



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/202158/>

Version: Submitted Version

---

**Preprint:**

Wilson, J.S., Fortier, L.-C., Fagan, R.P. et al. (Submitted: 2023) Molecular mechanism of bacteriophage tail contraction- structure of an S-layer-penetrating bacteriophage. [Preprint - bioRxiv] (Submitted)

<https://doi.org/10.1101/2023.08.04.551987>

---

© 2023 The Author(s). This preprint is made available under a CC-BY-NC-ND 4.0 International license. (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



1 **Molecular mechanism of bacteriophage tail contraction- structure of**  
2 **an S-layer-penetrating bacteriophage**

3

4 **Jason S. Wilson<sup>1,4</sup>, Louis-Charles Fortier<sup>2</sup>, Robert P. Fagan<sup>1,3,\*</sup>, Per A. Bullough<sup>1,3,\*</sup>**

5

6 <sup>1</sup>Molecular Microbiology, School of Biosciences, University of Sheffield, Sheffield, UK.

7 <sup>2</sup>Department of Microbiology and Infectious Diseases, Faculty of Medicine and Health Sciences,

8 Université de Sherbrooke, Sherbrooke, Québec, Canada. <sup>3</sup>The Florey Institute, University of

9 Sheffield, Sheffield, UK. <sup>4</sup>Present address: Biology Department, University of York, York, UK

10 \*e-mail: [p.bullough@sheffield.ac.uk](mailto:p.bullough@sheffield.ac.uk), [r.fagan@sheffield.ac.uk](mailto:r.fagan@sheffield.ac.uk)

11

## 12 **Abstract**

13 Viruses that infect bacteria (bacteriophages or phages) attach to the host cell envelope,  
14 inject their genetic material into the host cytosol and either persist as prophage or hijack the host  
15 machinery to produce progeny virions. Attachment is mediated through phage receptor binding  
16 proteins that are specific for different host cell surface molecules. A subset of phage, the  
17 myoviruses, possess contractile tails, the outer sheath of which contracts upon receptor binding,  
18 driving an inner tail tube through the cell envelope and delivering the phage genome into the host  
19 cytosol. The molecular details of phage tail contraction and mode of cell envelope penetration have  
20 remained poorly understood and were completely unknown for any phage infecting bacteria  
21 enveloped by a proteinaceous S-layer. Here we reveal the extended and contracted atomic  
22 structures of an intact contractile-tail phage that binds to and penetrates the protective S-layer of  
23 the Gram positive human pathogen *Clostridioides difficile*. Surprisingly, we find no evidence of  
24 the intrinsic enzymatic domains that other phages exploit in cell wall penetration, suggesting that  
25 sufficient energy is released upon tail contraction to penetrate the S-layer and the thick cell wall  
26 without enzymatic activity. However, it is also notable that the tail sheath subunits move less than  
27 those studied in related contractile injection systems such as the model phage T4. Instead, the  
28 unusually long tail length and flexibility upon contraction likely contribute towards the required  
29 free energy release for envelope penetration. Our results show that the principles of phage  
30 contraction and infection as determined in the model system of T4 are not universal. We anticipate  
31 that our structures will form a strong foundation to engineer *C. difficile* phages as therapeutics, and  
32 highlight important adaptations made in order to infect S-layer containing pathogens.

32 Phages are the most abundant biological entities on earth, yet only a tiny fraction have been  
33 characterised in any detail<sup>1</sup>. The most prevalent phage morphology described is that of an  
34 icosahedral capsid, containing the genome, attached to a tail<sup>2</sup>. The tail forms a conduit through  
35 which the genome is delivered into the host cell cytoplasm and is also associated with receptor  
36 binding functions<sup>3</sup>. Amongst phages, myoviruses have contractile tails of widely varying length,  
37 podoviruses have short non-contractile tails, while siphoviruses have long, flexible, non-  
38 contractile tails<sup>4</sup>.

39  
40 Myovirus phage tails and bacterial tailocins both belong to the group of so-called contractile  
41 injection systems (CISs)<sup>5</sup>. The most extensively studied myovirus structure is that of the Gram  
42 negative *Escherichia coli* phage, T4<sup>6-11</sup>. The T4 contractile tail shares many features in common  
43 with tailocins, although tailocins lack a capsid and do not inject DNA into their target<sup>12,13</sup>. For  
44 CISs, most detailed structural information only covers the initially extended tail conformation.  
45 However, contracted state structures have been reported at high resolution for *Pseudomonas* phage  
46 E217 and a tailocin, R2 pyocin<sup>14-16</sup>.

47  
48 Whilst other known myovirus structures share common tail features with T4, E217 and R2 pyocin,  
49 the molecular interactions stabilizing their contracted tail state are unknown<sup>16-19</sup>. Moreover, it is  
50 not known to what extent the subunit movements within the tail during contraction are the same  
51 across myoviruses or CISs in general. Whilst a number of different bacterial cell envelope types  
52 are targets for these phages, little is known about the penetration of the thick Gram positive  
53 peptidoglycan cell wall; furthermore there has been no explanation of the way phages are able to  
54 penetrate the additional protective proteinaceous S-layer found in the majority of eubacteria<sup>20</sup>. An

55 important question is whether S-layer penetrating contractile phages have adaptations to perform  
56 the extra work likely to be required to overcome this barrier.

57

58 Many bacterial species, including the notable Gram positive pathogen *Clostridioides difficile*,  
59 produce a proteinaceous S-layer that forms the outermost surface of the cell envelope. *C. difficile*  
60 infection (CDI) is the most common cause of antibiotic-associated diarrhoea, resulting in  
61 significant morbidity and mortality in susceptible populations<sup>21</sup>. Phages have been proposed as  
62 alternative antimicrobials to treat CDI<sup>22</sup>; naturally occurring phages can be used as a template for  
63 the design of such precision antimicrobials<sup>23,24</sup>. To advance designs, we need to describe the  
64 essential structural components and the mechanics involved in binding to the host cell and  
65 penetration of the host cell envelope. Here we describe the structure of  $\phi$ CD508, a contractile  
66 phage that infects clinically relevant strains of *C. difficile*<sup>25</sup>; its receptor is the major S-layer protein  
67 SlpA<sup>24,26,27</sup>.

68

69 We reveal the high resolution structure of the fully intact virion in both its extended and contracted  
70 form. We explain why the extent of contraction is significantly less than that reported for other  
71 phage structures. This has important implications for the mechanism of S-layer penetrating phages  
72 in general. We also highlight additional and significant key differences to previously characterised  
73 myoviruses: the baseplate of the tail is less complex; the needle tip is much more compact than  
74 that of other CISs and remarkably, appears to lack enzymatic functions, suggesting that the mode  
75 of envelope penetration may be purely mechanical.

76

77

## 78 Overall structure of $\phi$ CD508

79  $\phi$ CD508 is a myovirus<sup>25</sup>, with an overall tail length of 225 nm. Structures of *C. difficile* phage  
80  $\phi$ CD508 were determined by single particle EM of frozen hydrated samples (Fig. 1a, c); the  
81 resolution ranged from 2.6 Å to 4.0 Å for the extended phage, and 2.9 Å to 4.2 Å for the contracted  
82 form (Extended Data Table 1, Extended Data Fig. 1). In the extended state, 14 of the 17 predicted  
83 structural proteins identified by mass spectrometry were revealed and modelled at or close to their  
84 full chain length. A short section of the tape measure protein (gp59) was also modelled. Each  
85 protein chain was built *ab initio* where resolution was sufficient, or was first modelled by  
86 RosettaFold<sup>28</sup> or AlphaFold<sup>29</sup> and then fitted into the maps of both the extended and contracted  
87 phage (Fig. 1b). For additional structural details, refer to the Supplementary Information.

88

89 The virion can be divided into five distinct regions: the head, neck, tail, baseplate, and needle (Fig.  
90 1a, c). The structural proteins are encoded within a structural cassette (gp45-gp68) of the 74 gene  
91  $\phi$ CD508 genome (Fig. 1d).

92

93 The **head** (Fig. 2a) is formed of a near-complete T=7 icosahedral capsid<sup>30</sup>, 650 Å in diameter, and  
94 houses the genomic DNA. It is made up of 415 and 420 copies of two capsomer proteins, the major  
95 capsid protein gp49 and the capsid decoration protein gp48 respectively (Extended Fig. 2a ,b).  
96 Within the capsid of the extended phage, internal density shows layers of packaged DNA (Fig.  
97 2a). The contracted virion capsid reconstruction contains an identical arrangement of major capsid  
98 protein and capsid decoration protein but lacks internal density indicating that the DNA has been  
99 ejected from the head.

100

101 At the base of the capsid sits a unique opening, centred on what would be a fivefold symmetry axis  
102 in a complete icosahedron; it is made up of 12 copies of **portal** protein gp45 (Fig 2b; Extended  
103 Data Fig. 2g, h), forming an opening through which DNA is ejected from the capsid. Thus, there  
104 is a fivefold to twelvefold symmetry mismatch between the capsid and the portal (Extended Data  
105 Fig. 2j, k; Supplementary Information). The portal in turn connects to the **neck** (Fig. 2c; Extended  
106 Data Fig. 2i), which is made up of 3 protein complexes with a VIRFAM Type I neck  
107 arrangement<sup>31</sup>, comprising one dodecameric ring (gp50) and two hexameric rings (gp51 and gp53)  
108 (Fig. 2c). The symmetry is therefore reduced from twelvefold down to sixfold, travelling down the  
109 neck. The neck links the capsid to the tail (Fig. 2c).

110

111 The first complex in the neck is the **portal adaptor** gp50, which assembles into a twelvefold  
112 symmetric ring and acts to resolve a symmetry mismatch between the portal and tail (Fig. 2c,  
113 Extended Data Fig. 3a). Gp50 has no structural homology to other known phage neck protein  
114 structures (Extended Data Table 2). The **neck valve protein** gp51 forms a hexameric complex and  
115 represents the narrowest constriction in the neck lumen, ~23 Å in diameter at its minimum  
116 (Extended Data Fig. 3e). gp51 also has structural homology to other ‘stopper’ proteins such as  
117 SPP1 which is proposed to prevent DNA leakage from the capsid<sup>32</sup>. In our structure, strong density  
118 in the lumen of the channel weakens around the neck valve (Extended Data Fig. 2i), suggesting  
119 the neck valve may also constitute the point at which DNA is prevented from exiting through the  
120 tail of the  $\phi$ CD508 prior to contraction. Below the constriction, a weaker density is present which  
121 likely constitutes the tape measure protein gp59, although the enforced 6-fold symmetry in the  
122 reconstruction prevents identification or building of a model into this region (Extended Data Fig.  
123 2i).

124

125 The **tail** is attached to the neck and made up of a stack of 57 nested hexameric tail protein rings,  
126 with an inner ring of **tail tube protein** gp56 and an outer ring of **tail sheath protein** gp55 (Fig. 1,  
127 2d, 3). Each ring is offset from those above and below such that the extended tail assembly can be  
128 described as a six-stranded helix, with a twist of  $19^\circ$ , a rise of 39 Å, and an outer diameter of 230  
129 Å (Fig. 1b, 2d, 3a). This arrangement of tail proteins in the extended state is similar to other  
130 myovirus and phage tail-like particles. However, a notable feature is that the tail is considerably  
131 longer (225 nm) than those of other phages whose structures have been described (Fig. 5).

132

133 The main **tail tube** protein, gp56, has the conserved fold seen in other *Caudoviricetes* tail tubes<sup>33-</sup>  
134 <sup>36</sup>. However, gp56 lacks both the  $\alpha$ -loop and N-loop seen in other myovirus tail tube proteins, as  
135 well as the C-terminal loops seen in siphovirus tail tube proteins<sup>36</sup> (Extended Data Fig. 4d). This  
136 results in a more open packing between rings of the tail tube allowing for bending of the phage in  
137 the contracted state (Fig. 4).

138

139 The gp55 **tail sheath protein** is made up of three domains (Fig. 3c). The tail tube proximal domain  
140 I contains a loop consisting of residues 368 to 378, which is twice as long in  $\phi$ CD508 than in the  
141 sheath domains of pyocins, AFP, and PVC (6 residues and 3 residues respectively). For a typical  
142 ‘full’ contraction this loop (‘X-loop’) would interact with the linker of the N-terminal  $\beta$ -strand  
143 insertion of a neighbouring sheath protein (Fig. 3b, c), significantly reducing the possible range of  
144 motion of the linker compared to other known CIS structures (Extended Data Fig. 4e, f).

145

146 The tail is terminated by the **baseplate**. Importantly, the baseplate of  $\phi$ CD508 must be adapted for  
147 binding and penetration of the host cell S-layer. The baseplate is minimal compared with other  
148 structurally characterised phages<sup>11,17,18</sup> consisting of only 6 unique proteins (Fig. 2e). These are  
149 organised into a **hub** (Extended Data Fig. 6a) which connects to the tail and a **wedge** which forms  
150 a collar around the needle (Fig. 2e, Extended Data Fig. 6c). The baseplate is also expected to  
151 include the host attachment proteins gp67 and gp68, constituting the tail fibre and receptor binding  
152 proteins respectively, and which are essential for S-layer attachment<sup>27,37</sup>. Although these could not  
153 be sufficiently resolved in the EM maps for atomic modelling, low pass filtered maps show  
154 protrusions for the tail fibres (Extended Data Fig. 6f). Within the ring formed by the baseplate  
155 proteins lies the **needle** assembly; this is the other essential component presumed to be adapted for  
156 S-layer penetration, formed of two proteins, gp62 and gp63 providing a pointed end terminating  
157 the tail (Fig. 2e, Extended Data Fig. 6h). It is striking that we find no putative enzymatic domains  
158 within the baseplate/needle complex, unlike other comparable assemblies<sup>11,17,38</sup>. In urea contracted  
159 phage, we observe that the baseplate is lost from the tail, and it could not be resolved even at low  
160 resolution. Using the extended phage baseplate, we modelled what the contracted baseplate would  
161 look like. One spoke of the baseplate including the baseplate proximal sheath layer, baseplate hub  
162 and triplex proteins were aligned onto the contracted sheath. The rigid motion of these proteins  
163 would be sufficient to break the dimerisation interactions in the triplex proteins, as well as  
164 interactions between the triplex wing domains and the needle protein (Extended Fig. 6g). Breaking  
165 these contacts would be sufficient to allow clearance for the tail tube to pass through the baseplate.  
166  
167

168 **The  $\phi$ CD508 sheath does not contract as much as other CISs**

169 The overall RMSD between all proteins of the neck in the extended state compared to the  
170 contracted state is 0.34 Å, indicating that no large conformational change occurs in this region  
171 during phage contraction. After contraction and genome release, the neck-proximal tail sheath  
172 protein (gp55) remains attached to the neck-tail adaptor gp53 via its C-terminal insertion domain  
173 (Extended Data Fig. 3h, i). The neck-proximal sheath protein rotates to accommodate the  
174 contraction induced movements of the other sheath proteins in the tail below, pivoting around  
175 Gly<sup>260</sup> in the gp53 C-terminal linker (Extended Data Fig. 3i, k). The terminal sheath ring still  
176 maintains some contact with the tail tube. This reduced rotation is limited to the proximal layer,  
177 with the next layer below fully dissociating from the tail tube, similar to the arrangement observed  
178 for the main body of the contracted tail reconstruction.

179  
180 Contraction of  $\phi$ CD508 in urea reduces the tail length by just 20% whereas most CIS tails that  
181 have been described contract to ~50% of their extended length; the outer diameter of the sheath  
182 increases from 230 Å to 250 Å, and the inner diameter from 78 Å to 94 Å, sufficient to break  
183 contacts between gp55 sheath domain I helices and the tail tube gp56 (Fig. 3d). The contracted  
184 sheath retains the arrangement of gp55 subunits in a six-stranded helix, but with a decrease in  
185 helical rise from 39 Å to 32 Å, and a concomitant increase in twist from 19° to 34°. The contraction  
186 induced movement of the gp55 sheath proteins can be modelled as a rigid body pivoting motion  
187 about domain I (Extended Data Movie 1, Fig. 3a, b).

188  
189 Electron cryotomography of  $\phi$ CD508 virions incubated with S-layer fragments confirmed the 20%  
190 reduction in sheath length in the presence of the natural receptor (225 nm to 177 nm) (Fig. 4a). In

191 the urea contracted form of  $\phi$ CD508 the baseplate was lost, whereas the baseplate was retained  
192 when bound to the native S-layer receptor (Fig. 4b). Notably, at intermediate time points in the  
193 incubation we found that some fully contracted virions had not yet released their genome from the  
194 capsid (Fig. 4a). This was previously observed in images of the *Staphylococcus aureus* phage  
195  $\phi$ 812<sup>18</sup>.

196  
197 In both the extended and contracted form, the sheath ring retains integrity through its mesh  
198 network, similar to the sheath of other CISs<sup>12-14</sup> (Fig. 3b, d; Extended Data Movie 1). The overall  
199 RMSD between sheath protein monomers in the extended and contracted form is 0.63 Å,  
200 confirming a rigid body movement of these proteins. In helical reconstructions of contracted tail,  
201 the tail tube cannot be resolved due to a mismatch between the helical parameters of the sheath  
202 and tail tube. As the tail tube does not contract, the resulting mismatch in length drives the tail tube  
203 through the bottom of the baseplate. The measured 20% reduction (48 nm) in tail length is  
204 consistent with a decrease in helical rise from 39 Å to 32 Å for 57 rings of sheath protein.

205  
206 In the contracted phage tail, which lacks the packing interactions between the tail tube and the  
207 sheath, the bending stiffness (persistence length) is significantly reduced (Fig. 4a). Compared to  
208 other known contracted CIS structures, the more open packing of subunits both within the tail tube  
209 and tail sheath (Fig. 3, Extended Data Fig. 4) may allow the tail to be more flexible similar to the  
210 intermediate states of A511 contraction where flexibility is also observed<sup>17</sup>.

211

212

## 213 Discussion

214 Here we present a complete atomic model of the contractile bacteriophage,  $\phi$ CD508, in both  
215 extended and contracted states (Fig. 1).  $\phi$ CD508 is significant as it shows adaptations for binding  
216 to and penetration of the host cell S-layer, whereas all of the phage that have been structurally  
217 characterised previously, infect bacterial hosts that lack this widely encountered barrier<sup>6,15,17-19,35</sup>.  
218 The most striking difference between  $\phi$ CD508 and other known CIS structures is the much smaller  
219 relative contraction (Fig. 1, 5). There are other phages described in the literature for which reduced  
220 contraction has been observed, but not yet analysed<sup>39-41</sup>. Remarkably, these phages also infect  
221 Gram positive species that have S-layers. It is also notable that the tail tube protrusion length in  
222 these phages seems independent of total tail length (Fig. 5). This could represent an adaptation  
223 towards hosts with S-layers in general. To the best of our knowledge, for all other CISs studied in  
224 detail the tail tube protrusion length is correlated with the total tail length (Fig. 5). The other  
225 remarkable feature of  $\phi$ CD508 is its minimal baseplate with compact needle tip, and no enzymatic  
226 domains.

227

228 We reason that the reduced contraction of  $\phi$ CD508 is mediated by the sheath, as phages in which  
229 the baseplate has become detached remain in the same contracted state (Fig. 4), suggesting this  
230 state is stabilised by the sheath proteins alone and not influenced by partial rearrangement in the  
231 baseplate. The N-terminal linker forms part of the mesh that connects the sheath proteins to one  
232 another, and has been shown to act as a hinge maintaining links between subunits through the  
233 sheath contraction<sup>42,43</sup>. The increased length of the X-loop acts to restrict the motion of this N-  
234 terminal hinge compared with other CISs (Fig. 3c; Extended Data Fig. 4e). Further, the X-loop  
235 blocks complete contraction of the tail for  $\phi$ CD508 (Fig. 3); we modelled the  $\phi$ CD508 sheath

236 proteins into a more contracted conformation, based on pyocin trunk structures<sup>43</sup>, but increased  
237 contraction of  $\phi$ CD508 sheath proteins led to steric clashes between the X-loop and the  
238 neighbouring N-terminal linker as well as domain II of a neighbouring sheath (Extended Data Fig.  
239 4f).

240

241 The reduced compression between layers of sheath protein is not associated with a reduction in  
242 rotation. The helical twist between sheath disc layers increases from 19° to 34° during contraction,  
243 similar to other CISOs<sup>12,13,17,18,43</sup> (Fig. 5). A recent study showed that the rotation and translation of  
244 the tail sheath of pyocins occurs simultaneously during contraction<sup>42</sup>. Therefore, further  
245 contraction of  $\phi$ CD508 would lead to an over-rotation that cannot be accommodated by the mesh-  
246 like interactions between sheath proteins.

247

248 The disconnect between tail tube protrusion length and overall tail length in phages that infect S-  
249 layer coated bacteria is striking (Fig. 5) and may reflect the different energetics involved in  
250 penetrating the very diverse types of S-layers found in different species<sup>20</sup>, the need to  
251 accommodate different functionalities in the tape measure protein and the degree to which  
252 enzymatic attack contributes to the penetration process. For example, the phage tail-like Avidocins  
253 that also target *C. difficile* strains, have shorter tails than  $\phi$ CD508<sup>44</sup> but also incorporate predicted  
254 hydrolytic enzymes (Fig. 5).

255

256 Uniquely, here we were able to visualize exposure of the tail tube in the context of S-layer  
257 penetration. The major S-layer protein, SlpA, acts as the  $\phi$ CD508 receptor; we have previously  
258 shown that sequence variation within the main S-layer subunit SlpA (S-layer cassette type, SLCT)

259 correlates with phage infection spectrum and that heterologous expression of the SLCT-10 SlpA  
260 alone was sufficient to sensitise a normally resistant *C. difficile* strain to  $\phi$ CD508<sup>27</sup>. Tomograms  
261 of  $\phi$ CD508 bound to S-layer fragments (Fig. 4a, b) show how the phage interacts with the host  
262 during infection. Side views of  $\phi$ CD508 bound to the fragments show that the baseplate is held  
263 ~27 nm from the S-layer surface by the tail fibres (Fig. 4a, b). The thickness of the *C. difficile*  
264 cellular envelope is ~37 nm, and so combined with the ~27 nm distance from the S-layer, 48 nm  
265 of exposed tail tube would appear to be insufficient to reach the cell membrane and enter the  
266 cytoplasm. In T4 phage infection, the membrane is actively pinched outwards by 16 nm normal to  
267 its plane to meet the tail tube. It has been proposed that the tape measure protein gp29, containing  
268 a hydrophobic stretch, plays a role in this pinching to form a channel through which DNA is  
269 injected<sup>45</sup>. As our S-layer fragments lack membrane (Fig. 4), we cannot test for a similar pinching  
270 phenomenon, however, the  $\phi$ CD508 tape measure protein gp59 contains at least three predicted  
271 transmembrane stretches and may play a similar role to T4 gp29 (Extended Data Fig. 8). We also  
272 imaged phages contracted on whole cells; although the cells were too thick to visualise the  
273 membrane state upon phage adsorption and contraction, we did not observe any additional  
274 contraction compared with phages contracted on S-layer fragments. This suggests indeed that the  
275 membrane must be drawn closer to the end of the tail tube by some as yet unidentified mechanism.  
276  
277 A search of genome sequences of phages infecting bacteria with S-layers suggests that the  
278 particularly compact  $\phi$ CD508-form of the needle, may represent a widespread, but previously  
279 unreported adaptation (Extended Data Fig. 9b, c). The terminus of the  $\phi$ CD508 needle is indeed  
280 an apex domain as already reported in other phages<sup>46</sup>; it displays a conserved hexahistidine motif  
281 coordinating a putative iron ion (Extended Fig. 6, 7). However, the notable feature of  $\phi$ CD508,

282 and the other S-layer penetrating phages we have analysed, is that the needle lacks the otherwise  
283 widely conserved  $\beta$ -helix (Extended Data Fig. 9a). This  $\beta$ -helix has been implicated in piercing  
284 the outer membrane of Gram negative bacteria<sup>46</sup>, but is also found widely in CISs targeting Gram  
285 positive bacteria and eukaryotes<sup>13,46,47</sup>. The compact structure of the  $\phi$ CD508 type of needle is  
286 perhaps an adaptation towards easing it across the S-layer barrier, although the role of the needle  
287 apex itself in infection remains unclear<sup>48</sup>.

288

289 We propose a model for  $\phi$ CD508 phage infection (Fig. 4c). Initially, the phage binds to SlpA on  
290 the surface of the *C. difficile* cell<sup>49</sup> via receptor binding domains on the tail fibres. It is likely that  
291 binding is initially reversible and that the phage conducts a ‘random walk’ across the host cell  
292 surface (Fig. 4c (i), (ii)). These initially transient interactions have been observed or proposed for  
293 a number of other phages including those infecting *C. difficile*<sup>50-53</sup>. Only when the phage is  
294 committed to irreversible binding to the SlpA receptor does a structural change take place in the  
295 baseplate leading to tail sheath contraction, but only by 20% (Fig. 4c (ii), (iii)). It is tempting to  
296 speculate that the contraction drives the compact needle proteins through the surface of the cell  
297 envelope. However, the details of how this would happen are far from clear. Firstly, we do not  
298 know if the needle is driven through the very closely packed SlpA lattice itself or through one of  
299 the lattice defects/grain boundaries in the natural S-layer<sup>26</sup>; the latter might require less force but  
300 how weak points in the S-layer could be detected is unknown.

301

302 Once the S-layer barrier has been breached, the cell wall must then be penetrated. It is remarkable  
303 that  $\phi$ CD508 appears to have no enzymatic means of doing this, so that it would seem that  
304 mechanical work must be performed at this step. The energetics of this remain to be determined

305 although it should be noted that the cell wall is permeated by pores and cavities that may be weak  
306 points<sup>54</sup>.

307

308 The final barrier is the cell membrane. As the contraction length is insufficient for the tail tube to  
309 reach the cytoplasm, the membrane may bulge outwards to meet the tail tube in a process  
310 analogous to T4 phage infection<sup>45</sup>. The compressed DNA in the capsid might expel the trimeric  
311 tape measure protein gp59, which then forms a membrane channel via the predicted  
312 transmembrane helices of each polypeptide and through which DNA can be injected or drawn  
313 (Extended Data Fig. 8). It is also possible that local damage to the cell wall results in the membrane  
314 bulging as a consequence of the cell turgor pressure.

315

316 The detailed mechanism by which the potential energy of the extended state is converted into work  
317 to penetrate the host envelope is unknown for most CISs. Imaging of contraction intermediates of  
318 phages indicates that contraction is triggered at the baseplate and propagates as a wave along the  
319 tail<sup>17,55,56</sup>. It has been proposed that long-lived intermediates represent a stage in infection of Gram  
320 positive bacteria at which digestion of the cell wall is carried out by enzymes located at the tail  
321 tube tip<sup>17</sup>. In  $\phi$ CD508 we have not identified any intermediates in which the tail region proximal  
322 to the head remains in the uncontracted conformation<sup>17</sup>. This is consistent with the hypothesis that  
323 penetration of the S-layer and cell wall does not involve any enzymatic activity.

324

325 Fraser *et al.* modelled and experimentally measured the energetics and forces generated through  
326 the contraction of an R-type pyocin<sup>42</sup>. Interfacial sheath-sheath subunit interactions dominated the  
327 energetics; the free energy difference between the extended and contracted states was found to be

328 largely enthalpic. The model suggests that because of the wavelike nature of the contraction, CISs  
329 that are longer than the observed ‘contraction wavelength’ generate similar forces. This raises the  
330 question of why the tail of  $\phi$ CD508 is relatively long and why, although the sheath-sheath  
331 rearrangement is not as extensive as that in other CISs, sufficient force is still generated. It is  
332 possible that the ‘contraction wavelength’<sup>42</sup> is longer in  $\phi$ CD508 so that force is maintained over  
333 a longer distance. Moreover, the  $\phi$ CD508 tail becomes very flexible in the contracted state (Fig.  
334 4a) possibly adding an entropic component to the total free energy of contraction, thus  
335 compensating for the relatively smaller sheath-sheath interfacial rearrangement.

336

337 It is clear that  $\phi$ CD508 is a member of a novel class of contractile tail phages with reduced  
338 contraction (Fig. 5), including phages infecting other S-layer producing bacteria. Our CryoEM  
339 structures show that although many common features are shared between  $\phi$ CD508 and other CIS  
340 structures, there are some very significant differences. In order to develop phages as novel  
341 antimicrobials, and to engineer phages targeting clinically important strains of *C. difficile*,  
342 characterisation of the advantage these adaptations confer and the detailed interactions with SlpA  
343 will be vital.

344

## 345 **References**

- 346 1 Hatfull, G. F. Dark matter of the biosphere: the amazing world of bacteriophage  
347 diversity. *J Virol* **89**, 8107-8110 (2015). <https://doi.org/10.1128/JVI.01340-15>
- 348 2 Ackermann, H. W. 5500 Phages examined in the electron microscope. *Arch Virol* **152**,  
349 227-243 (2007). <https://doi.org/10.1007/s00705-006-0849-1>
- 350 3 Hardy, J. M., Dunstan, R. A., Lithgow, T. & Coulibaly, F. Tall tails: cryo-electron  
351 microscopy of phage tail DNA ejection conduits. *Biochem Soc Trans* **50**, 459-422W  
352 (2022). <https://doi.org/10.1042/BST20210799>
- 353 4 Fokine, A. & Rossmann, M. G. Molecular architecture of tailed double-stranded DNA  
354 phages. *Bacteriophage* **4**, e28281 (2014). <https://doi.org/10.4161/bact.28281>
- 355 5 Taylor, N. M. I., van Raaij, M. J. & Leiman, P. G. Contractile injection systems of  
356 bacteriophages and related systems. *Mol Microbiol* **108**, 6-15 (2018).  
357 <https://doi.org/10.1111/mmi.13921>
- 358 6 Chen, Z. *et al.* Cryo-EM structure of the bacteriophage T4 isometric head at 3.3-Å  
359 resolution and its relevance to the assembly of icosahedral viruses. *Proc Natl Acad Sci U*  
360 *S A* **114**, E8184-E8193 (2017). <https://doi.org/10.1073/pnas.1708483114>
- 361 7 Fokine, A. *et al.* Molecular architecture of the prolate head of bacteriophage T4. *Proc*  
362 *Natl Acad Sci U S A* **101**, 6003-6008 (2004). <https://doi.org/10.1073/pnas.0400444101>
- 363 8 Fokine, A. *et al.* The molecular architecture of the bacteriophage T4 neck. *J Mol Biol*  
364 **425**, 1731-1744 (2013). <https://doi.org/10.1016/j.jmb.2013.02.012>
- 365 9 Leiman, P. G. *et al.* Morphogenesis of the T4 tail and tail fibers. *Virol J* **7**, 355 (2010).  
366 <https://doi.org/10.1186/1743-422X-7-355>

- 367 10 Sun, L. *et al.* Cryo-EM structure of the bacteriophage T4 portal protein assembly at near-  
368 atomic resolution. *Nat Commun* **6**, 7548 (2015). <https://doi.org:10.1038/ncomms8548>
- 369 11 Taylor, N. M. *et al.* Structure of the T4 baseplate and its function in triggering sheath  
370 contraction. *Nature* **533**, 346-352 (2016). <https://doi.org:10.1038/nature17971>
- 371 12 Desfosses, A. *et al.* Atomic structures of an entire contractile injection system in both the  
372 extended and contracted states. *Nat Microbiol* **4**, 1885-1894 (2019).  
373 <https://doi.org:10.1038/s41564-019-0530-6>
- 374 13 Jiang, F. *et al.* Cryo-EM structure and assembly of an extracellular contractile injection  
375 system. *Cell* **177**, 370-383 e315 (2019). <https://doi.org:10.1016/j.cell.2019.02.020>
- 376 14 Ge, P. *et al.* Action of a minimal contractile bactericidal nanomachine. *Nature* **580**, 658-  
377 662 (2020). <https://doi.org:10.1038/s41586-020-2186-z>
- 378 15 Li, F. *et al.* High-resolution cryo-EM structure of the *Pseudomonas* bacteriophage E217.  
379 *Nat Commun* **14**, 4052 (2023). <https://doi.org:10.1038/s41467-023-39756-z>
- 380 16 Yang, F. *et al.* Fine structure and assembly pattern of a minimal myophage Pam3. *Proc*  
381 *Natl Acad Sci U S A* **120**, e2213727120 (2023). <https://doi.org:10.1073/pnas.2213727120>
- 382 17 Guerrero-Ferreira, R. C. *et al.* Structure and transformation of bacteriophage A511  
383 baseplate and tail upon infection of *Listeria* cells. *EMBO J* **38** (2019).  
384 <https://doi.org:10.15252/embj.201899455>
- 385 18 Novacek, J. *et al.* Structure and genome release of Twort-like Myoviridae phage with a  
386 double-layered baseplate. *Proc Natl Acad Sci U S A* **113**, 9351-9356 (2016).  
387 <https://doi.org:10.1073/pnas.1605883113>
- 388 19 Wang, Z. *et al.* Structure of *Vibrio* phage XM1, a simple contractile DNA injection  
389 machine. *Viruses* **15**, 1673 (2023). <https://doi.org:10.3390/v15081673>

- 390 20 Fagan, R. P. & Fairweather, N. F. Biogenesis and functions of bacterial S-layers. *Nat Rev*  
391 *Microbiol* **12**, 211-222 (2014). <https://doi.org/10.1038/nrmicro3213>
- 392 21 Guery, B., Galperine, T. & Barbut, F. *Clostridioides difficile*: diagnosis and treatments.  
393 *BMJ* **366**, 14609 (2019). <https://doi.org/10.1136/bmj.l4609>
- 394 22 Kortright, K. E., Chan, B. K., Koff, J. L. & Turner, P. E. Phage therapy: a renewed  
395 approach to combat antibiotic-resistant bacteria. *Cell Host Microbe* **25**, 219-232 (2019).  
396 <https://doi.org/10.1016/j.chom.2019.01.014>
- 397 23 Gebhart, D. *et al.* A modified R-type bacteriocin specifically targeting *Clostridium*  
398 *difficile* prevents colonization of mice without affecting gut microbiota diversity. *mBio* **6**  
399 (2015). <https://doi.org/10.1128/mBio.02368-14>
- 400 24 Kirk, J. A. *et al.* New class of precision antimicrobials redefines role of *Clostridium*  
401 *difficile* S-layer in virulence and viability. *Sci Transl Med* **9** (2017).  
402 <https://doi.org/10.1126/scitranslmed.aah6813>
- 403 25 Sekulovic, O., Garneau, J. R., Neron, A. & Fortier, L. C. Characterization of temperate  
404 phages infecting *Clostridium difficile* isolates of human and animal origins. *Appl Environ*  
405 *Microbiol* **80**, 2555-2563 (2014). <https://doi.org/10.1128/AEM.00237-14>
- 406 26 Lanzoni-Mangutchi, P. *et al.* Structure and assembly of the S-layer in *C. difficile*. *Nat*  
407 *Commun* **13**, 970 (2022). <https://doi.org/10.1038/s41467-022-28196-w>
- 408 27 Royer, A. L. M. *et al.* *Clostridioides difficile* S-layer protein A (SlpA) serves as a general  
409 phage receptor. *Microbiol Spectr* **11**, e0389422 (2023).  
410 <https://doi.org/10.1128/spectrum.03894-22>

- 411 28 Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-  
412 track neural network. *Science* **373**, 871-876 (2021).  
413 <https://doi.org/10.1126/science.abj8754>
- 414 29 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**,  
415 583-589 (2021). <https://doi.org/10.1038/s41586-021-03819-2>
- 416 30 Caspar, D. L. & Klug, A. Physical principles in the construction of regular viruses. *Cold*  
417 *Spring Harb Symp Quant Biol* **27**, 1-24 (1962).  
418 <https://doi.org/10.1101/sqb.1962.027.001.005>
- 419 31 Lopes, A., Tavares, P., Petit, M. A., Guerois, R. & Zinn-Justin, S. Automated  
420 classification of tailed bacteriophages according to their neck organization. *BMC*  
421 *Genomics* **15**, 1027 (2014). <https://doi.org/10.1186/1471-2164-15-1027>
- 422 32 Lhuillier, S. *et al.* Structure of bacteriophage SPP1 head-to-tail connection reveals  
423 mechanism for viral DNA gating. *Proc Natl Acad Sci U S A* **106**, 8507-8512 (2009).  
424 <https://doi.org/10.1073/pnas.0812407106>
- 425 33 Arnaud, C. A. *et al.* Bacteriophage T5 tail tube structure suggests a trigger mechanism for  
426 Siphoviridae DNA ejection. *Nat Commun* **8**, 1953 (2017).  
427 <https://doi.org/10.1038/s41467-017-02049-3>
- 428 34 Kizziah, J. L., Manning, K. A., Dearborn, A. D. & Dokland, T. Structure of the host cell  
429 recognition and penetration machinery of a *Staphylococcus aureus* bacteriophage. *PLoS*  
430 *Pathog* **16**, e1008314 (2020). <https://doi.org/10.1371/journal.ppat.1008314>
- 431 35 Zheng, W. *et al.* Refined cryo-EM structure of the T4 tail tube: exploring the lowest dose  
432 limit. *Structure* **25**, 1436-1441 e1432 (2017). <https://doi.org/10.1016/j.str.2017.06.017>

- 433 36 Zinke, M. *et al.* Architecture of the flexible tail tube of bacteriophage SPP1. *Nat Commun*  
434 **11**, 5759 (2020). [https://doi.org:10.1038/s41467-020-19611-1](https://doi.org/10.1038/s41467-020-19611-1)
- 435 37 Phetruen, T. *et al.* Receptor binding protein of prophage reversibly recognizes the low-  
436 molecular weight subunit of the surface-layer protein SlpA in *Clostridioides difficile*.  
437 *Front Microbiol* **13**, 998215 (2022). [https://doi.org:10.3389/fmicb.2022.998215](https://doi.org/10.3389/fmicb.2022.998215)
- 438 38 Arisaka, F., Kanamaru, S., Leiman, P. & Rossmann, M. G. The tail lysozyme complex of  
439 bacteriophage T4. *Int J Biochem Cell Biol* **35**, 16-21 (2003).  
440 [https://doi.org:10.1016/s1357-2725\(02\)00098-5](https://doi.org/10.1016/s1357-2725(02)00098-5)
- 441 39 Feyereisen, M. *et al.* Isolation and characterization of *Lactobacillus brevis* phages.  
442 *Viruses* **11** (2019). [https://doi.org:10.3390/v11050393](https://doi.org/10.3390/v11050393)
- 443 40 Kaliniene, L. *et al.* Molecular analysis of *Arthrobacter* myovirus vB\_ArtM-ArV1: we  
444 blame it on the tail. *J Virol* **91** (2017). [https://doi.org:10.1128/JVI.00023-17](https://doi.org/10.1128/JVI.00023-17)
- 445 41 Klumpp, J. *et al.* The odd one out: *Bacillus* ACT bacteriophage CP-51 exhibits unusual  
446 properties compared to related Spounavirinae W.Ph. and Bastille. *Virology* **462-463**, 299-  
447 308 (2014). [https://doi.org:10.1016/j.virol.2014.06.012](https://doi.org/10.1016/j.virol.2014.06.012)
- 448 42 Fraser, A. *et al.* Quantitative description of a contractile macromolecular machine. *Sci*  
449 *Adv* **7** (2021). [https://doi.org:10.1126/sciadv.abf9601](https://doi.org/10.1126/sciadv.abf9601)
- 450 43 Ge, P. *et al.* Atomic structures of a bactericidal contractile nanotube in its pre- and  
451 postcontraction states. *Nat Struct Mol Biol* **22**, 377-382 (2015).  
452 [https://doi.org:10.1038/nsmb.2995](https://doi.org/10.1038/nsmb.2995)
- 453 44 Gebhart, D. *et al.* Novel high-molecular-weight, R-type bacteriocins of *Clostridium*  
454 *difficile*. *J Bacteriol* **194**, 6240-6247 (2012). [https://doi.org:10.1128/JB.01272-12](https://doi.org/10.1128/JB.01272-12)

- 455 45 Hu, B., Margolin, W., Molineux, I. J. & Liu, J. Structural remodeling of bacteriophage  
456 T4 and host membranes during infection initiation. *Proc Natl Acad Sci U S A* **112**,  
457 E4919-4928 (2015). [https://doi.org:10.1073/pnas.1501064112](https://doi.org/10.1073/pnas.1501064112)
- 458 46 Browning, C., Shneider, M. M., Bowman, V. D., Schwarzer, D. & Leiman, P. G. Phage  
459 pierces the host cell membrane with the iron-loaded spike. *Structure* **20**, 326-339 (2012).  
460 [https://doi.org:10.1016/j.str.2011.12.009](https://doi.org/10.1016/j.str.2011.12.009)
- 461 47 Spinola-Amilibia, M. *et al.* The structure of VgrG1 from *Pseudomonas aeruginosa*, the  
462 needle tip of the bacterial type VI secretion system. *Acta Crystallogr D Struct Biol* **72**,  
463 22-33 (2016). [https://doi.org:10.1107/S2059798315021142](https://doi.org/10.1107/S2059798315021142)
- 464 48 Miller, J. M., Knyazhanskaya, E. S., Buth, S. A., Prokhorov, N. S. & Leiman, P. G.  
465 Function of the bacteriophage P2 baseplate central spike Apex domain in the infection  
466 process. *bioRxiv* (2023). [https://doi.org:10.1101/2023.02.25.529910](https://doi.org/10.1101/2023.02.25.529910)
- 467 49 Whittle, M. J. *et al.* A Novel bacteriophage with broad host range against *Clostridioides*  
468 *difficile* ribotype 078 supports SlpA as the likely phage receptor. *Microbiol Spectr* **10**,  
469 e0229521 (2022). [https://doi.org:10.1128/spectrum.02295-21](https://doi.org/10.1128/spectrum.02295-21)
- 470 50 Islam, M. Z. *et al.* Molecular anatomy of the receptor binding module of a bacteriophage  
471 long tail fiber. *PLoS Pathog* **15**, e1008193 (2019).  
472 [https://doi.org:10.1371/journal.ppat.1008193](https://doi.org/10.1371/journal.ppat.1008193)
- 473 51 Moldovan, R., Chapman-McQuiston, E. & Wu, X. L. On kinetics of phage adsorption.  
474 *Biophys J* **93**, 303-315 (2007). [https://doi.org:10.1529/biophysj.106.102962](https://doi.org/10.1529/biophysj.106.102962)
- 475 52 Quiberoni, A., Guglielmotti, D., Binetti, A. & Reinheimer, J. Characterization of three  
476 *Lactobacillus delbrueckii* subsp. *bulgaricus* phages and the physicochemical analysis of

477 phage adsorption. *J Appl Microbiol* **96**, 340-351 (2004). <https://doi.org/10.1046/j.1365->  
478 [2672.2003.02147.x](https://doi.org/10.1046/j.1365-2672.2003.02147.x)

479 53 Thanki, A. M. *et al.* Unravelling the links between phage adsorption and successful  
480 infection in *Clostridium difficile*. *Viruses* **10** (2018). <https://doi.org/10.3390/v10080411>

481 54 Pasquina-Lemonche, L. *et al.* The architecture of the Gram-positive bacterial cell wall.  
482 *Nature* **582**, 294-297 (2020). <https://doi.org/10.1038/s41586-020-2236-6>

483 55 Donelli, G., Guglielmi, F. & Paoletti, L. Structure and physico-chemical properties of  
484 bacteriophage G. I. Arrangement of protein subunits and contraction process of tail  
485 sheath. *J Mol Biol* **71**, 113-125 (1972). [https://doi.org/10.1016/0022-2836\(72\)90341-5](https://doi.org/10.1016/0022-2836(72)90341-5)

486 56 Moody, M. F. Sheath of bacteriophage T4. 3. Contraction mechanism deduced from  
487 partially contracted sheaths. *J Mol Biol* **80**, 613-635 (1973). <https://doi.org/10.1016/0022->  
488 [2836\(73\)90200-3](https://doi.org/10.1016/0022-2836(73)90200-3)

489

490 **Data Availability**

491 CryoEM maps will be deposited with the Electron Microscopy Data Bank and coordinates will be  
492 deposited with the Protein Data Bank. The complete  $\phi$ CD508 genome sequence is available  
493 through the National Center for Biotechnology Information database under the accession number  
494 OR295560.

495

496 **Acknowledgements**

497 We thank Svetomir Tzokov of the Faculty of Science Electron Microscopy Facility, University of  
498 Sheffield, for assistance with EM and Adelina Acosta Martin of the Faculty of Science Mass  
499 Spectrometry Facility for assistance with MS analysis. We acknowledge Diamond Light Source  
500 for access and support of the cryo-EM facilities at the UK national electron Bio-Imaging Centre  
501 (eBIC), proposal EM19832, funded by the Wellcome Trust, MRC and BBSRC, Yun Song and  
502 Vinod Kumar Vogirala for help with data collection and Mathew Arnold for help with data  
503 processing. We are most grateful for helpful discussions with Rebekkah Menday, David Rice,  
504 Rebecca Corrigan and Indrajit Lahiri.

505

506 This work was funded by BBSRC grant BB/P02002X/1 to PAB and RPF and The University of  
507 Sheffield's Imagine:Imaging Life programme.

508

509 For the purpose of open access, the authors have applied a Creative Commons Attribution (CC  
510 BY) licence to any Author Accepted Manuscript version arising.

511

512

513 **Author contributions**

514 JSW designed the study, carried out all experiments, collected and analysed data, wrote and revised  
515 the manuscript. LCF designed experiments, analysed data, and revised the manuscript. RPF  
516 designed the study, analysed data, supervised the study, wrote and revised the manuscript. PAB  
517 designed the study, analysed and interpreted EM data, supervised the study, wrote and revised the  
518 manuscript.

519

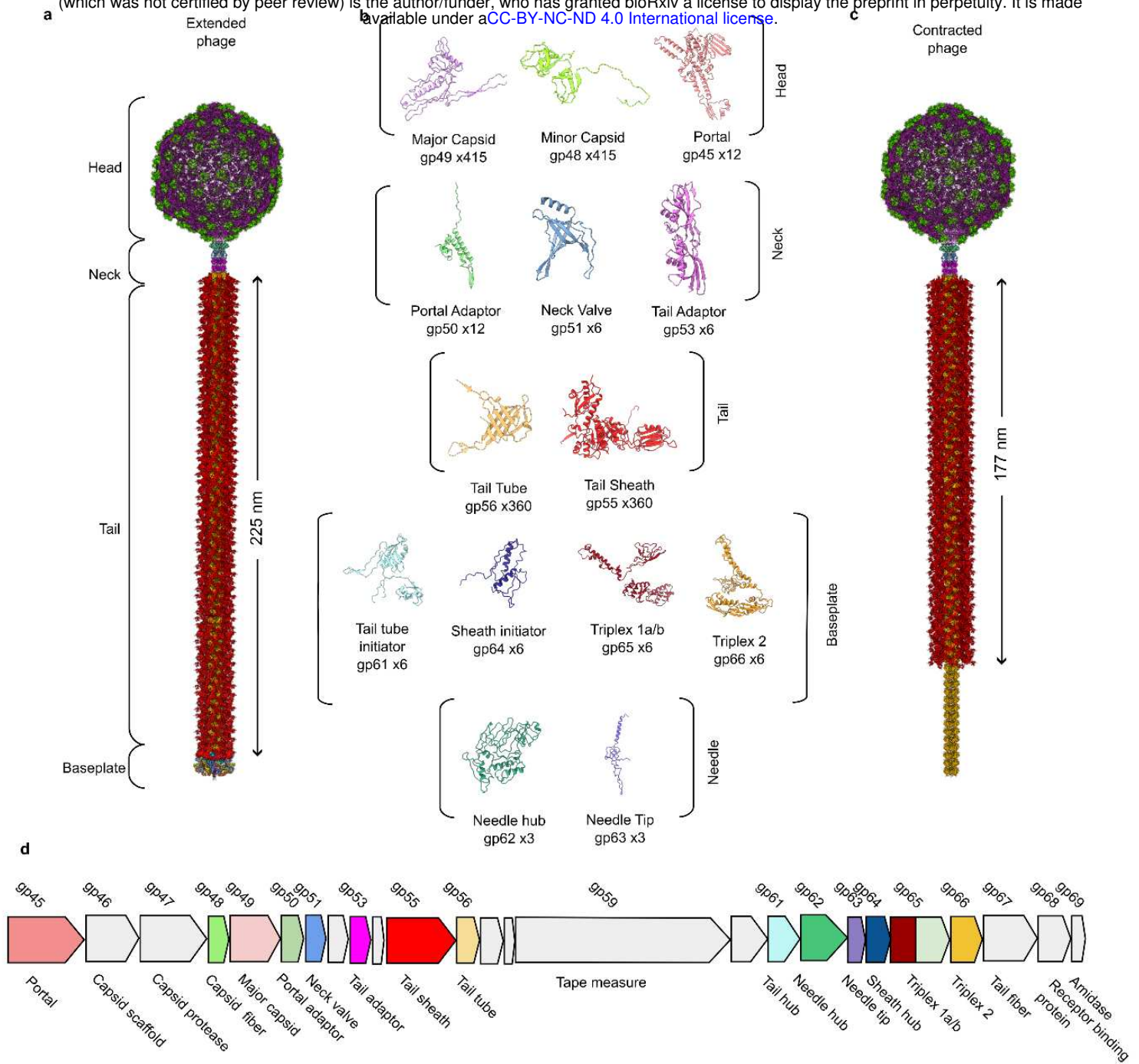
520 **Competing interests**

521 The authors declare no competing interests.

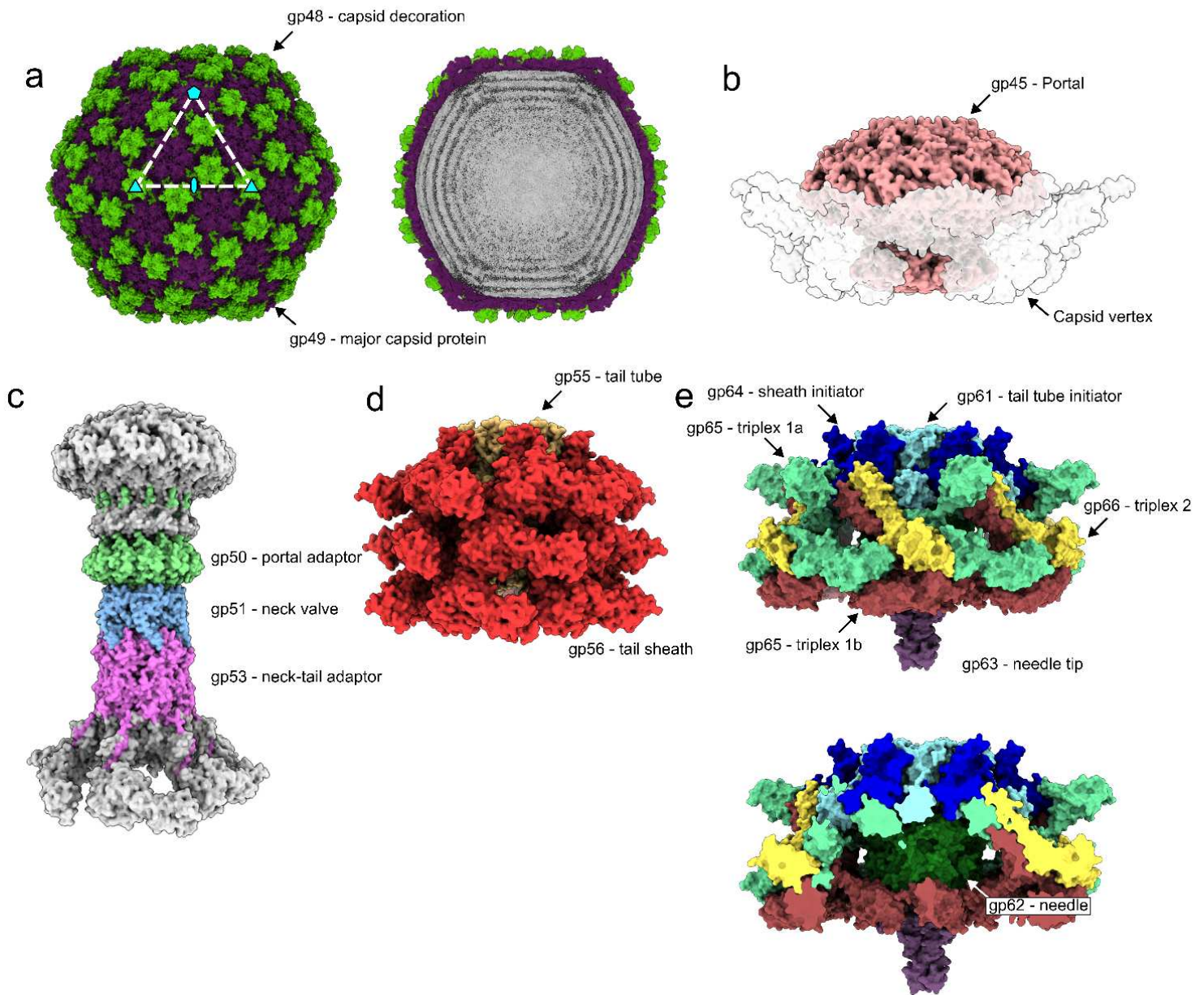
522

523 **Materials and correspondence**

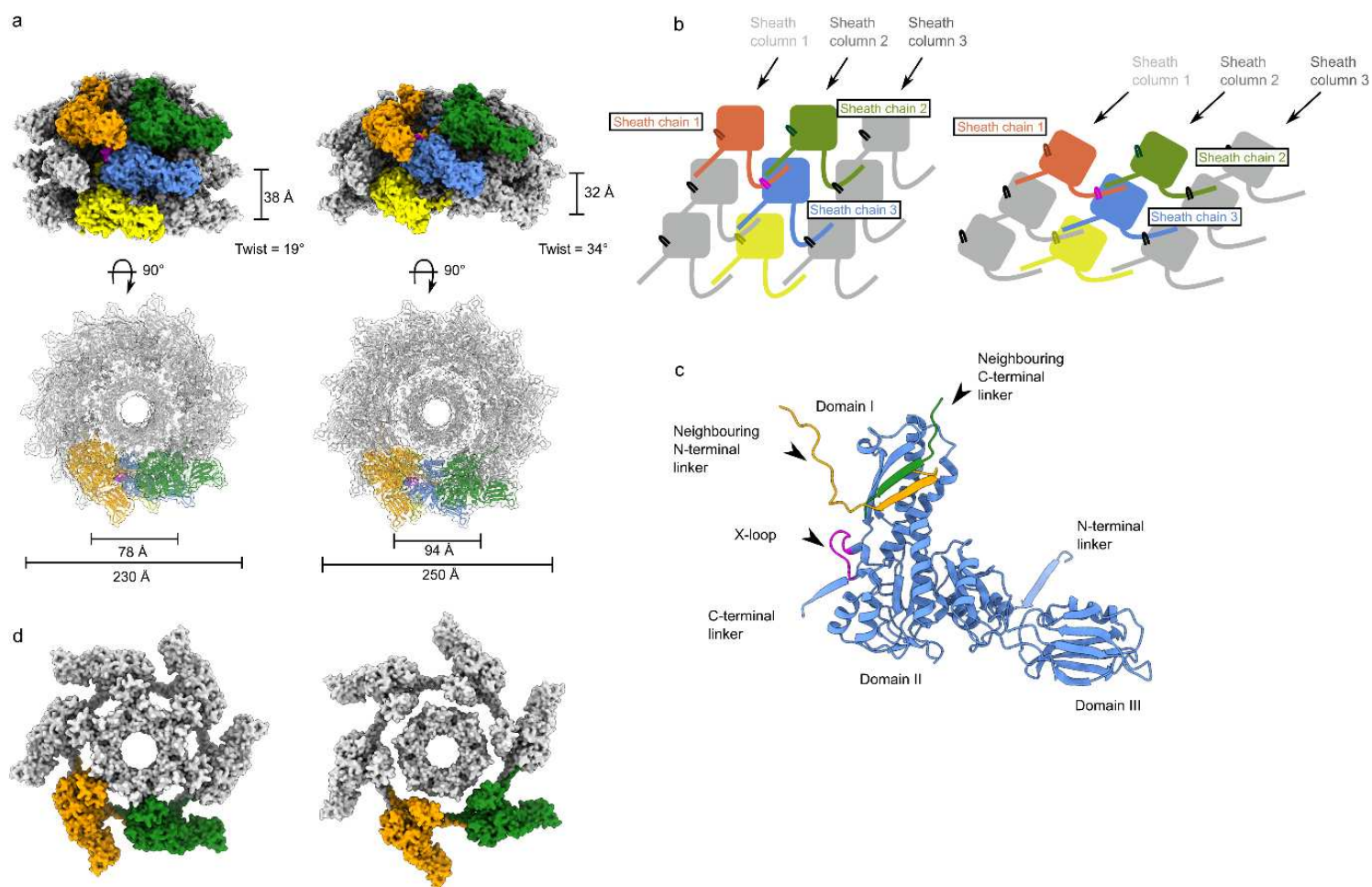
524 Requests should be addressed to Per Bullough (email:p.bullough@sheffield.ac.uk) or Robert  
525 Fagan (email:r.fagan@sheffield.ac.uk).



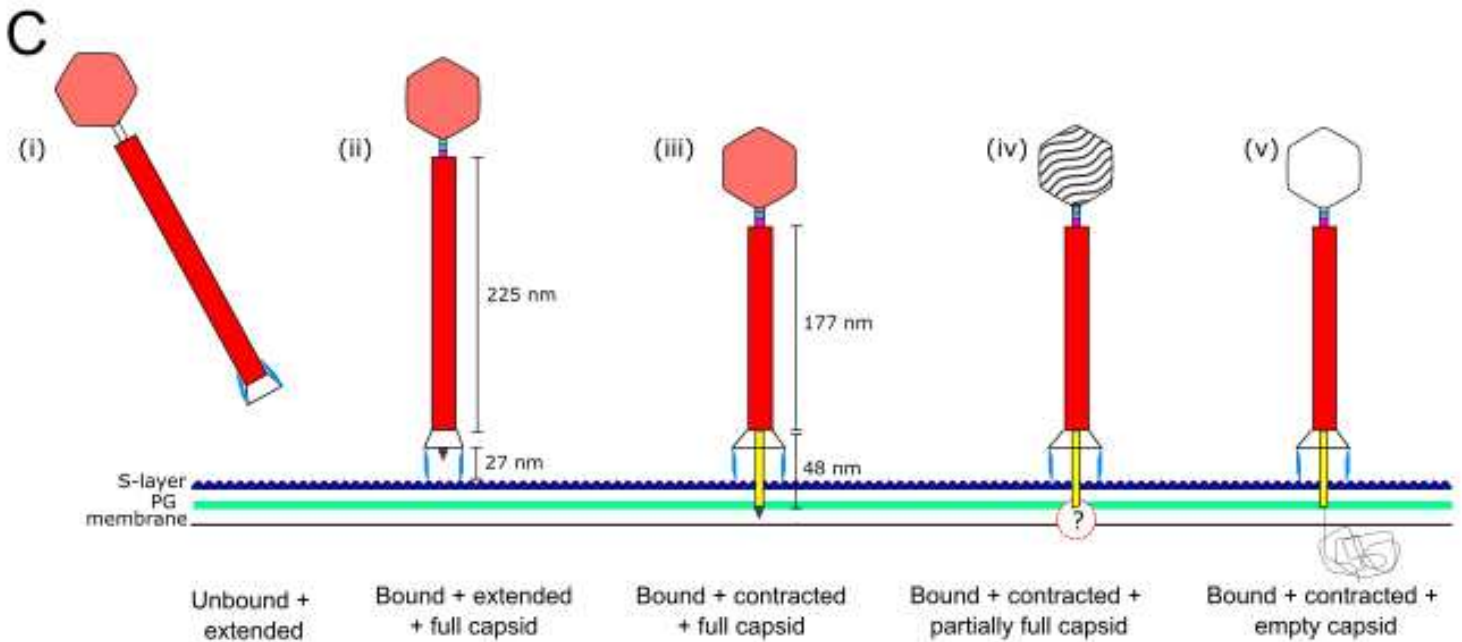
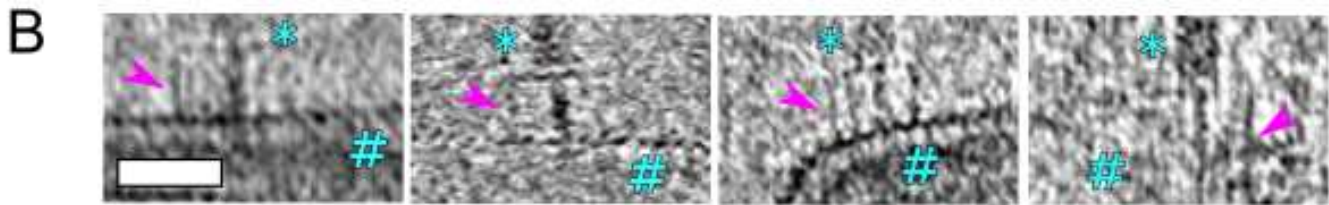
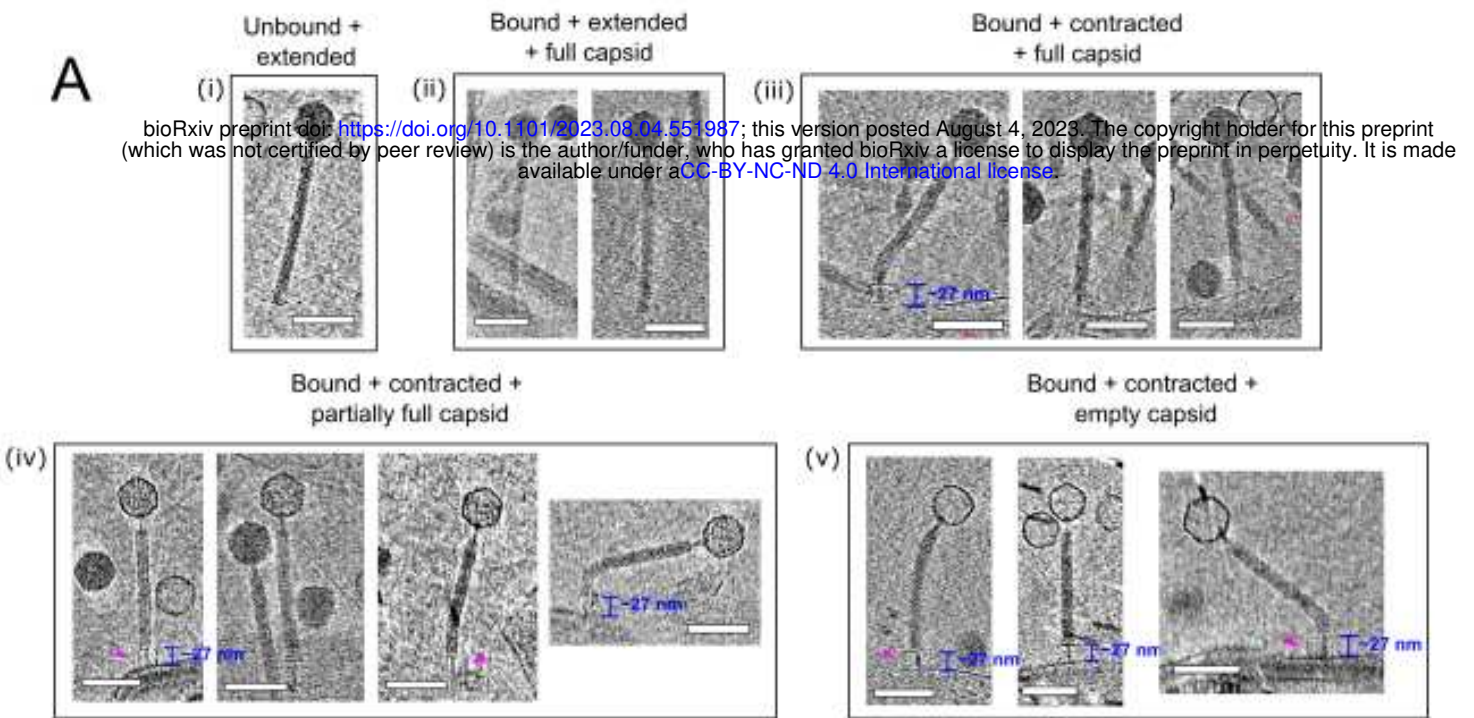
**Figure 1 | Structural overview of  $\phi$ CD508. a.** Composite model of entire extended  $\phi$ CD508 phage generated from overlapping features of each model. **b.** Gallery of proteins built into cryoEM maps of  $\phi$ CD508, with protein name, gene product number, and copy number in entire extended phage. **c.** Composite model of entire contracted  $\phi$ CD508 phage generated from overlapping features of each model. **d.** Genome organisation of  $\phi$ CD508 structural cassette consisting of gene products 45 to 68. Proteins built into CryoEM maps are coloured as in the gallery, and proteins, where no model could be built, are coloured grey.



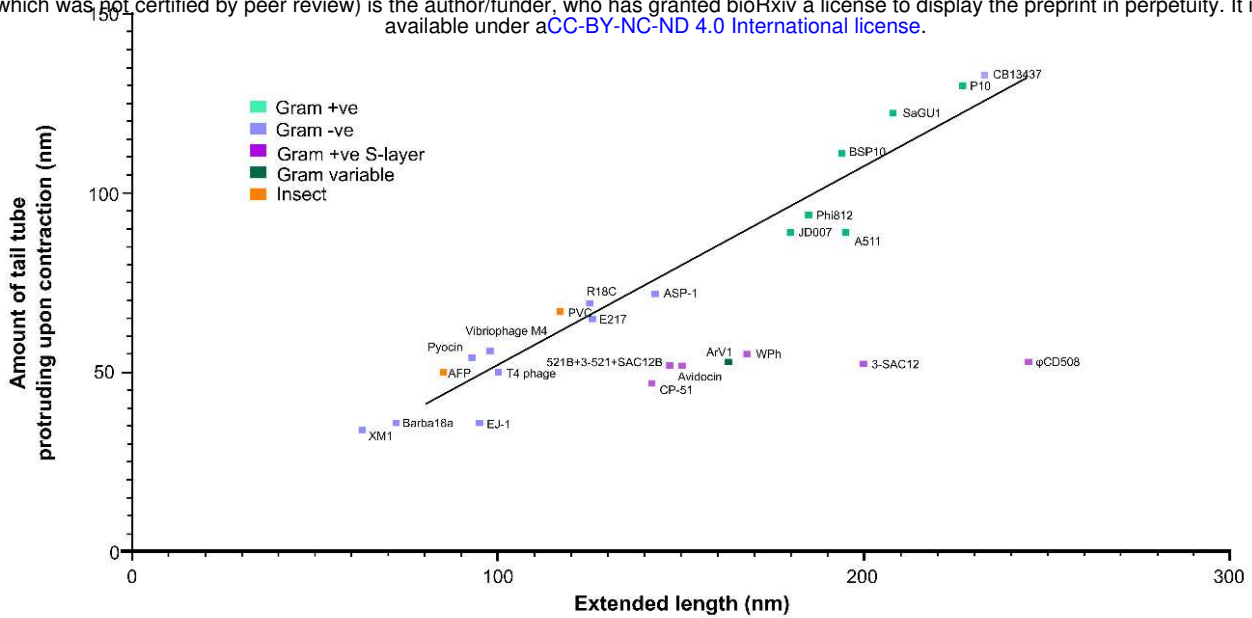
**Figure 2 | Structural assemblies of  $\phi$ CD508.** Shaded surface representation of  $\phi$ CD508 assemblies as determined by single particle analysis. **a.** Capsid consisting of gp48 and gp49. Slice through extended phage capsid shows DNA layers (right). **b.** Portal protein dodecamer within the unique pentameric capsid vertex, shown with a transparent surface. **c.** Neck proteins shaded by protein identity (portal adaptor = green, neck valve = blue, neck-tail adaptor = pink), with portal and sheath proteins shown in grey for context. **d.** Three layers of sheath (red) and tail tube protein (orange) in the extended state. **e.** Baseplate and needle assembly with hub (blue and cyan), wedge (yellow, mint, and brown), and needle (green and purple).



**Figure 3 | Reduced contraction of the  $\phi$ CD508 sheath.** **a.** Surface rendering of the extended tail (left) and contracted tail (right), showing helical parameters of sheath proteins and inner and outer sheath dimensions. Four interwoven chains are coloured (orange, green, blue, and yellow), and the X-loop for the blue chain is shown in magenta. **b.** Schematic of mesh network between sheath layers coloured as in a. **c.** Cartoon representation of a single sheath protein showing domain architecture and the X-loop region, coloured as in a. Also shown are the C-terminal linker and N-terminal linker that form the extended mesh between neighbouring chains. **d.** A Surface rendering of the top layer of the sheath and tail tube proteins as in a. In the extended state (left), the sheath forms extensive interactions between the sheath and tail tube. In the contracted form (right), the modelled tail tube is able to pass through the sheath ring.



**Figure 4 | Model for  $\phi$ CD508 reduced contraction.** **a.** Gallery of  $\phi$ CD508 bacteriophage bound to S-layer fragments from cryoelectron tomograms. The order of images follows the predicted model of phage infection, being free extended phage (i), attachment to the S-layer surface (ii), contraction but no DNA release from the capsid (iii), partial emptying of the capsid (iv), and empty capsids (v). Stages ii-v show examples where the phage is able to bend once contracted. Scale bar = 100 nm **b.** Highlighted baseplates (denoted with \*) attached to S-layer fragments (denoted with #), with tail fibers (gp67-gp68) visible and highlighted with magenta arrows. **c.** Schematic model of  $\phi$ CD508 as in a.



b

Phage/PTLP	Extended length	Contracted length	Contraction amount	Length of needle protruding	Extended rise/twist	Contracted rise/twist	Contraction method	Reference
$\phi$ CD508	225 nm	177 nm	22%	48 nm	38.8 Å, 19.2°	31.6 Å, 33.9°	3 M Urea, S-layer, heat-induced	This paper
phi812	185 nm	94 nm	49%	94 nm	38.9 Å, 21.4°	18.9 Å, 30.7°	3 M Urea	[17]
T4 phage	100 nm	50 nm	50%	50 nm	40.6 Å, 17.2°	16.4 Å, 32.9°	4 M Urea	[10]
A511	195 nm	106 nm	46%	89 nm	N/A	N/A	2 M Urea	[18]
Pyocin	93 nm	39 nm	58%	54 nm	38.4 Å, 18.3°	16.2 Å, 33.1°	Naturally	[14]
AFP	85 nm	35 nm	59%	50 nm	40.7 Å, 20.0°	17.4 Å, 31.7°	3 M Gu-HCl	[13]
PVC	117 nm	50 nm	57%	67 nm	39.3 Å, 19.9°	17.0 Å, 31.4°	Naturally	[12]
XM1	63 nm	29 nm	54%	34 nm	N/A	N/A	Naturally	[19]
E217	125 nm	60 nm	52%	65nm	40.8 Å, 31.3°	22.1Å, 30.0°	pH	[15]

**Figure 5 | Comparison of contraction parameters in CISS.** a. Graph of the extended tail length vs. the amount of tail tube protruding upon contraction for a number of contractile bacteriophage and phage tail-like particles (PTLPs) from literature. Each particle is coloured based on the bacterial/insect host type. ‘Conventional’ virions fit well along a trendline ( $R^2 = 0.975$ ). b. A table of tail parameters for bacteriophage and PTLPs for which structural data are available.