDL'S DI TES YFENNET TO COVAS AVAAS AVAALP SWITAYEKA RADLER AVAANCISS. POULA         SK HINT CDL'S DI TES YFENNET TO COVAS AVAAS AVAALP SWITAYEKA RADLER AVAANCISS. POULA         SK HINT CDL'S DI TES YFENNET TO COVAS AVAAS AVAALP SWITAYEKA RADLER AVAACHANCISS. POULA         SK HINT CDL'S DI TES SYFENNET TO COVAS AVAAS AVAALP SWITAYE RADLER AVAALPS WITAYEKA RADLE AV	INSP8         IPPU         SK HINT CDLS DI TES YEENNET TE SVAR AMAALPSWIAYEK RADLEEN (KNDVS. PQLLK SK HINT CDLS DI TES YEENNET TE SVAR AMAALPSWIAYEK RADLEEN (KNDVS. PQLLK SK HINT CDLS DI TES YEENNET TE SVAR AMAALPSWIAYEK RADLEEN (KNDVS. PQLLK SK HINT CDLS DI TES YEENNET TE SVAR AMAALPSWIAYER RADCHERN (KNDVS. PQLLK SK HINT CDLS DI TES YEENNET TE SVAR AMAALPSWIAYER RADCHERN (KNDVS. PQLLK SK HINT CDLS DI TES YEENNET TE SVAR AMAALPSWIAYER RADCHERN (KNDVS. PQLLK SK HINT CDLS DI TES YEENNET TE TOTS VAR AMAALPSWIAYER RADCHERN (KNDVS. PQLLK SK HINT CDLS DI TES YEENNET TE TOTS VAR AMAALPSWIAYER RADCHERN (KNDVS. PQLLK SK HINT CDLS DI TES YEENNET TE TOTS VAR AMAALPSWIAYER RADCHERN (KNDVS. PQLLK SK HINT CDLS DI TES YEENNET TE TOTS VAR AMAALPSWIAYER RADCHERN (KNDVS. EVVLK SARS-GVZ         SY AN AMAALPSWIAYER RADCHERN (KNDVS. PQLLK SK HINT CND AMAALPSWIAYER (KNDVS. KNDVS. KNDVS. KNDVS. KNDVS. KNDVS. KNDVS. KNDVS. KNDSCHING, KNDVS. SARS-GVZ         SY AN AMAALPSWIAYER RADCHERN (KNDVS. KNDVS. KNDVS. KNDVS. SARS-GVZ         SY AN AMAALPSWIAYER RADCHERN (KNDVS. KNDVS. SARS-GVZ         SY AN AMAALPSWIAYER RADCHERN (KNDVS. KNDVS. SARS-GVZ         SY AN AMAALPSWIAYER RADCHERN (KNDVS. KNDVS. SARS-GVZ         SY AN AMAALPSWIAYER KANDYK KNDVS. SARS-GVZ <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>											
FIP       SK HNTCD15D       SK SY E MY TILG SVA AVAALPS WIA LEVA RADLEE AKKNOVS. POILS TES WY E WITT ILC SVA AVAALPS WIALEK ARADLEE AKKNOVS. POILS SK HS FC10D       DOULS YF E MNS ILC SVA AVAALPS WIALEK ARADLEE AKKNOVS. POILS SK HS FC10D       DOILS SK SS FC10D </th <th>Image: Strate in the strate</th> <th></th> <th></th> <th></th> <th>nsp</th> <th>8</th> <th>47</th> <th></th> <th></th>	Image: Strate in the strate				nsp	8	47					
FIPU       SKINT (D LG D LG T LG ST LG N, T LG SV AGALAALP S NIA LEA RAD LE CANKEND VS. FOULT SKINT (D LG B T LG ST LG N, ALALP S NIA LEA RAD LE CANKEND VS. FOULT SKINT (D LG B T LG ST LG N, ALALP S NIA LEA RAD LE CANKEND VS. FOULT SKINT (C LG CAN LG V FF NN ST LG SV AG ST VG LY ALALP S NIA LEA RAD LE CANKEND VS. FOULT SKINT (C LG CAN LG V FF NN ST LG SV AG ST VG LP SV VIET AR COLVEN AVANES FOULT SKINT (C LG CAN LG V FF NN ST LG SV AG ST VG LP SV VIET AR COLVEN AVANES FOULT SKINT (C LG CAN LG V FF NN ST LG SV AG SV VG LP SV AFF AF C CAN VG AM CG SL FV VIET SARS-GV/2         SO GAV D L N F C CE ML D NR AT LG ALAS FF S LP SV AA AT A C CAY CE AM CG SC COV SARS-GV/2       SO GAV D L N FT C CE ML D NR AT LG ALAS FF S LP SV AA AT A C CAY CE AVAN CG SL FV VIET SARS-GV/2         IFV       CTKKA FN TAS SEP FE RE ASVOKK D KK AG AAAS KK KA ANA VD RKS KT VS AM IS LF G ML K (T KKA FN TAS SEP FE RE ASVOKK D KK A CAAAS KK KA ANA VD RKS KT VS AM IS LF G ML K (T KKA FN T AS SEP FE RE ASVOKK D KK AG AAAS KK KA ANA VN RKS KV VS AM IS LF G ML K (T KKA FN T AS SEP FE AS TO KK T D K AM CAAAS KK KA ANA VN RKS KV VS AM IS LF G ML K (T KKA FN T AS SEP FE AS TO KK T D K AM CAAAS KK KA ANA VN RKS KV VS AM IS LF G ML K (T KKA NV AS SEP D R LS TO KK T D KAAA CAN KK AT AVN RKS KV VS AM IS LF G ML K (T KKA NV AS SEP TR D AJAN OK K D D KM T D M KK AAA AN NKRS KV SAM SE LF G ML K (T KK S LN VA S SEP TR D AJAN OK L D K M D OM T O M KK AAA AN AKK ST VS AM OT M F G M IK (K KS LN VA S SEP TR D AJAM OK L D K M D OM T O M KK AAA SE D KRA KY VS AM C T M F G M IK (K KS LN VA S SEP TR D AJAM OK L D K M D OM T O M KK AAA SE D KRA KY V SAM OT M F G M K (K KS LN VA S SEP TR D AJAM OK L D K M D OM T O M KK AAA SE D KRA KY V SAM OT M F G M K (K KS LN VA S SEP TR D AJAM OK L D K M D OM T O M KK AAA SE D KRA KY V SAM OT M F G M K (K KS LN VA S SEP TR D AJAM OK L D K M D OM T O M KK AAA SE D K M M D O M T O M KK AAA SE D K M O M M T O M K A A SE D K M	FIP       SKHNIGDUBE TESTENTITIOSY ANALYSKIALPS NIALES ARADUEENAKNOVS. POILA NEWSON 238         HOOV229E       SKHSDPGLG DIE BIESTENTITIOSY ANALPS NIALES ARADUEENAKNOVS. POILA PEDU         SKNSCHUDER TESTENTITIOSY ANALPS NIALES ARADUEENAKNOVS. POILA SKNSCHUDER TESTENTITIOSY AS TWORDS TA ROEYENAKNOVS. POILA SKNSCHUDER TESTENTITIOSY AS TWORDS TA ROEYENAKNOVS. POILA SKNSCHUDER TESTENTITIOSY AS TWORDS TA ROEYENAKNOVS. POILA SKNSCHUDER TESTENTITIOSY AS TWORDS TESTEN SYNAATATA ORAVGENVING SELEVILS SKNSCHUDER TESTENTITIOSY AS THORAGONAL TO ANALYSKAT ORAVGENVING SELEVILS SKNSCHUDER TESTENTITIOSY AS THORAGONAL TESTENTIAL SKNOVANGES EVVILS SKNSCHUDER TESTENTIAL SKNOKKUDKUDEKA TA SKNY SKAN ORAVGENVING SELEVILS SKNSCHUDER TESTENTIAL SKNOKKUDKUDEKA TA SKNY SKNY SKNY SKNY SKNY SKNY SKNY SKNY		<b>a</b> kuum al						011			
HOW 229E       SK H S D F GLO DL VDS YFENDS I G SVA S SFV GM P S F VA YET ROOK YENA GOL NA GSS. POIL R         PEDU       SK H S D F GLO DL VDS YFENDS I G SVA S SFV GM P S F VA YET ROOK YENA GOL NA GSS. POIL R         SK H S D F GLO DL VDS YFENDS I G SVA S SFV GM P S F VA YET ROOK A Y GEAMING GD S D VI K         SARS-GOV2       SM G G A VDI NK L CEEM LD NR A TLO ALAS PS SL P S YAA YA TA O E A YE OA VAN GOS. E VVI K         SARS-GOV2       SW G G A VDI NK L CEEM LD NR A TLO ALAS PS I P S YAA YA TA O E A YE OA VAN GOS. E VVI K         SARS-GOV2       SW G G A VDI NK L CEEM LD NR A TLO ALAS PS I P S YAA YA TA O E A YE OA VAN GOS. E VVI K         SARS-GOV2       SW G G A VDI NK L CEEM LD NR A TLO ALAS PS I P S YAA YA TA O E A YE OA VAN GOS. E VVI K         SARS-GOV2       SW G G A VDI NK L CEEM LD NR A TLO ALAS YK A SRAVD RK CI V SA M HS L E G GLI K         TGEV       OF KKA MUNKA SF PE RE AS VOKKI D KM E O ALAS YK A SRAVD RK CI V SA M HS L E G GLI K         PEDV       OF KKA MUNKA SF D R E S SU OK KI I NK ME O ALAA MYK E A RAVN RK CI V SA M HS L E G GLI K         PEDV       OF KKA MUNKA SF D R E S SU OK KI I NK ME O ALAA MYK K RANG NK SK VY SA M HS L E G GLI K         SARS-GOV2       AL OK A VI T KA NA KK SF D R B SA NYK L D RAME O ALAO MYK E A RAVN RK K VY SAM O' M E G ALA         SARS-GOV2       KL K S LNYKK SE PD R DAAM OR L E KM D O AM TO MYK CAR SE D K RA WY SAM O' M E G ALA         SARS-GOV2       KL K S SVN TH I E OO RINO VLD S TH PALASA TR WY VI TP NL E VILS KWR E N NY WE AG A LI W M E G ALAO MYK YA SA SU M T I L CAR NO VLD K K W	How 228E       Skiks proloud by Vp Skips Phois IE GSVA Skip Volp S FV AY BETR ROY ELEMENT ANGES. POIN Skiks proloud by Skips Phois Mark Solve TY Volp Sy Vivol PS VY VIPS NROY ELEMENT AND SOLVE SKIPS POIN Skips proloud by Skips Phois Mark Solve TY Volp Sy Vivol PS VY VIPS NROY ELEMENT AND SOLVE SKIPS POINT Skips proloud by Skips Phois Mark Solve TY Volp Sy Vivol PS VY VIPS NROY ELEMENT AND SOLVE SKIPS POINT Skips proloud by Skips Phois Mark Solve TY Volp Sy Vivol PS VY VIPS NROY ELEMENT AND SAVE SKIPS VIA Skips proloud by Skips Phois Mark Solve Ty Mark Solve Skips Vaa Var Noope Skips PS VAA VAR Solve Skips PS VAA VAR Skips PS VAA	FIPV TGEV	SKHNTCI SKHNTCI	DLSDLIESYF	ENTTILOSVA	SAYAALPSW Sayaalpswi	LAYEKARADL IALEK <mark>a</mark> radl	EEAKKNDVS.P EEAKKNDVS.P	QILK			
PEUV       IN IS A FOLD DE LES IF DUNS MED SIVE AT LEVELA TY LE LEN OUT EQUING GELPOUNG SER OUT SARS-COV         SARS-COV       SING A VDINK L'EE MLDNRA TO ALLAS EFF SLD SYAAFAT OEAAVEORAMOS GELPOUNG GELPOUN	PEUV MRRS-Cov       Sind Safe Journal Strate Strate Strate Lead Octation Strate S	HCoV-229E	SKHSDFO	GLGDLVDSYF	ENDSILOSVA	SFVGMPSFV	VAYETARQEY	ENAVANGSS.P	ÕIIK			
SARS-Cov2 SARS-Cov2 SHOGAVDINK CEEMLDNRATCAIAS FSSDPSYAAFATAOEAYEONANGDS. EVVLX SARS-Cov2 SHOGAVDINK CEEMLDNRATCAIAS FSSDPSYAAFATAOEAYEONANGDS. EVVLX SARS-Cov2 SHOGAVDINK CEEMLDNRATCAIAS FSSDPSYAAFATAOEAYEONANGDS. EVVLX SARS-Cov2 SHOGAVDINK CEEMLDNRATCAIAS FSSDPSYAAFATAOEAYEONANGDS. EVVLX SARS-Cov2 SHOGAVDINK CEEMLDNRATCAIAS FYSSDPSYAAFATAOEAYEONAYEONAYEONA SHOGAVDINK CEEMLDNRATCAIAS FYSSDPSYAAFATAOEAYEONAYEONAYEONAYEONAYEONAYEONAYEONAY	SARS-COV SARS-COV2 SING GA VDINKLCEE MLD NRA TOGALA EFFS SLPSYAAFATROEAYECAVANGDS. EVVLK SARS-COV2 SING GA VDINKLCEE MLD NRA TOGALA SING SING SING SING SING SING SING SING	MERS-CoV	TFSGNVI	DLDALASDIF	DTPSVLQATI	SEFSHLATFA	AELEA <mark>A</mark> QKAY	QEAMDSGDTSP	QVLK			
37       47       57       67       77       87         FIPY       CIMTRAC NIAKS BFERRASVOKIS DKMARO ANASMYK ARAVDRKSKIV SAMHSLINEGNIK         TGEV       CIMTRAC NIAKS BFERRASVOKIS DKMARO ANASMYK ARAVDRKSKIV SAMHSLINEGNIK         HGV.228E       CIMTRAC NIAKS BFERRASVOKIS DKMARO ANASMYK ARAVDRKSKIV SAMHSLINEGNIK         PEDV       CIMTRAC NIAKS BFDRESSVOKIS DKMARO ANASMYK ARAVDRKSKIV SAMHSLINEGNIK         MERS.COV       ALOKKAMN VANA BFDRESSVOKIS DRMARO ANASMYK ARAVDRKSKIV SAMHSLINEGNIK         SARS-COV       ALOKKAMN VANA BFDRESSVOKIS DRMARO ANASMYK ARAVINKSKIV SAMHSLINEGNIK         SARS-COV       ALOKKANN IAKN AVJEK DKANA RILERMAD CAM TSMAK KAKIV SAMOTM LEGULA         SARS-COV       ALOKKANN IAKN AVJEK DKANA RILERMAD CAM TSMAK KAKIV SAMOTM LEGULA         SARS-COV       KLKKSINVANSEFDRDAAMORELEKMAD CAM TOMYK CAREDKRAKIV TSAMOTM LEGULA         SARS-COV       KLMKSINVANSEFDRDAAMORELEKMAD CAM TSMAK CARESED KRAKIV TSAMOTM LEGULA         SARS-COV       KLMKSINVANSEFDRDAAMORELEKMAD CAM TSMAK CARESED KRAKIV TSAMOTM LEGULA         MERS-COV       KLMKSIN HINNARD CVPTS VITAN SARATE VVVITED LEVISK KINCOEN VITAN SARATE VVVITED LEVISK KINCOEN VITAN SARATE VVVITED LEVISK KINCOEN VITAN SARATE VTIAN SARATE VTIAN SARATE VTIAN SARATE VTIAN SARATE VVVITED SARATE VTIAN SARATE VTIAN SARATE VTIAN SARATE VVITED SARATE VTIAN SARATE	37       47       57       67       77       87         FIPV       UTKACNIAXSEFEREASVQKUDKWAEOAAASWIXKEARAVDRKSKIVSAMHSLUPGMLK         PGV       UTKAANIAXSEFEREASVQKUDKWAEOAAASWIXKEARAVDRKSKIVSAMHSLUPGMLK         PGV       UTKAANNAXAEFDRESSVQKUTKUKKEOAAASWIXKEARAVDRKSKIVSAMHSLUPGMLK         PGV       UTKAANNAXAEFDRESSVQKUTKUKKEOAAASWIXKEARAVDRKSKIVSAMHSLUPGMLK         PGV       UTKAANNVAXSEFDRESSVQKUTKUKKEOAAASWIXKEARAVNRKSTVVSAMHSLUPGMLK         PEV       UTKAANNVAXSEFDRESSVQKUTKUKKEOAAAASWIXKEARAVNRKSTVVSAMHSLUPGMLK         SARS-Cov       XIXKANVAXSEFDREASVOKUKUKKEARAVNRKSONSANNTHSLUPGMLK         SARS-Cov       XIXKSINVAXSEFDREAANDRKULEKMIDOAMTSMIXKOARAVNRKSTVSAMOTMEFTMLR         SARS-Cov2       KIKKSINVAXSEFDREAANDRKILEKMIDOAMTOWXKOARSEDERAKVTSAMOTMEFTMLR         SARS-Cov2       KIKKSINVAXSEFDREAAMORKILEKMIDOAMTOWXKOARSEDERAKVTSAMOTMEFTMLR         PEV       NDAAMORKILEKMIDOAMTOWXKOARSEDERAKVTSAMOTMEFTMLR         PEV       KIDMSSVDTILNAARDOCVPISVIPATSARATUVVIPDISEVKAMVDORVHYACAINSI NAKKAKAVTSAMOTMETMISESVOKUKUNAKSEEDENTINAANDENTIT         MERS-Cov2       KIDMSSVDTILNAARDOCVPISVIPATSARATUVVIPDISEVKAMVDORVHYACAINSI SARS-Cov2       SUDTILNAARDOCVPISVIPATSARATUVVIPDISENSUKKONVERAVYAASALUSIS         MERS-Cov2       KIDNDALNNINARDOCVPISVIPATSARATUVVIPDISENSUKKONVERAVYAASALUSIS       SILSANOTHISENSUKKONVERAVIPANASALUSIS         SARS-Cov2       VEVKDANGSISTINANANGOLINIT       SITTAAKKAVVIPDINIT	SARS-CoV	SMQGAVI Smogavi	DINKLCEEML INBLCEEML	DNRATLOALA	SEFSSLPSYA	AAFATAQEAY AAYATAOFAY	EQAVANGDS.E	VVLK			
37       47       57       67       77       87         FIPV       OLTKACNIAKSEFEREASVOKK DKM DKM EOAAASVYKE ARAVDRKSKIVSA HISLIFEGNLK         TGEV       OLTKAFDIAKSEFEREASVOKK DKM DKM EOAAASVYKE ARAVDRKSKIVSA HISLIFEGNLK         HCOV220       OLTKAFDIAKSEFEREASVOKK LDKM EOAAASVYKE ARAVDRKSKIVSA HISLIFEGNLK         PEOV       OLKKAMOVKA EFPREASTORK LDKM EOAAASVYKE ARAVDRKSKIVSA HISLIFEGNLK         PEOV       OLKKAMOVKA SEPPREASTORK LDKM EOAAAOVYKE ARAVDRKSKIVSA HISLIFEGNLK         MERS-COV       AVOKAVTEKKAN YEKKAKAVKAR LERMD OAMTAOVYKO ARABEDKKAR VYSAMOTMLEGUIK         SARS-COV2       KKSLNVKKSEPPREASTORK LDRMAE OAAAOVYKE ARAVNKKS VYSAMOTMLEGUIK         SARS-COV2       KKSLNVKKSEPPREASTORK LDRMAE OAAAOVYKE ARAVNKKSKVYSAMOTMLEGUIK         SARS-COV2       KKSLNVKKSEPPREASTORK VLD STILDAMO CANTOVYKO ARABEDKKAR VYSAMOTMLEGUIK         SARS-COV2       KKSLNVKKSEPPREASTORK VLD STILDAAMORK LEKMD OAMTOVYKO ARABEDKKAR VYSAMOTMLEGUIK         YEKKSLNVKKSEPPREASTORK VLD STILDAAMORK VLD STILDAAANT KLKNVV VD DEVKKAR VYSAMOTMLEGUIK         YEKKSLNVKKSEPPREASTORK VLD STILDAAANORK VLD STILDAAANT KLKNVV VD DEVKAR VYSAMOTMLEGUIK         YEKKSLNVKKSEPPREASTORK VLD STILDAAANORK VLD STILDAAANT KLKNVV VD DEVKAR VYSANOTMLEGUIK         YEKKSLNVKKSEPPREASTORK VLD STILDAAANORK VLD STILDAAAANT KLKNVV VD DEVKAR VYSANOTMLEGUIK         YEKKSLNVKKSEPPREASTORK VLD STILDAAANORV VD STILDAAAANT KLKNVV VD DESTKKINGENVKYKAGANOTMLEGUIK         YEKKSLNVKKSLNVKSLNKKSLNVKKKAKAAANYK	37       47       57       67       77       87         FIPV       CM TKAC NIANSEFEREASVOKK DKKA PO AAAS WX FARAVDRKSKIVSA HISLEGMLK         TGEV       CJ TKAF NIANSEFEREASVOKK DKKA DKKA PO AAAS WX FARAVDRKSKIVSA HISLEGMLK         HCW229       CJ KKAF NIANSEFEREASVOKK DKKA DKKA PO AAAS WX FARAVDRKSKIVSA HISLEGMLK         PEDV       CJ KKAN VAXAS FDRESSVOKK DR NK ME CAAAASWX FERAVDRKSKIVSA HISLEGMLK         MERS.Cov2       CK KKAN VAXAS FDRESSVOKK DR NK ME CAAAASWX FERAVDRKSKIVSA MHSLEGMLK         SARS.Cov2       KIKKS INVAXSEFDREASTOR STORE         97       107       17       127       137       47         FIPV       KIKKS INVAXSEFDREASTORE       STAFAVNRKSKIVSA ANOTMEFMIK       STAFAVNRKSKI SINGALASTORE       STAFAVNRKSKI SINGALASTORE         SARS.Cov2       KIKKS INVAXSEFDREASTORE       STAFAVNRKSKI SINGALASTORE       STAFAVNRKSKI SINGALASTORE         SARS.Cov2       KIKKS INVAXSEFDREASTORE       STAFAVNRKSKI SINGALASTORE       STAFAVNRKSKI SINGALASTORE         97       107       117       127       37       47         FIPV       KIKKS INVAXSEFDREASTORE       STAFAVNRKSKI SINGALASTORE       STAFAVNKKSKI SINGALASTORE         SARS.Cov2       SVDT LINANDANG VLEDASTIF       STAFAVNRKSKI SINGALASTORE       STAFAVNKKSKI SINGALASTORE         MESK.Cov4       KIKKS SINT HID ARNOVE <td< th=""><th>3AK3-00V-2</th><th>DIQUAVE</th><th></th><th></th><th></th><th></th><th></th><th></th></td<>	3AK3-00V-2	DIQUAVE									
37       47       57       67       77       87         FIPV       OFTKACNIAKSEFEREASVOKKIDKAREOAAASVYKEARAVDRKISKIVSA HISLIFGMLK         HC0-V229E       OFTKACNIAKSEFEREASVOKKIDKAREOAAASVYKEARAVDRKISKIVSA HISLIFGMLK         PDV       OFKKAANVARA BEPDRESVOKKIDKAREOAAASVYKEARAVDRKISKIVSA HISLIFGMLK         PDV       OFKKAANVARA BEPDRESVOKKIDKAREOAAASVYKEARAVVRKISKIVSA HISLIFGMLK         PDV       OFRHAMIVAKSEPDREASVOKTIDKAREOAAAAVKEARAVVRKISKIVSA MHSLIFGMLK         SARS-COV2       UKKASLIVAKSEPDREASTORKLERMADOAMISVKAARAVKEARAVNRKISKIVSA MHSLIFGMLK         SARS-COV2       KUDKAVKIJAKNAYEKDKAVARKLERMADOAMISVKAARAVKARAVKAANOTMESKUVSA MHSLIFGMLK         SARS-COV2       KUDKSVDTILKKEPDREAAMORKLEKMADOAMITOKYKOAABEDKKARVYSAMOTMETTILEGULK         SARS-COV2       KUDKSSVDTILOARNOKVEPDSVIDSVIDANOAMITOKYKOAABEDKARAVYSAMOTMETTILE         YEV       KLDMNSSVDTILNKANOVVEPDSVIDSVIDANOAMITOKYKOAABEDKARAVYSAMOTMETTILE         HCOV222E       RDMSSVDTILNKANOVVEPDSVIDSVIDANOAMITOKYKOAABEDKRAVYSAMOTMETTILE         YEV       KLDMSSVDTILNKANOVVEPDSVIDSVIDANOAMITOKYKOAABEDKARAVTSAMOTMETTILKAANOTMETTILKAAANOANEVEPDSVIKANOVVEPDSVIDAVAARAVKAAANOTMETSKINTSAMOTMETTILKAANOANEVEPDSVIDANOANEVERSKINTYSAMOTMETTILKAANOANEVEPSSANOANEVENTIL         YEV       KLDMSSVDTILNIKKOVEPDSVIDSVIDAVAARAVKAANOANEVENDENVERSKINTYSAMOTMETTILKAANOANEVENDENVENTILKAANOANEVENDENVENTILKAANOANEVENDENVENTILKAANOANEVENDENVENTILKAANOANEVENDENVENTILKAANOANOANEVENTILKAANOANEVENTILKAANOANOANEVENTILKAANOANEVENTILKAANOANEVENTILKAANOANEVENTILKAANOANEVENTILKA	37       47       57       67       77       87         FIPY       CID TKACNIAKSEFEREASVOKI DKVAECAAAS VXKARANDRKSKIVSAMHSLIFGMIK         HCW-220       CIKAAFNIAKSEFEREASVOKI DKVAECAAASVXKARANDRKSKIVSAMHSLIFGMIK         HCW-20       CIKAAFNIAKSEFEREASVOKI DKVAECAAASVXKARANDRKSKIVSAMHSLIFGMIK         HCW-20       CIKAAFNIAKSEFEREASVOKI DKVAECAAASVXKARANDKASKIVSAMHSLIFGMIK         HCW-20       CIKAAFNIAKSEFEREASTORSIDAKIDKVECAAASVXKARANTKEKKIVSAMHSLIFGMIK         HCW-20       CIKAAFNIAKSEFEREASTORSIDAKIDKVECAAASVXKARANTKEKKIVSAMHSLIFGMIK         HCKS-COV       CIKHAANVIASEFEREASTORSIDAKAAVKEKAANTKEKKIVSAMHSLIFGMIK         MERS-COV       KUKSINVASEFEREASTORSIDAAMORKILEKMADOAMTSVKOARAENKKAKIVSAMCTMIFTULR         SARS-COV       KUKSINVASEFEREASTORSIDAAMORKILEKMADOAMTONKOARSEDKKAKIVSAMCTMIFTULR         SARS-COV       KUKSINVAKSEFEREASTORSULEKMADOAMTONKOARSEDKKAKIVSAMCTMIFTULR         SARS-COV       KUKSINVAKSEFEREASTORSULEKMADOAMTONKOARSEDKKAKIVSAMCTMIFTULR         PEDV       97       107       117       127       147         HCW-2222       KUKSINVAKSEFEREASTORSULEKMADOAMTONKOARSEDKKAKIVSAMCVINFERMITULR       100 XKASINTTILDARKOVUNTULASTONYKAASTONYKAASTONYKANOTKAKIVSANOTMIFTULR       100 XKASINTTILDARKOVUNTULASTONYKAASTONYKAASTONYKAGANOTMIFTULR         HCW-2222       KUDMSSVDTILNMAROOANSVILLEKMADOATKUVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVV											
37       47       57       67       77       87         FIPU       OLTIKACNITAKSEFEREASVOKKIDKME OAAASMYKE ARAVDRKSKIVSAHSLIFGULK         TGEV       OLTIKACNITAKSEFEREASVOKKIDKME OAAASMYKE ARAVDRKSKIVSAHSLIFGULK         HC0V222E       OLKKAPNIAKSEFEREASVOKKIDKME OAAASMYKE ARAVDRKSKIVSAMHSLIFGULK         PEDV       OLKKAPNIAKSEFEREASVOKKIDKME OAAASMYKE ARAVDRKSKIVSAMHSLIFGULK         PEDV       OLKKAPNIAKSEFEREASVOKKIDKME OAAASMYKE ARAVDRKSKIVSAMHSLIFGULK         MERS-COV       ALOKAVNIAKSEFEREASVOKKIDRME OAAASMYKE ARAVDRKSKIVSAMHSLIFGULK         SARS-COV2       KLKKSLNVAKSEFEREASVOKKIDRME OAAASMYKE ARAVDRKSKIVSAMHSLIFGULK         SARS-COV2       KLKSLNVAKSEFEREASVOKKIDAMORALERME OAAASMYKE ARAVDRKSKIVSAM SKUVSAM SLIFGULK         SARS-COV2       KLKSLNVAKSEFEREASVOKKIDAMORALERME OAAMTSMYKORARAEDKAKIVSAM OTMLETKILT         YACKANDAKASEFEREASVOKKIDAMORALERME OAAATSMYKORARAEDKAKIVSAM OTMLETKILT       SARS-COV2         YACKANDAKASEFEREASVOKKIDAMORALERME OAAATSMYKORARAEDKAKIVSAM SKUVSAM SLIFGULK       SARS-COV2         YACKANDAKASEFEREASVOKKIDAMORATION ON TOWN OA SUNTALITYTELT       SARS-COV2         YACKANDAKASEFEREASVENKING OAMTROWN ON TOWN SONT SUNTACONSTANCE ON THE TALK         YACKANDAKASEFEREASVENKING ON THE TALKAND OAMTROWN ON TOWN ON THE TALK         YACKANTAKASEFEREASVENKING       SANSOVY SANSONT THE TALKANDELNITY THE SITCERT         YACKANTTI TOOT TOT TOT TOT TOT TOWN OA SUNTALONG ON THE TALKAND THE TALKANDY TH	37       47       57       67       77       87         FIPV         TGEV       OLT KACNIANSEFFRASVOKI DKMAE OAAASMYK ARAVDRKSKIVSAMHSLIFGMLK         HOW 2000         UKAANVAR SEFFRASVOKI DKMAE OAAASMYK ARAVDRKSKIVSAMHSLIFGMLK         HOW 2000         OLT KACNIANSEFFRASVOKI DKMAE OAAASMYK ARAVDRKSKIVSAMHSLIFGMLK         HOW 2000         OLT KACNIANSEFFRASVOKI DKMAE OAAASMYK ARAVDRKSKIVSAMHSLIFGMLK         HOW 2000         OLT KACNIANSEFFRASVOKI DKMAE OAAASMYK ARAVDRKSKIVSAMHSLIFGMLK         MENNE VERSKIVSA         SARSCOV         VENKANNYK SEFDREASVOKI DKMAE OAAASMYK ARAVDRKKIVSAMHSLIFGMLK         SARSCOV         VENKAVNYK SEFDREASVOKI DKMAE OAAASMYK ARAVNKKSKVVSAMHSLIFGMLK         SARSCOV         VENKAVNYK SEFDREASVOKI DKMAE OAAASMYK ARAVNKKSKVVSAMHSLIFGMLK         SARSCOV         VENKAVNYK SEFDREASVOKI DKMAR VERMOR         PON KUK KAVNYK SEFDREASVOKU DKMAR VERMOR         VENKAVNYK KSENDVNETUK         VENKAVNYK KSENDVNETUK         VENKAVNYK KSENDVNETUK         VENKAVNYK KSENDVNENNEV <td cols<="" th=""><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th></td>	<th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>										
Image: Structure of the st	FIV       1/1       1/2       1	ſ	27	47	57	67	77	07				
FIPV       01 T KAP TT TAKE DE BARAGUOKKLDK ME CAAAS MYKE RRAVDRAKES T USAM HS LD CALK         HCOV229E       01 KKAP TT TAKE DE BARAGUOKKLDK ME CAAAS MYKE RRAVDRAKES T USAM HS LD CALK         HCOV229E       01 KKAP TT TAKE DE BARAGUOKKLDK ME CAAAS MYKE RRAVDRAKS T USAM HS LD CALK         MERS-COV       ALQKAMU LAKNA ZED RE SIVOKKI DRME CAAAS MYKE RRAVNRKSKUVSAM HS LD CALK         MERS-COV       ALQKAMU LAKNA YEKD KAVA RKLERMAD OAM TSMYKOAR AED KKAKIVSAM OTMLFGUIK         SARS-COV       KIKKSLNVAKSEPD RD AAMORKLERMAD OAM TSMYKOAR AED KKAKIVSAM OTMLFGUIK         SARS-COV2       KIKKSLNVAKSEPD RD AAMORKLERMAD OAM TOM YKOAR SED KRAKVTSAM OTMLFTMLR         SARS-COV2       KIKKSLNVAKSEPD RD AAMORKLERMAD OAM TOM YKOAR SED KRAKVTSAM OTMLFTMLR         TGEV       KIKKSLNVAKSEPD RD AAMORKLERMAD OAM TOM YKOAR SED KRAKVTSAM OTMLFTMLR         TGEV       KIKSUNT HIE QARNEVLED IS IIPAASATRI VUVIP NLEVIS KVROENNVH YAGATIN TIL         PEDV       RIDMSSVDTTILNMARNEVLED IS IIPAASATRI VUVIP DHESKIK KMVDENVH YAGATIN TIL         PEDV       RLDMSSVDTTILNMARNEVLED IS VIPAITSAARIVVVVVVD DHESKVKMMU SECVHAGUTUNT         PEDV       RLDMSSVDTTILNMARNEVLESTINELAKKIRVVVID SITCERTINET         SARS-COV2       KIDNDLENG KUVHIKEVTAANE LINIT WE SITCERTINET         SARS-COV2       KIDNDLANG HYHIKEVTAANE LINIT WE SITCERTINET         SARS-COV2       KIDNDLENG KUVHIKEVTAANE LINIT WE SITCERTINET         SARS-COV2       KIDND ALNNI INNARD COVE IS VIPAUSA	Import       OIT TKAPT IT NS DE BERAA'S VOKKLDKING CAAAS MYKE ARAVDRKISKI VOKM HS LIFGGULK         HGOV229E       OIT KKAPT IT NS DE BERAA'S VOKKLDKING CAAAS MYKE ARAVDRKISKI VOKM HS LIFGGULK         HGOV229E       OIT KKAPT IT NS DE BERAA'S VOKKLDKING CAAAS MYKE ARAVDRKISKI VOKM HS LIFGGULK         MERS-COV       ALOKAMU KABE FDRESSUVOKKI NRMAE CAAAAS MYKE ARAVDRKISKI VOKM HS LIFGGULK         MERS-COV       ALOKAMU JAKNA YEKD KAVARKI LERMAD CAMT SMYKE ARAVNRKISKUV SAM OTMLFGUIK         SARS-COV       KIKKSLNVAKSE FDRDAAMORKLEKMAD CAMT OM YKOAR SED KRAKUT SAM OTMLFTMLR         SARS-COV2       KIKKSLNVAKSE FDRDAAMORKLEKMAD CAMT OM YKOAR SED KRAKUT SAM OTMLFTMLR         SARS-COV2       KIKKSLNVAKSE FDRDAAMORKLEKMAD CAMT OM YKOAR SED KRAKUT SAM OTMLFTMLR         FIPV       KLKSLNVAKSE FDRDAAMORKLEKMAD CAMT OM YKOAR SED KRAKUT SAM OTMLFTMLR         SARS-COV2       KLDMS SUNT TI DOARNG VIE JES II PAAAS ATRE I VUVTP NLE VLSKVR QENNYH YAGA IN TI         HCOV228       RIDMS SUNT TI DOARNG VIE JES II PAAAS ATRE I VUVTP DIE VLSKVR QENNYH YAGA IN TI         HCOV228       RIDMS SUNT TI LMARRO CVIE JES II PAAAS ATRE I VUVTP DIE SEVKENNYH YAGA IN TI         HCOV28       RIDMS SUNT TI LMARRO CVIE JES II PAAS ATRE I VUVTP DIE VUN VY AGAY NU GEV YAGY AGAI NI TI         HCOV28       RIDMS SUNT TI LMARRO CTIE JES I PAAS ATRE I VUVIP DIS VIRAUQRE CVIA YAGY AGAI NI TI         HCOV28       RIDMS SUNT TI LMARRO CTIE JES I PAAS ATRE I VUVIP DIS VIRAUQRE CVIA YAGY AGAI NI TI         HCOV28<								CMIK			
HCov229E OF KKAMNVAKAEPDRESSVOKKINRMAEQAAAAWYKEARAVNRKSKVVSAMHSLEDGULR PEDV OF HAMNVAKSEPDRESSTORKLORMAEQAAAOWYKEARAVNRKSKVVSAMHSLEDGULR MERSCOV ALOKAVNIAKNAYEKDKAVARKLERMADQAMTSMYKQARARVNRKSKVVSAMHSLEDGULR SARSCOV2 KUKSLNVAKSEPDRDAAMORKLEKMADQAMTOMYKQARSEDDKRAKVTSAMOTMEPTMLR SARSCOV2 KUKSLNVAKSEPDRDAAMORKLEKMADQAMTOMYKQARSEDDKRAKVTSAMOTMEPTMLR SARSCOV2 KUKSLNVAKSEPDRDAAMORKLEKMADQAMTOMYKQARSEDKRAKVTSAMOTMEPTMLR SARSCOV2 KUKSSVNTIIEQARNOVLPISIIPAASATRIJVVITPNLEVLSKVRQENNVHYAGAIMSI GEV KDMSSVNTIIEQARNOVLPISIIPAASATRIJVVITPSLEVESKIRQENNVHYAGAIMTI HCOV239E RLDMSSVDTILNDARNOVLPISIIPAASATRIJVVITPSLEVESKIRQENNVHYAGAIMTI HCOV239E RLDMSSVDTILNIARNOVLPISIIPAASATRIJVVITPSLEVESKIRQENNVHYAGAIMTI MERSCOV KLDNDSVDTILNIARNOCUPISVIPAARATVVVVPDHDSVTMMVDGEVHYAGIMTI SARSCOV2 KLDNDVLNGIISNARNOCIPISVIPAUAARTKINVVIPDFTVMNOVVTYPSLNYAGAIMDI SARSCOV2 KLDNDALNNIINNARDGCVPINIIPLITTAAKIMVVVPDYTYSLNYAGAIMDI SARSCOV2 KLDNDALNNIINNARDGCVPINIIPLITCERTTKLC PEDV IDLKDNDGKVVHVKEVTAANELNIIWPJSITCERTTKLC PEDV IDLKDNDGKVVHVKEVTAANELNIIWPJSITCERTTKLC BSII(3ub0) HCOV239E QEVKDANGSVHVKEVTAANELNIIWPJSITCERTVKLC BSII(T1) MERSCOV OVVDAJSKIVULKEVTAANELNIIWPJVLCCRASTSAVKLC SARSCOV QOVVDAJSKIVQLSEINMDNSPNLAWPILVTALRANS.AVKLC BSII(T2)	HCov222E       OLKKAMNUVKAEFDRESUVOKIINRMAEOAAAOMYKEARAVNRKISTVUOAMESLIEGMLE PEDV         MERS-Cov       ALOKAVNIAKNAYEKDKAVARKLERMADOAMTOMYKEARAVNRKISTVUSAMUSLIEGMLE ALOKAVNIAKNAYEKDKAVARKLERMADOAMTOMYKOARSEDKKAKUVSAMOTMLEGNIK SARS-Cov2         SARS-Cov2       KIKKSLNVAKSEFDRDAAAORKLEKMADOAMTOMYKOARSEDKRAKVTSAMOTMLEGNIK KIKKSLNVAKSEFDRDAAAORKLEKMADOAMTOMYKOARSEDKRAKVTSAMOTMLETMLR KIKKSLNVAKSEFDRDAAAORKLEKMADOAMTOMYKOARSEDKRAKVTSAMOTMLETMLR KIKKSLNVAKSEFDRDAAMORKLEKMADOAMTOMYKOARSEDKRAKVTSAMOTMLETMLR KIKKSLNVAKSEFDRDAAMORKLEKMADOAMTOMYKOARSEDKRAKVTSAMOTMLETMLR KIKKSLNVAKSEFDRDAAMORKLEKMADOAMTOMYKOARSEDKRAKVTSAMOTMLETMLR KIKKSLNVAKSEFDRDAAMORKLEKMADOAMTOMYKOARSEDKRAKVTSAMOTMLETMLR KIKSLNVAKSEFDRDAAMORKLEKMADOAMTOMYKOARSEDKRAKVTSAMOTMLETMLR KIKSLNVAKSEFDRDAAMORKLEKMADOAMTOMYKOARSEDKRAKVTSAMOTMLETMLR HCOV228         PEDV       RIDMSSVDTTI LOQARNGVEJSITPAISAARIVVVVVDEDSSVKIRMVDGVVYAGVVTI SARSCOV         KIDNDKSVDTI LNARRGUVEJSVIPAVSARKIVNVIPDSTVNRIQREGCVHYAGTINNI MERSCOV       KIDNDKSVDTI LNARRGUVEJSVIPAVSARKIVVIPDSTVNRIQREGCVHYAGTINNI SARSCOV2         KIDNDKSVDTI LNARRDCUVEJSVIPAVSARKIMVVIPDVNTYNOVVVEDSLNVAGALWDI SARSCOV2       KIDNDALNNI INNARDCUVEJSVIPAVSARKIMVVVEDVVVVEDSLNVAGALWDI SARSCOV2         KIDNDALNNI INNARDCUVEJSVIPAVSARKIMVVVEDVTYKNOUCGTPTYASALMEI SARSCOV2       KIDNDALNNI INNARDCUPINI IPLITAAKIMVVVEDVTYKNOUGNTETYASALMEI BSII (3ub0)         BSII (11)       BSII (12)       BSII (12)         MERSCOV       VUVNDANSHVIKKSSUVYAQNESISSISVIPALAWELVUTALRANS.AVKLO         SARSCOV2       QUVVDADSKIVUSSUV.VDSNENITWEVTALRANS.AVKLO	TGEV	QLTKAF	NIAKSEFERE NIAKSDFERE	ASVQKKLDK	IAEQAAASMY	KEARAVDRKS KEARAVDRKS	KIVSAMHSLLF	GMLK			
MERS-COV       ALOKAVNITAKNAYEKD KAVARKLERMAD OAMT SMYKOARAEDKKAKIVSANOTMEGOIK         SARS-COV       KLKKSINVAKSEFDRDAAMORKLEKMAD OAMT OMYKOARSEDKRAKVTSANOTMETMLR         SARS-COV2       KLKSINVAKSEFDRDAAMORKLEKMAD OAMT OMYKOARSEDKRAKVTSANOTMETMLR         SARS-COV2       KLKSINVAKSEFDRDAAMORKLEKMAD OAMT OMYKOARSEDKRAKVTSANOTMETMLR         SARS-COV2       KLKSINVAKSEFDRDAAMORKLEKMAD OAMT OMYKOARSEDKRAKVTSANOTMETMLR         SARS-COV2       KLMSSUNT IEQARNEVLELKMAD OAMT OMYKOARSEDKRAKVTSANOTMETMLR         FIPV       KLDMSSUNT IEQARNEVLELKMAD OAMT OMYKOARSEDKRAKVTSANOTMETMLR         YOT       107       117       127       137       147         FIPV       KLDMSSUNT IEQARNEVLELKMAD OAMT OMYKOARSEDKRAKVTSANOTMETMLR       147       147         YEEV       KLDMSSUNT IEQARNEVLESITENAASATREVUTTENSLEVENSKIRKUTSANOTMETAGAINSTI         PEV       REDMSSUDT LINLAKDEVIPLSITIPAUSAATREVUTTENSLEVENSKIRQENNUHYAGAINTI         MERS-COV       KLDMSSUDT LINLAKDEVUPLSVIPAUSAATREVUTENSLIKERVUNDUVTENSLIKAGAINTI         MERS-COV       KLDMSSUDT LINLAKDEVIPLSVIPAUSAATREVUTENSLIKERVUNDUVTENSLEGENNUHYAGAINTI         MERS-COV       KLDMSSUDT KLNKEVTAANELNITWPLSVIPLSTICERT       TKLO         SARS-COV2       KLDMSSUDTENHKEVTAANELNITWPLSTICERT       TKLO         SARS-COV2       KLDNDALNNINNARDECUPINTENDESTERT       SARSCOV       SSII (11)         MERS-COV       UVKDANGS	MERS-Cov       ALOKAVNIAXNAYEKDKAVARKLERMADOAMT SVYKOARAEDKKAR IVSAMOTMLEGUTK         SARS-Cov       EVKKSLNVAKSEFDRDAAMORKLEKMADOAMT OMYKOARSEDKRARVTSAMOTMLETTLE         SARS-Cov2       EVKKSLNVAKSEFDRDAAMORKLEKMADOAMT OMYKOARSEDKRARVTSAMOTMLETTLE         SARS-Cov2       EVKSSLNVAKSEFDRDAAMORKLEKMADOAMT OMYKOARSEDKRARVTSAMOTMLETTLE         SARS-Cov2       EVKSSLNVAKSEFDRDAAMORKLEKMADOAMT OMYKOARSEDKRARVTSAMOTMLETTLE         SARS-Cov2       EVKSSLNVAKSEFDRDAAMORKLEKMADOAMT OMYKOARSEDKRARVTSAMOTMLETTLE         FIPV       KLDMSSVNTILEOARNOVLPLSITERATOMYKOARSEDKRARVTSAMOTMLETTLE         HCOV239E       RLDMSSVNTILEOARNOVLPLSITERATOMYKOARSEKUVVPDHDSFVKMMVDGFVHXAGAINTT         HCOV239E       RLDMSSVDTILNIAROVVPLSVTPATSAARTVVVVPDHDSFVKMMVDGFVHXAGVTTL         MERS-Cov       KLDNDSVDTILNIAROVPLSVTPAVSAARKUNVVVPDJSDISSKNKTCOGTTFVASALVET         MERS-Cov       KLDNDALNNINNARDGCVPLSVTPAVSARKUNVVPDYDYSLSVTCOGNTETVASALVET         SARS-Cov2       KLDNDALNNINNARDGCVPLNTTELINITWPLSTTCERT         SARS-Cov2       KLDNDGKNVHLKEVTAONAESLSWEJVGCERT         FIPV       VEVKDANGSHVHLKEVTAONAESLSWEJVGCERT         SARS-Cov2       CUVKDNDGKNVHVKKSDV.VDSNENLINTWPLSTCERT         SARS-Cov2       CUVKDANGSHVHKKEVTAONAESLSWEJUGCERT         SARS-Cov2       CUVVDADSKIVUVKSDV.VDSNENLAWPLIVTALRANS.AVKLO         SARS-Cov2       CUVVDADSKIVUVSETNONSPNLAWPLIVTALRANS.AVKLO	HCoV-229E	QLKKAM OTB <b>HAM</b>	NVAKAEFDRE NVAKSEFDRE	SSVQKKINRM ASTORKIDBM	IAEQAAAAMYI IAEQAAAAMYI	KEARAVNRKS KEARAVNRKS	KVVSAMHSLLF KVVSAMHSLLF	GMLR			
SARS-CoV SARS-CoV-2 KLKKSLNVAKSEFDRDAAMORKLEKMADOAMTOUVNOARSEDKRAKVTSAMOTMETMLR SARS-CoV-2 KLKKSLNVAKSEFDRDAAMORKLEKMADOAMTOUVNOARSEDKRAKVTSAMOTMETMLR FIPV TGEV KLDMSSVNTIIEQARNGVLPISIIPAASATREIVVTRPNLEVUSKVRQENNVHYAGAINST HCOV-229E RLDMSSVNTIIEQARNGVLPISIIPAASATREIVVTPNLEVUSKVRQENNVHYAGAINTI PCOV KLDMSSVDTILNMARNGVVPISVIPATSAARLVVVTPSSKIRQENNVHYAGAINTI PCOV RLDMSSVDTILNMARNGVVPISVIPATSAARLVVVTPSKIRQENNVHYAGAINTI MERS-CoV KLDNDVLNGIISNARNGCIPISVIPATSAARLVVVTPDTVKNTQCFVHYAGAINTI SARS-CoV KLDNDVLNGIISNARNGCIPISVIPATSAARLVVVTPDTVMNQVVTYSLNYAGAINTI SARS-CoV KLDNDVLNGIISNARNGCIPINIIPLTTAAKLMVVIPDTVMNQVTYSLNYAGAINDI SARS-COV-2 KLDNDALNNIINNARDGCVPINIIPLTTAAKLMVVVPDYGTYKNTCOGNTFTVASALWEI SARS-COV-2 KLDNDALNNIINNARDGCVPINIIPLTTAAKLMVVVPDYGTYKNTCOGNTFTVASALWEI BSII(3ub0) HCOV-229E PEDV NERS-COV QUVVDAGSNVHLKEVTAANELNITWPISITCERTIVKLO PEDV IDIKDNDGKVVHLKEVTAANELNITWPISITCERTIVKLO BSII(11) BSII(11) BSII(12)	SARS-Cov2 KLKSLNVANSEFDRDAAMORKLEKMADOAMTOWYKARSEDKRAKVTSAMOTMETTULK SARS-Cov2 KLKSLNVANSEFDRDAAMORKLEKMADOAMTOWYKARSEDKRAKVTSAMOTMETTULK 97 107 117 127 137 147 FIPV KLDMSSVNTTEOARN GVLPLSIEMDAAMORKLEKMADOAMTOWYKARSEDKRAKVTSAMOTMETTULK FIPV KLDMSSVNTTEOARN GVLPLSIEPAASATREIVVYCARSEDKRAKVTSAMOTMETTULK HCOV229E RLDMSSVDTTEOARN GVLPLSIEPAASATREIVVYTPNLEVLSKVRQENNVHYAGAINSE HCOV229E RLDMSSVDTTLNARNGGVPLSVEPATSAARTVVVPDHDSFVKMMVDGFVHYAGAINSE MERS-CoV KLDNSVDTILNIARNGGVPLSVEPATSAARTVVVPDHDSFVKMVDGFGVHYAGVTTUL MERS-COV KLDNDKULNIKGVOPLSVEPATSAARTVVVPDHDSFVKMVDGFGVHYAGAINTE SARS-COV2 KLDNDALNNIINNARDGCVPLNIEPLCASNKERVVIPDFTVWNQVVTYPSLNYAGAINTE SARS-COV2 KLDNDALNNIINNARDGCVPLNIEPLTAAKIMVVVPDYSICOGNTFTYASALWEI SARS-COV2 KLDNDALNNIINNARDGCVPLNIEPLTAAKIMVVVPDYGTYKNTCDGNTFTYASALWEI MERS-COV2 KLDNDALNNIINNARDGCVPLNIEPLTAAKIMVVVPDYGTYKNTCDGNTFTYASALWEI SARS-COV2 KLDNDALNNIINNARDGCVPINIEPLTAAKIMVVVPDYGTYKNTCDGNTFTYASALWEI SARS-COV2 KLDNDALNNIINNARDGCVPINIEPLTAAKIMVVVPDYGTYKNTCDGNTFTYASALWEI SARS-COV2 KLDNDALNNIINNARDGCVPINIEPLTAAKIMVVVPDYGTYKNTCDGNTFTYASALWEI SARS-COV2 QQVVDADSKIVOLSEISMDNSPNLAWPIIVTALRANS.AVKLO BS II (3ub0) BS II (11) BS II (12)	MERS-CoV	ÃLQ <b>kav</b> i	IAKNAYEKD	KAVARKLER	IADQAMT SMY I	KQ <b>ARA</b> E <b>DKK</b> A	KIVSAMQTMLF	GMIK			
97       107       117       127       137       147         FIPV       KLDMSSVNTIIEQARNGVLPISIIPAASATRLIVVITPNLEVLSKVRQENNVHYAGAIWSI         HCov229E       RLDMSSVDTILNQARNGVVPISIIPAASATRLVVITPNLEVISKVRQENNVHYAGAIWSI         PEDV       RLDMSSVDTILNARNGVVPISVIPATSAARLVVVIPDHDSVIFSKIRQENNVHYAGAIWSI         MERS-Cov       KLDMSSVDTILNARNGVVPISVIPATSAARLVVVIPDHDSVIPAVSATKLNIVTSDIDSYNRIQREGCVHYAGTIWNI         MERS-Cov       KLDNDVLNGIINNARNGCVPISVIPATSAARLVVVIPDHTVWNQVVTYSLNXAGAIWDI         SARS-Cov       KLDNDVLNGIINNARDGCVPINIPLTIAAKLMVVIPDYTVKNTCOGTFTYASALWEI         SARS-Cov       KLDNDALNNIINNARDGCVPINIPLTTAAKLMVVVPDYGTYKNTCOGNTFTYASALWEI         SARS-Cov       KLDNDALNNIINNARDGCVPINIPLTAAKLMVVVPDYGTYKNTCOGNTFTYASALWEI         SARS-Cov       KLDNDGKVHIKEVTAANELNITWPISITCERT         HCov229E       QEVKDNDGKNVHIKEVTAANELNITWPISITCERT         PEDV       DIKDNDGKNVHIKDVTKENQEILVWIDIGCERT         MERS-Cov       TVINNVDNEIVKEVTAQNAESLSWPIVLGCERT         MERS-Cov       TVINNVDNEIVKSDV, VDSNENLIWPIVLGCERTIVKLQ         BS II (T1)       BS II (T2)         SARS-Cov2       QQVVDADSKIVCLSEINMDNSPNLAWPIIVTALRANS.AVKLQ	97       107       117       127       137       147         FIPV       KLDMSSVNTILEQARNGVLPISIIPAASATRIVVITPNLEVISKVRQENNVHVAGAINST         HCov229E       RLDMSSVDTILNQARNGVLPISVIPATSAARLVVVVPDHDSFVKMMVDGFVHVAGAINTT         PEDV       RLDMSSVDTILNLAKDGVVPISVIPATSAARLVVVVPDHDSFVKMMVDGFVHVAGAINTT         MERS-Cov       KLDNDVLNGTISNARNGCIPUSVIPATSAARLVVVVPDHDSFVKMMVDGFVHVAGVVTT         SARS-Cov       RLDMSQVDTILNLAKDGVVPISVIPATSAARLVVVVPDHDSFVKMVDGFVHVAGVVTT         SARS-Cov       RLDNDVLNGTISNARNGCIPUSVIPATSAARLVVVVPDHDSFVKMVDGFVHVAGVVTT         SARS-Cov       KLDNDVLNTSINNARNGCIPUSVIPATSAARLVVVVPDHVSVPDVGTYKNTCOGTTFTVASALWET         SARS-Cov       KLDNDALNNTINNARDGCVPINTIPLITAAKLMVVVPDYGTYKNTCOGNTETVASALWET         FIPV       VEVKDANGSEVHLKEVTAANELNITWPISTCERT         YEVKDANGSEVHLKEVTAANELNITWPISTCERT       TKLQ         PEDV       IDTKDNDGKVVHVKEVTAQNAESLSWPIVLGCERT         PEDV       IDTKDNDGKVVHVKEVTAQNAESLSWPIVLGCERT         MERS-Cov       QQVVDADSKTVQLSEISMDNSPNLAWPILTVTALRANS.AVKLQ         SARS-Cov2       QQVVDADSKTVQLSEINMDNSPNLAWPILVTALRANS.AVKLQ	SARS-CoV SARS-CoV-2	KLKKSL KLKKSL	NVAKSEFDRD NVAKSEFDRD	AAMORKLEKM AAMORKLEKM	IADQAMTQMY IADQAMTQMY	KQARSEDKRA Koarsedkra	KVTSAMQTMLF KVTSAMQTMLF	TMLR TMLR			
97       107       117       127       137       147         FIPV       KLDMSSVNTIIEQARNGVLPLSIIPAASATRLIVVITPNLEVLSKVRQENNVHYAGAIWSTI         HCov-229E       RLDMSSVNTIIEQARNGVVPLSIIPAASATRLVVVTPNLEVLSKVRQENNVHYAGAIWTI         PEDV       RLDMSSVDTILNMARNGVVPLSVIPATSAARLVVVVPDHDSFVKMMVDGFVHYAGAIWTI         MERS-Cov       KLDNDVLNGTISNARNGCIPLSVIPATSAARLVVVVPDHDSFVKMMVDGVVYAGAIWTI         MERS-Cov       KLDNDVLNGTISNARNGCIPLSVIPAVSATKLNVTSDIDSVNRIQREGCYHYAGTIWNI         MERS-Cov       KLDNDVLNGTISNARNGCIPLSVIPAVSATKLNVVIPDFTVWNQVVTPSLNYAGAIWDI         SARS-Cov-2       KLDNDALNNIINNARDGCVPINIELTSAKKLMVVVPDYTTYSLOGNTFTTYASALWEI         SARS-Cov-2       KLDNDALNNIINNARDGCVPINIELTTAAKLMVVVPDYTTYKNTCDGNTFTYASALWEI         MERS-Cov       KLDNDALNNIINNARDGCVPINIELTTAAKLMVVVVPDYGTYKNTCDGNTFTYASALWEI         SARS-Cov-2       QEVKDANGAQVHLKEVTAANELNITWPISITCERTTKLO         PEDV       DIKKDNGKVVHVKEVTAANELNITWPISITCERTTKLO         PEDV       DIKKDNGKVVHVKEVTAANELNITWPISITCERTTKLO         MERS-Cov-2       QEVKDNDGKNVHLKEVTAQNAESLSWPIVUGCERTVKLO         BSII(3ub0)       BSII(T1)         MERS-Cov       TVINNVDNEKVKSDV.VVSNENDENTWPIVEGERTVKLO         SARS-Cov-2       QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLO         SARS-Cov-2       QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLO	97       107       117       127       137       147         FIPV       KLDMSSVNTTIEOARNGVLPISIIPAASATRLIVVITPNLEVLSKVRQENNVHYAGAIWSI         HCov-229E       RLDMSSVDTILLNOARNGVVPISVIPATSAARLVVVVPDHDSFVKMMVDGFVHYAGAIWTI         PEDV       RLDMSSVDTILLNAARNGVVPISVIPATSAARLVVVVPDHDSFVKMMVDGFVHYAGAIWTI         MERS-CoV       KLDNDVLNGIISNARNGCIPISVIPAVSATKLNIVTSDIDSYNRIQREGCVHYAGTIWNI         MERS-CoV       KLDNDVLNGIISNARNGCIPISVIPAVSATKLNIVTSDIDSYNRIQREGCVHYAGTIWNI         SARS-CoV       KLDNDALNNIINNARDGCVPINIPLICASNKLRVVIPDYNTYKNTCDGTTFTYASALWEI         SARS-CoV-2       KLDNDALNNIINNARDGCVPINIPLITAAKLMVVIPDYNTYKNTCDGNTFTYASALWEI         FIPV       VEVKDANGSAVHLKEVTAANELNITWPISITCERTTKLO         MERS-CoV-2       KLDNDALNNIINNARDGCVPINIPLITCERTTKLO         BS II (3ub0)       BS II (3ub0)         HCov-229E       QEVKDDNGGRVHLKEVTAANELNITWPISITCERTTKLO         PEDV       IDIKDNDGKVVHVKEVTAONAESLSWPIVLGCERIVKLO         BS II (T1)       BS II (T1)         MERS-CoV       QVVDADSKIVQLSEISMDSPNLAWPIIVTALRANS.AVKLO         SARS-CoV-2       QVVDADSKIVQLSEISMDNSPNLAWPIIVTALRANS.AVKLO	0/11/0 001/2						~				
97 107 117 127 137 147 FIPV KLDMSSVNTIIEQARNGVLPISTIPAASATRLIVVTPNLEVLSKVRQENNVHYAGAIWST HCoV-229E RLDMSSVNTIIEQARNGVLPISTIPAASATRLIVVTPNLEVLSKVRQENNVHYAGAIWST HCoV-229E RLDMSSVDTILNMARNGVVPISVIPATSAARLVVVVPDHDSFVKMMVDGFVHYAGAIWST MERS-CoV KLDNDVLNGTISNARNGCIPISVIPAVSATKLNIVTSDIDSYNRIQEGCVHYAGIWNT MERS-CoV KLDNDVLNGTISNARNGCIPISVIPLCASNKLRVVIPDFTVWNQVVTYPSINYAGALWDT SARS-CoV KLDNDALNNIINARDGCVPINIIPLTTAAKLMVVVPDYSTVKNCCDGTTFTYASALWEI SARS-CoV KLDNDALNNIINNARDGCVPINIIPLTTAAKLMVVVPDYGTYKNTCDGTTFTYASALWEI SARS-CoV KLDNDALNNIINNARDGCVPINIIPLTTAAKLMVVVVPDYGTYKNTCDGNTFTYASALWEI SARS-COV KLDNDGKVYHLKEVTAANELNITWPISITCERTTKLO PEDV IDIKDNDGKNVHLKEVTAANELNITWPISITCERTVKLO SARS-COV QQVVDADSKIVQLSEINMDNSPNLAWPIIVTALRANS.AVKLO SARS-COV QQVVDADSKIVQLSEINMDNSPNLAWPIIVTALRANS.AVKLO	97       107       117       127       137       147         FIPV       KLDMSSVNTIIEQARNGVLPLSIIPAASATRLIVVIPNLEVLSKVRQENNVHYAGAIWST         HCOV-229E       RLDMSSVNTIIEQARNGVVPLSVIPASAARLVVVVPDHDSFVKMMVDGEVNYAGAIWTI         PEDV       RLDMSSVDTILNARNGVVPLSVIPATSAARLVVVVPDHDSFVKMMVDGEVNYAGAIWTI         MERS-CoV       KLDNDVLNGFISNARNGCVPLSVIPATSAARLVVVVPDHDSFVKMMVDGEVNYAGAIWTI         MERS-CoV       KLDNDVLNGFISNARNGCVPLSVIPATSAARLVVVVPDHDSFVKMMVDGFVHYAGAIWTI         MERS-CoV       KLDNDVLNGFISNARNGCVPLSVIPAVSATKLNIVTSDIDSMRTQREGCVHYAGTIWNI         MERS-CoV       KLDNDVLNGTISNARNGCVPLSVIPAVSATKLNVVIPDFVKMNVVTPDSLNYAGAIWDI         SARS-CoV2       KLDNDALNNIINNARDCVPLWEVTAANELNITWPLSTCERTTKLO         FIPV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLO         MERS-COV2       KLDNDALNNIINNARDCVPLWEVTAANELNITWPLSITCERTTKLO         SARS-COV2       KLDNDGKVHLKEVTAANELNITWPLSITCERTTKLO         PEDV       IDIKDNDGKVHVKEVTAANELNITWPLSITCERTTKLO         PEDV       IDIKDNDGKVHVKEVTAANELNITWPLSITCERTTKLO         BS II (3ub0)       BS II (11)         MERS-COV       VEVKDANGSHVHIKEVTAANELNITWPLSITCERTVKLO         PEDV       IDIKDNDGKVHVKEVTAANELNITWPLSITCERTVKLO         BS II (T1)       BS II (T1)         MERS-COV       VVNDDGKVHVKEVTAANELSENTENTWPLIVTALRANS.AVKLO <td< th=""><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th></td<>											
97       107       117       127       137       147         FIPV         TGEV       KLDMSSVNTIIEQARNGVLPISIEPAAISATRIJVVITPNLEVLSKVRQENNVHVAGAIWSTI         HONSSVDTILNDARNGVLPISIEPAAISATRIVVVITPNLEVESKIRQENNVHVAGAIWST         HONSSVDTILNDARNGVLPISVEPATSAARIVVVITPNLEVESKIRQENNVHVAGAIWST         PEDV         RLDMSSVDTILNLAKDGVVPISVEPATSAARIVVITPSLEVFSKIRQENNVHVAGAIWTI         PEDV         RLDMSSVDTILNLAKDGVVPISVEPATSAARIVVITPSLEVFSKIRQENVHVAGAIWTI         MERSCOV         KLDNDVLNGTISNARNGCIPISVEPATSAARIVVITPSLEVFSKIRQEGCVHVAGTIWNI         MERSCOV         KLDNDVLNGTISNARNGCIPISVEPATSAARIVVITPSLEVFSKIRQENVUVAGFVHVAGTIWNI         MERSCOV         SARS-Cov2         KLDNDVLNGTISNARNGCIPISVEPATSAARIVVITPSLEVFSKIRQENVUVAGFVHVAGTIWNI         SARS-Cov2         MERSCOV KLDNDVLNGTISNARNGCIPISVEPATSAARIVVITPSLEVESTICE         SARS-Cov2         MERSCOV         SARS-Cov2         SARS-Cov2         SARS-Cov2         SARS-Cov2         SARS-COV2         SARS-COV2         SARS-CO	97 107 117 127 137 147 FIPV FIPV FIPV FIPV FIEV KLDMSSVNTI I E QARNGVLP LSI I PAASATRI I IVVTPNLEVLSKVRQENNVH YAGA IWSI HCoV-229E RLDMSSVDTI LN LAKDGVVPI SVIPATSAARLVVVVPDH DSFVKMMVDGFVH YAGA IWTI PEDV RLDMSSVDTI LN LAKDGVVPI SVIPATSAARLVVVVPDH DSFVKMMVDGFVH YAGA IWTI MERS-CoV KLDND VLNG I ISNARNGCI PI SVIPATSAARLVVVVPDH DSFVKMMVDGFVH YAGA IWTI MERS-CoV KLDND VLNG I ISNARNGCI PI SVIPATSAARLVVVV IPDFTVWN QVVTYPSIN YAGA IWDI SARS-CoV KLDND ALNNI INNARDGCVPI NI IPLITAAKIMVVI PDFTVWN QVTYPSIN YAGA IWEI SARS-CoV KLDND ALNNI INNARDGCVPI NI IPLITAAKIMVVI PDFTVKN COGNTFT YASALWEI SARS-COV KLDND ALNNI INNARDGCVPI NI IPLITAAKIMVVI VPDYGTYKNTCOGNTFT YASALWEI SARS-COV KLDND ALNNI INNARDGCVPI NI IPLITAAKIMVVI VPDYGTYKNTCOGNTFT YASALWEI SARS-COV KLDND ALNNI INNARDGCVPI NI IPLITAAKIMVVI VPDYGTYKNTCOGNTFT SARS-COV VEVK DANGSHVHLKEVTAANELNI TWPLSITCERTTKLO HCOV-229E QEVKDNDGKNVHLKDVTKEN QEILVWPLILTCERVVKLO PEDV I DIKDNDGKVVHVKEVTAQNAESLSWPLVGCERIVKLO SARS-COV QQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANS.AVKLO SARS-COV QQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANS.AVKLO											
FIPV       KLDMSSVNTIICQARNGVLPLSITPAASATRIVVIPNLEVISKVRQENNVHYAGAIWSI         HCoV-229E       RLDMSSVNTILDQARNGVLPLSITPAASATRIVVIPNLEVISKVRQENNVHYAGAIWTI         PEDV       RLDMSSVDTILNARNGVVPLSVIPATSAARIVVVVPDHDSFVKMMVDGFVHYAGAIWTI         MERS-COV       KLDNSVDTILNAKDGVVPLSVIPAVSATKINIVTSDIDSYNRIORGCVHYAGAIWNI         MERS-COV       KLDNDVLNGISNARNGCIPLSVIPAVSATKINIVTSDIDSYNRIORGCVHYAGAIWDI         SARS-COV2       KLDNDALNNIINNARDGCVPINIPLSVIPAVSATKINIVTSDIDSYNRIORGCVHYAGAIWDI         SARS-COV2       KLDNDALNNIINNARDGCVPINIPLSVIPAVSATKINIVTSDIDSYNRIORGCVTYSAIWEI         FIPV       VEVKDANGAQVHLKEVTAANELNITVPLSITCERTTKLO         GEV       VEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLO         HCov.229E       QEVKDNDGKNVHLKEVTAANELNITWPLSITCERTTKLO         PEDV       VEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLO         MERS-COV2       QEVKDNDGKNVHLKDVTKENQEILVWPLUCGERTVKLO         BS II (3ub0)       BS II (11)         MERS-COV       VVINNVDNEVKSSDV.VDSNENITWPLVLECTRASTSAVKLO         PEDV       ID IKDNDGKVVHVKSSDV.VDSNENITWPLVLECTRASTSAVKLO         SARS-COV2       QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLO         SARS-COV2       QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLO	FIPV TGEV       KID MSSVNTTI E QARNGVLPISTIPAASATRI IVVTPNLEVISKVRQENNVH VAGAIWST KLDMSSVNTI I D QARNGVLPISTIPAASATRI VVVTPNLEVISKVRQENNVH VAGAIWTT HCoV-229E         PEDV       RLDMSSVDTILNMARNGVVPISVIPATSAARIVVVVPDHDSFVKMMVDGFVH VAGAIWTT PEDV         PEDV       RLDMSSVDTILNLAKDGVVPISVIPATSAARIVVVVPDHDSFVKMMVDGFVH VAGAIWTT SARS-CoV         KLDNDVLNGTISNARNGCIPISVIPAVSATKINIVTSDIDSYNRIQREGCVH VAGTIWNT SARS-CoV       KLDNDVLNGTISNARNGCIPISVIPAVSATKINIVTSDIDSYNRIQREGCVH VAGTIWNT SARS-CoV-2         KLDNDALNNTINNARDGCVPINTIPLSTCERTTKLQ TGEV       VEVKDANGAQVHLKEVTAANELNITWPISTCERTTKLQ VEVKDANGSHVHLKEVTAANEINITWPISTCERTTKLQ DNDALNNTINNARDGCVPINTIPLTTAAKIMVVVPDYGTYKNTCDGNTFTVASALWEI SARS-CoV-2         KLDNDALNNTINNARDGCVPINTIPLSTCERTTKLQ TGEV       DSII(3ub0)         HCov-229E       QEVKDNDGKNYHLKEVTAANEINITWPISTCERTTKLQ DIKDNDKVYVEVKEVAQANAESLSWEVLICERRYVKLQ         BSII(11)       BSII(11)         MERS-COV-2       QUVVDADSKIVUSSDV.VDSNENLTWPIVLCCTRASTSAVKLQ         BSII(11)       BSII(12)	I	97	107	117	127	137	147				
TGEV HCoV-229E       KLD MSSVDTILDQARNGVVPLSTIPAASATRLVVITPSLEVFSKIRQENNVHYAGAIWTI PEDV         PEDV       RLD MSSVDTILNMARNGVVPLSVIPATSAARLVVVVPDHDSFVKMMVDGFVHYAGVWTL PEDV         MERS-CoV       KLD NDSVDTILNLAKDGVVPLSVIPATSAARLVVVVPDHDSFVKMMVDGFVHYAGVWTL SARS-CoV         KLD ND VLNGIISNARNGCIPLSVIPLCASNKLRVVIPDFVWNQVVYPSLNYAGALWDI SARS-CoV       KLD NDALNNIINNARDGCVPLNIPLTAAKLMVVIPDYNTYKNTCDGTTFTYASALWEI SARS-CoV-2         157       167       177       187         FIPV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ TGEV       EVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ PEDV       BS II (3ub0)         HCov.229E       QEVKDNDGKNVHLKEVTAANELNITWPLSITCERTTKLQ PEDV       BS II (3ub0)       BS II (11)         MERS-CoV       QUVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ SARS-CoV-2       QUVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ       BS II (T1)	TGEV       KLD MSSVNTIIDQARNGVVPLSIIDAASATRLVVITPSLEVFSKIRQENNVHVAGAIWTI         HCoV-229E       RLD MSSVDTILNMARNGVVPLSVIPATSAARLVVVPDHDSFVKMMVDGFVHVAGAIWTI         PEDV       RLD MSSVDTILNMARNGVVPLSVIPATSAARLVVVPDHDSFVKMMVDGFVHVAGVVTL         MERS-CoV       KLD NDVLNGIISNARNGCIPLSVIPATSARLVVVPDFVVVPDHDSFVKMMVDGFVHVAGTIWNI         MERS-CoV       KLD NDVLNGIISNARNGCIPLSVIPAVSATKLNIVTSDIDSVNRIQREGCVHVAGTIWNI         SARS-CoV       KLD NDALNNIINNARDGCVPLNIPSVIPAVSATKLNIVTSDIDSYNRIQREGCVHVAGTIVIN         SARS-CoV-2       KLD NDALNNIINNARDGCVPLNIPSVIPAKAKLMVVVPDYSLNXAGALWDI         SARS-CoV-2       KLD NDALNNIINNARDGCVPLNIPLTAAKLMVVVPDYSLNXAGALWEI         SARS-CoV-2       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ         HCoV-229E       QEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLQ         PEDV       IDIKDNDGKNVHLKEVTAQNAESISWPLVECTRASTSAVKLQ         PEDV       IDIKDNDGKVVHVKEVTAQNAESISWPLVECTRASTSAVKLQ         MERS-CoV       VVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ         SARS-CoV-2       QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ	L FIPV	KLDMSS	VNTIIEQARN	GVLPLSIIPA	ASATRIIVV	I <b>P</b> NLEVLSK <b>V</b>	RQEN <b>NV</b> H <b>YAGA</b>	IWSI			
HCoV229E       RLDMSSVDTILNLAKDGVVPLSVIPAVSATKLVVVVPDHDSVRMVVDGEVHYAGVVTL PEDV         RLDMSSVDTILNLAKDGVVPLSVIPAVSATKLNIVTSDIDSYNRIQREGCVHYAGVVTL MERS-CoV         KLDNDVLNGISNARNGCIPLSVIPAVSATKLNIVTSDISVRRIQREGCVHYAGVVTL SARS-CoV         KLDNDALNNIINNARDGCVPLSVIPAVSATKLNIVTSDISVRRIQREGCVHYAGALWDI SARS-CoV-2         KLDNDALNNIINNARDGCVPLNIPLTAAKLMVVIPDYNTYKNTCDGTTFTYASALWEI SARS-CoV-2         FIPV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ TGEV         VEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLQ PEDV         ILKADNGKVVHVKEVTAQNAESLSWPLVGCERIVKLQ SARS-CoV-2         QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ SARS-CoV-2	HCoV229E       RIDMSSVDTILNIAKDGVVPISVIPATSVARLVVVPDHDSVIPATSAARLVVVPDHDSVNR MVDGEVHYAGVVWIL         PEDV       RIDMSSVDTILNIAKDGVVPISVIPATSVARLVVTDIDSVNR IQEEGVHYAGVVWIL         MERS-CoV       KLDNDVLNGIISNARNGCIPISVIPACASNKLNIVTSDISVR IQEEGVHYAGVUWIL         SARS-CoV       KLDNDALNNIINNARDGCVPINIPLSVIPACASNKLNVVIPDFTVWNQVVTYPSLNVAGALWDI         SARS-CoV-2       KLDNDALNNIINNARDGCVPINIPLTAAKLMVVIPDYTYKNTCDGTTFTYASALWEI         SARS-CoV-2       KLDNDALNNIINNARDGCVPINIPLTAAKLMVVVPDYGTYKNTCDGNTFTYASALWEI         FIPV       VEVKDANGAQVHLKEVTAANELNITWPISITCERTTKLQ         TGEV       VEVKDANGSHVHLKEVTAANELNITWPISITCERTTKLQ         Hcov229E       QEVKDNDGKNVHLKEVTAQNAESISWPIVLGCERIVKLQ         PEDV       IDIKDNDGKVVHVKEVTAQNAESISWPIVLGCERIVKLQ         MERS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPIVTALRANS.AVKLQ         SARS-CoV-2       QQVVDADSKIVQLSEISMDNSPNLAWPIVTALRANS.AVKLQ	TGEV	KLDMSS	VNTIIDQARN	GVLPLSIIPA	ASATRLVVI	T <b>P</b> SLEV <b>F</b> SK <b>I</b>	R Q E N N V H YAGA	IWTI			
MERS-CoV       KLDNDVLNGIISNARNGCIPLSVIPLCASNKLRVVIPDFTVWNQVVTYPSLNYAGALWDI         SARS-CoV       KLDNDALNNIINNARDGCVPLNIIPLTTAAKLMVVIPDYNTYKNTCDGTTFTYASALWEI         SARS-CoV-2       KLDNDALNNIINNARDGCVPLNIIPLTTAAKLMVVVPDYGTYKNTCDGTTFTYASALWEI         SARS-CoV-2       KLDNDALNNIINNARDGCVPINIIPLTAAKLMVVVPDYGTYKNTCDGNTFTYASALWEI         SARS-CoV-2       KLDNDALNNIINNARDGCVPINIIPLTAAKLMVVVPDYGTYKNTCDGNTFTYASALWEI         FIPV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ         TGEV       VEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLQ         PLOV       VEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLQ         PEOV       VEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLQ         MERS-CoV       QEVKDNDGKNVHLKEVTAQNAESLSWPLVGCERIVKLQ         BS II (3ub0)       BS II (T1)         MERS-CoV       TVINNVDNEIVKSSDV.VDSNENITWPLVECTRASTSAVKLQ         SARS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ         SARS-CoV-2       QQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANS.AVKLQ	MERS-CoV       KLD NDVLNGTISNARNGCIPLSVIPLCASNKLRVVIPDFTVWNQVVTYPSLNYAGALWDI         SARS-CoV       KLD NDALNNIINNARDGCVPLNIPLTTAAKLMVVIPDYNTYKNTCDGTTFTYASALWEI         SARS-CoV-2       KLD NDALNNIINNARDGCVPLNIPLTTAAKLMVVVPDYGTYKNTCDGNTFTYASALWEI         SARS-CoV-2       KLD NDALNNIINNARDGCVPINIPLTAAKLMVVVPDYGTYKNTCDGNTFTYASALWEI         FIPV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ         TGEV       VEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLQ         PLOV-229E       QEVKDNDGKNVHLKEVTAANELSLSWPLVLGCERIVKLQ         PEV       IDIKDNDGKVVHVKEVTAQNAESLSWPLVLGCERIVKLQ         SARS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ         SARS-CoV-2       QQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANS.AVKLQ	HCov-229E PEDV	RLDMSS	VDTILNMARN VDTILNLAKD	GVVPLSVIPA GVVPLSVIPA	VSATKLNIV:	IS <b>D</b> IDS <b>Y</b> NR <b>I</b>	MVDGFVHYAGV QREGCVHYAGT	IWNI			
SARS-CoV-2 KLDNDALNNIINNÄRDGOVFINIISLIITAAKLAVVVIPDYGTYKNICDGNIFITAAKLAVI SARS-CoV-2 KLDNDALNNIINNÄRDGOVFINIIPLITAAKLAVVVIPDYGTYKNICDGNIFITAAKLAVI FIPV VEVKDANGAQVHLKEVTAANELNIIPLITAAKLAVVVVPDYGTYKNICDGNIFTYASALWEI TGEV VEVKDANGAQVHLKEVTAANELNIIWPLSITCERTTKLO HCoV-229E QEVKDNDGKNVHLKDVTKENQEILVWPLSITCERTVKLO PEOV IDIKDNDGKNVHLKDVTKENQEILVWPLILTCERVVKLO MERS-CoV TVINNVDNEIVKSSDV.VDSNENITWPLVLGCERIVKLO SARS-CoV QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLO SARS-CoV-2 QQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANS.AVKLO	SARS-CoV-2 KIDNDALNNI INNARDGOVFINI ISLITIAAKDWVVIFDINI IKNICOGI FITAAALWEI SARS-CoV-2 KIDNDALNNI INNARDGOVFINI IPLTAAKLWVVVPDYGTYKNTCDGNTFTYASALWEI FIPV VEVKDANGAQVHLKEVTAANELNI IPLTAAKLWVVVPDYGTYKNTCDGNTFTYASALWEI TGEV VEVKDANGSHVHLKEVTAANELNI TWPLSITCERTTKLQ PEOV UEVKDANGSHVHLKEVTAANELNI TWPLSITCERTTKLQ PEOV IDIKDNDGKNVHLKEVTAANELNI TWPLSITCERTTKLQ MERS-CoV IVINNVDNEIVKSSDV.VDSNENI TWPLVLGCERIVKLQ SARS-CoV QQVVDADSKI VQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ SARS-CoV-2 QQVVDADSKI VQLSEINMDNSPNLAWPLIVTALRANS.AVKLQ	MERS-CoV	KLDNDVI	LNGIISNARN	GCIPLSVIP1	CASNKLRVV	IPDFTVWNQV	VTYPSLN <b>YAGA</b>				
157       167       177       187         FIPV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ         TGEV       VEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLQ         Hcov-229E       QEVKDNDGKNVHLKEVTAANELNITWPLSITCERTTKLQ         PEDV       IDIKDNDGKVVHVKEVTAQNAESLSWPLVLGCERTVKLQ         MERS-CoV       TVINNVDNEIVKSSDV.VDSNENLTWPLSITCERTVKLQ         SARS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ	157       167       177       187         FIPV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ         TGEV       VEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLQ         HCov-229E       QEVKDNDGKNVHLKEVTAANELNITWPLSITCERVVKLQ         PEDV       IDIKDNDGKNVHVKEVTAQNAESISVPLVLGCERIVKLQ         MERS-CoV       TVINNVDNEIVKSSDV.VDSNENLTWPLVLECTRASTSAVKLQ         SARS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ	SARS-CoV-2	KLDNDAI	LNNIINNARD	GCVPLNIIPI	TTAAKLMVV	<b>PD</b> YGT <b>Y</b> KNT	CDGN <b>T</b> FT <b>YASA</b>	LWEI			
157       167       177       187         FIPV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ         TGEV       VEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLQ         HCoV-229E       QEVKDNDGKNVHLKEVTAANELNITWPLSITCERTTKLQ         PEDV       IDIKDNDGKNVHLKEVTAANELNITWPLSITCERTTKLQ         MERS-COV       IVINNVDNEIVKSSDV.VDSNENITWPLVLGCERIVKLQ         SARS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ         SARS-CoV-2       QQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANS.AVKLQ	157       167       177       187         FIPV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ         TGEV       VEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLQ         HCoV-229E       QEVKDNDGKNVHLKEVTAANELNITWPLSITCERTVKLQ         PEDV       IDIKDNDGKNVHVKEVTAQNAESISWPLVLGCERIVKLQ         MERS-CoV       TVINNVDNEIVKSSDV.VDSNENITWPLVLECTRASTSAVKLQ         SARS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPLVTALRANS.AVKLQ											
157       167       177       187         FIPV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ         TGEV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ         MCOV 229E       QEVKDNDGKNVHLKEVTAANELNITWPLSITCERTTKLQ         PEOV       IDIKDNDGKVVHVKEVTAQNAESLSWPLVLGCERTVKLQ         MERS-CoV       TVINNVDNEIVKSSDV.VDSNENITWPLVLGCERTVKLQ         SARS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ         SARS-CoV-2       QQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANS.AVKLQ	157       167       177       187         FIPV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ         TGEV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ         HCoV-229E       QEVKDNDGKNVHLKEVTAANELNITWPLSITCERTTKLQ         PBOV       IDIKDNDGKVVHVKEVTAQNAESLSWPLVLGCERTVKLQ         MERS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPLVLGCERTVKLQ         SARS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPLVLGCERTASTSAVKLQ         SARS-CoV-2       QQVVDADSKIVQLSEISMDNSPNLAWPLVTALRANS.AVKLQ											
157       167       177       187         FIPV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ         rgev       VEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLQ         Hcov-229E       QEVKDNDGKNVHLKEVTAANELNITWPLSITCERTVKLQ         PEDV       IDIKDNDGKVVHVKEVTAQNAESLSWPLVLGCERIVKLQ         MERS-cov       TVINNVDNEIVKSSDV.VDSNENITWPLVLCCTRASTSAVKLQ         SARS-cov2       QQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANS.AVKLQ	157       167       177       187         FIPV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ         TGEV       VEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLQ         HCoV-229E       QEVKDNDGKNVHLKEVTAANELNITWPLSITCERTVKLQ         PDV       IDIKKDNDGKNVHLKEVTAQNAESLSWPLVLGCERIVKLQ         MERS-CoV       TVINNVDNEIVKSSDV.VDSNENITWPLVLGCERIVKLQ         SARS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ         SARS-CoV-2       QQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANS.AVKLQ											
FIPV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ         TGEV       VEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLQ         HCov-229E       QEVKDNDGKNVHLKDVTKENQEILVWPLILTCERVVKLQ         PEDV       IDIKDNDGKVVHVKEVTAQNAESLSWPLVLGCERIVKLQ         MERS-CoV       TVINNVDNEIVKSSDV.VDSNENLTWPLVLCCTRASTSAVKLQ         SARS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ	FIPV       VEVKDANGAQVHLKEVTAANELNITWPLSITCERTTKLQ         TGEV       VEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLQ         HCov-229E       QEVKDNDGKNVHLKEVTAANELNITWPLSITCERTVKLQ         PEDV       IDIKDNDGKNVHLKDVTKENQEILVWPLILTCERVVKLQ         MERS-CoV       TVINNVDNEIVKSSDV.VDSNENITWPLVLGCERIVKLQ         SARS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ         SARS-CoV-2       QQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANS.AVKLQ		157	167	177	187						
TGEV       VEVKDANGSHVHLKEVTAANELNITWPLSITCERTTKLQ       DSIN (SUDD)         HCoV-229E       QEVKDNDGKNVHLKDVTKENQEILVWPLILTCERVVKLQ       DSIN (SUDD)         PEDV       IDIKDNDGKVVHVKEVTAQNAESLSWPLVLGCERIVKLQ       BSIN (T1)         MERS-CoV       TVINNVDNEIVKSSDV.VDSNENITWPLVLECTRASTSAVKLQ       BSIN (T2)         SARS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ       BSIN (T2)	TGEV       VEVKDANGSHVHLKEVTAANELNITWPISITCERTTKLQ         HCoV-229E       QEVKDNDGKNVHLKDVTKENQEILVPPIITCERVVKLQ         PEDV       IDIKDNDGKVVHVKEVTAQNAESISWPIVLGCERIVKLQ         MERS-CoV       VINNVDNEIVKSSDV.VDSNENLTWPIVLECTRASTSAVKLQ         SARS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPIIVTALRANS.AVKLQ         SARS-CoV-2       QQVVDADSKIVQLSEINMDNSPNLAWPIIVTALRANS.AVKLQ	FIPV	• VE <b>V</b> K <b>DA</b> 1	NGAQ <b>VHL</b> K <b>EV</b>	TAANELNIT	PLSITCERT	T <mark>KLQ</mark>		h0)			
PEDV       IDIKDNDGKVVHVKEVTAQNAESISWPIVLGCERIVKLQ       BSII (T1)         MERS-CoV       TVINNVDNEIVKSSDV.VDSNENITWPIVLECTRASTSAVKLQ         SARS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLQ         SARS-CoV-2       QQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANS.AVKLQ	PEDV       IDIKDNDGKVVHVKEVTAQNAESISWPIVLGCERIVKLQ       BS II (T1)         MERS-CoV       TVINNVDNEIVKSSDV.VDSNENITWPIVLECTRASTSAVKLQ         SARS-CoV       QQVVDADSKIVQLSEISMDNSPNLAWPIIVTALRANS.AVKLQ         SARS-CoV-2       QQVVDADSKIVQLSEINMDNSPNLAWPIIVTALRANS.AVKLQ		VEVKDAN	NGSH <b>VHL</b> KEV GKNVHLKDV	TAANELNLTW	PLSITCERT. PLTLTCERV	T <mark>KLQ</mark> VKLQ		50)			
MERS-COV TVINNVDNEIWKSSDV. VDSNENLTWPLVLECTRASTSAVKLO SARS-COV QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLO SARS-COV2 QQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANS.AVKLO	MERS-COV IVINNVDNE IVKSSDV. VDSNENLIWPLVLECIRASISAVKLO SARS-COV QQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANS.AVKLO SARS-COV-2 QQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANS.AVKLO	PEDV	IDIKDNI	GKV <b>VHV</b> K <b>EV</b>	TAQNAESLSW	PLVLGCERI	VKLQ	BS II (T1	)			
SARS-COV-2 QQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANS.AVKLO	SARS-COV-2 QQVVDADSKIVQLSEINMDNSPNLAWPLIVTALRANS.AVKLO	MERS-CoV SARS-CoV	TVINNVI 00VVDAI	DNEIVKSSDV DSKIVOLSEI	. VDSNENLTW SMDNSPNLAW	PLVLECTRAS PLIVTALRAN	STSAVKLQ NS.AVKLO	BS II (T2	2)			
		SARS-CoV-2	QQ <b>V</b> V <b>DAI</b>	SKI <b>VÕL</b> S <b>EI</b>	NMDNSPNLAW	PLIVTAL RA	NS.A <mark>VKLÕ</mark>		'			
		FIPV TGEV HCoV-229E PEDV MERS-CoV SARS-CoV/2	157 VEVKDAN VEVKDAN QEVKDNI IDIKDNI TVINNVI QQVVDAI	167 NGAQ <b>VHLKEV</b> NGSH <b>VHLKEV</b> OGKN <b>VHLKDV</b> OGKV <b>VHV</b> KEV ONEIVKSSDV OSKIVQLSEI OSKIVQLSEI	$\begin{array}{c} 177\\ \textbf{T} \textbf{A} \textbf{A} \textbf{N} \textbf{E} \perp \textbf{N} \textbf{I} \top \textbf{K}\\ \textbf{T} \textbf{A} \textbf{A} \textbf{N} \textbf{E} \perp \textbf{N} \textbf{L} \top \textbf{K}\\ \textbf{T} \textbf{K} \textbf{E} \textbf{N} \textbf{Q} \textbf{E} \perp \textbf{L} \forall \textbf{M}\\ \textbf{T} \textbf{A} \textbf{Q} \textbf{N} \textbf{A} \textbf{E} \textbf{S} \textbf{L} \textbf{S}\\ \textbf{V} \textbf{D} \textbf{S} \textbf{N} \textbf{E} \textbf{N} \textbf{L} \top \textbf{K}\\ \textbf{S} \textbf{M} \textbf{D} \textbf{N} \textbf{S} \textbf{P} \textbf{N} \textbf{L} \textbf{A} \textbf{X}\\ \end{array}$	187 IPLSITCERT IPLSITCERT IPLILTCERV IPLVLGCERI IPLVLECTRAS IPLIVTALRAN	TKLQ TKLQ VKLQ STSAVKLQ NS.AVKLQ NS.AVKLQ	BS II (3u BS II (T1 BS II (T2	b0) ) ?)			

Figure S 15: Sequence alignment of nsp7 and nsp8 from seven CoV species. The multiple sequence alignment of nsp7-8 sequences of the seven tested CoVs is genrated with Clustal Omega [43] and converted by ESPript [44] <u>http://espript.ibcp.fr</u> using the amino acid sequences without C-terminal linkers and His<sub>6</sub> as input (Table S 6).
Highlighted are conserved (black) and semi-conserved (bold) sequences. Further highlighted are molecular contacts at binding site (BS) II in the SARS-CoV nsp7+8 (2:2) heterotetramer candidate structures T1 (green) and T2 (orange) (subcomplexes of pdb 2ahm) as well as in the FIPV nsp7+8 (2:2) heterotrimer (pink, pdb 3ub0).
Contacts (VDW radius -0.4 Å) was analyzed with ChimeraX [45].





766 Figure S 16: Conserved stretch of nsp8 has various binding contexts. Candidate complexes involving similar 767 conserved residues (red) in the nsp8 BS II are shown here for (A) the SARS-CoV nsp7+8 (2:2) heterotetramer 768 candidate structures T1 and (B) T2 (subcomplexes of pdb 2ahm) as well as (C) the FIPV nsp7+8 (2:2) heterotrimer 769 (pdb 3ub0). (D) Cryo electron microscopy structure of the nsp7+8+12+13 (1:2:1:1) polymerase complex (pdb 6xez) 770 [29]. The two subunits of nsp8 (green), bind as a monomer nsp8a and as a nsp7+nsp8b heterodimer to the nsp12 771 RdRp. From there, nsp8a and nsp8b extend and interact as a facet with the RNA duplex and also individually 772 with one subunit of the nsp13 helicase. The molecular contacts between nsp8 and nsp12/13 are mediated by 773 exactly the same amino acids in the nsp8 BS II that are involved in nsp7+8 complex formation. (E) Zoom into the 774 interacting region at nsp8a and nsp8b BS II and its amino acids in contact (VDW radius -0.4 Å) with the nsp12 775 thumb domain (brown), nsp13.1 and nsp13.2 (blue). All side chains of residues involved in contacts are displayed. 776 Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National 777 778 Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National 779 Institute of Allergy and Infectious Diseases [45].

780

# 781 **1. References**

- Cui, J., F. Li, and Z.-L. Shi, Origin and evolution of pathogenic coronaviruses. Nature Reviews Microbiology, 2019. 17(3): p. 181-192.
- Zaki, A.M., et al., *Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia.* New England Journal of Medicine, 2012. 367(19): p. 1814-1820.
- 7863.Drosten, C., et al., Identification of a novel coronavirus in patients with severe acute787respiratory syndrome. New England Journal of Medicine, 2003. 348(20): p. 1967-1976.
- Novel, C.P.E.R.E., *The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China.* Zhonghua liu xing bing xue za zhi= Zhonghua liuxingbingxue zazhi, 2020. 41(2): p. 145.

- 5. Gorbalenya, A.E., Severe acute respiratory syndrome-related coronavirus-The species and
  its viruses, a statement of the Coronavirus Study Group. bioRxiv, 2020.
- Dong, E., H. Du, and L. Gardner, *An interactive web-based dashboard to track COVID-19 in real time.* The Lancet infectious diseases, 2020. 20(5): p. 533-534.
- 7957.Pedersen, N.C., An update on feline infectious peritonitis: virology and796immunopathogenesis. The Veterinary Journal, 2014. 201(2): p. 123-132.
- Vlasova, A., et al., *Porcine coronaviruses*, in *Emerging and Transboundary Animal Viruses*.
   2020, Springer. p. 79-110.
- 7999.Snijder, E.J., E. Decroly, and J. Ziebuhr, The Nonstructural Proteins Directing800Coronavirus RNA Synthesis and Processing. Adv Virus Res, 2016. 96: p. 59-126.
- 10. van Hemert, M.J., et al., SARS-coronavirus replication/transcription complexes are
  membrane-protected and need a host factor for activity in vitro. PLoS Pathog, 2008. 4(5): p.
  e1000054.
- 804 11. Shannon, A., et al., *Favipiravir strikes the SARS-CoV-2 at its Achilles heel, the RNA*805 *polymerase.* bioRxiv, 2020.
- Subissi, L., et al., One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. Proceedings of the National Academy of Sciences of the United States of America, 2014. 111(37): p. E3900-E3909.
- Kirchdoerfer, R.N. and A.B. Ward, *Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors*. Nat Commun, 2019. **10**(1): p. 2342.
- 812 14. Gao, Y., et al., Structure of the RNA-dependent RNA polymerase from COVID-19 virus.
  813 Science, 2020. 368(6492): p. 779-782.
- Peng, Q., et al., *Structural and Biochemical Characterization of the nsp12-nsp7-nsp8 Core Polymerase Complex from SARS-CoV-2.* Cell Reports, 2020. **31**(11): p. 107774.
- 816 16. Hillen, H.S., et al., *Structure of replicating SARS-CoV-2 polymerase*. bioRxiv, 2020.
- 817 17. Ferron, F., et al., *Structural and molecular basis of mismatch correction and ribavirin*818 *excision from coronavirus RNA*. Proceedings of the National Academy of Sciences, 2018.
  819 115(2): p. E162-E171.
- 820 18. Gordon, C.J., et al., *The antiviral compound remdesivir potently inhibits RNA-dependent*821 *RNA polymerase from Middle East respiratory syndrome coronavirus.* J Biol Chem, 2020.
  822 295(15): p. 4773-4779.
- 19. Krichel, B., et al., *Processing of the SARS-CoV pp1a/ab nsp7-10 region*. Biochem J, 2020.
- 20. Zhai, Y., et al., *Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer*. Nat Struct Mol Biol, 2005. **12**(11): p. 980-6.
- Xiao, Y., et al., Nonstructural proteins 7 and 8 of feline coronavirus form a 2:1 heterotrimer
  that exhibits primer-independent RNA polymerase activity. J Virol, 2012. 86(8): p. 4444-54.
- Li, S., et al., *New nsp8 isoform suggests mechanism for tuning viral RNA synthesis*. Protein
  Cell, 2010. 1(2): p. 198-204.
- 830 23. Konkolova, E., et al., *Structural analysis of the putative SARS-CoV-2 primase complex.*831 Journal of Structural Biology, 2020. **211**(2): p. 107548.
- Biller, J., et al., Structural mass spectrometry goes viral. Advances in virus research,
  2019. 105: p. 189-238.
- 834 25. Erba, E., L. Signor, and C. Petosa, *Exploring the structure and dynamics of macromolecular complexes by native mass spectrometry*. Journal of Proteomics, 2020: p. 103799.

- 837 26. Wang, G., A.J. Johnson, and I.A. Kaltashov, *Evaluation of electrospray ionization mass*838 spectrometry as a tool for characterization of small soluble protein aggregates. Analytical
  839 chemistry, 2012. 84(3): p. 1718-1724.
- Ziebuhr, J., *The coronavirus replicase*, in *Coronavirus replication and reverse genetics*. 2005,
  Springer. p. 57-94.
- 842 28. Hegyi, A. and J. Ziebuhr, *Conservation of substrate specificities among coronavirus main*843 *proteases.* J Gen Virol, 2002. 83(Pt 3): p. 595-9.
- Chen, J., et al., Structural basis for helicase-polymerase coupling in the SARS-CoV-2
  replication-transcription complex. Cell, 2020.
- 84630.Akabayov, B., et al., Impact of macromolecular crowding on DNA replication. Nature847communications, 2013. 4(1): p. 1-10.
- 848 31. Von Brunn, A., et al., Analysis of intraviral protein-protein interactions of the SARS
  849 coronavirus ORFeome. PloS one, 2007. 2(5): p. e459.
- 850 32. Falke, S., (Doctoral Thesis) Coronaviral Polyprotein Nsp7-10 : Proteolytic Processing and
  851 Dynamic Interactions within the Transcriptase/Replicase Complex. 2014, Staats- und
  852 Universitätsbibliothek Hamburg.
- Xue, X., et al., Production of authentic SARS-CoV M pro with enhanced activity: application *as a novel tag-cleavage endopeptidase for protein overproduction*. Journal of molecular
  biology, 2007. 366(3): p. 965-975.
- 85634.van den Heuvel, R.H., et al., Improving the performance of a quadrupole time-of-flight857instrument for macromolecular mass spectrometry. Anal Chem, 2006. 78(21): p. 7473-83.
- 85835.Marty, M.T., et al., Bayesian deconvolution of mass and ion mobility spectra: from binary859interactions to polydisperse ensembles. Analytical chemistry, 2015. 87(8): p. 4370-4376.
- Strohalm, M., et al., *mMass data miner: an open source alternative for mass spectrometric data analysis.* Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry, 2008. 22(6): p. 905-908.
- 86437.Franke, D., C.M. Jeffries, and D.I. Svergun, Correlation Map, a goodness-of-fit test for one-865dimensional X-ray scattering spectra. Nature methods, 2015. 12(5): p. 419.
- 866 38. Franke, D., A.G. Kikhney, and D.I. Svergun, *Automated acquisition and analysis of small*867 *angle X-ray scattering data.* Nuclear Instruments and Methods in Physics Research
  868 Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2012.
  869 689: p. 52-59.
- Bayesian assessment of protein molecular mass from solution X-ray scattering data. Scientific reports, 2018. 8(1): p. 1-13.
- 40. Svergun, D., C. Barberato, and M.H. Koch, *CRYSOL-a program to evaluate X-ray*solution scattering of biological macromolecules from atomic coordinates. Journal of applied
  crystallography, 1995. 28(6): p. 768-773.
- Konarev, P.V., et al., *PRIMUS: a Windows PC-based system for small-angle scattering data analysis.* Journal of applied crystallography, 2003. 36(5): p. 1277-1282.
- Petoukhov, M.V., et al., *New developments in the ATSAS program package for small-angle scattering data analysis.* Journal of applied crystallography, 2012. 45(2): p. 342-350.
- 43. Sievers, F., et al., *Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega*. Molecular systems biology, 2011. 7(1): p. 539.
- 44. Robert, X. and P. Gouet, *Deciphering key features in protein structures with the new*882 *ENDscript server*. Nucleic acids research, 2014. 42(W1): p. W320-W324.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.30.320762. this version posted October 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY-NC-ND 4.0 International license.

45. Goddard, T.D., et al., UCSF ChimeraX: Meeting modern challenges in visualization and
analysis. Protein Science, 2018. 27(1): p. 14-25.