Spindle-shaped archaeal viruses evolved from rod-shaped ancestors to package a larger genome

Fengbin Wang, Virginija Cvirkaite-Krupovic, Matthijn Vos, Leticia Beltran, Mark A.B. Kreutzberger, Jean-Marie Winter, Zhangli Su, Jun Liu, Stefan Schouten, Mart Krupovic, et al.

To cite this version:

Fengbin Wang, Virginija Cvirkaite-Krupovic, Matthijn Vos, Leticia Beltran, Mark A.B. Kreutzberger, et al.. Spindle-shaped archaeal viruses evolved from rod-shaped ancestors to package a larger genome. Cell, 2022, 185 (8), pp.1297-1307.e11. 10.1016/j.cell.2022.02.019 . pasteur-03620791

HAL Id: pasteur-03620791
https://pasteur.hal.science/pasteur-03620791

Submitted on 26 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1

Spindle-shaped archaeal viruses

evolved from rod-shaped ancestors

to package a larger genome

5

6   Fengbin Wang[1,†], Virginija Cvirkaite-Krupovic[2,†], Matthijn Vos[3], Leticia C. Beltran[1], Mark A.B.
7   Kreutzberger[1], Jean-Marie Winter[3], Zhangli Su[1], Jun Liu[4], Stefan Schouten[5],
8   Mart Krupovic[2,*] and Edward H. Egelman[1,*,a]
9
10   [1]Department of Biochemistry and Molecular Genetics
11   University of Virginia
12   Charlottesville, VA 22903, U.S.A.
13
14   [2]Institut Pasteur, Université de Paris
15   CNRS UMR6047, Archaeal Virology Unit, 75015 Paris, France
16
17   [3]NanoImaging Core Facility, Centre de Ressources et Recherches Technologiques (C2RT),
18   Institut Pasteur, 75015 Paris, France
19
20   [4]Department of Microbial Pathogenesis & Microbial Sciences Institute,
21   Yale University School of Medicine
22   New Haven, CT 06510, U.S.A.
23
24   [5]NIOZ Royal Netherlands Institute for Sea Research
25   Department of Marine Microbiology and Biogeochemistry
26   Texel, The Netherlands
27
28   † These authors contributed equally to this paper
29   ª Lead Contact
30   * Correspondence to mart.krupovic@pasteur.fr, egelman@virginia.edu

Summary

Spindle- or lemon-shaped viruses infect archaea in diverse environments. Due to the highly

pleomorphic nature of these virions, which can be found with cylindrical tails emanating from

the spindle-shaped body, structural studies of these capsids have been challenging. We have

determined the atomic structure of the capsid of Sulfolobus monocaudavirus 1, a virus that

infects hosts living in nearly boiling acid. A highly hydrophobic protein, likely integrated into the

host membrane before the virions assemble, forms seven strands that slide past each other in

both the tails and the spindle body. We observe the discrete steps that occur as the tail tubes

expand, and these are due to highly conserved quasi-equivalent interactions with neighboring

subunits maintained despite significant diameter changes. Our results show how helical

assemblies can vary their diameters, becoming nearly spherical to package a larger genome,

and suggest how all spindle-shaped viruses have evolved from archaeal rod-like viruses.

Introduction

Capsids are a hallmark of viruses, which distinguishes them from all other types of mobile genetic elements (Forterre et al., 2014; Raoult and Forterre, 2008). Over billions of years, viruses have 'invented' capsids on multiple independent occasions from non-homologous and structurally unrelated proteins (Krupovic et al., 2019). Nevertheless, the majority of known viruses package their genomes into icosahedral or filamentous helical protein capsids (also called nucleocapsids if surrounded by an additional layer, e.g., lipid membrane) (Krupovic and Koonin, 2017; Sevvana et al., 2021). Viruses infecting archaea notoriously deviate from this general paradigm by producing virions with unique, odd-shaped morphologies, which are not observed among bacterial or eukaryotic viruses (Dellas et al., 2014; Prangishvili et al., 2017). Such archaea-specific virion architectures resemble droplets, champagne bottles or spindles (Baquero et al., 2020). Viruses with spindle-shaped (or lemon-shaped) virions are particularly common in diverse extreme and moderate environments and infect a wide range of archaeal lineages from the phyla Crenarchaeota, Euryarchaeota and Thaumarchaeota as well as Asgardarchaeota, a group of archaea widely considered to represent the closest archaeal relatives of eukaryotes (Medvedeva et al., 2021). Because of the broad distribution in Archaea, it has been suggested that spindle-shaped viruses were associated with the last archaeal common ancestor, and possibly, even the last universal cellular ancestor (LUCA) (Krupovic et al., 2020).

Based on virion characteristics and genomic relationships, most spindle-shaped viruses fall into two groups (Krupovic et al., 2014). One of the groups comprises smaller spindle-shaped viruses classified into families *Fuselloviridae* (e.g., Sulfolobus spindle-shaped virus 1 [SSV1]),

3

67 *Halspiviridae* (e.g., haloarchaeal virus His1) and *Thaspiviridae* (e.g., Nitrosopumilus spindle-

68 shaped virus 1 [NSV1]) as well as several unclassified viruses. These viruses encode homologous

69 major capsid proteins (MCP) containing two hydrophobic, potentially membrane-spanning

70 domains and are extruded from the host cell through a budding-like mechanism without

71 causing cell lysis (Kim et al., 2019; Quemin et al., 2016). Attempts to determine the structures

72 of SSV1 (Stedman et al., 2015) and His1 (Hong et al., 2015) by cryo-EM did not yield atomic

73 models and provided little insight into virion organization. The second group includes members

74 of the family *Bicaudaviridae*, which have larger virions and genomes. A characteristic feature of

75 this virus assemblage is the presence of 'tails' emanating from one or both pointed ends of the

76 spindle. Notably, Acidianus two-tailed virus (ATV) and Sulfolobus monocaudavirus 1 (SMV1)

77 have been reported to develop tails extracellularly, with the only requirement for this

78 transformation being the incubation at temperatures close to those of the natural habitats (75-

79 90°C) (Haring et al., 2005; Uldahl et al., 2016). It has been suggested that virus-encoded MoxR-

80 type AAA+ ATPase and a von Willebrand domain A-containing cochaperone play an active role

81 in this process by hydrolyzing ATP stored within the virions (Scheele et al., 2011). However, the

82 exact mechanism underlying this virion morphogenesis outside of the host cells has not been

83 elucidated. Other bicaudaviruses, namely, Acidianus tailed spindle virus (ATSV) and Sulfolobus

84 tengchongensis spindle-shaped viruses 1 and 2 (STSV1 and STSV2, respectively) are released

85 from the cells with tails extending from one of the virion poles and do not undergo further

86 transformation (Erdmann et al., 2014; Hochstein et al., 2016; Xiang et al., 2005). Remarkably,

87 upon infection, STSV2 and SMV1 block normal cell division, transforming the host cell into a

88 giant virion-producing factory, which is up to 20 times larger compared to non-infected cells

89   (Liu et al., 2021). The soluble structural protein repeatedly identified as the major component

90   of bicaudavirus virions (Erdmann *et al.*, 2014; Hochstein *et al.*, 2016; Prangishvili et al., 2006;

91   Xiang *et al.*, 2005) is unrelated to that of the smaller spindle-shaped viruses from the first

92   group, leading to the suggestion that the two groups of viruses are evolutionarily distinct

93   (Krupovic *et al.*, 2014). The protein has been crystallized for ATV and ATSV and shows a four-

94   helix bundle fold (Goulet et al., 2010; Hochstein et al., 2018).

95   Filamentous and spherical capsids (displaying helical and icosahedral symmetries,

96   respectively) embody the simplest virion designs dictated by the principle of 'genetic economy'

97   inherent to viral genomes, whereby relatively large capsids are built from a small number of

98   distinct protein subunits interacting with each other in a way that is identical throughout the

99   capsid (Crick and Watson, 1956). For icosahedral viruses, Caspar and Klug proposed that

100   interactions between the capsomers are quasi-equivalent (rather than identical) and developed

101   the quasi-equivalence theory which allows predicting the arrangement of capsomers in larger

102   icosahedral capsids (Caspar and Klug, 1962). How complex capsids, such as those of spindle-

103   shaped viruses, are organized and how they have evolved remains unclear. While two

104   theoretical models have been postulated for the geometric basis of the spindle shape in

105   archaeal viruses (Perotti et al., 2016; Perotti et al., 2019), we show that they bear no relation to

106   the actual structure of these viruses. Similarly, we can show that the crystal structures for the

107   putative MCPs in both ATV (Goulet *et al.*, 2010) and ATSV (Hochstein *et al.*, 2018) were not of

108   the MCPs, but unrelated viral proteins. Since spindle-shaped viruses appear to have evolved

109   from archaeal rod-like viruses, our results explain why spindle-shaped viruses have thus far only

110   been observed to infect archaea.

111
Results

113   Cryo-EM images of SMV1 (Figure 1A) show that the spindle-shaped body of the virion is

114   continuous with the tubular tails extending out from the body. As observed for ATSV (Hochstein

115   *et al.*, 2018), the periodicity in the spindle body is the same as that in the tails, strongly

116   suggesting a seamless transition between the tails and the spindles. While this periodicity was

117   described as ~40 Å for ATSV (Hochstein *et al.*, 2018), we found a 29 Å periodicity in SMV1

118   (Figure 1A). The presence or absence of tails can be modulated by prolonged incubation at high

119   temperatures (Haring *et al.*, 2005; Uldahl *et al.*, 2016). However, the heat-treated SMV1

120   population remained very heterogeneous with tails displaying variable lengths and diameters.

121   We noticed that increase in pH from 6 to 10 results in nearly uniform transformation of spindle-

122   shaped virions into tubular structures and is accompanied by the release of material present

123   inside of the viral particle (Figure S1A). Fractionation of the SMV1 particles on 5-20% sucrose

124   gradient followed by analysis of the different gradient fractions showed that the viral genome

125   was released in the form of amorphous aggregates along with proteins present in the lumen of

126   the viral particles, whereas tubular structures were largely devoid of DNA (Figure S1B). Thus,

127   given that the tubular tails must have helical symmetry, and understanding the tails would

128   reveal the structure of the more complicated spindle body, most of the structural studies in this

129   paper were performed on the capsids which were transformed into a tubular state (see

130   Methods). Nevertheless, extensive analysis of millions of segments revealed a nearly

131   continuous variation in the diameter of the tail tubes. Progress was only made by starting from

132   a dataset containing six million particles extracted from 16,401 micrographs, allowing us to use

133   multiple 2D classification cycles to find nearly homogeneous subsets of constant diameter. This

134   classification approach showed that rather than being continuously variable, there were

135   discrete changes in diameter. We were then able to generate three-dimensional

136   reconstructions at a near-atomic level of resolution for 12 discrete diameters, ranging from 270

137   to 335 Å (Figure 1B), each having a different helical symmetry, but each with seven strands

138   (Table 1; Figure S2A). The α-helices in the map had a clear hand and showed unambiguously

139   that the seven strands were left-handed, as had been determined at low-resolution by cryo-

140   electron tomography and sub-tomogram averaging. The resolution of the best volume (Table 1,

141   symmetry 7, ~3.5 Å, Figure 1C) would be more than adequate to build a full-atomic model *ab*

142   *initio* when given the protein sequence.

143   To determine the identity of the MCP, SMV1 proteins were separated by SDS-PAGE and

144   the major protein band of ~14 kDa was subjected to trypsin digestion followed by mass

145   spectrometry analysis (Figure S3). The dominant protein in the band, gp11 (W0UU99), was the

146   homolog of the MCPs previously identified in all other bicaudaviruses (Erdmann *et al.*, 2014;

147   Hochstein *et al.*, 2016; Prangishvili *et al.*, 2006; Xiang *et al.*, 2005). However, it was impossible

148   to thread the sequence of this putative MCP (gp11) through the density map, trying both

149   possible N-to-C orientations. The absence of density in the map for a bulky side chain might be

150   explained by disorder of that residue, but no threading could account for the pattern of bulky

151   side chains seen in the map. Further, while the atomic structures of the ATV (PDB ID: 3FAJ) and

152   ATSV (PDB ID: 5EQW) putative MCPs each displayed a compact four-helix bundle with a short

153   fifth helix projecting out, the SMV1 capsid protein was clearly composed of two long helices

154   connected by a short turn. We therefore searched for other possible candidates. A C_α trace of

7

155   the density map suggested that the MCP is ~90 amino acids long. Assuming that there might be

156   disordered residues at either the N- or C-terminus that were not present in the density map, we

157   focused on SMV1 proteins containing between 90 and 150 residues. There were 17 such

158   candidates (out of 96 proteins), but none could be successfully threaded through the map.

159   We next performed the N-terminal sequencing of the proteins in the major gel band.

160   Two different five-residue N-terminal peptides were identified: one from a minor species,

161   VEDYF, corresponding to the N-terminus of gp11 (the putative MCP that did not fit the map),

162   and one from the major species, VTFGT, that did not correspond to the N-terminus from any of

163   the 96 ORFs in the SMV1 genome. A search of the ORFs, found this pentapeptide uniquely in

164   SMV1 gp03 (W0UUV5), a hypothetical 157-residue sequence annotated as a membrane

165   protein. Removing 59 N-terminal residues would place the pentapeptide at the N-terminus of a

166   processed protein, and the resulting 98-residue protein fit the density map perfectly (Figure

167   1D). The 98-residue protein contains no useful tryptic digestion sites (one at the C-terminus

168   would create a three-residue peptide), explaining why peptides from this protein were never

169   detected after tryptic digestion and mass spectrometry. The processing is likely performed by a

170   cellular protease, because SMV1 and other bicaudaviruses do not encode identifiable

171   proteases. Notably, however, gp03 does not contain a canonical signal sequence or

172   recognizable signal peptidase cleavage site (Figure S4A). The fit revealed that additional density

173   facing the outside of the capsid, near the turn between the two helices, was most likely due to

174   heavy N-linked glycosylation of Asn53. Since the predicted MW for the 98-residue MCP is

175   10.1kDa, we suggest that extensive glycosylation of this single residue can explain why the

176   protein runs in a gel with an apparent MW of ~14kDa. The glycosylation on the outside of the

177   virion may play a role in protecting the capsid against very aggressive conditions, as was shown

178   for extensively glycosylated pili of hyperthermophilic and acidophilic archaea (Wang et al.,

179   2020; Wang et al., 2019). To further validate this, we used trifluoromethanesulfonic acid

180   (TFMS), which removes both N- and O-linked glycans from proteins (Sojar and Bahl, 1987).

181   Conditions were found under which a significant shift to lower mass could be obtained for a

182   silver-stained band that ran at ~13 kDa before acid treatment and at ~11 kDa after acid

183   treatment. These two bands were excised from the gel and analysed by mass spectrometry,

184   which found that the gp03 can only be detected in the lower band. Interestingly, the protein

185   gp11 was detected in both bands (Figure S3B-D). Thus, we designate SMV1 gp03 as the actual

186   MCP. Given the abundance of gp11 in the virions of all known bicaudaviruses and lack of any

187   discerneable contribution to capsid formation, we hypothesize that this protein functions as a

188   nucleocapsid protein which compacts and, potentially, protects the viral DNA. Notably, in ATV,

189   gp42, one of the two ATV-encoded homologs of SMV1 gp11, has been shown to act as a strong

190   inhibitor of the host RNA polymerase (Sheppard et al., 2016), consistent with the possibility

191   that this soluble protein is injected into the cell upon viral genome delivery.

192

193   *Kinship among all spindle-shaped viruses*

194   We have imaged ATV by cryo-EM (Figure 2A) and carried out a similar analysis. As with

195   SMV1, ATV is formed from 7-start helical strands, and the subunit contains two α-helices linked

196   by a short turn. The repeat of the 7-start helices in ATV is 32 Å compared with the 29 Å in

197   SMV1. Although the resolution we have obtained for ATV (~6 Å) is much worse than for SMV1,

198   there is no ambiguity in building a homology model for the ATV gp59 (YP_319890) based upon

199   the SMV1 results, and this model fits well into the reconstructed density (Figure 2C). The crystal

200   structure for the putative MCP in ATV, gp62 (PDB id: 3FAJ)(Goulet *et al.*, 2010), cannot be fitted

201   into the density, and clearly is the structure of a different protein. Thus, we designate gp59 as

202   the actual MCP of ATV.

203   Homologs of SMV1 gp03 and ATV gp59 are conserved in all bicaudaviruses, including

204   STSV1 and STSV2 (Figure S4B). Based upon the results for SMV1 and ATV, everything suggests

205   that the actual MCP of ATSV is ATSV_F163, rather than ATSV_D135 for which an x-ray crystal

206   structure was determined (Hochstein *et al.*, 2018). The N-glycosylation sequon (NxS/T, where x

207   is any amino acid) located at the beginning of the second α-helix is conserved in all homologs

208   (Figure S4B), suggesting that glycosylation is important for virion assembly and/or stability.

209   Features of the bicaudavirus MCPs are remarkably similar to those of the MCPs of group I

210   spindle-shaped viruses, such as fusellovirus SSV1, halspivirus His1 and thaspivirus NSV1 (Figure

211   S4C). Notably, the MCP of SSV1 is encoded as a pre-protein which is proteolytically processed at

212   the N-terminus (Iverson et al., 2017; Quemin et al., 2015), as in the case of SMV1. Furthermore,

213   MCPs of fuselloviruses also contain the conserved N-glycosylation sites at the position

214   equivalent to that in bicaudaviruses and SSV1 MCP has been indeed shown to be glycosylated

215   (Quemin *et al.*, 2015). Finally, fuselloviruses, halspiviruses and thaspiviruses were all shown to

216   exist in tubular form, either upon virion egress or following genome ejection (Hong *et al.*, 2015;

217   Kim *et al.*, 2019; Quemin *et al.*, 2016). In light of this evidence, combined with the results of

218   bipartite network analysis showing that bicaudaviruses and group I spindle-shaped viruses

219   share several signature genes (Iranzo et al., 2016), we suggest that all archaeal spindle-shaped

220   viruses have evolved from a common ancestor and use similar principles of virion organization

221   (Figure 3C). However, details in both the protein subunits and the virion morphology differ, as

222   would be expected from evolutionary divergence. For instance, viruses STSV1 and STSV2 have

223   only one, relatively short tail and do not undergo morphological changes outside of the cell

224   (Erdmann *et al.*, 2014; Xiang *et al.*, 2005); ATSV has a very long and narrow tail of uniform

225   length (Hochstein *et al.*, 2018); whereas ATV and SMV1 analyzed in the present paper can

226   "grow" tails from both pointed ends of the spindle (Haring *et al.*, 2005; Uldahl *et al.*, 2016).

227   Strikingly, this emerging virus lineage might extend beyond spindle-shaped viruses to also

228   include rod-shaped viruses of the family *Clavaviridae* (Mochizuki et al., 2010). In particular, a

229   high resolution experimentally-determined structure exists for a rod-like clavavirus APBV1

230   virion (Ptchelkine et al., 2017), which is constructed from highly hydrophobic subunits with a

231   helical-hairpin structure and an N-glycosylation site at the same position as in SMV1 and other

232   bicaudaviruses (Figure 3B). Although MCPs of APBV1 and SMV1 do not display significant

233   sequence similarity, which is not unexpected due to high sequence divergence and sparse

234   sampling of archaeal viruses, the interactions between the MCP subunits in the two groups of

235   viruses are very similar, full of hydrophobic interactions (Figure 3B).

236

237   *How the virions expand*

238   The subunit model from "symmetry 7" of SMV1 (Figure 1D) fits the other 11 volumes

239   with no modifications or perturbations. In addition, changes in the local packing of the subunit

240   were almost infinitesimally small in the diameter range from 270 to 335 Å (Figure 4,

241   Supplemental Movies 1 and 2), showing how the diameter could change significantly while at

242   the same time preserving a highly conserved local environment for each subunit. The RMSD

243   between two adjacent subunits in the same strand comparing the smallest and the largest

244   diameter was only 0.3 Å (Figure 4A), while the RMSD between subunits in adjacent strands was

245   only 0.4 Å (Figure 4B) when comparing the smallest and largest diameter.

246   Helical symmetry can be defined as a rise per subunit, a rotation per subunit, and a

247   possible C_n rotational symmetry. Two of the 12 structures had a C_7 rotational symmetry (Table

248   1, symmetries 1 and 8), so the rise per subunit in each strand is the same as the helical rise. For

249   the other symmetries in Figure 4C we have multiplied the helical rise by seven to yield the rise

250   per subunit within a strand. The determined diameters for these 12 points fall into a small

251   range, considering that the spindle bodies can have a maximum diameter of ~1,200 Å. Because

252   the 29 Å periodicity remains constant in both spindle body and tail, and this comes from the

253   repeat of the 7-start helices which have a fixed 203 Å pitch (7 x 29 Å) in both the tails and the

254   spindle body, the number of subunits per strand in one turn will be u, which is 203 Å divided by

255   the rise per subunit along a strand. We have found for the 12 symmetries reconstructed that u

256   is incremented by seven for each larger diameter (which is an increment of one per 29 Å),

257   explaining why the changes in diameter, D, are discrete, and not continuous. Let p be the path

258   length along a 7-start strand per 203 Å turn, then:

259        $p^2 = (\pi D)^2 + 203^2$

260   We have shown that the local packing does not change as the diameter changes (Figure

261   4, Supplemental Movies 1,2), so the spacing between subunits along a strand, q, can be taken

262   as a constant as the diameter changes, and $p = uq$. From the observed tubes, $q=18.8$ Å. Given

263   this, the rise per subunit and the diameter can be easily calculated for each successive value of

264   *u*. Therefore, we can extrapolate the results from the tubes to the larger diameters that would

265   be found in the spindles (Figure 4C, Supplemental Movie 3).

266   How can the subunit structure explain the remarkable ability of these strands to slide

267   past each other in discrete steps? An analysis of the protein sequence (Figure 5A) shows that

268   most of the residues are quite hydrophobic. A plot of transmembrane probabilities for the

269   sequence indicates a strong prediction for two transmembrane helices (Figure 5B). This can be

270   seen in terms of the actual structure, where there is a hydrophobic region in the center of the

271   subunit that is ~35 Å thick, with polar regions below and above it (Figure 5C). Therefore, these

272   capsid proteins would not be soluble as monomers and must be integral to the host membrane

273   before the virions are formed. When packed into virions (Figure 5D), the hydrophobic regions

274   are completely buried and the polar regions are facing either the outside or the lumen. Notably,

275   the interface between neighboring 7-start strands is almost completely hydrophobic, with

276   minimal interactions between the hydrophilic regions. This will allow these strands to slide by

277   each other while still excluding solvent (Supplemental Movies 4,5). Consistent with this,

278   treatment of the virions with detergent, such as N-lauryl sarcosine, leads to rupture of the

279   virions (Figure S5A), much as such treatment would destabilize a membrane.

280   The extensive hydrophobicity of the capsid and likely insertion of the protein subunits

281   into the membrane prior to virion assembly motivated us to look for possible lipids associated

282   with the virions (Figure S5B). The distribution of lipids in the SMV1 virions was not the same as

283   that found in the host, suggesting that these lipids were selectively incorporated rather than

284   being contaminants from cellular debris. In bacterial mating pili, the subunits also have a

285   helical-hairpin structure (Costa et al., 2016; Zheng et al., 2020) and these subunits exist as

13

286   integral membrane proteins before polymerization and after depolymerization. A single lipid

287   molecule has been observed tightly bound to each subunit in the three published structures.

288   However, we found no ordered lipid molecules in the SMV1 tube reconstructions. We can

289   exclude the possibility that disordered lipid molecules (that we do not visualize) are present

290   between the hydrophobic helices as the space between these helices is too narrow to

291   accommodate a lipid. Since the tube segments used for three-dimensional reconstruction were

292   carefully screened to only include the most ordered and regular ones, we are thus left with the

293   possibility that the lipids may be found in the many dislocations or irregular regions in the

294   virions, in particular, where additional subunits are added to helical turns as the diameter

295   becomes larger. Alternatively, lipids could be associated with the specialized structures that are

296   present at the termini of SMV1 virions.

297

298

Discussion

299   **Discussion**

300   We have shown how a highly hydrophobic subunit can assemble into tubes with a

301   variable diameter, and how these tubes can seamlessly expand and contract to form a spindle-

302   shaped body for a virion that maintains the same quasi-equivalent contacts between subunits

303   that are present in the tubes. The hydrophobic effect is, by definition, non-specific (Hillyer and

304   Gibb, 2016; Newberry and Raines, 2019), and the hydrophobic surfaces of the helices (Figure 5)

305   allow them to slide past each other (Supplemental Movies 4,5) while still maintaining the

306   integrity of the capsid and excluding very acidic solvent in the environment from the interior of

307   the virion. Our expectation is that the interior of the virions is close to the same near-neutral

308   pH as the cytoplasm of the host (Baker-Austin and Dopson, 2007) and that is why the capsid

309   must be impermeable to acidic solvent. While the hydrophobic effect alone is non-specific,

310   specificity in the intersubunit contacts is provided by the helix-helix contacts that exist between

311   neighboring subunits which dictate that discrete changes in diameter are required to maintain

312   these quasi-equivalent contacts (Supplemental Movies 4,5). The notion of quasi-equivalence

313   was first introduced by Caspar and Klug to explain how icosahedral viruses could be assembled

314   from many copies of a single protein (Caspar and Klug, 1962). If one only had 60 copies of such

315   a capsid protein a perfect icosahedron could be assembled so that every subunit was in an

316   equivalent environment. It was also known that it was impossible to place more than 60 copies

317   of a protein on the surface of a sphere such that all would be in strictly equivalent

318   environments. But a capsid built from only 60 protein subunits would be quite small, and it had

319   been observed that larger icosahedral viruses contained many more than 60 copies of the

320   capsid protein. The solution was found in the possibility of quasi-equivalence (Caspar and Klug,

321   1962), wherein multiples of 60 protein subunits could be assembled into a spherical shell

322   maintaining not identical, but quasi-equivalent contacts. This has now been amply confirmed

323   with high resolution structures for many icosahedral viruses (Damodaran et al., 2002; Johnson

324   and Olson, 2021; Johnson and Speir, 1997), although it has also been shown that certain

325   icosahedral viruses can violate quasi-equivalence (Rayment et al., 1982; Stehle et al., 1994).

326   However, the extension of quasi-equivalence theory to the radial expansion of helical viruses

327   has not been envisioned.

328   Here we show that the quasi-equivalence principle explaining the ability of icosahedral

329   viruses to expand the number of subunits needed to encapsulate larger genomes has a striking

330   parallel in helical viruses. In APBV1 or other rod-like viruses, there will be a stoichiometric ratio

331   between the number of capsid proteins and the number of bases or base pairs encapsidated. As

332   pointed out by Caspar and Klug (Caspar and Klug, 1962), the length of tobacco mosaic virus

333   virions is simply determined by the length of the single-stranded RNA genome given the

334   stoichiometric ratio of three bases per protein subunit. In filamentous bacteriophage, the

335   packaging of DNA involves charged residues in the lumen of the phage, and thus modifications

336   of these residues can change the length of the virions while the length of the genome is

337   unaltered (Hay and Lithgow, 2019). Thus, there is a linear relation between the length of the

338   virion and the size of the genome, even though the actual stoichiometry can be modified. In

339   contrast, for spherical viruses the number of capsid subunits will be proportional to the surface

340   area, which is the radius squared, while the size of the genome encapsidated will scale as the

341   radius cubed. Thus, the ratio of base pairs per capsid subunit will scale as the radius, becoming

342   more efficient in genome packaging the larger the capsid.

343   With hydrophobic subunits having the ability to slide past each other, it is easy to

344   imagine how a rod-like virus such as APBV1 could evolve into a spindle-shaped virus (Figure 3C)

345   while still maintaining all of the local interactions that hold the capsid together. We suggest

346   that the minimum energy conformation of the spindle-shaped capsids is actually a tube, as

347   evidenced by the fact that when the genome is released, the SMV1 spindles relax into tubes

348   (Figure S1). The same spindle-to-tube transformation has been observed upon genome ejection

349   for a group I spindle-shaped virus His1 (Hong *et al.*, 2015). It is therefore the pressure of the

350   genome that expands the capsid that would normally be rod-like into a spindle shape. Thus, the

351   unusual spindle morphology results from the radial expansion of a rod.

352   Archaeal and bacterial gas vesicles (Li and Cannon, 1998; Pfeifer et al., 2002; Strunk et

353   al., 2011) have some remarkable similarity in morphology to spindle-shaped viruses. The gas

354   vesicles are made of a single major structural protein of ~8 kDa and have a hydrophobic interior

355   surface. However, whereas the capsid proteins contain two hydrophobic α-helices, those of gas

356   vesicles from both bacteria and archaea have a α-β-β-α structure (Knitsch et al., 2017; Strunk *et*

357   *al.*, 2011), and the β-strands are predicted to mediate subunit-subunit interactions. Finally,

358   unlike the capsid proteins which are predicted to contain two transmembrane domains, the

359   proteins of gas vesicles are predicted to have none. Thus, we conclude that the superficial

360   similarity between spindle-shaped viruses and gas vesicles is a result of convergent evolution.

361   However, given the hydrophobic nature of both structures, we predict that changes in diameter

362   described for both are mediated by these hydrophobic interactions.

363   There has been great interest in understanding the role of BAR domain proteins (Frost

364   et al., 2008), ESCRT-III (Endosomal Sorting Complexes Required for Transport) (Pfitzner et al.,

365 2021) and dynamin family proteins (Antonny et al., 2016) in membrane remodeling. All of these

366 proteins form helical tubes of variable diameters that have been shown to constrict as part of

367 their role in processes such as membrane fission and vesiculation. Structural studies of these

368 assemblies have been extremely difficult due to the variability of tube diameters, and we still

369 do not have any models at an atomic level of detail for the elementary steps in constriction.

370 Similarly, keratin filaments, assembled from α-helical coiled-coils, have also been revealed to

371 form variable diameter tubes *in situ* (Weber et al., 2021). While we do not suggest any

372 homology between these proteins and the SMV1/ATV capsid proteins, we suggest that the

373 general principle underlying the sliding of the strands of SMV1 proteins past each other while

374 maintaining the structural integrity of the capsid might extend to other filamentous assemblies

375 with variable diameters. The great advances in cryo-EM that have allowed us to determine in

376 atomic detail the structure of SMV1 capsid will likely allow a similar level of understanding for

377 other assemblies in the near future.

378

379 *Limitations of the Study*

380 We have determined a high resolution structure for only a single spindle-shaped virus, SMV1,

381 and determined a lower resolution structure for a second spindle-shaped virus, ATV. These

382 results have been combined with an existing high resolution structure for the clavavirus APBV1

383 (Ptchelkine *et al.*, 2017) and homology models of the major capsid proteins from other spindle-

384 shaped viruses to show the likely common ancestry and suggest how the spindle-shaped viruses

385 have evolved from rod-like viruses such as APBV1. Future high resolution studies of other rod-

386 like and spindle-shaped viruses will greatly extend our understanding of this evolution. Just as

387   knowing the morphology and genetics of apes and humans suggests common ancestry,

388   determining the intermediates that existed in this evolutionary process gives one a much more

389   profound understanding of biological history.

390

405

406

407

408 **Author Contributions**

409 M.K. initiated and developed the project. F.W. performed microscopy, image analysis,

410 reconstructed the SMV1 and ATV tubes, and generated and refined the atomic models. V.C-K.

411 cultured the viruses, developed the conditions for tube formation, and performed sample

412 preparation and fractionation analyses. M.V. and J-M.W. performed microscopy. Z.S. analyzed

413 data and developed graphical representations. L.C.B. did the deglycosylation of the MCP.

414 M.A.B.K., M.V. and J-M.W. collected the tomographic data. J.L. analyzed the tomographic data

415 and generated the three-dimensional reconstruction. S.S. analyzed the lipids. E.H.E. obtained

416 funding, analyzed data and supervised the research. F.W., M.K. and E.H.E. wrote the manuscript

417 with input from all authors.

418
419 **Declaration of Interests**
420
421 The authors declare no competing interests.
422
423

Figure Legends

Figure 1. Cryo-EM of bicaudavirus SMV1

A, Representative cryo-EM of the Sulfolobus monocaudavirus 1 (SMV1). A 29 Å periodicity can be seen in both spindle-shaped bodies and tails. Scale bar, 50 nm.

B, Cryo-EM reconstructions of the SMV1 tails made from 12 different diameters. The top view models of the largest (symmetry 12) and smallest (symmetry 1) diameters are shown.

C, A surface of the 3.5 Å resolution reconstruction of symmetry 7. The left-handed 7-start protofilaments and 29 Å spacing between them are labeled.

D, Ribbon model of a single MCP protein (left). The map quality of the subunit is shown on the right. The arrow points to the N-linked glycosylation associated with Asn53.

Figure 2. Cryo-EM of the Acidianus two-tailed virus (ATV)

A, Representative cryo-EM of ATV. A 32 Å periodicity can be seen in both spindle-shaped bodies and tails, indicated by black arrows. Scale bar, 50 nm.

B, Surface of the cryo-EM reconstruction of the ATV tails for one diameter, with the 7-start protofilaments labeled.

C, The homology model of ATV fit into the cryo-EM map in side and top views.

Figure 3. Spindle-shaped viruses evolved to carry larger genomes

A, Single subunit structures of seven MCPs. The structures of the APBV1 and SMV1 MCPs have been determined by high-resolution cryo-EM (PDB 5OXE and 7ROC). The other five MCP structures are homology models based on the SMV1 structure (see Methods).

B, Comparison of hydrophobic interactions of APBV1 and SMV1.

C, The estimated internal volumes of seven viruses plotted against genome size.

Figure 4. The local contacts between capsid subunits are conserved as the diameter changes

A, Interactions along the 7-start protofilament. One protofilament from the tube is highlighted on the left. Two subunits in one protofilament from symmetry 1 and 12 are aligned (center) and compared (right).

B, Interactions between adjacent 7-start protofilaments. Two protofilaments are highlighted on the left. The contacts between three subunits, two on the bottom strand (S_a and S_b) and one on the top (S'_a) are compared in two views (center and right).

C, The rise per subunit along a strand for 12 different symmetries is plotted versus the diameter of the tubes (orange spheres), showing a nearly linear relationship. The blue dashed line shows the extrapolation to both smaller and larger diameters.

Figure 5. The SMV1 MCP has clustered lipophilicity

A, Amino acid composition of the SMV1 MCP.

B, TMHMM prediction of transmembrane probabilities for the SMV1 MCP sequence.

C, The spatial distribution of hydrophobic and hydrophilic residues in a single MCP.

D, Different views of the SMV1 tubes colored by lipophilicity.

Table 1. Helical symmetry and model statistics for SMV1 and ATV

Symmetry ID	SMV1												ATV
	1	2	3	4	5	6	7	8	9	10	11	12	---
Helical symmetry													
Point group	C7	C1	C1	C1	C1	C1	C1	C7	C1	C1	C1	C1	C1
Helical rise (Å)	5.15	0.72	0.70	0.68	0.67	0.66	0.64	4.38	0.61	0.60	0.59	0.58	1.08
Helical twist (°)	-9.0	50.2	-155.5	-104.1	101.7	153.1	-52.6	-7.7	50.4	-155.3	-103.9	101.8	-156.0
Map resolution (Å)													
Model:map FSC (Å, 0.5)	5.1	4.8	4.3	4.1	4.1	3.9	3.7	3.8	3.7	3.8	4.0	4.3	---
Map:map FSC (Å, 0.143)	4.4	4.3	3.8	3.8	3.8	3.8	3.7	3.7	3.4	3.5	3.6	4.0	6.3
Refinement and Model validation													
Bond lengths RMSD (Å)	0.004	0.004	0.006	0.005	0.006	0.004	0.003	0.004	0.006	0.007	0.005	0.004	---
Bond angles RMSD (°)	0.658	0.640	0.770	0.841	0.671	0.699	0.635	0.703	0.681	0.735	0.651	0.658	---
Clashscore	12.4	9.9	12.0	9.3	10.0	8.4	9.0	10.9	9.2	11.0	8.7	7.4	---
Ramachandran Favored (%)	96.7	95.7	94.6	95.7	94.6	95.7	93.4	95.0	93.4	94.6	94.6	95.7	---
Ramachandran Outlier (%)	0	0	0	0	0	0	0	0.2	0	0	0	0	---
RSCC	0.86	0.87	0.87	0.87	0.88	0.87	0.87	0.88	0.87	0.87	0.86	0.87	---
Deposition ID													
PDB (model)	7RO2	7RO3	7RO4	7RO5	7RO6	7ROB	7ROC	7ROD	7ROE	7ROG	7ROH	7ROI	---
EMDB (map)	24585	24586	24587	24588	24589	24590	24591	24592	24593	24594	24595	24597	24596

473

STAR Methods

475

RESOURCE AVAILABILITY

477

Lead Contact

479 Further information and requests for resources and reagents should be directed to and will be

480 fulfilled by the lead contact, Edward H. Egelman, egelman@virginia.edu

481 **Materials Avalability**
482
483 • This study did not generate any new unique reagents.

484
485 **Data and Code Availability**

486 • The SMV1 maps were deposited in the Electron Microscopy Data Bank (EMDB) with

487 entry codes EMD-24585, EMD-24586, EMD-24587, EMD-24588, EMD-24589, EMD-

488 24590, EMD-24591, EMD-24592, EMD-24593, EMD-24594, EMD-24595, and EMD-24597

489 and the respective atomic models were deposited in the Protein Data Bank (PDB) with

490 entry codes 7RO2, 7RO3, 7RO4, 7RO5, 7RO6, 7ROB, 7ROC, 7ROD, 7ROE, 7ROG, 7ROH

491 and 7ROI. The ATV map was deposited in the EMDB with entry code EMD-24596.

492
493 • This paper does not report original code.
494

495 • Any additional information required to reanalyze the data reported in this paper is

496 available from the lead contact upon request

497
498
499

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Archaeal strains and growth conditions

Sulfolobus monocaudavirus 1, SMV1, was propagated in *Sulfolobus islandicus* CRISPR deletion

mutant delta C1C2 (Gudbergsdottir et al., 2011). The host cell culture was grown in Sulfolobus

medium supplemented with 0.2% tryptone, 0.1% yeast extract, 0.2% sucrose and 0.002% uracil

(Zillig et al., 1993). The culture was started from −80°C stock. The cells were grown in 50 ml of

medium at 76°C with shaking. After 24 h of incubation, the cell culture was diluted 20 times in

pre-warmed (76°C) medium and the growth was continued until it reached an OD_{600} of ~0.2.

Then the cells were infected with SMV1 stock and incubation continued at 76°C with agitation

for 24h. 24 h post infection the cells were removed by centrifugation in Sorval 3000 rotor, 7000

rpm, 20 min, 15°C, virus-containing supernatant was collected and virions precipitated with

PEG600 (10.5% w/v) and NaCl (5.8% (w/v) for 2h at room temperature. The PEG-precipitate was

then pelleted in a Sorval 3000 rotor, 9000 rpm, 30 min, 15°C and the resulting pellet was re-

suspended in one tenth of the original volume in 20 mM Tris-acetate (pH6) buffer.

METHODS DETAILS

Virus production and treatments

To purify the virus, solid CsCl was added to the virus suspension to a final concentration of 45%

(w/v). The samples were run in a Beckman SW41 rotor, 38,000 rpm, at least 16 h, 15°C. After the

run, the opalescent band corresponding to SMV1 virions was collected. The purified SMV1 was

stored in CsCl at 4°C until used.

522   For SMV1 virion dissociation, CsCl purified virions were diluted 50 times in 20 mM Tris-acetate

523   (pH6) buffer and pelleted down by centrifugation in a Beckman 45Ti rotor, 35,000 rpm, 2 h, 15°C.

524   The resultant pellet was re-suspended in 20mM Tris-acetate (pH6) buffer and used for

525   dissociation. The SMV1 virions were dissociated by addition of 0.1% (v/v) NaOH and incubation

526   at 37°C for 1 h. Then non-treated (control) and NaOH-treated virions were loaded on top of the

527   linear 5-20 % sucrose gradients and run in a Beckman SW41 rotor, 24,000 rpm, 20 min, 15°C.

528   After the run, 12 fractions from the top of each tube were collected and analyzed in 4-12% Bis-

529   Tris acrylamide gels. After the run the gels were stained for DNA with ethidium bromide and for

530   proteins with InstantBlue. The contents of the collected fractions were also visualized by negative

531   staining TEM. For that, 10 µl of the fraction of interest was adsorbed onto copper grids with

532   carbon-coated Formvar films and negatively stained with 2.0% (w/v) uranyl acetate for 20 s. The

533   samples were observed under a FEI Tecnai BioTwin 120 microscope operated at 120 kV.

534

535   **Cryo-electron microscopy and image processing**

536   The SMV1 sample was applied to glow-discharged lacey carbon grids and vitrified using a

537   Vitrobot Mark IV (Thermo Fisher). Grids were imaged at on a Titan Krios (300 keV, Thermo

538   Fisher) with a K3 camera (Gatan). 16,401 micrographs were collected under electron counting

539   mode at 1.1 Å per pixel, using a defocus range of 1–2 µm with ~50 electrons/$Å^2$ distributed into

540   40 fractions. Motion correction and CTF estimation were done in cryoSPARC (Punjani et al.,

541   2017; Rohou and Grigorieff, 2015; Zheng et al., 2017). A total of seven million particles were

542   auto-picked by "Filament Tracer" with a shift of 13 pixel, and non-virion bad particles were

543   removed by 2D classification. About six million SMV1 particles remained having a range of

544   diameters. From this six million particle dataset, 12 relatively homogeneous subsets were

545   sorted out by iterative 2D classifications. The possible helical symmetries were calculated from

546   an averaged power spectrum for each subset, generated from the raw particles. For each

547   subset, the actual helical symmetry was determined in cryoSPARC by trial and error, until the

548   hand of α-helices and amino acid side chains were seen (Egelman, 2000; Punjani et al., 2020).

549   The resolution of each reconstruction was estimated by both Map:Map FSC and Model:Map

550   FSC. The final volumes were then sharpened with a negative B-factor automatically estimated

551   in cryoSPARC, and the statistics are listed in Table 1.

552   For ATV the same approach was used, with 2.8M particle images collected using a shift of 9

553   pixels between adjacent boxes.

554

555   **Model building of SMV1 MCP**

556   The density corresponding to a single SMV1 MCP was segmented from the experimental cryo-

557   EM density using Chimera (Pettersen et al., 2004). Possible MCP sequences were threaded

558   through the density using DeepTracer (Pfab et al., 2021). Only one sequence, W0UUV5, also

559   detected by N-terminal sequencing, could be threaded through the map, matching the bulky

560   side chain densities and the glycosylation site. This model was adjusted manually in Coot

561   (Emsley and Cowtan, 2004) and real-space refined in PHENIX (Afonine et al., 2018). Using the

562   determined helical symmetry, a filament model was generated in Chimera and refined against

563   the full cryo-EM map using PHENIX real-space refinement. MolProbity was used to evaluate the

564   quality of the filament model (Williams et al., 2018). The refinement statistics are shown in

565   Table 1.

566

Cryo-electron tomography

For cryo-electron tomography, CsCl-purified virions were pelleted by ultracentifugation

(Beckman Type 50.2 Ti rotor, 32000 rpm, 2h, 15 °C), resuspended in water and incubated at 90

°C for one week. Heat-treated SMV1 particles were mixed with CsCl-purified rudivirus SIRV2 as

a control, since the helical hand had been determined previously (DiMaio et al., 2015). Data

were collected on a Titan Krios microscope (TFS) operating at 300kV in EFTEM mode using a

bioquatum/K3 (Gatan) energy filter and camera with a 20 eV slit. Tilt series were recorded at a

magnification of 26kx corresponding to 3.4 Å/pixel. A tilt range of +/-60° was set with a 3° tilt

increment and a total dose of 120 e$^-$/Å^2 with a dose symmetric tilting scheme using the TFS

tomography software with defocus values between -4 and -6 micron defocus.

A total of 935 segments were manually picked along the long tails of SMV1 from 8×

binned tomograms by "tomopick" in the tomography package i3 (Winkler, 2007). Sub-

tomograms of SMV1 were first extracted from the 8× binned tomograms, and then aligned in

the i3 software package. After initial alignment, sub-tomograms were extracted from unbinned

tomograms for further refinement. Multivariate statistical analysis (MSA) implemented in the I3

package was then used for 3D classification and sub-tomogram averaging. As a control, a total

of 200 segments were manually picked from the tubes of SMV2 from the same 8× binned

tomograms. The similar procedure was used to generate a 3D averaged structure of the SIRV2

tube without applying any helical symmetry. The averaged structure of SMV1 is a left-handed

helix with ~29 Å periodicity, while the averaged structure of SIRV2 is a right-handed helix with

~43 Å periodicity.

588

Homology modeling of ATV, SSV1, His1 and ATSV MCP

The MCP homologs in bicaudaviruses were identified by PSI-BLAST (Altschul et al., 1997)

searches (E-value cutoff of 0.05) queried with the sequence of SMV1 MCP (YP_009008070)

against the NCBI protein database restricted to members of the family *Bicaudaviridae*

(taxid:423358). The sequences were aligned using PROMALS3D (Pei and Grishin, 2014). MCPs of

SSV1 and His1 were determined experimentally (Pietila et al., 2013; Quemin *et al.*, 2015). The

MCP structures of ATV, SSV1, His1 and ATSV were first predicted by AlphaFold2 (Jumper et al.,

2021). Then the predicted structures were flexibly aligned to SMV1 by FATCAT (Li et al., 2020).

For ATV, the predicted MCP structure was then docked into the 6 Å resolution cryo-EM map

and real-space refined in PHENIX (Afonine *et al.*, 2018) with helical symmetry imposed.

599

N-terminal sequencing of SMV1 by Edman Degradation

The analysis was performed on an ABI Procise 494 sequencer. The sample was placed on a

PVDF membrane for the Edman degradation with a cyclic procedure where residues were

cleaved off one at a time and identified by chromatography. There were three steps in each

cycle. In step 1 the PITC reagent was coupled to the N-terminal amino group under alkaline

conditions. In step 2 the N-terminal residue was cleaved in acidic media. In step 3, the PITC-

coupled residue was transferred to a flask, converted to a PTH-residue and identified by HPLC

chromatography. The cycle was then started again for the identification of the next N-terminal

residue.

609

SMV1 deglycosylation reaction and tandem mass spectrometry

Briefly, 50 μl of concentrated SMV1 were lyophilized. Then, 150 μl of TFMS was added to the

tube and the mixture was incubated at 4 °C for 24 h. Following this, 150 μl of 60% pyridine

solution, cooled to ~15 °C with a methanol dry ice bath, was added to the reaction tube,

neutralizing the TFMS acid. Using a Slide-A-Lyzer dialysis cassette with a 2,000 Da protein

molecular weight cut-off (Thermo Fisher Scientific), the reaction solutions were removed from

the sample with overnight dialysis at 4 °C into Tris/HCl (pH 8) buffer. After dialysis, small

amounts of aggregates—presumably deglycosylated SMV1—were observed in the cassette.

Centrifugation for 15 min at 4 °C and 20,000 x g was used to pellet the aggregates. The pellet

was resuspended with 50 μl of Tris buffer (pH 8). SDS-PAGE was then performed at a constant

voltage of 120 V using 16.5% precast polyacrylamide Mini-PROTEAN Tris-Tricine precast gels

(Bio-Rad) and Tris/Tricine/SDS running buffer (Bio-Rad). For each sample, 15 μl of boiled sample

in SDS was added to separate wells in the gel, and 5 μl of Precision Plus Protein Dual Xtra

protein standard (Bio-Rad) was used as a marker. Silver staining was performed using a Pierce

Silver Stain for Mass Spectrometry kit (Thermo Fisher Scientific). The detected gel bands were

then excised, and mass spectrometry (Alphalyse) was performed to analyze the composition of

each band. The protein samples were reduced and alkylated with iodoacetamide, and

subsequently digested with chymotrypsin. The resulting peptides were concentrated by Speed

Vac lyophilization. The peptides were dissolved in 0.1% formic acid and injected on a Dionex

Ultimate 3000 nano-LC system (Thermo Scientific) coupled to a Bruker Maxis Impact QTOF mass

spectrometer for MS/MS analysis. Each sample underwent a 30-minute-long gradient run on

the instruments. As an external control of the instrument, tryptic BSA was analysed together

29

632   with the samples. The results were run against the Sulfolobus monocaudavirus SMV1 proteins

633   from the UniProt.

634

635   **Analysis of SMV1 and host cell lipids**

636   For lipid analyses, CsCl-purified virions were pelleted by ultracentifugation (Beckman Type 50.2

637   Ti rotor, 35000 rpm, 2h, 15 °C), resuspended in 20 mM Tris-acetate (pH6) and run on 5-20 %

638   (wt/vol) sucrose gradient (Beckman SW32 Ti rotor, 24000 rpm, 30 min, 15 °C). The opalescent

639   virus-containing band was collected, pelleted as above and resuspended in 20 mM Tris-acetate

640   (pH6). The cellular and viral lipids were analyzed by UHPLC-MSD at NIOZ using an Agilent 1290

641   Infinity II ultra-high performance LC coupled to a 6230 Agilent MSD in selected ion mode as

642   described (Besseling et al., 2020).

643
644
645

Supplemental Figures

Figure S1. Related to Figure 1. The pH shift to induce SMV1 to form long tails

A, Cryo-EM of SMV1 before and after the NaOH (pH=10) treatment. Scale bars are 50 nm. The arrow indicates a tube containing a viral genome, while all of the other tubes in this image are empty.

B, SDS-PAGE and negative staining analysis of different fractions from the 5-20% sucrose gradient which were loaded with non-treated virus (left) and NaOH-treated virus (right). The non-treated virus migrated to the bottom of the gradient, where most of the protein and DNA were found (fractions 8-12), whereas following NaOH treatment, the viral DNA was found in fraction 1 in the form of amorphous aggregates, and the virion tubes were enriched in fraction 3.

Figure S2. Related to Figure 1 and Table 1. Cryo-EM volumes and helical symmetries for all 12 reconstructions of SMV1

A, The symmetry ID, helical rise, helical twist and point group symmetry are labeled.

B, The "gold-standard" half map FSCs were calculated in cryoSPARC using a 0.143 criterion.

Figure S3. Related to Figure 1. TFMS deglycosylation analysis of SMV1

A, SDS-PAGE with Coomassie blue staining of SMV1. The MW marker lane M is labeled on the left. A single band at ~ 13-14 kDa is observed for SMV1.

B, SMV1 samples on silver-stained SDS-PAGE before and after deglycosylation using TFMS. M indicates the molecular weight markers lane.

C, The band indicated by blue arrowhead in (B) was cut and used for tandem mass spectrometry analysis. Protein W0UU99 was detected.

D, The band indicated by red arrowhead in (B) was cut and used for tandem mass spectrometry analysis. Both W0UU99 and the correct capsid protein, W0UUv5, were detected.

Figure S4. Related to Figure 3. Sequence analysis and alignments

A, SignalP (Almagro Armenteros et al., 2019) analysis did not reveal any signal peptides or signal peptidase cleavage sites in the cleaved N-terminus.

B, The N-terminal regions which are processed in the mature proteins are not shown. The start site experimentally determined for SMV1 is indicated with an arrowhead. The secondary structure determined for the SMV1 protein is shown above the alignment, with the red ribbon representing α-helices. The N-glycosylation seqon is boxed, with the target Asn residue indicated with an asterisk. Each sequence is identified with the corresponding GenBank accession number followed by the virus name.

C, The predictions of transmembrane helices within the MCPs of seven viruses were made using the TMHMM server v2.0.

Figure S5. Related to Figure 5. Sensitivity to detergent and presence of lipids in SMV1 virions

 A, Transmission electron micrographs of negatively stained SMV1 virions treated with 0.1% N-lauryl sarcosine. Bars, 200 nm. **B,** Distribution of the lipid species identified in SMV1 host,

31

689   *Saccharolobus islandicus*, and SMV1 virions. GDGT, glycerol dibiphytanyl glycerol tetraether

690   lipids; GTGT, glycerol tribiphytanyl glycerol tetraether lipids. Numbers following GDGT or GTGT

691   represent the number of cyclopentane rings present in the lipid species.

692

693

694   **Supplementary Movie 1. Related to Figure 4.**
695

696   An alignment of one subunit (yellow) between two adjacent diameters: symmetry 7 (cyan) and
697   symmetry 8 (magenta).
698

699   **Supplementary Movie 2. Related to Figure 4.**
700

701   An alignment of one subunit (yellow) between two very different diameters: symmetry 2
702   (orange) and symmetry 12 (green).
703

704   **Supplementary Movie 3. Related to Figure 4.**
705

706   The helical symmetry established for the tubes and extrapolated to the spindle-shaped bodies
707   (Fig. 4C) has been used to construct a model for the full capsid. The seven strands of subunits
708   are each shown in a different color. Although discrete subunits are not seen in this visualization,
709   98,000 copies of the capsid protein have been used to generate this model. The basis of the
710   model is that the local packing throughout the structure is fixed with a spacing of 2.7 Å
711   between subunits along a strand, and the helical pitch of each strand is also fixed at 203 Å. As
712   the diameter increases due to internal pressure from the packaged genome, the strands slide
713   past each other.
714

715   **Supplementary Movie 4. Related to Figure 5.**
716

717   The role of hydrophobic (lipophilic) surfaces in the sliding of SMV1 strands past each other. This
718   animation shows the continuous transformation of the same number of total subunits from the
719   narrowest diameter tube to the widest diameter tube of the 12 that we have reconstructed. As
720   the tube becomes wider, more subunits are added per turn to the left-handed 7-start helices.
721   The hydrophobic surfaces (gold) at the interfaces between the strands allow for this sliding to
722   take place, at the same time that solvent is excluded from passing through the capsid.
723

724   **Supplementary Movie 5. Related to Figure 5.**
725   The top view of Supplemental movie 4.

726

References

728 Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., and Adams,

729 P.D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr

730 D Struct Biol *74*, 531-544. 10.1107/S2059798318006551.

731 Almagro Armenteros, J.J., Tsirigos, K.D., Sonderby, C.K., Petersen, T.N., Winther, O., Brunak, S.,

732 von Heijne, G., and Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using

733 deep neural networks. Nat Biotechnol *37*, 420-423. 10.1038/s41587-019-0036-z.

734 Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J.

735 (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.

736 Nucleic Acids Res *25*, 3389-3402.

737 Antonny, B., Burd, C., De Camilli, P., Chen, E., Daumke, O., Faelber, K., Ford, M., Frolov, V.A.,

738 Frost, A., Hinshaw, J.E., et al. (2016). Membrane fission by dynamin: what we know and what

739 we need to know. EMBO J *35*, 2270-2284. 10.15252/embj.201694613.

740 Baker-Austin, C., and Dopson, M. (2007). Life in acid: pH homeostasis in acidophiles. Trends in

741 microbiology *15*, 165-171. 10.1016/j.tim.2007.02.005.

742 Baquero, D.P., Liu, Y., Wang, F., Egelman, E.H., Prangishvili, D., and Krupovic, M. (2020).

743 Structure and assembly of archaeal viruses. Advances in virus research *108*, 127-164.

744 10.1016/bs.aivir.2020.09.004.

745 Besseling, M.A., Hopmans, E.C., Bale, N.J., Schouten, S., Damste, J.S.S., and Villanueva, L. (2020).

746 The absence of intact polar lipid-derived GDGTs in marine waters dominated by Marine Group

747 II: Implications for lipid biosynthesis in Archaea. Sci Rep *10*, 294. 10.1038/s41598-019-57035-0.

748 Caspar, D.L., and Klug, A. (1962). Physical principles in the construction of regular viruses. Cold

749 Spring Harb. Symp. Quant. Biol. *27*, 1-24.

750 Costa, T.R., Ilangovan, A., Ukleja, M., Redzej, A., Santini, J.M., Smith, T.K., Egelman, E.H., and

751 Waksman, G. (2016). Structure of the Bacterial Sex F Pilus Reveals an Assembly of a

752 Stoichiometric Protein-Phospholipid Complex. Cell *166*, 1436-1444 e1410.

753 10.1016/j.cell.2016.08.025.

754 Crick, F.H., and Watson, J.D. (1956). Structure of small viruses. Nature *177*, 473-475.

755 10.1038/177473a0.

756 Damodaran, K.V., Reddy, V.S., Johnson, J.E., and Brooks, C.L., 3rd (2002). A general method to

757 quantify quasi-equivalence in icosahedral viruses. J. Mol. Biol. *324*, 723-737. 10.1016/s0022-

758 2836(02)01138-5.

759 Dellas, N., Snyder, J.C., Bolduc, B., and Young, M.J. (2014). Archaeal Viruses: Diversity,

760 Replication, and Structure. Annu Rev Virol *1*, 399-426. 10.1146/annurev-virology-031413-

761 085357.

762 DiMaio, F., Yu, X., Rensen, E., Krupovic, M., Prangishvili, D., and Egelman, E.H. (2015). Virology.

763 A virus that infects a hyperthermophile encapsidates A-form DNA. Science *348*, 914-917.

764 10.1126/science.aaa4181.

765   Egelman, E.H. (2000). A robust algorithm for the reconstruction of helical filaments using single-

766   particle methods. Ultramicroscopy *85*, 225-234.

767   Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta

768   Crystallogr.D.Biol.Crystallogr. *60*, 2126-2132.

769   Erdmann, S., Chen, B., Huang, X., Deng, L., Liu, C., Shah, S.A., Le Moine Bauer, S., Sobrino, C.L.,

770   Wang, H., Wei, Y., et al. (2014). A novel single-tailed fusiform Sulfolobus virus STSV2 infecting

771   model *Sulfolobus* species. Extremophiles *18*, 51-60. 10.1007/s00792-013-0591-z.

772   Forterre, P., Krupovic, M., and Prangishvili, D. (2014). Cellular domains and viral lineages.

773   Trends in microbiology *22*, 554-558. 10.1016/j.tim.2014.07.004.

774   Frost, A., Perera, R., Roux, A., Spasov, K., Destaing, O., Egelman, E.H., De, C.P., and Unger, V.M.

775   (2008). Structural basis of membrane invagination by F-BAR domains. Cell *132*, 807-817.

776   Goulet, A., Vestergaard, G., Felisberto-Rodrigues, C., Campanacci, V., Garrett, R.A., Cambillau,

777   C., and Ortiz-Lombardia, M. (2010). Getting the best out of long-wavelength X-rays: de novo

778   chlorine/sulfur SAD phasing of a structural protein from ATV. Acta crystallographica. Section D,

779   Biological crystallography *66*, 304-308. 10.1107/S0907444909051798.

780   Gudbergsdottir, S., Deng, L., Chen, Z., Jensen, J.V., Jensen, L.R., She, Q., and Garrett, R.A. (2011).

781   Dynamic properties of the Sulfolobus CRISPR/Cas and CRISPR/Cmr systems when challenged

782   with vector-borne viral and plasmid genes and protospacers. Mol Microbiol *79*, 35-49.

783   10.1111/j.1365-2958.2010.07452.x.

784 Haring, M., Vestergaard, G., Rachel, R., Chen, L., Garrett, R.A., and Prangishvili, D. (2005).

785 Virology: independent virus development outside a host. Nature *436*, 1101-1102.

786 10.1038/4361101a.

787 Hay, I.D., and Lithgow, T. (2019). Filamentous phages: masters of a microbial sharing economy.

788 EMBO reports *20*. 10.15252/embr.201847427.

789 Hillyer, M.B., and Gibb, B.C. (2016). Molecular Shape and the Hydrophobic Effect. Annu Rev

790 Phys Chem *67*, 307-329. 10.1146/annurev-physchem-040215-112316.

791 Hochstein, R., Bollschweiler, D., Dharmavaram, S., Lintner, N.G., Plitzko, J.M., Bruinsma, R.,

792 Engelhardt, H., Young, M.J., Klug, W.S., and Lawrence, C.M. (2018). Structural studies of

793 Acidianus tailed spindle virus reveal a structural paradigm used in the assembly of spindle-

794 shaped viruses. Proc. Natl. Acad. Sci. U.S.A. *115*, 2120-2125. 10.1073/pnas.1719180115.

795 Hochstein, R.A., Amenabar, M.J., Munson-McGee, J.H., Boyd, E.S., and Young, M.J. (2016).

796 Acidianus Tailed Spindle Virus: a New Archaeal Large Tailed Spindle Virus Discovered by

797 Culture-Independent Methods. J Virol *90*, 3458-3468. 10.1128/JVI.03098-15.

798 Hong, C., Pietila, M.K., Fu, C.J., Schmid, M.F., Bamford, D.H., and Chiu, W. (2015). Lemon-

799 shaped halo archaeal virus His1 with uniform tail but variable capsid structure. Proc. Natl. Acad.

800 Sci. U.S.A. *112*, 2449-2454. 10.1073/pnas.1425008112.

801 Iranzo, J., Koonin, E.V., Prangishvili, D., and Krupovic, M. (2016). Bipartite Network Analysis of

802 the Archaeal Virosphere: Evolutionary Connections between Viruses and Capsidless Mobile

803 Elements. J Virol *90*, 11043-11055. 10.1128/JVI.01622-16.

804   Iverson, E.A., Goodman, D.A., Gorchels, M.E., and Stedman, K.M. (2017). Genetic Analysis of the

805   Major Capsid Protein of the Archaeal Fusellovirus SSV1: Mutational Flexibility and

806   Conformational Change. Genes (Basel) *8*. 10.3390/genes8120373.

807   Johnson, J.E., and Olson, A.J. (2021). Icosahedral virus structures and the protein data bank. J

808   Biol Chem *296*, 100554. 10.1016/j.jbc.2021.100554.

809   Johnson, J.E., and Speir, J.A. (1997). Quasi-equivalent viruses: a paradigm for protein

810   assemblies. J. Mol. Biol. *269*, 665-675. 10.1006/jmbi.1997.1068.

811   Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K.,

812   Bates, R., Zidek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction

813   with AlphaFold. Nature. 10.1038/s41586-021-03819-2.

814   Kim, J.G., Kim, S.J., Cvirkaite-Krupovic, V., Yu, W.J., Gwak, J.H., Lopez-Perez, M., Rodriguez-

815   Valera, F., Krupovic, M., Cho, J.C., and Rhee, S.K. (2019). Spindle-shaped viruses infect marine

816   ammonia-oxidizing thaumarchaea. Proc. Natl. Acad. Sci. U.S.A. *116*, 15645-15650.

817   10.1073/pnas.1905682116.

818   Knitsch, R., Schneefeld, M., Weitzel, K., and Pfeifer, F. (2017). Mutations in the major gas vesicle

819   protein GvpA and impacts on gas vesicle formation in *Haloferax volcanii*. Mol Microbiol *106*,

820   530-542. 10.1111/mmi.13833.

821   Krupovic, M., Dolja, V.V., and Koonin, E.V. (2019). Origin of viruses: primordial replicators

822   recruiting capsids from hosts. Nature reviews. Microbiology *17*, 449-458. 10.1038/s41579-019-

823   0205-6.

824   Krupovic, M., Dolja, V.V., and Koonin, E.V. (2020). The LUCA and its complex virome. Nature

825   reviews. Microbiology *18*, 661-670. 10.1038/s41579-020-0408-x.

826   Krupovic, M., and Koonin, E.V. (2017). Multiple origins of viral capsid proteins from cellular

827   ancestors. Proc. Natl. Acad. Sci. U.S.A. *114*, E2401-E2410. 10.1073/pnas.1621061114.

828   Krupovic, M., Quemin, E.R., Bamford, D.H., Forterre, P., and Prangishvili, D. (2014). Unification

829   of the globally distributed spindle-shaped viruses of the Archaea. J Virol *88*, 2354-2358.

830   10.1128/JVI.02941-13.

831   Li, N., and Cannon, M.C. (1998). Gas vesicle genes identified in *Bacillus megaterium* and

832   functional expression in *Escherichia coli*. Journal of bacteriology *180*, 2450-2458.

833   Li, Z., Jaroszewski, L., Iyer, M., Sedova, M., and Godzik, A. (2020). FATCAT 2.0: towards a better

834   understanding of the structural diversity of proteins. Nucleic Acids Res *48*, W60-W64.

835   10.1093/nar/gkaa443.

836   Liu, J., Cvirkaite-Krupovic, V., Baquero, D.P., Yang, Y., Zhang, Q., Shen, Y., and Krupovic, M.

837   (2021). Virus-induced cell gigantism and asymmetric cell division in archaea. Proc. Natl. Acad.

838   Sci. U.S.A. *118*. 10.1073/pnas.2022578118.

839   Medvedeva, S., Sun, J., Yutin, N., Koonin, E.V., Nunoura, T., Rinke, C., and Krupovic, M. (2021).

840   Viruses of Asgard archaea. bioRxiv, 2021.2007.2029.453957. 10.1101/2021.07.29.453957.

841   Mochizuki, T., Yoshida, T., Tanaka, R., Forterre, P., Sako, Y., and Prangishvili, D. (2010). Diversity

842   of viruses of the hyperthermophilic archaeal genus Aeropyrum, and isolation of the Aeropyrum

843   pernix bacilliform virus 1, APBV1, the first representative of the family *Clavaviridae*. Virology

844   *402*, 347-354. 10.1016/j.virol.2010.03.046.

845   Newberry, R.W., and Raines, R.T. (2019). Secondary Forces in Protein Folding. ACS Chem Biol

846   *14*, 1677-1686. 10.1021/acschembio.9b00339.

847   Pei, J., and Grishin, N.V. (2014). PROMALS3D: multiple protein sequence alignment enhanced

848   with evolutionary and three-dimensional structural information. Methods Mol Biol *1079*, 263-

849   271. 10.1007/978-1-62703-646-7_17.

850   Perotti, L.E., Dharmavaram, S., Klug, W.S., Marian, J., Rudnick, J., and Bruinsma, R.F. (2016).

851   Useful scars: Physics of the capsids of archaeal viruses. Phys Rev E *94*, 012404.

852   10.1103/PhysRevE.94.012404.

853   Perotti, L.E., Zhang, K., Rudnick, J., and Bruinsma, R.F. (2019). Kirigami and the Caspar-Klug

854   construction for viral shells with negative Gauss curvature. Phys Rev E *99*, 022413.

855   10.1103/PhysRevE.99.022413.

856   Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and

857   Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis.

858   J.Comput.Chem. *25*, 1605-1612.

859   Pfab, J., Phan, N.M., and Si, D. (2021). DeepTracer for fast de novo cryo-EM protein structure

860   modeling and special studies on CoV-related complexes. Proc. Natl. Acad. Sci. U.S.A. *118*.

861   10.1073/pnas.2017525118.

862   Pfeifer, F., Gregor, D., Hofacker, A., Pl √ ∂sser, P., and Zimmermann, P. (2002). Regulation of gas

863   vesicle formation in halophilic archaea. J. Mol. Microbiol. Biotechnol. *4*, 175-181.

864   Pfitzner, A.K., Moser von Filseck, J., and Roux, A. (2021). Principles of membrane remodeling by

865   dynamic ESCRT-III polymers. Trends in cell biology. 10.1016/j.tcb.2021.04.005.

866   Pietila, M.K., Atanasova, N.S., Oksanen, H.M., and Bamford, D.H. (2013). Modified coat protein

867   forms the flexible spindle-shaped virion of haloarchaeal virus His1. Environmental microbiology

868   *15*, 1674-1686. 10.1111/1462-2920.12030.

869   Prangishvili, D., Bamford, D.H., Forterre, P., Iranzo, J., Koonin, E.V., and Krupovic, M. (2017). The

870   enigmatic archaeal virosphere. Nature reviews. Microbiology *15*, 724-739.

871   10.1038/nrmicro.2017.125.

872   Prangishvili, D., Vestergaard, G., Haring, M., Aramayo, R., Basta, T., Rachel, R., and Garrett, R.A.

873   (2006). Structural and genomic properties of the hyperthermophilic archaeal virus ATV with an

874   extracellular stage of the reproductive cycle. J. Mol. Biol. *359*, 1203-1216.

875   10.1016/j.jmb.2006.04.027.

876   Ptchelkine, D., Gillum, A., Mochizuki, T., Lucas-Staat, S., Liu, Y., Krupovic, M., Phillips, S.E.V.,

877   Prangishvili, D., and Huiskonen, J.T. (2017). Unique architecture of thermophilic archaeal virus

878   APBV1 and its genome packaging. Nat. Commun. *8*, 1436. 10.1038/s41467-017-01668-0.

879   Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for

880   rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290-296.

881   10.1038/nmeth.4169.

882   Punjani, A., Zhang, H., and Fleet, D.J. (2020). Non-uniform refinement: adaptive regularization

883   improves single-particle cryo-EM reconstruction. Nat. Methods *17*, 1214-1221.

884   10.1038/s41592-020-00990-8.

885   Quemin, E.R., Chlanda, P., Sachse, M., Forterre, P., Prangishvili, D., and Krupovic, M. (2016).

886   Eukaryotic-Like Virus Budding in Archaea. mBio *7*. 10.1128/mBio.01439-16.

887   Quemin, E.R., Pietila, M.K., Oksanen, H.M., Forterre, P., Rijpstra, W.I., Schouten, S., Bamford,

888   D.H., Prangishvili, D., and Krupovic, M. (2015). Sulfolobus Spindle-Shaped Virus 1 Contains

889   Glycosylated Capsid Proteins, a Cellular Chromatin Protein, and Host-Derived Lipids. J Virol *89*,

890   11681-11691. 10.1128/JVI.02270-15.

891   Raoult, D., and Forterre, P. (2008). Redefining viruses: lessons from Mimivirus. Nature reviews.

892   Microbiology *6*, 315-319. 10.1038/nrmicro1858.

893   Rayment, I., Baker, T.S., Caspar, D.L., and Murakami, W.T. (1982). Polyoma virus capsid

894   structure at 22.5 A resolution. Nature *295*, 110-115. 10.1038/295110a0.

895   Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from

896   electron micrographs. J. Struct. Biol. *192*, 216-221. 10.1016/j.jsb.2015.08.008.

897   Scheele, U., Erdmann, S., Ungewickell, E.J., Felisberto-Rodrigues, C., Ortiz-Lombardia, M., and

898   Garrett, R.A. (2011). Chaperone role for proteins p618 and p892 in the extracellular tail

899   development of Acidianus two-tailed virus. J Virol *85*, 4812-4821. 10.1128/JVI.00072-11.

900   Sevvana, M., Klose, T., and Rossmann, M.G. (2021). Principles of Virus Structure. Encyclopedia

901   of Virology, 257-277. 10.1016/B978-0-12-814515-9.00033-3.

902   Sheppard, C., Blombach, F., Belsom, A., Schulz, S., Daviter, T., Smollett, K., Mahieu, E., Erdmann,

903   S., Tinnefeld, P., Garrett, R., et al. (2016). Repression of RNA polymerase by the archaeo-viral

904   regulator ORF145/RIP. Nat. Commun. *7*, 13595. 10.1038/ncomms13595.

905   Sojar, H.T., and Bahl, O.P. (1987). Chemical deglycosylation of glycoproteins. Meth. Enzymol.

906   *138*, 341-350.

907   Stedman, K.M., DeYoung, M., Saha, M., Sherman, M.B., and Morais, M.C. (2015). Structural

908   insights into the architecture of the hyperthermophilic Fusellovirus SSV1. Virology *474*, 105-

909   109. 10.1016/j.virol.2014.10.014.

910   Stehle, T., Yan, Y., Benjamin, T.L., and Harrison, S.C. (1994). Structure of murine polyomavirus

911   complexed with an oligosaccharide receptor fragment. Nature *369*, 160-163.

912   10.1038/369160a0.

913   Strunk, T., Hamacher, K., Hoffgaard, F., Engelhardt, H., Zillig, M.D., Faist, K., Wenzel, W., and

914   Pfeifer, F. (2011). Structural model of the gas vesicle protein GvpA and analysis of GvpA

915   mutants in vivo. Mol Microbiol *81*, 56-68. 10.1111/j.1365-2958.2011.07669.x.

916   Uldahl, K.B., Jensen, S.B., Bhoobalan-Chitty, Y., Martinez-Alvarez, L., Papathanasiou, P., and

917   Peng, X. (2016). Life Cycle Characterization of Sulfolobus Monocaudavirus 1, an Extremophilic

918   Spindle-Shaped Virus with Extracellular Tail Development. J Virol *90*, 5693-5699.

919   10.1128/JVI.00075-16.

920   Wang, F., Baquero, D.P., Su, Z., Beltran, L.C., Prangishvili, D., Krupovic, M., and Egelman, E.H.

921   (2020). The structures of two archaeal type IV pili illuminate evolutionary relationships. Nat.

922   Commun. *11*, 3424. 10.1038/s41467-020-17268-4.

923   Wang, F., Cvirkaite-Krupovic, V., Kreutzberger, M.A.B., Su, Z., de Oliveira, G.A.P., Osinski, T.,

924   Sherman, N., DiMaio, F., Wall, J.S., Prangishvili, D., et al. (2019). An extensively glycosylated

925   archaeal pilus survives extreme conditions. Nat Microbiol *4*, 1401-1410. 10.1038/s41564-019-

926   0458-x.

927   Weber, M.S., Eibauer, M., Sivagurunathan, S., Magin, T.M., Goldman, R.D., and Medalia, O.

928   (2021). Structural heterogeneity of cellular K5/K14 filaments as revealed by cryo-electron

929   microscopy. eLife *10*. 10.7554/eLife.70307.

930   Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V.,

931   Keedy, D.A., Hintze, B.J., Chen, V.B., et al. (2018). MolProbity: More and better reference data

932   for improved all-atom structure validation. Protein Science *27*, 293-315. 10.1002/pro.3330.

933   Winkler, H. (2007). 3D reconstruction and processing of volumetric data in cryo-electron

934   tomography. J. Struct. Biol. *157*, 126-137. https://doi.org/10.1016/j.jsb.2006.07.014.

935   Xiang, X., Chen, L., Huang, X., Luo, Y., She, Q., and Huang, L. (2005). *Sulfolobus tengchongensis*

936   spindle-shaped virus STSV1: virus-host interactions and genomic features. J Virol *79*, 8677-8686.

937   10.1128/JVI.79.14.8677-8686.2005.

938   Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017).

939   MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron

940   microscopy. Nat. Methods *14*, 331-332. 10.1038/nmeth.4193.

941   Zheng, W., Pena, A., Low, W.W., Wong, J.L.C., Frankel, G., and Egelman, E.H. (2020).

942   Cryoelectron-Microscopic Structure of the pKpQIL Conjugative Pili from Carbapenem-Resistant

943   *Klebsiella pneumoniae*. Structure *28*, 1321-1328 e1322. 10.1016/j.str.2020.08.010.

944   Zillig, W., Kletzin, A., Schleper, C., Holz, I., Janekovic, D., Hain, J., Lanzendörfer, M., and

945   Kristjansson, J.K. (1993). Screening for Sulfolobales, their plasmids and their viruses in *Icelandic*

946   *solfataras*. Systematic and Applied Microbiology *16*, 609-628.

947

TABLE FOR AUTHOR TO COMPLETE

*Please upload the completed table as a separate document. **Please do not add subheadings to the key resources table.** If you wish to make an entry that does not fall into one of the subheadings below, please contact your handling editor. **Any subheadings not relevant to your study can be skipped.** (**NOTE:** For authors publishing in Cell Genomics, Cell Reports Medicine, Current Biology, and Med, please note that references within the KRT should be in numbered style rather than Harvard.)*

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Archaeal and virus strains		
SMV1	Uldahl et al., 2016	N/A
ATV	Prangishvili et al., 2006	N/A
Saccharolobus islandicus ΔC1C2	Gudbergsdottir et al., 2011	N/A
Acidianus covivator AA9	Prangishvili *et al.*, 2006	N/A
Chemicals, peptides, and recombinant proteins		
Bacto Tryptone	Gibco	Cat #: 211705
Bacto yeast extract	Gibco	Cat #: 212750
Sucrose	Sigma-Aldrich	Cat #: S0389
PEG6000	Sigma-Aldrich	Cat #: 81260
NaCl	Sigma-Aldrich	Cat #: S9888
Trisma base	Sigma-Aldrich	Cat #: T1503
Acetic acid	Sigma-Aldrich	Cat #: A6283
CsCl	Euromedex	Cat #: EU0770
Ethidium bromide	Eurobio Scientific	Cat #: GEPBET02-AF
InstantBlue	Abcam	Cat #: ab119211
Deposited data		

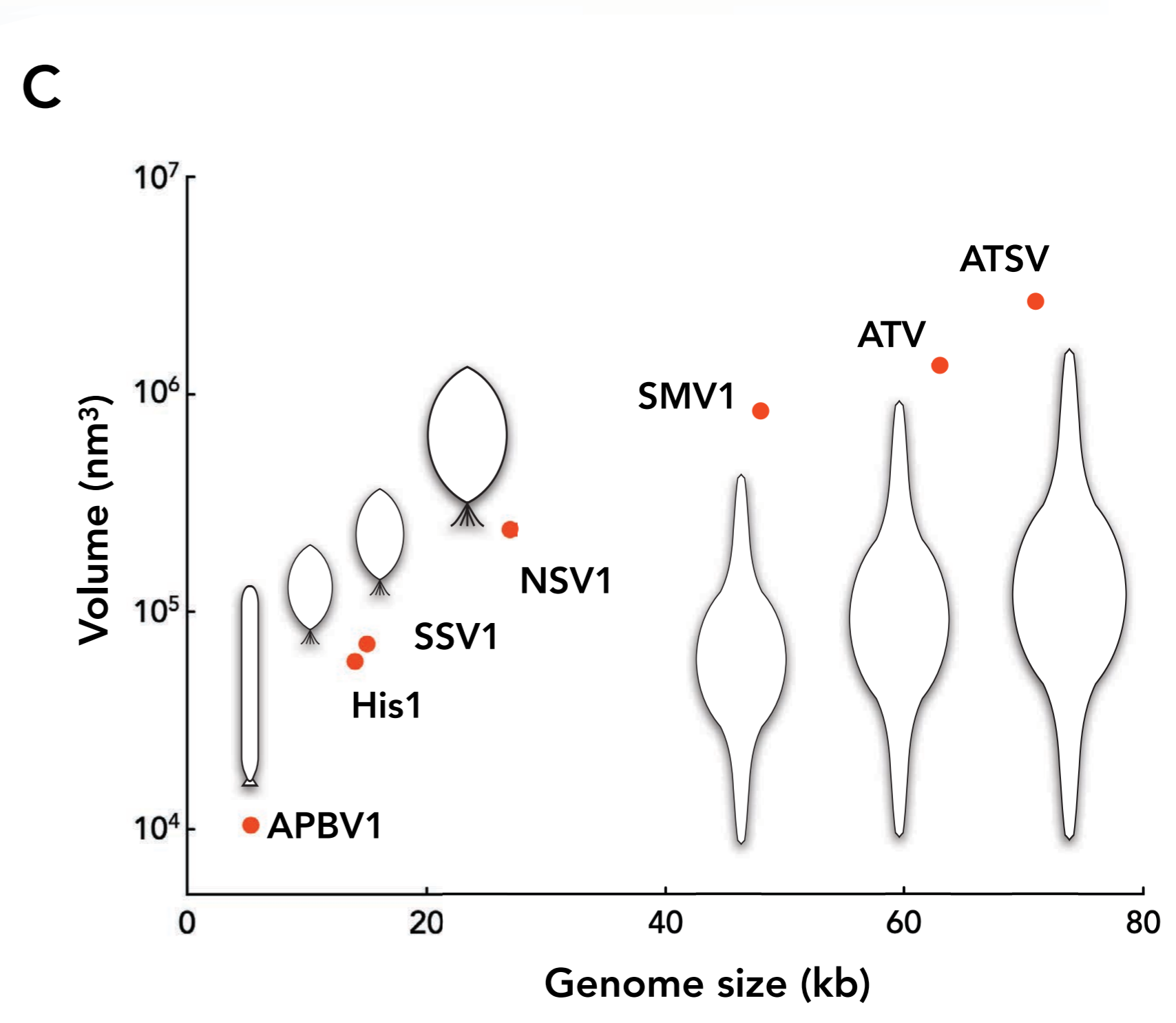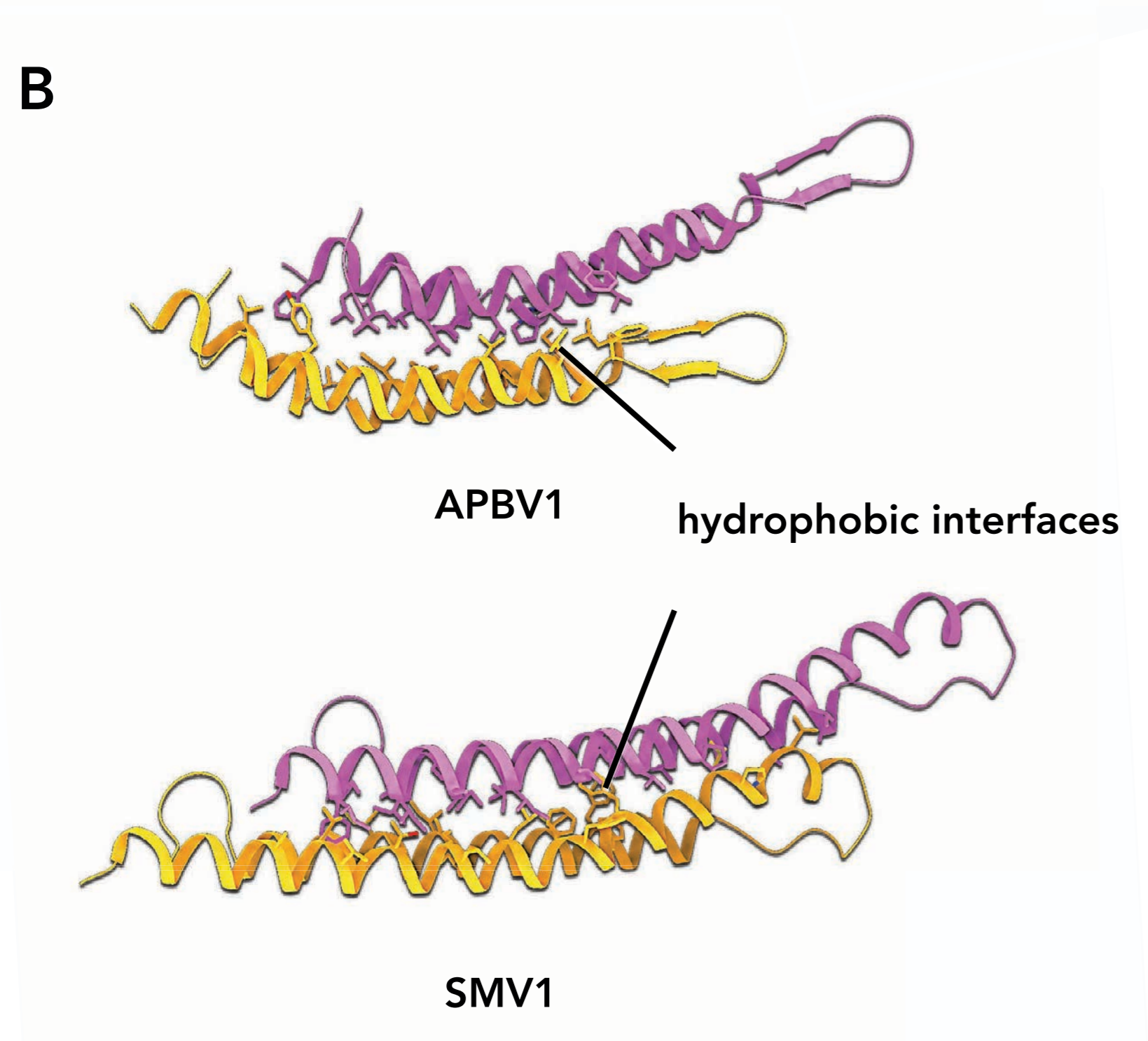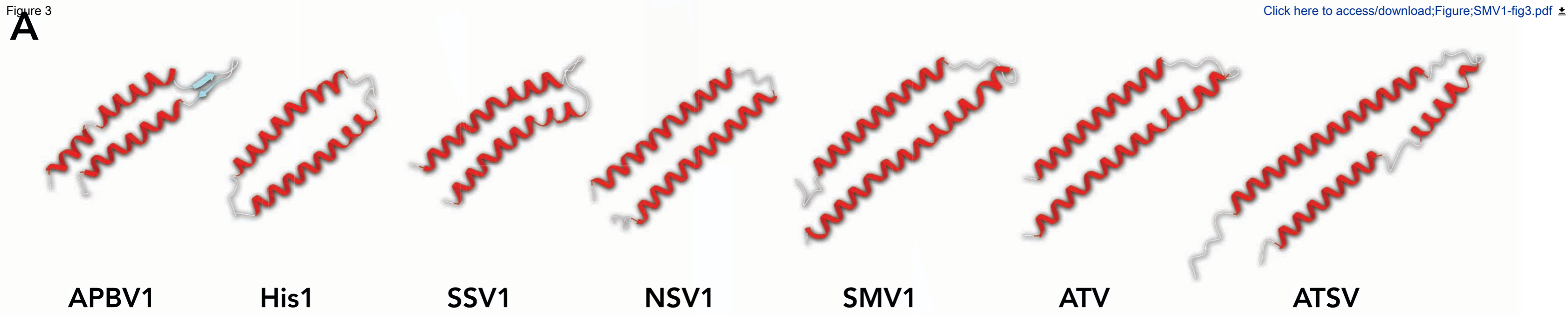| SMV1 Cryo-EM maps	This paper	EMDB: EMD-24585, EMD-24586, EMD-24587, EMD-24588, EMD-24589, EMD-24590, EMD-24591, EMD-24592, EMD-24593, EMD-24594, EMD-24595, EMD-24597
SMV1 Cryo-EM atomic models	This paper	PDB: 7RO2, 7RO3, 7RO4, 7RO5, 7RO6, 7ROB, 7ROC, 7ROD, 7ROE, 7ROG, 7ROH, 7ROI
ATV Cryo-EM map	This paper	EMDB: EMD-24596
Software and algorithms		
cryoSPARC	Punjani et al., 2017	https://cryosparc.com
Chimera	Pettersen et al., 2004	https://www.cgl.UCSF.edu/chimera
DeepTracer	Pfab et al., 2021	https://deeptracer.uw.edu
Coot	Emsley and Cowtan, 2004	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot
Phenix	Afonine et al., 2018	https://phenix-online.org
MolProbity	Williams et al., 2018	http://molprobity.biochem.duke.edu
Psi-Blast	Altschul et al., 1997	https://www.ebi.ac.uk/Tools/sss/psiblast/
PROMALS3D	Pei and Grishin, 2014	http://prodata.swmed.edu/promals3d/promals3d.php
AlphaFold2	Jumper et al., 2021	https://alphafold.ebi.ac.uk
FATCAT	Li et al., 2020	http://fatcat.godziklab.org/

Figure 1

A

29Å

29Å

B

Symmetry

335

- 12
- 11
- 10
- 9
- 8
- 7
- 6
- 5
- 4
- 3
- 2
- 1

270

Diameter (Å)

symmetry 12

symmetry 1

C

29 Å

7
6
5
4
3
2
1

Left-handed 7-start

symmetry 7

D

N-helix

C-helix

N

C

MCP subunit

N-linked glycosylation

~3.5 Å cryo-EM map

Figure 2

A

B

7
6
5
4
3
2
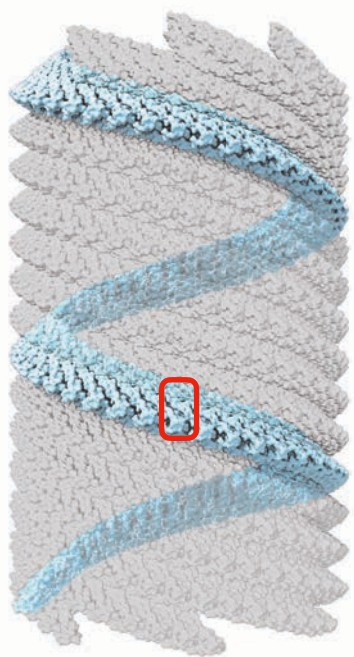**Left-handed 7-start**
1

⊢——— 180 Å ———⊣

C

↻ **90°**

ATV map and model

Figure 3

Figure 4

A
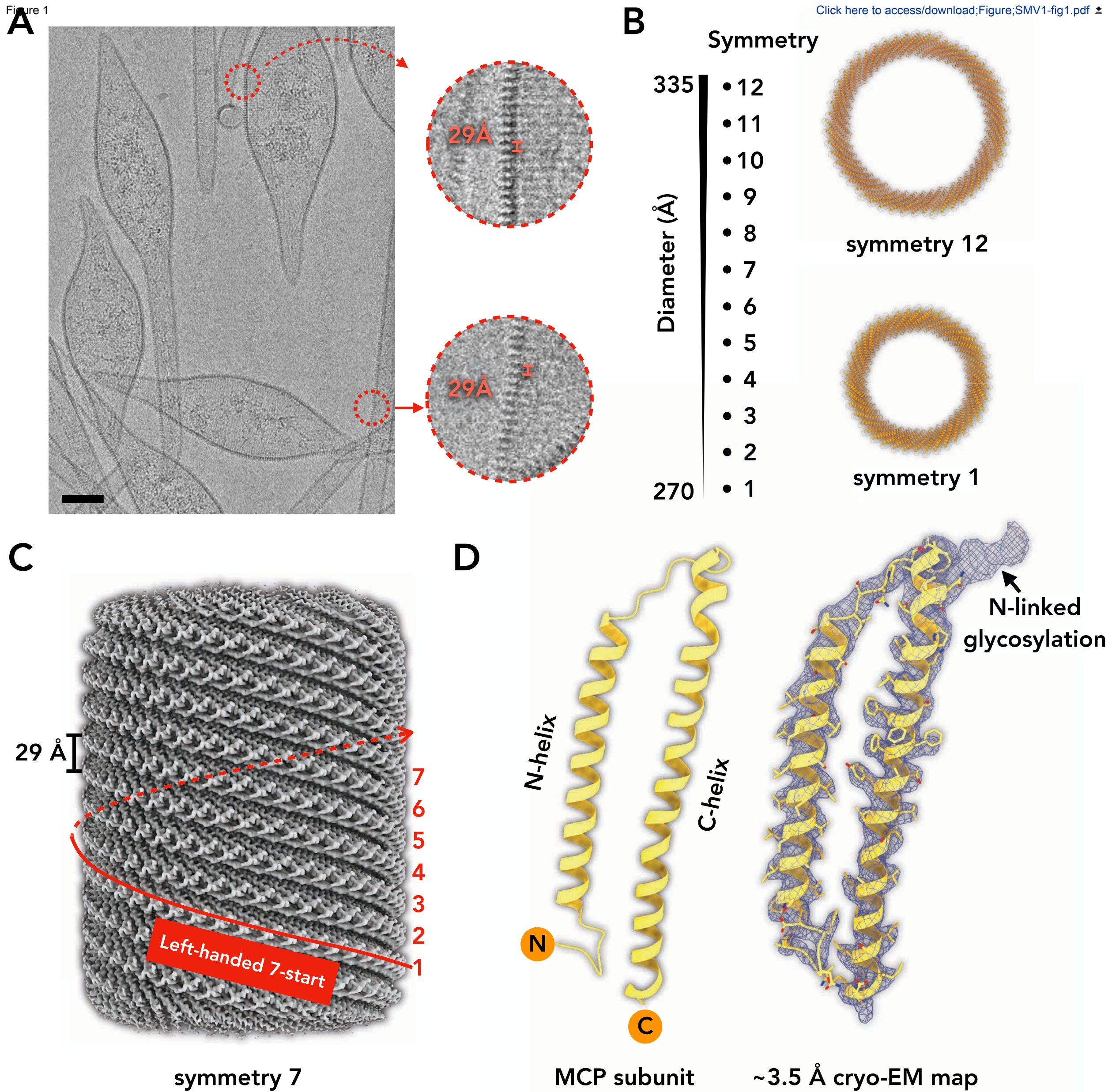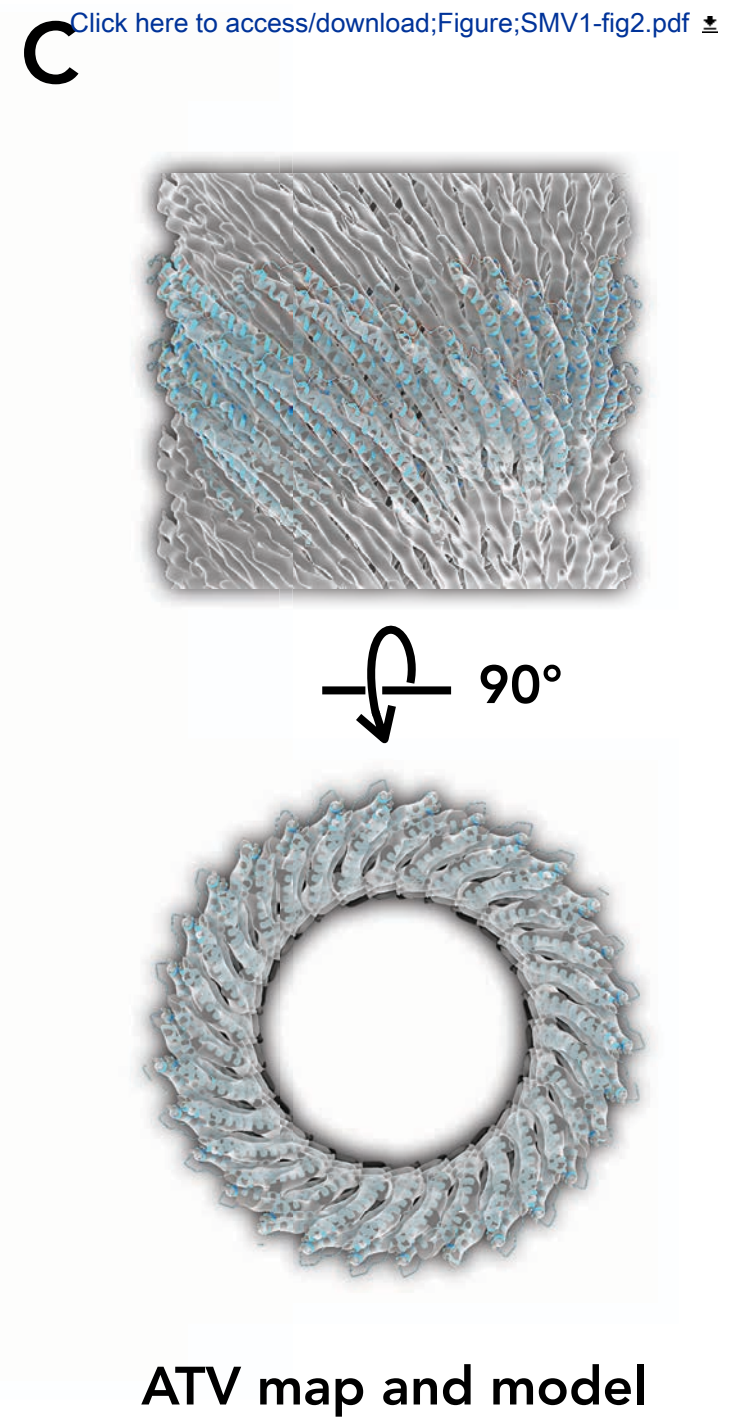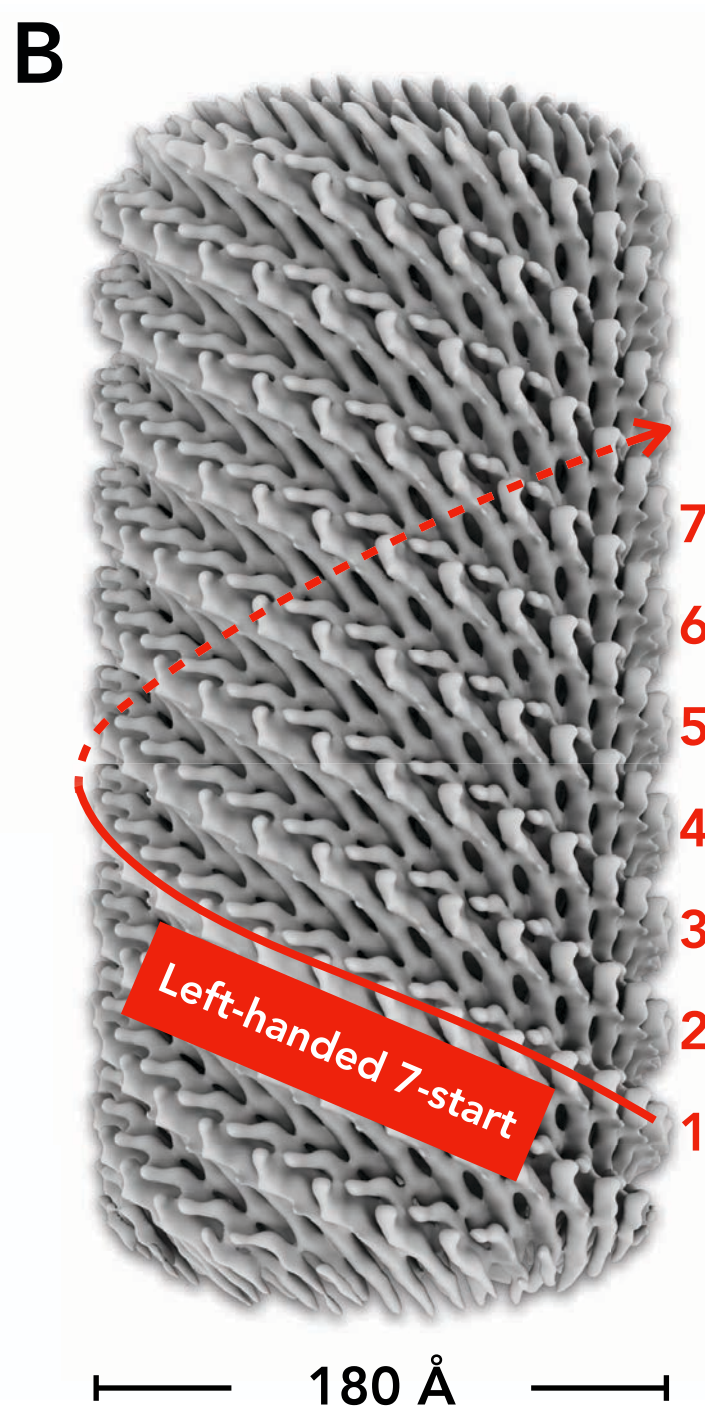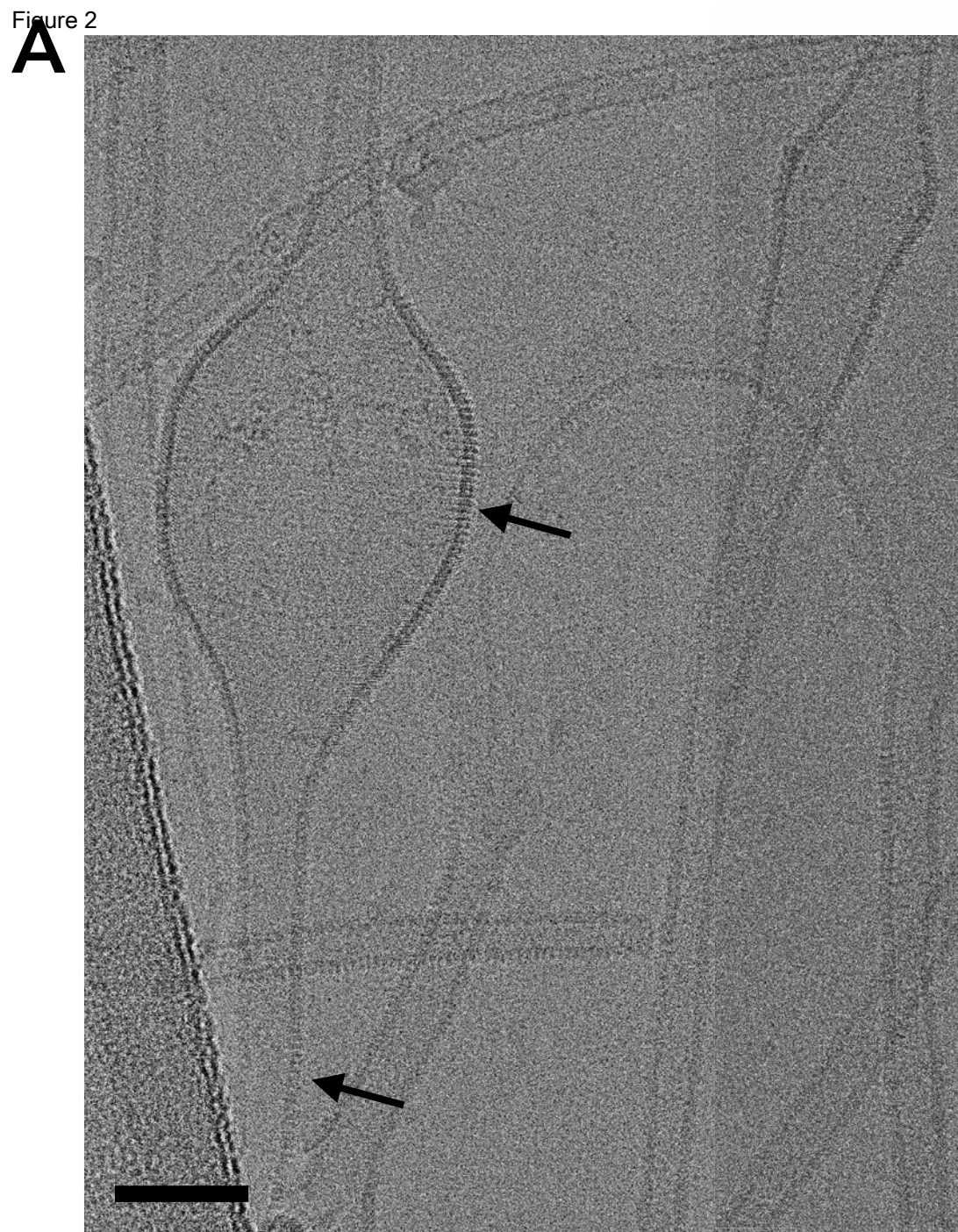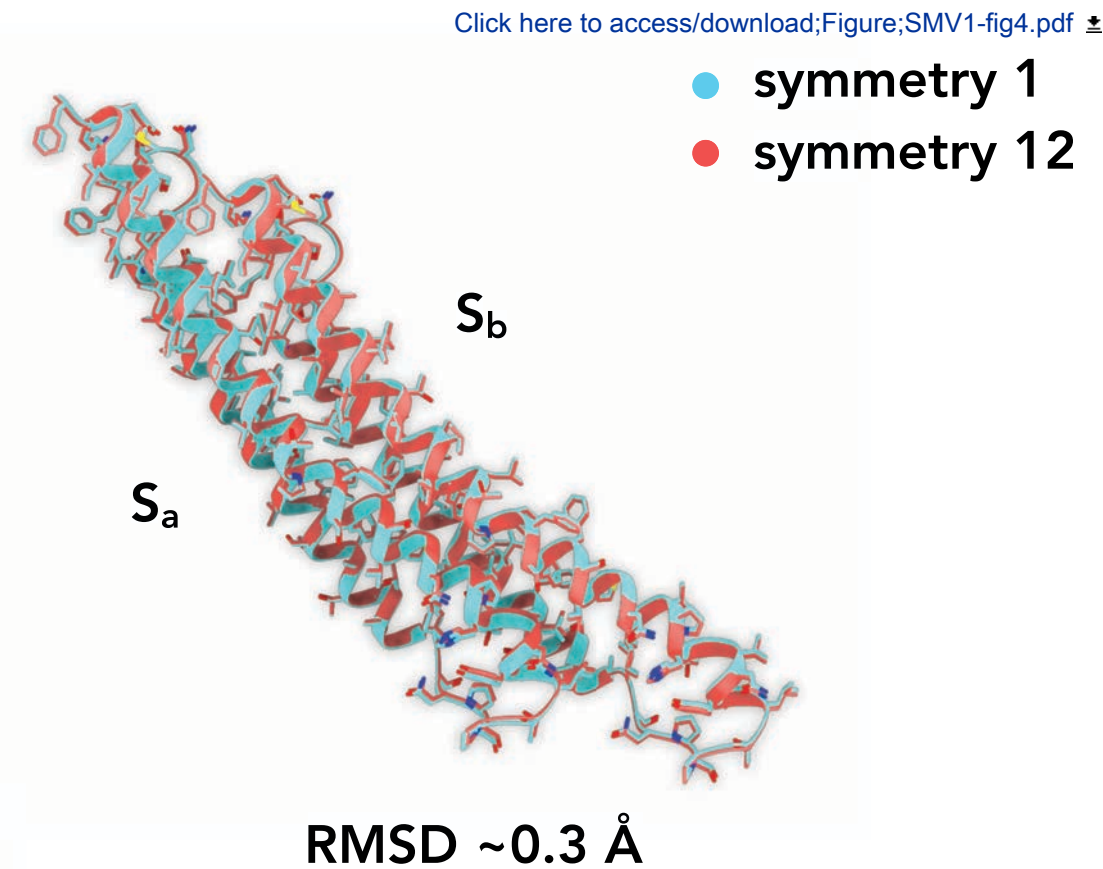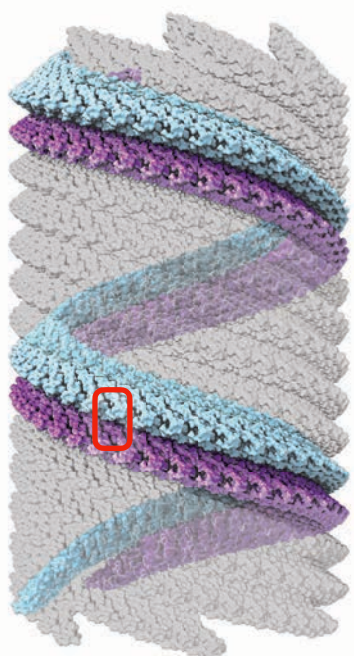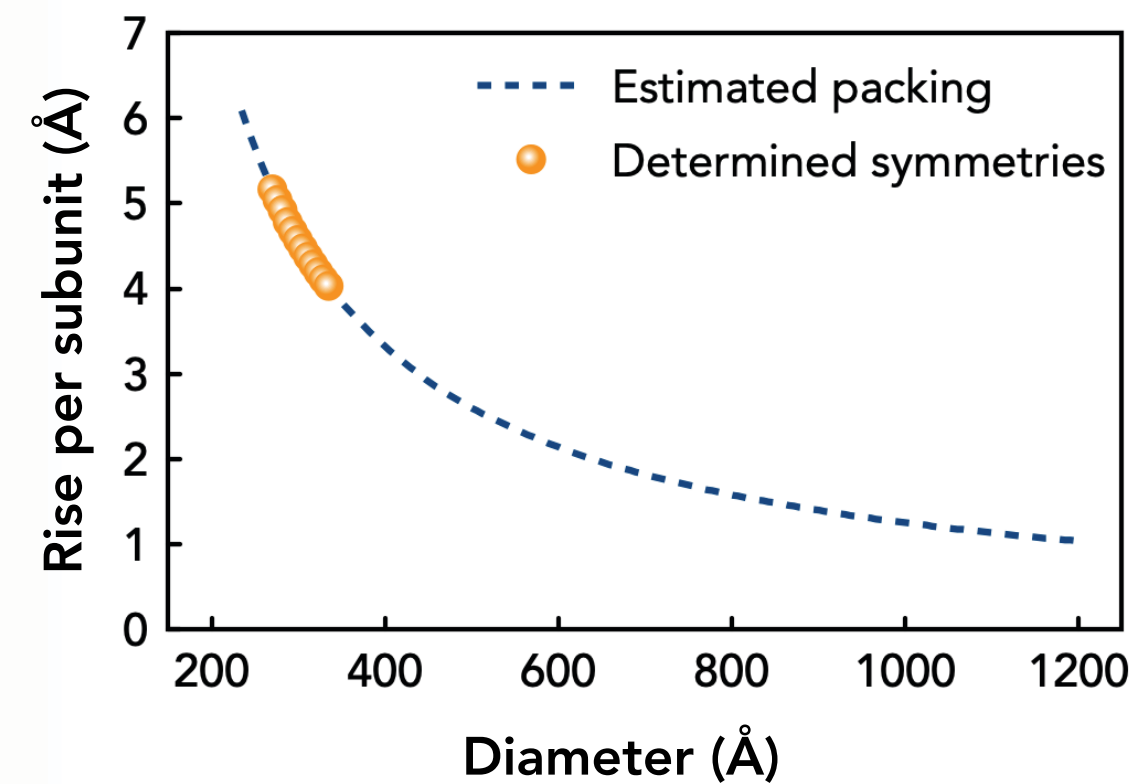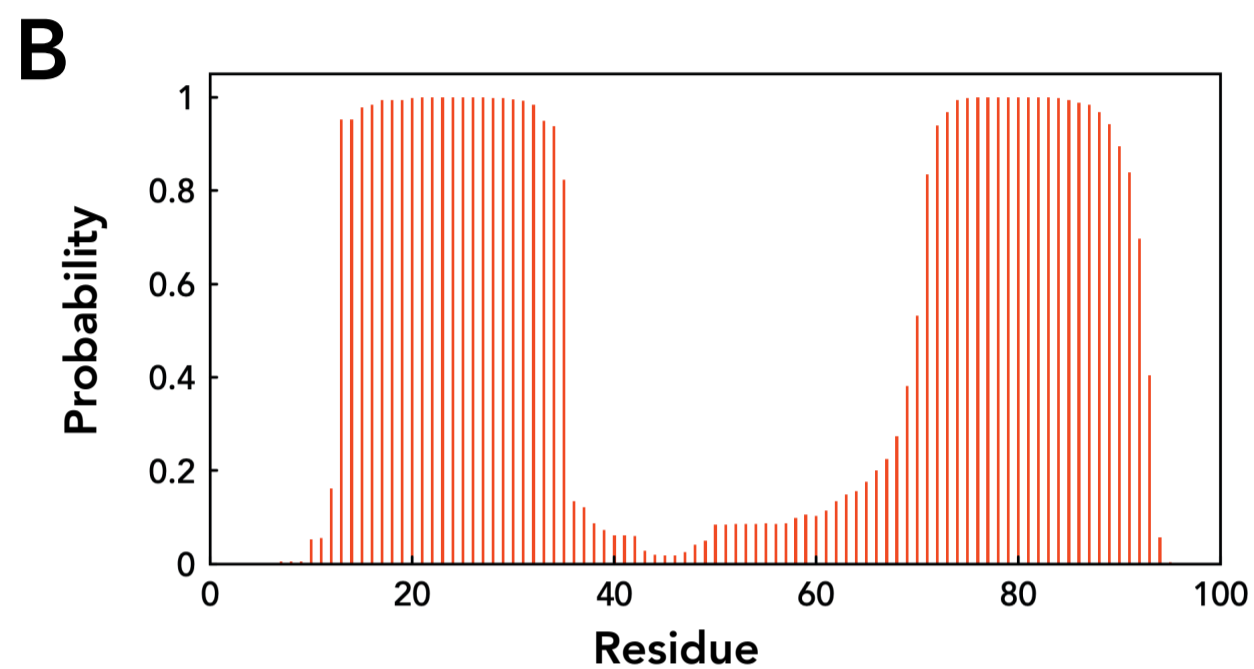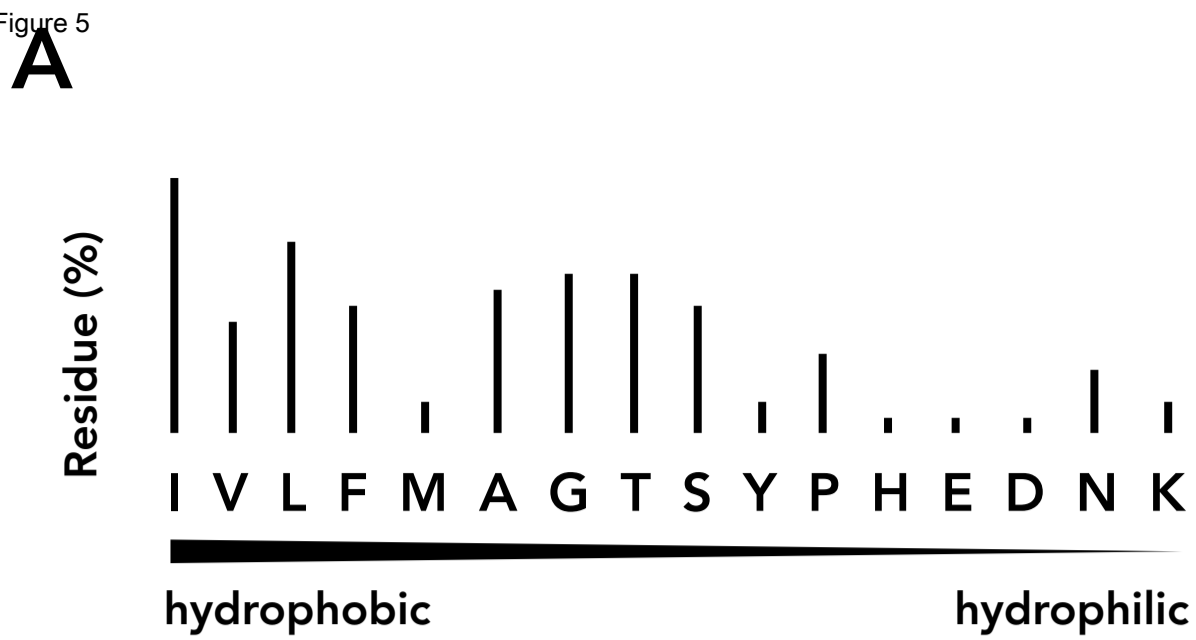
Interactions along the 7-start

Aligned by two subunits

RMSD ~0.3 Å

- symmetry 1
- symmetry 12

S_a

S_b

B
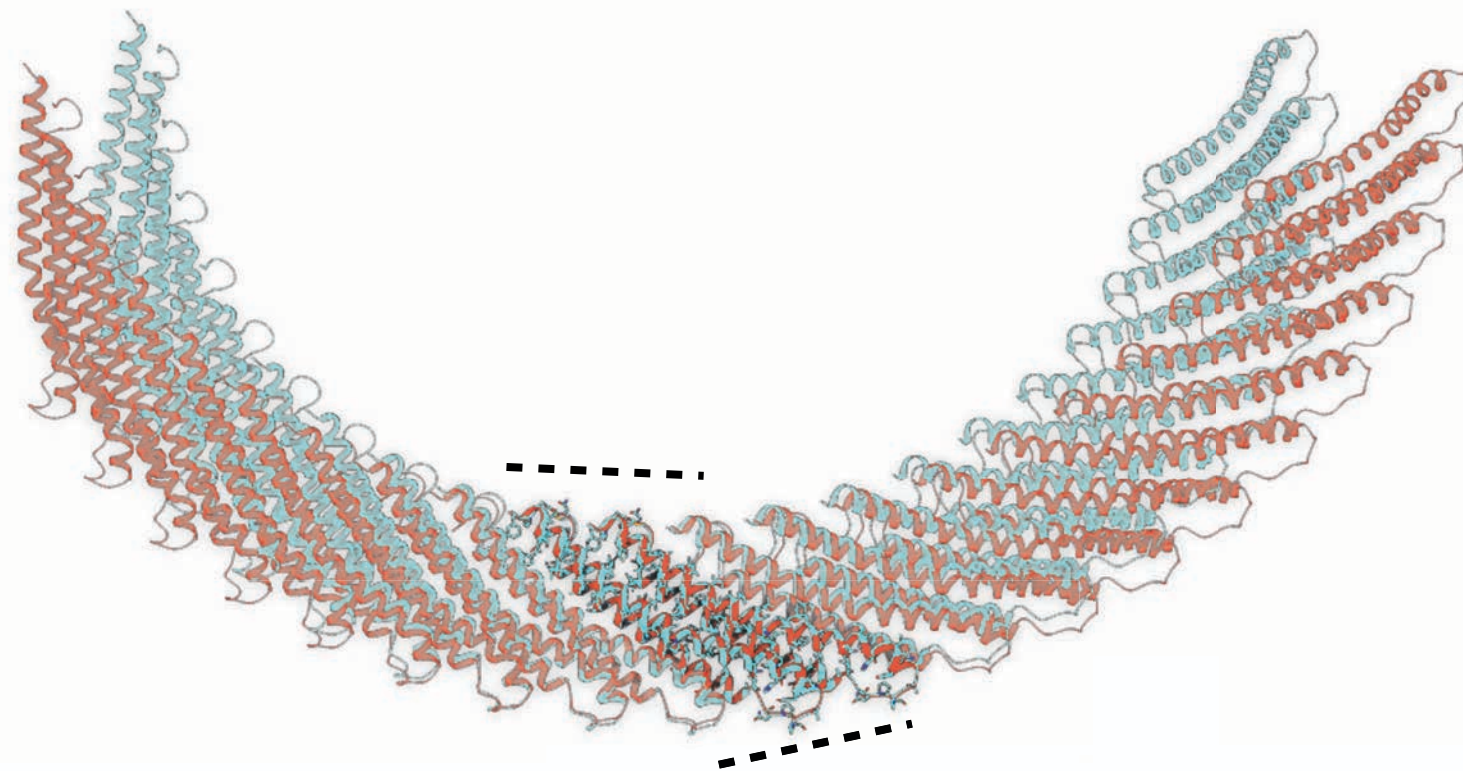
Interactions between 7-start protofilaments

S'_a

S_a

S_b

80°

RMSD ~0.4 Å

C

- - - Estimated packing
- Determined symmetries

Rise per subunit (Å)

Diameter (Å)

Figure 5

A

Residue (%)

I V L F M A G T S Y P H E D N K

hydrophobic   hydrophilic

B

TMHMM transmembrane probabilities

C

hydrophilic

hydrophobic

hydrophilic

surface lipophilicity

−20   +20

D

top view

bottom view

side view (surface)

side view (lumen)

A

pH = 10
→
37°C, 3h

B

Non-treated SMV1   NaOH-treated SMV1

Sucrose gradient fractions   F1 2 3 4 5 6 7 8 9 10 11 12   F1 2 3 4 5 6 7 8 9 10 11 12

MM   F1 2 3 4 5 6 7 8 9 10 11 12   MM   F1 2 3 4 5 6 7 8 9 10 11 12

F1

F9

F12

F1

F2

F3

A

symmetry 1	symmetry 2	symmetry 3	symmetry 4
5.15 Å	0.72 Å	0.70 Å	0.68 Å
-9.0 °	50.2 °	-155.5 °	-104.1 °
C7	C1	C1	C1

symmetry 5	symmetry 6	symmetry 7	symmetry 8
0.67 Å	0.66 Å	0.64 Å	4.38 Å
101.7 °	153.1 °	-52.6 °	-7.7 °
C1	C1	C1	C7

symmetry 9	symmetry 10	symmetry 11	symmetry 12
0.61 Å	0.60 Å	0.59 Å	0.58 Å
50.4 °	-155.3 °	-103.9 °	101.8 °
C1	C1	C1	C1

B

C

Id#	Id	Protein name	Coverage		MW	Score
1.1	A0A4W2BXP1_BOBOX	Peptidase S1 domain-containing protein OS=Bos indicus x Bos taurus OX=30522 PE=4 SV=1	▭▮▮▮▭	33.6 %	24068	451
1.2	W0UU99_9VIRU	Conserved archaeal virion protein OS=Sulfolobus monocaudavirus SMV1 OX=1351702 PE=4 SV=1	▭▭▮	22.9 %	17302	402

D

Id#	Id	Protein name	Coverage		MW	Score
2.1	W0UU99_9VIRU	Conserved archaeal virion protein OS=Sulfolobus monocaudavirus SMV1 OX=1351702 PE=4 SV=1	▭▭▮	22.9 %	17302	326
2.2	A0A4W2DWI3_BOBOX	Peptidase S1 domain-containing protein OS=Bos indicus x Bos taurus OX=30522 PE=4 SV=1	▭▮▮▮▭	25.2 %	31157	325
2.3	W0UUV5_9VIRU	Hypothetical membrane protein OS=Sulfolobus monocaudavirus SMV1 OX=1351702 PE=4 SV=1	▭▮▭	11.5 %	16152	126

A

B

C