

This is a repository copy of *Cryo-EM structure in situ reveals a molecular switch that safeguards virus against genome loss*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/159724/>

Version: Accepted Version

Article:

Bayfield, Oliver W, Steven, Alasdair C and Antson, Alfred A orcid.org/0000-0002-4533-3816 (2020) Cryo-EM structure in situ reveals a molecular switch that safeguards virus against genome loss. *eLife*. e55517. ISSN 2050-084X

<https://doi.org/10.7554/eLife.55517>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

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 including the URL of the record and the reason for the withdrawal request.

Cryo-EM structure *in situ* reveals a molecular switch that safeguards virus against genome loss

Oliver W. Bayfield^{a,b}, Alasdair C. Steven^b, and Alfred A. Antson^a

^aYork Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5DD, United Kingdom.

^bLaboratory of Structural Biology Research, National Institute of Arthritis Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD 20892, U.S.A.

Abstract

The portal protein is a key component of many double-stranded DNA viruses, governing capsid assembly and genome packaging. Twelve subunits of the portal protein define a tunnel, through which DNA is translocated into the capsid. It is unknown how the portal protein functions as a gatekeeper, preventing DNA slippage, whilst allowing its passage into the capsid, and how these processes are controlled. A cryo-EM structure of the portal protein of thermostable virus P23-45, determined *in situ* in its procapsid-bound state, indicates a mechanism that naturally safeguards the virus against genome loss. This occurs via an inversion of the conformation of the loops that define the constriction in the central tunnel, accompanied by a hydrophilic–hydrophobic switch. The structure also shows how translocation of DNA into the capsid could be modulated by a changing mode of protein–protein interactions between portal and capsid, across a symmetry-mismatched interface.

portal protein | virus assembly | DNA packaging | cryo-EM | bacteriophage | symmetry mismatch

Correspondence: oliver.bayfield@york.ac.uk, fred.antson@york.ac.uk

Introduction

Tailed bacteriophages constitute the majority of viruses in the biosphere (Bergh et al., 1989; Michaud et al., 2018) and are a significant component of the human microbiome (Shkoporov and Hill, 2019). During assembly, these viruses translocate their genomic double-stranded DNA through a portal protein that occupies a single vertex of an icosahedral capsid. A similar mechanism is employed by the evolutionarily related herpesviruses (McElwee et al., 2018; Trus et al., 2004). Structural information about the portal protein is important not only for deducing the mechanism of capsid assembly (Chen et al., 2011), but also for understanding molecular events associated with genome translocation into preformed capsids (Mao et al., 2016; Sun et al., 2015, 2008), and genome ejection during infection (Wu et al., 2016). Although structures of isolated portal proteins, without the native capsid environment, have been determined to near-atomic resolution by X-ray crystallography and cryo-electron microscopy (Lebedev et al., 2007; Lokareddy et al., 2017; Simpson et al., 2000; Sun et al., 2015), a number of observations

34 concerning these structures have yet to be rationalised in the context of the portal's many functional
35 roles, including: the variable diameter of the central tunnel and flexibility of tunnel loops (Lebedev et al.,
36 2007; Simpson et al., 2000; Sun et al., 2015); the symmetry mismatch between the portal and capsid
37 vertex (12-fold versus 5-fold) (Simpson et al., 2000; Sun et al., 2008); and the portal's role in DNA
38 translocation (Harvey, 2015; Ray et al., 2010). The influence of the properties of the internal tunnel, and
39 how these can be modulated by external factors to coordinate DNA translocation, remains unclear.
40 Cryo-EM studies on mesophilic herpesviruses characterised the shape of the portal protein tunnel in the
41 mature virion and showed how DNA can be locked inside (Liu et al., 2019; McElwee et al., 2018).
42 However, there are no detailed structural data on portal proteins *in situ* for unexpanded capsids, primed
43 for DNA packaging. Moreover, it has proven difficult to derive accurate models for tunnel loops of the
44 portal protein, such as in the case of tailed bacteriophage ϕ 29, where the flexible nature of the tunnel
45 loops prevented their observation in a crystal structure (Simpson et al., 2000) and also in cryo-EM
46 structures of the procapsid and mature capsid (Xu et al., 2019).

47 To gain knowledge about the structure of the dynamic DNA tunnel, we utilised a thermostable
48 bacteriophage, P23-45. Thermophilic viruses must package their genomes under extreme temperature,
49 imposing additional challenges compared to their mesophilic counterparts. This *Thermus thermophilus*
50 bacteriophage is one of the few viruses for which conditions for packaging DNA into capsids *in vitro*
51 have been established, and where isolated empty capsids were demonstrated to be competent at
52 packaging DNA (Bayfield et al., 2019). Previous cryo-EM reconstructions of procapsids (unexpanded)
53 and mature capsids (expanded), in which icosahedral symmetry was imposed, have revealed the
54 extent of conformational changes that the major capsid proteins undergo upon capsid maturation.
55 During the transition, the capacity of the capsid nearly doubles (Bayfield et al., 2019). In this study, the
56 structure of the portal protein *in situ*, and analysis of the reconstruction of the unexpanded procapsid
57 without imposing icosahedral symmetry, reveal substantive conformational differences in the structure
58 of the portal protein (Bayfield et al., 2019). The most remarkable difference, induced *in situ*, is an
59 inversion in the conformation of tunnel loops of the portal protein. The tunnel loop inversion "switches"
60 the surface properties at the tunnel's constriction from hydrophobic to hydrophilic and creates a wider
61 opening. These observations indicate that the capsid shell plays a role in defining the conformation and
62 properties of the portal protein, modulating DNA translocation into capsid.

63 **Results**

64 **Structure of the *in situ* procapsid portal.** P23-45 procapsids were purified from lysates of infected
65 *Thermus thermophilus* cells (Figure 1A). The procedures used for cryo-EM data collection and
66 computing the icosahedrally-averaged reconstruction were described earlier (Bayfield et al., 2019). The
67 *in situ* structure of the portal protein within the procapsid was determined by localised reconstruction of

68 portal-containing vertices to a resolution of 3.7 Å by averaging around the 12-fold symmetric axis
69 (Supplementary File 1, Figure 1-figure supplement 1) (Ilca et al., 2015). The portal protein oligomer is a
70 ring of 12 subunits (Figure 1B, C, Video 1), with each subunit folded into Crown, Wing, Stem, Clip, and
71 Tunnel loop domains (Figure 1D). Most amino acid side-chains were clearly resolved in the map
72 (Figure 1E), enabling construction of an accurate atomic model (PDB 6QJT). Comparison with the
73 crystal structure of the portal protein from the closely related phage G20c (PDB 6IBG, 99.3% sequence
74 identity) reveals several significant structural differences: notably, in the positions of the Crown and
75 Wing domains and in the conformation of the tunnel loops (Figure 2, Video 2). In the *in situ* structure,
76 the C-terminal Crown domain (residues 377–436) is shifted upwards away from the main body by ~5 Å
77 (Figure 2A,B), and twisted by ~13° around the central axis (Figure 2C, Video 2). The Wing domain
78 pivots ~8° downwards, towards the Clip (Figure 2B, Video 2). Although the two portal proteins compared
79 are from different phage, they have closely related sequences. The most conservative substitution,
80 I328V, is located in the tunnel, and two additional conservative substitutions, S189N and S367G, are
81 located at the outer surface of the wing in residues with solvent-exposed side chains. Such mutations
82 are unlikely to bring about the observed differences in conformation.

83 **Differences between the portal conformations in the *in situ* and crystal structures.** The most
84 pronounced conformational differences seen in the *in situ* structure are in the tunnel loops (Figure 2).
85 The tunnel diameter at its most constricted part is wider by ~8 Å (Figure 2E, F). Hydrophobic residues
86 V325 and I330 are no longer exposed to the tunnel as they are in the crystal structure, and are replaced
87 by polar residues Q326 and N329 due to inversion in the tunnel loop conformation (Figure 2D).
88 Residues 330–335, which protrude into the tunnel and are part of the longest helix in the crystal
89 structure, instead adopt an extended loop conformation *in situ* (Figure 2D), facilitating the tunnel loop
90 remodelling. These modifications alter the shape and surface properties of the tunnel, which widens and
91 changes from hydrophobic to hydrophilic (compare Figs. 2E, F).

92 The first N-terminal residue that could be reliably modelled in the *in situ* reconstruction was Leu26
93 (Figure 2D), in common with the crystal structure (PDB 6IBG). Mass spectrometry detected N-terminal
94 residues of the portal protein subunits (Figure 1-figure supplement 2), indicating that the 25-amino acid N-
95 terminal segment is present in at least some chains, but adopts variable conformations. Although the
96 first residue with a defined conformation points toward the interior of the capsid in P23-45, it cannot be
97 ruled out that the flexible N-terminal segment folds back and contributes to portal-capsid interactions.

98 **Portal–Capsid interactions.** The portal–capsid interface is spacious, with only relatively small surface
99 areas of the portal’s, within the wing and clip domains, engaged in interactions with the capsid (Figure
100 3A,B). Fitting of the C12-symmetrised portal reconstruction, presented here, into the asymmetric
101 procapsid reconstruction (Bayfield et al., 2019), reveals the details of the portal-capsid interactions at

102 this symmetry-mismatched interface. In the asymmetric reconstruction of the procapsid, the portal
103 protein appears 12-fold symmetric (Bayfield et al., 2019). Residues 185–189 (β -hairpin loop) of the
104 portal Wing are involved in interactions with the capsid (Figure 3B). These loops may pivot downwards
105 to make closer contact with the capsid inner wall, in select chains (Figure 3B,C). Such adjustments in
106 specific subunits of the portal protein would not be resolved in a symmetrically averaged structure;
107 however, the bridging regions observed between the capsid and the portal in the asymmetric procapsid
108 reconstruction suggests local deviations from C12 symmetry are possible. The portal Wing (β -hairpin
109 loop) is in close proximity to residues 24-30 and 337-340 of the major capsid protein (Figure 3B).
110 Interactions of the portal Clip likely involve portal protein residues 263-275 (α -helix and adjacent loop
111 within the Clip domain) interacting with the major capsid protein residues 119-127 and 357-358 (Figure
112 3D). In common with ϕ 29 (Simpson et al., 2000), portal-capsid interactions are mediated by residues of
113 both polar and hydrophobic character. The portal–capsid symmetry mismatch means that only select
114 portal chains make contact with the capsid: these are chains A-C-(E/F)-H-J at the Wing (Figure 3E), and
115 chains C-E-(G/H)-J-L at the Clip (Figure 3F).

116 Discussion

117 **Procapsid assembly primes the portal for packaging.** The *in situ* structure of the portal protein differs
118 from the crystal structure globally, in changes in domain positions, and locally, in conformational
119 changes such as the inversion of the tunnel loop. Structural data indicate how the changes on these two
120 levels are linked:

- 121 1. Assembly of capsid proteins around the portal stabilises a $\sim 8^\circ$ rotational adjustment in Wing
122 domains (Figure 2B).
- 123 2. As the Wing domain pivots, the C-terminal helix of the Crown domain that interacts with the Wing,
124 slides, facilitating a $\sim 5 \text{ \AA}$ shift of the Crown towards the capsid centre (Figure 2B).
- 125 3. Movement of the Crown creates space between the Wing and Crown, which allows remodelling of
126 tunnel loops, facilitated by an unfolding of a 5-residue segment of the long helix (residues 331-335,
127 Figure 2D) within the Wing domain.
- 128 4. The loop remodelling “switches” the properties of the tunnel surface from hydrophobic to hydrophilic,
129 causing the tunnel to “open” at its most constricted part (Figure 2 E,F)
- 130 5. Reversal of the Crown and tunnel movements (steps 4 to 1 above) would cause the tunnel to revert
131 to the “closed” state, as shown schematically on Figure 4.

132 It is reasonable to assume that the two conformational states observed in structural studies, reflect
133 energetically preferred states of the portal protein that are utilised during DNA translocation. The switch
134 between the open and closed states, resulting in alteration of surface properties of the internal tunnel
135 may therefore have a role in the packaging mechanism. The observed conformational differences

136 between the two portal protein states are consistent with the normal mode analysis (Bayfield et al.,
137 2019), suggesting that a dynamic equilibrium exists between these two states. Analogous
138 conformational changes in a central tunnel, involving a hydrophobic–hydrophilic “switch” in surface
139 properties, have been proposed to play a key mechanistic role in other systems, for example the GroEL
140 molecular chaperone, where ATP-induced changes facilitate protein refolding (Mayhew et al., 1996;
141 Weissman et al., 1996; Xu et al., 1997).

142 It is important to consider how the two portal states are related and how they may participate in the DNA
143 translocation mechanism. Whereas the ~ 22 Å-wide hydrophilic tunnel observed *in situ* would allow the
144 passage of B-form and potentially even the wider A-form DNA (Harvey, 2015; Ray et al., 2010) into the
145 capsid, the more restrictive tunnel diameter of ~ 14 Å observed in the crystal structure requires the tunnel
146 loops to protrude towards the DNA grooves, involving changes in the tunnel loop conformations
147 (Lebedev et al., 2007).

148 **Mechanism preventing DNA slippage during translocation.** Based on structural observations, we
149 propose the following mechanism (Figure 4). At the start of a packaging cycle, the Crown is protruding
150 towards capsid and hence the tunnel is open for DNA translocation (Figure 4) and its internal surface is
151 hydrophilic (Figure 2F). As shown for the $\phi 29$ system, DNA is expected to be translocated in bursts
152 followed by dwell intervals, serving to reset the motor (Chistol et al., 2012; Moffitt et al., 2009). When the
153 packaging driving force is removed, as when the motor is resetting to bind ATP (Feiss and Rao, 2012),
154 or when the motor fully detaches in preparation for tail docking (Cuervo et al., 2019), the risk of the
155 genome leaking from the capsid increases. This risk is highest when the pressure inside the head is at
156 its greatest, when the head becomes fuller. In this instance, the portal tunnel can act to negate this risk,
157 constricting to prevent DNA from slipping out (Figure 4). In this scenario, the tunnel loops engage with
158 DNA to prevent its slippage, in a manner analogous to a ratchet. This would be caused by the downward
159 movement of the Crown, pushing on the tunnel loops. Such a mechanism is consistent with variation in
160 the length of packaging dwell periods, which become longer as the capsid fills, as observed for the $\phi 29$
161 system (Chistol et al., 2012; Liu et al., 2014; Moffitt et al., 2009), and with the arresting DNA slippage, as
162 observed in “single molecule” experiments for T4 (Ordyan et al., 2018).

163 As a result of this synergy in the movement of portal Crown domain and tunnel loops, the closed state
164 could be induced more readily by a higher internal pressure pushing on the Crown domain, which builds
165 as the capsid fills with DNA, or by the occasional slippage of DNA which could interact transiently with
166 the Crown (“snagging”). The role of the tunnel loops in engaging with the DNA, particularly during the
167 late stages of packaging, is supported by the observations that tunnel loop deletions allow DNA to
168 escape from the capsid in phages $\phi 29$ and T4 (Grimes et al., 2011; Padilla-Sanchez et al., 2014).
169 Overall, this describes a packaging mechanism that is naturally safeguarded against genome loss by the
170 portal protein. When fully packaged, DNA can be held in place by the constricted tunnel (Liu et al., 2019;

171 McElwee et al., 2018) and by tail factors that completely block DNA exit (Cuervo et al., 2019). During
172 infection and DNA ejection, bacterial cell surface binding is likely able to influence the conformation of
173 the phage tail and consequently the portal protein, inducing a more open conformation needed for DNA
174 escape. Due to the nature of portal-capsid interactions and the attendant symmetry mismatch, discussed
175 below, the portal ratcheting mechanism could be active, regardless of the capsid expansion state.

176 **The portal's high order of symmetry reconciles a symmetry mismatch.** The C12-symmetric portal is
177 accommodated in a C5-symmetric penton cavity at one capsid vertex, despite the attendant symmetry
178 mismatch. Comparison of the C12 portal reconstruction presented here with the asymmetric procapsid
179 reconstruction (Bayfield et al., 2019), reveals how this is achieved. The $\sim 8^\circ$ rotational adjustment of the
180 Wing position, bringing it closer to the capsid wall, may assist in forming close portal–capsid contacts,
181 whereas the portal's Clip external diameter is already well matched to the aperture of the capsid's
182 penton hole (i.e. the space vacated if one complete penton is removed), so that close interactions can
183 be made. However, the symmetry mismatch creates a problem in that the same pairs of interacting
184 residues at the portal–capsid interface are not consistently aligned around all subunits, and could be
185 offset by as much as ~ 2 nm in the case of P23-45. This misalignment occurs both at the portal Wing and
186 at the Clip, where the portal and capsid make contact in the asymmetric capsid reconstruction. The
187 sparsity of connected portal–capsid regions indicates that the total surface area of interaction is small,
188 and the residues involved in such interactions are hence also restricted in number and positioning.

189 The high order of symmetry of the portal helps to mitigate these problems. Its 12-fold symmetry is
190 advantageous in that regions of the portal which can interact with the capsid are repeated at a
191 correspondingly high frequency, which reduces the distance between the mismatched interacting
192 residues. The remaining distance can easily be closed by pivoting of flexible loops towards the capsid,
193 such as at the β -hairpin loops of the portal Wing (residues 185–189). These loops are in equivalent
194 positions in $\phi 29$ (Xu et al., 2019). As a result, only minimal, localised deviations from ideal C12
195 symmetry are needed to make interactions with the capsid. The portal can therefore utilise the same
196 few residues to interact around its circumference, which contrasts with the situation that would exist if
197 the portal possessed C6, C3, or other lower symmetries matching that of tail components. The
198 symmetry mismatch of the interaction is a general feature amongst all tailed bacteriophages and
199 related viruses, including herpesviruses (Liu et al., 2019; McElwee et al., 2018). In the case of
200 bacteriophage $\phi 29$ prohead (Mao et al., 2016; Xu et al., 2019), one of the structural roles of the pRNA
201 appears to be equivalent to that of the capsid protein P-domain, in interacting with the outside of the
202 portal Clip, with the $\phi 29$ capsid protein P-domain, instead making contact with an N-terminal segment
203 of the portal protein.

204 At the interface between the portal and capsid vertex, with respective C12 and C5 symmetries,
205 interactions will repeat with a periodicity of $360^\circ/60 = 6^\circ$, as similarly suggested by Hendrix prior to the
206 determination of portal structures (Hendrix, 1978). Rotation of the portal with respect to the capsid by
207 only 6° would therefore create an equivalent global register, with rotations less than 6° generating non-
208 superposable registers of the whole capsid particle (Video 3). Different portal–capsid registers will have
209 different energies of interaction, and hence equivalent angular registers are expected to be
210 energetically equivalent. A comparable symmetry mismatch is observed between the portal and internal
211 core of bacteriophage T7 (C12 *versus* C8) (Cerritelli et al., 2003), where the mismatched interactions
212 may facilitate the detachment of core proteins. As neither detachment of the portal nor its rotation with
213 respect to the capsid (Baumann et al., 2006), appear to play a role in capsid maturation, the effect of
214 symmetry mismatch in the capsid vertex is to permit flexibility at the portal–capsid interface, allowing
215 the portal and capsid to undergo independent conformational changes, whilst ensuring stable
216 interaction of the portal protein with the tail.

217 **Conclusions**

218 Accommodation of the portal protein dodecamer in the procapsid involves conformational adjustments.
219 Interaction between the portal and the capsid shell alters the relative positions of domains, in particular
220 the Wing and Crown, and causes remodelling in the tunnel loops that define the most constricted part
221 of the axial tunnel. The unique conformation of the portal *in situ* demonstrates that the capsid plays a
222 role determining portal conformation, allowing DNA to pass through the tunnel, whilst the portal has the
223 ability to modulate packaging activity and slippage by switching its tunnel properties so that it can
224 engage and disengage with DNA. Whilst portal proteins across other double-stranded DNA viruses
225 (with a terminase motor) may deviate from the classical domain arrangement observed in P23-45, all
226 such viruses face the same basic challenge of safeguarding against genome loss. With regards to
227 portal-capsid interactions, the adoption of 12-fold symmetry by the portal, rather than a symmetry
228 matching that of the capsid vertex, is likely a consequence of the independent evolution of head and tail
229 assemblies, which has selected the matching of symmetries between the portal (12-fold) and the tail (6-
230 fold). This study posits that the problem of mismatched portal–capsid interactions is resolved by the
231 large number of subunits constituting the portal protein, which minimises distances between interacting
232 regions across a spacious interface.

233 **Methods**

234 **Cryo-EM data processing and model building.** From 38,044 extracted particles used in the
235 reconstruction of the unexpanded icosahedral capsid (EMD-4447) (Bayfield et al., 2019), subparticles
236 centred on each vertex were extracted from each capsid particle, and aligned on the z-axis (Ilca et al.,
237 2015). After 3D classifications without imposing symmetry or changing orientations in RELION (Scheres,

238 2012), a class containing 10,025 particles and exhibiting clear portal features was selected for
239 subsequent 3D refinement in RELION, with C12 symmetrical averaging. The atomic model was built
240 using the crystal structure PDB 6IBG as a starting model, with modification to domain positions and to
241 individual amino acids, including side-chain conformations, introduced in Coot (Emsley and Cowtan,
242 2004). Cycles of model rebuilding were followed by real-space refinement in PHENIX (Adams et al.,
243 2010). Resolution was assessed using the gold-standard method using the FSC 0.143 criterion.
244 Refinement and model statistics are presented in Supplementary File 1. Rendering of figures and
245 structure analyses were performed in UCSF Chimera (Pettersen et al., 2004) and ChimeraX (Goddard
246 et al., 2018).

247 **Liquid Chromatography–Mass Spectrometry.** Capsids of P23-45 in the unexpanded state were
248 purified as previously described (Bayfield et al., 2019), and digested with enzyme Glu-C, followed by
249 liquid chromatography tandem mass spectrometry. A 20 µl aliquot (20 µg protein) was reduced with DDT
250 and alkylated with iodoacetamide. Digestion was performed for ~18 hours at 37 °C using sequencing
251 grade Glu-C (Promega). Peptides were analysed by nanoHPLC-MS/MS over a 65 min acquisition with
252 elution from a 50 cm C18 PepMap column onto an Orbitrap Fusion Tribrid mass spectrometer via an
253 EasyNano ionisation source. LC-MS/MS chromatograms were analysed using PEAKS-Studio X (Tran et
254 al., 2019). Peaks were picked and searched against the combined *Thermus thermophilus* and *Thermus*
255 phage P23-45 proteomes. Database searching required Glu-C specificity with one site of non-specificity
256 per peptide identity allowed. Expected cleavage is C-terminal to Glu, a lower rate of cleavage C-terminal
257 to Asp is also known to occur. Peptide matches were filtered to achieve a false discovery rate of <1%.

258 **Data deposition.** Cryo-EM reconstruction (EMD-4567) and atomic coordinates (PDB 6QJT) have been
259 deposited with the wwPDB (www.wwpdb.org).

260 **Acknowledgements**

261 The authors thank Emma Hesketh and Rebecca Thompson at the Astbury Centre, University of Leeds,
262 for assistance with cryo-EM data collection and helpful discussion; Adam Dowle at the Biology
263 Technology Facility, University of York, for assistance with mass spectrometry. This work was supported
264 by Wellcome Trust–National Institutes of Health Studentship 103460 (to O.W.B.), by the Intramural
265 Research Program of National Institute of Arthritis and Musculoskeletal and Skin Diseases (to A.C.S.)
266 and by Wellcome Trust fellowship 206377 (to A.A.A.).

267

268 **References**

269

270 Adams PD, Afonine P V, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ,
271 Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson
272 JS, Terwilliger TC, Zwart PH. 2010. PHENIX: A comprehensive Python-based system for
273 macromolecular structure solution. *Acta Crystallogr Sect D Biol Crystallogr* **66**:213–221.

274 doi:10.1107/S0907444909052925
275 Baumann RG, Mullaney J, Black LW. 2006. Portal fusion protein constraints on function in DNA
276 packaging of bacteriophage T4. *Mol Microbiol* **61**:16–32. doi:10.1111/j.1365-2958.2006.05203.x
277 Bayfield OW, Klimuk E, Winkler DC, Hesketh EL, Chechik M, Cheng N, Dykeman EC, Minakhin L,
278 Ranson NA, Severinov K, Steven AC, Antson AA. 2019. Cryo-EM structure and in vitro DNA
279 packaging of a thermophilic virus with supersized T=7 capsids. *Proc Natl Acad Sci* 201813204.
280 doi:10.1073/PNAS.1813204116
281 Bergh Ø, Børsheim KY, Bratbak G, Haldal M. 1989. High abundance of viruses found in aquatic
282 environments. *Nature* **340**:467–468. doi:10.1038/340467a0
283 Cerritelli ME, Trus BL, Smith CS, Cheng N, Conway JF, Steven AC. 2003. A second symmetry
284 mismatch at the portal vertex of bacteriophage T7: 8-Fold symmetry in the procapsid core. *J Mol*
285 *Biol* **327**:1–6. doi:10.1016/S0022-2836(03)00117-7
286 Chen D-H, Baker ML, Hryc CF, DiMaio F, Jakana J, Wu W, Dougherty M, Haase-Pettingell C, Schmid
287 MF, Jiang W, Baker D, King J a, Chiu W. 2011. Structural basis for scaffolding-mediated assembly
288 and maturation of a dsDNA virus. *Proc Natl Acad Sci U S A* **108**:1355–60.
289 doi:10.1073/pnas.1015739108
290 Chistol G, Liu S, Hetherington CL, Moffitt JR, Grimes S, Jardine PJ, Bustamante C. 2012. High degree
291 of coordination and division of labor among subunits in a homomeric ring ATPase. *Cell* **151**:1017–
292 1028. doi:10.1016/j.cell.2012.10.031
293 Cuervo A, Fàbrega-Ferrer M, Machón C, Conesa JJ, Fernández FJ, Pérez-Luque R, Pérez-Ruiz M,
294 Pous J, Vega MC, Carrascosa JL, Coll M. 2019. Structures of T7 bacteriophage portal and tail
295 suggest a viral DNA retention and ejection mechanism. *Nat Commun* **10**:1–11.
296 doi:10.1038/s41467-019-11705-9
297 Emsley P, Cowtan K. 2004. Coot: Model-building tools for molecular graphics. *Acta Crystallogr Sect D*
298 *Biol Crystallogr* **60**:2126–2132. doi:10.1107/S0907444904019158
299 Feiss M, Rao VB. 2012. Viral Molecular Machines In: Rossmann MG, Rao VB, editors. *Advances in*
300 *Experimental Medicine and Biology, Advances in Experimental Medicine and Biology*. Boston, MA:
301 Springer US. pp. 489–509. doi:10.1007/978-1-4614-0980-9
302 Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE. 2018. UCSF
303 ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**:14–25.
304 doi:10.1002/pro.3235
305 Grimes S, Ma S, Gao J, Atz R, Jardine PJ. 2011. Role of ϕ 29 Connector Channel Loops in Late-Stage
306 DNA Packaging. *J Mol Biol* **410**:50–59. doi:10.1016/j.jmb.2011.04.070.Role
307 Harvey SC. 2015. The Scrunchworm Hypothesis: Transitions Between A-DNA and B-DNA Provide the
308 Driving Force for Genome Packaging in Double-Stranded DNA Bacteriophages. *J Struct Biol*
309 **189**:1–8. doi:10.1109/EMBC.2016.7590696.Upper
310 Hendrix RW. 1978. Symmetry mismatch and DNA packaging in large bacteriophages. *Proc Natl Acad*
311 *Sci U S A* **75**:4779–4783. doi:10.1073/pnas.75.10.4779
312 Ilca SL, Kotecha A, Sun X, Poranen MM, Stuart DI, Huiskonen JT. 2015. Localized reconstruction of
313 subunits from electron cryomicroscopy images of macromolecular complexes. *Nat Commun* **6**:1–8.
314 doi:10.1038/ncomms9843
315 Lebedev AA, Krause MH, Isidro AL, Vagin AA, Orlova E V, Turner J, Dodson EJ, Tavares P, Antson AA.
316 2007. Structural framework for DNA translocation via the viral portal protein. *EMBO J* **26**:1984–
317 1994. doi:10.1038/sj.emboj.7601643
318 Liu S, Chistol G, Hetherington CL, Tafoya S, Aathavan K, Schnitzbauer J, Grimes S, Jardine PJ,
319 Bustamante C. 2014. A viral packaging motor varies its DNA rotation and step size to preserve
320 subunit coordination as the capsid fills. *Cell* **157**:702–713. doi:10.1016/j.cell.2014.02.034
321 Liu Y-T, Jih J, Dai X, Bi G-Q, Zhou ZH. 2019. Cryo-EM structures of herpes simplex virus type 1 portal
322 vertex and packaged genome. *Nature*. doi:10.1038/s41586-019-1248-6
323 Lokareddy RK, Sankhala RS, Roy A, Afonine P V, Motwani T, Teschke CM, Parent KN, Cingolani G.
324 2017. Portal protein functions akin to a DNA-sensor that couples genome-packaging to icosahedral
325 capsid maturation. *Nat Commun* **8**:14310. doi:10.1038/ncomms14310
326 Mao H, Reyes-Aldrete E, Sherman MB, Woodson M, Atz R, Grimes S, Jardine PJ, Morais MC. 2016.

327 Structural and Molecular Basis for Coordination in a Viral DNA Packaging Motor. *Cell Rep*
328 **14**:2017–2029. doi:10.1017/CBO9781107415324.004

329 Mayhew M, Da Silva ACR, Martin J, Erdjument-Bromage H, Tempst P, Hartl FU. 1996. Protein folding in
330 the central cavity of the GroEL-GroES chaperonin complex. *Nature* **379**:420–426.
331 doi:10.1038/379420a0

332 McElwee M, Vijayakrishnan S, Rixon F, Bhella D. 2018. Structure of the herpes simplex virus portal-
333 vertex. *PLoS Biol* **1**–15. doi:10.1371/journal.pbio.2006191

334 Michaud JM, Thompson LR, Kaul D, Espinoza JL, Richter RA, Xu ZZ, Lee C, Pham KM, Beall CM,
335 Malfatti F, Azam F, Knight R, Burkart MD, Dupont CL, Prather KA. 2018. Taxon-specific
336 aerosolization of bacteria and viruses in an experimental ocean-atmosphere mesocosm. *Nat*
337 *Commun* **9**:1–10. doi:10.1038/s41467-018-04409-z

338 Moffitt JR, Chemla YR, Athavan K, Grimes S, Jardine PJ, Anderson DL, Bustamante C. 2009.
339 Intersubunit coordination in a homomeric ring ATPase. *Nature* **457**:446–50.
340 doi:10.1038/nature07637

341 Ordyan M, Alam I, Mahalingam M, Rao VB, Smith DE. 2018. Nucleotide-dependent DNA gripping and
342 an end-clamp mechanism regulate the bacteriophage T4 viral packaging motor. *Nat Commun* **9**.
343 doi:10.1038/s41467-018-07834-2

344 Padilla-Sanchez V, Gao S, Kim HR, Kihara D, Sun L, Rossmann MG, Rao VB. 2014. Structure-Function
345 Analysis of the DNA Translocating Portal of the Bacteriophage T4 Packaging Machine. *J Mol Biol*
346 **426**:1019–1038. doi:10.1371/journal.pone.0178059

347 Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF
348 Chimera - A visualization system for exploratory research and analysis. *J Comput Chem* **25**:1605–
349 1612. doi:10.1002/jcc.20084

350 Ray K, Sabanayagam CR, Lakowicz JR, Black LW. 2010. DNA crunching by a viral packaging motor:
351 Compression of a procapsid-portal stalled Y-DNA substrate. *Virology* **398**:224–232.
352 doi:10.1016/j.virol.2009.11.047

353 Scheres SHW. 2012. RELION: Implementation of a Bayesian approach to cryo-EM structure
354 determination. *J Struct Biol* **180**:519–530. doi:10.1016/j.jsb.2012.09.006

355 Shkoporov AN, Hill C. 2019. Bacteriophages of the Human Gut: The “Known Unknown” of the
356 Microbiome. *Cell Host Microbe* **25**:195–209. doi:10.1016/j.chom.2019.01.017

357 Simpson A, Tao Y, Leiman P, Badasso M, He Y, Jardine PJ, Olson N, Morais M, Grimes S, Anderson D,
358 Baker T, Rossmann M. 2000. Structure of the bacteriophage ϕ 29 DNA packaging motor. *Nature*
359 **408**:745–750. doi:10.1038/35047129

360 Sun L, Zhang X, Gao S, Rao PA, Padilla-Sanchez V, Chen Z, Sun S, Xiang Y, Subramaniam S, Rao VB,
361 Rossmann MG. 2015. Cryo-EM structure of the bacteriophage T4 portal protein assembly at near-
362 atomic resolution. *Nat Commun* **6**:7548. doi:10.1038/ncomms8548

363 Sun S, Kondabagil K, Draper B, Alam TI, Bowman VD, Zhang Z, Hegde S, Fokine A, Rossmann MG,
364 Rao VB. 2008. The structure of the phage T4 DNA packaging motor suggests a mechanism
365 dependent on electrostatic forces. *Cell* **135**:1251–62. doi:10.1016/j.cell.2008.11.015

366 Tran NH, Qiao R, Xin L, Chen X, Liu C, Zhang X, Shan B, Ghodsi A, Li M. 2019. Deep learning enables
367 de novo peptide sequencing from data-independent-acquisition mass spectrometry. *Nat Methods*
368 **16**:63–66. doi:10.1038/s41592-018-0260-3

369 Trus BL, Cheng N, Newcomb WW, Homa FL, Brown JC, Steven AC, Homa L, Brown JC, Steven AC.
370 2004. Structure and Polymorphism of the UL6 Portal Protein of Herpes Simplex Virus Type 1. *J*
371 *Virol* **78**:12668–12671. doi:10.1128/JVI.78.22.12668

372 Weissman JS, Rye HS, Fenton WA, Beechem JM, Horwich AL. 1996. Characterization of the active
373 intermediate of a GroEL-GroES-mediated protein folding reaction. *Cell* **84**:481–490.
374 doi:10.1016/s0092-8674(00)81293-3

375 Wu W, Leavitt JC, Cheng N, Gilcrease EB, Motwani T, Teschke CM, Casjens SR, Steven AC. 2016.
376 Localization of the houdinisome (Ejection proteins) inside the bacteriophage P22 virion by
377 bubblegram imaging. *MBio* **7**:1–10. doi:10.1128/mBio.01152-16

378 Xu J, Wang D, Gui M, Xiang Y. 2019. Structural assembly of the tailed bacteriophage ϕ 29. *Nat Commun*
379 **10**:2366. doi:10.1038/s41467-019-10272-3

380 Xu Z, Horwich A, Sigler PB. 1997. The crystal structure of the asymmetric GroEL-GroES-(ADP)₇
381 chaperonin complex. *Nature* **388**:741–750. doi:10.1038/41944
382
383
384
385
386

387 **Figure Legends**

388 **Figure 1.** Structure of the portal protein *in situ*. (A) Cryo-electron micrograph of P23- 45 procapsids,
389 scale bar 50 nm. (B) Cryo-EM reconstruction map with one subunit coloured red, scale bar 50 Å, and
390 same for (C) but rotated 90°, viewed along 12-fold axis. (D) Ribbon diagram of one portal protein subunit.
391 (E) Regions of the map and corresponding atomic models with residue numbers.

392 **Figure 2.** Comparison with the crystal structure. (A) Single subunit of the *in situ* structure is in blue and
393 an apposing chain from the crystal structure is in yellow. (B) Superposition of single subunits, exposing
394 structural differences between the crystal structure and the *in situ* structure. The curved arrow indicates
395 the pivoting of the Wing domain by ~8° in the *in situ* structure. (C) The two dodecamers overlaid, viewed
396 from Crown (top domain in A), along the tunnel axis. Dodecamers are superposed based on residues
397 26–376 (Clip, Stem, and Wing), revealing a ~13° rotation of the Crown domain about the tunnel axis. (D)
398 Overlay of *in situ* (blue) and crystal structure (yellow), ribbon diagram, with side-chains shown. (E) Van
399 der Waals surface of the crystal structure (PDB 6IBG) showing tunnel loop-constricted region, with
400 tunnel colouring by the hydrophobicity on the Kyte-Doolittle scale where white is hydrophobic and brown
401 is hydrophilic, and same for (F) but for the *in situ* structure (PDB 6QJT). Diameters of most constricted
402 part of tunnels measured from Van der Waals surfaces are shown.

403 **Figure 3.** Portal–capsid interactions. (A) Sections through the capsid reconstruction perpendicular to the
404 portal tunnel axis, at three different heights as denoted on (C) by dotted lines. (B) Interactions between
405 the portal Wing and capsid. Portal protein subunit making interactions with the capsid is in pink. Portal
406 subunits not making interactions are in blue. (C) Ribbon diagram of the *in situ* portal protein fitted into
407 the procapsid map. (D) Interactions between the portal Clip and capsid. Portal protein subunit making
408 closest interactions with the capsid is in green. (E) Subunits of the portal protein interacting with the
409 capsid by their Wing regions are in magenta, labelled clockwise. (F) Subunits of the portal protein
410 interacting via their Clip are in green. View is from the center of the portal with chains labelled as in (E).

411 **Figure 4.** Mechanism of portal tunnel closure. Left - the open state where the Crown (blue) is elevated,
412 facilitating partial retreat of the tunnel loops (terracotta) toward the crown, widening the tunnel. Right –
413 the closed state where the Crown is depressed into the body of the portal protein, facilitating closure of
414 the tunnel where tunnel loops adopt a conformation that extends into the tunnel.

415

416

417 **Supplementary Legends**

418 **Figure 1-figure supplement 1.** FSC curve for the portal protein reconstruction. Fourier Shell Correlation is plotted
419 as function of spatial frequency. Dotted lines denote resolution estimate at FSC=0.143 according to the gold-
420 standard method.

421 **Figure 1-figure supplement 2.** Mass spectrometry analysis of the portal protein from unexpanded capsids. Blue
422 bars beneath the sequence denote regions for which peptides were detected.

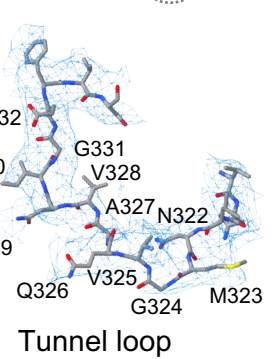
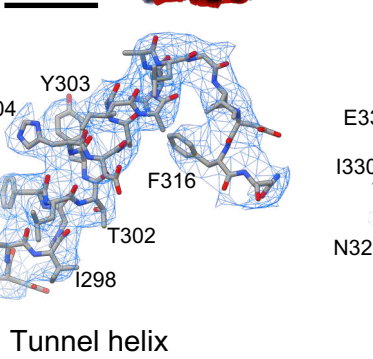
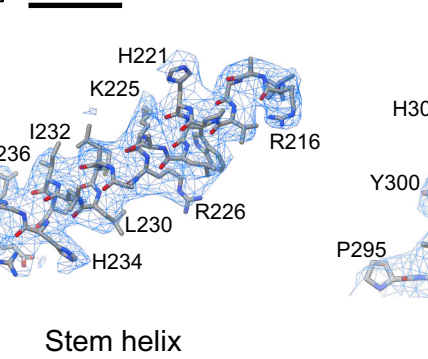
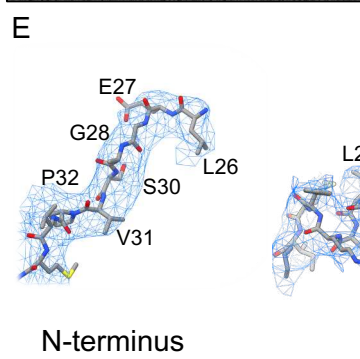
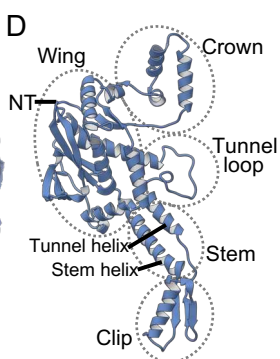
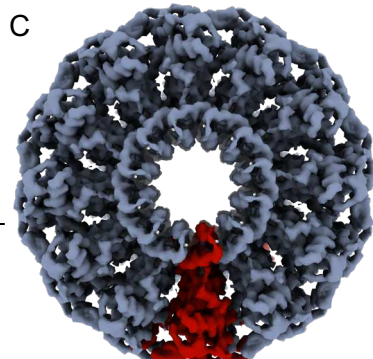
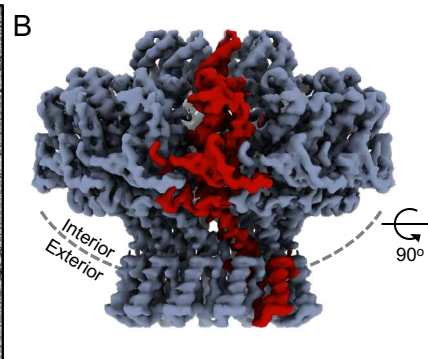
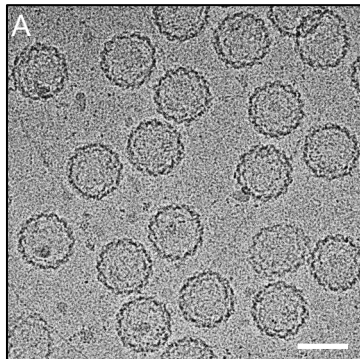
423 **Video 1.** Reconstruction of the *in situ* portal. Surface rendering, first viewed perpendicular to the tunnel axis, then
424 viewed along the tunnel axis.

425 **Video 2.** Morph between the *in situ* structure (first) and crystal structure (second). Ribbon diagram, first viewed
426 perpendicular to the tunnel axis, then viewed along the tunnel axis, then rotated back to initial view with two
427 apposing chains displayed.

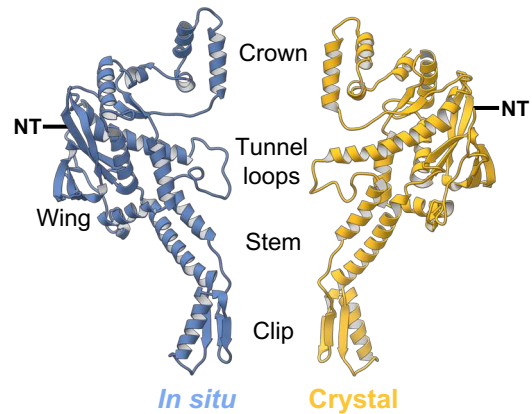
428 **Video 3.** Portal–capsid registers. One-degree step change in portal register (inner 12-fold circle) with respect to
429 capsid vertex (outer 5-fold circle), beginning with “0°”. Portal register “6°” is superposable on register “0°” by 144°
430 rotation of the whole capsid (i.e. rotating both inner and outer circles together).

431 **Supplementary File 1.** Cryo-EM Data Collection and Refinement Statistics.

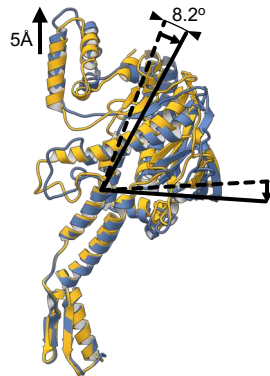
432



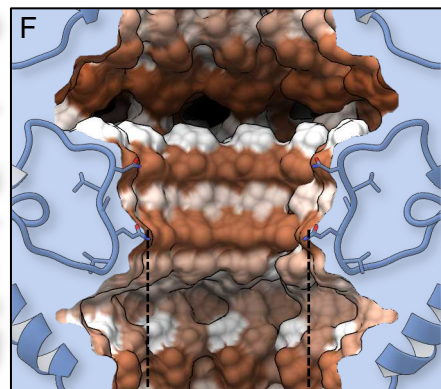
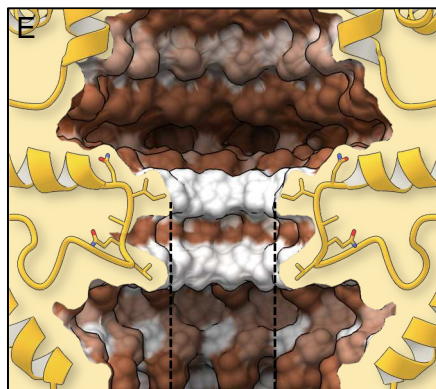
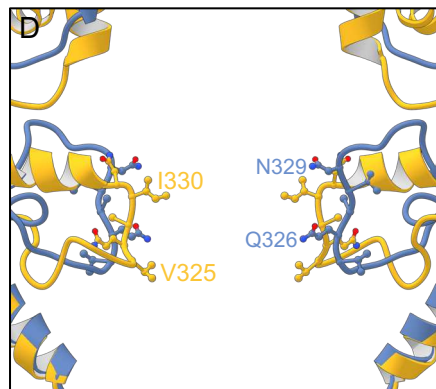
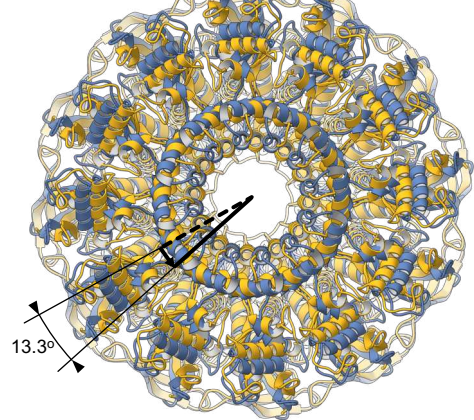
A



B

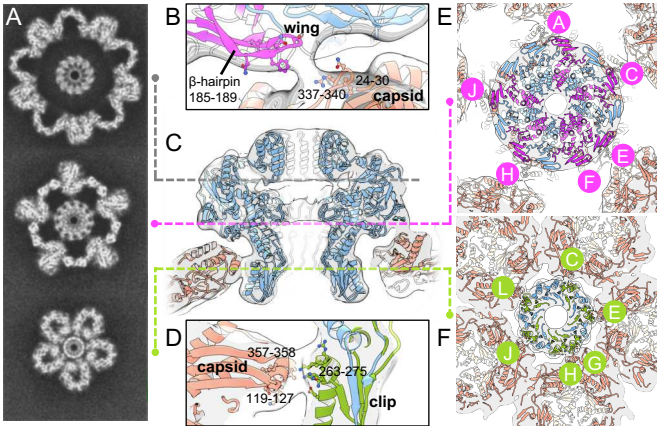


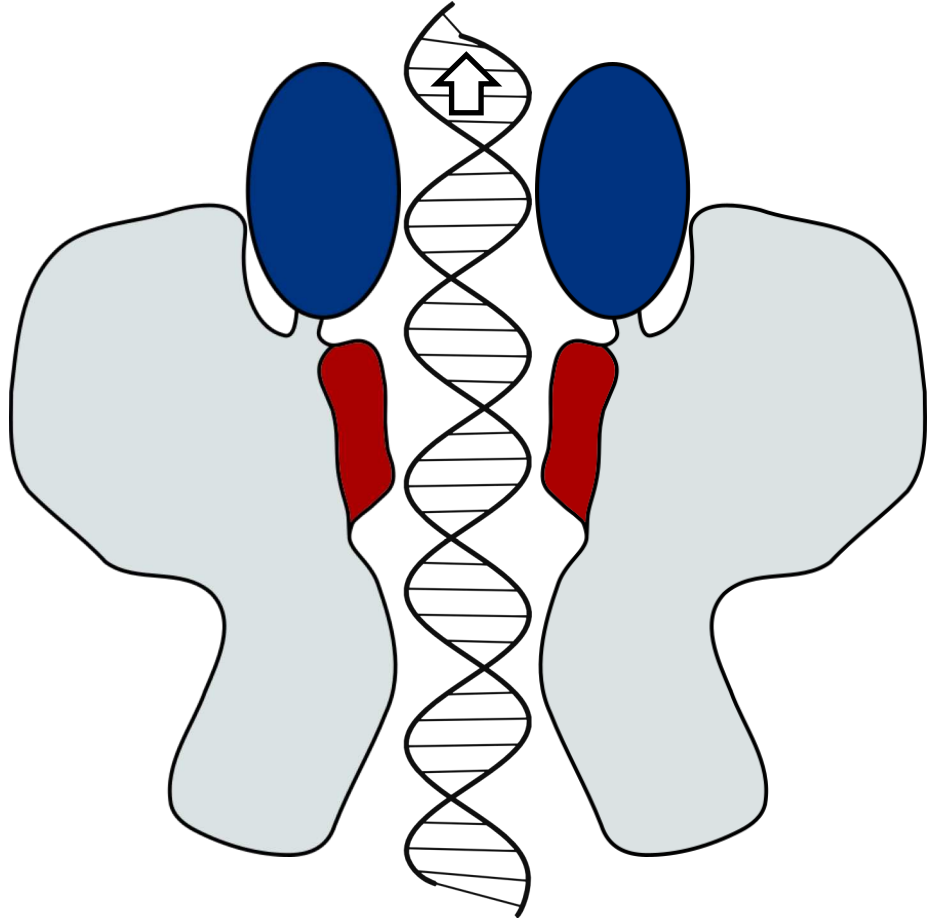
C



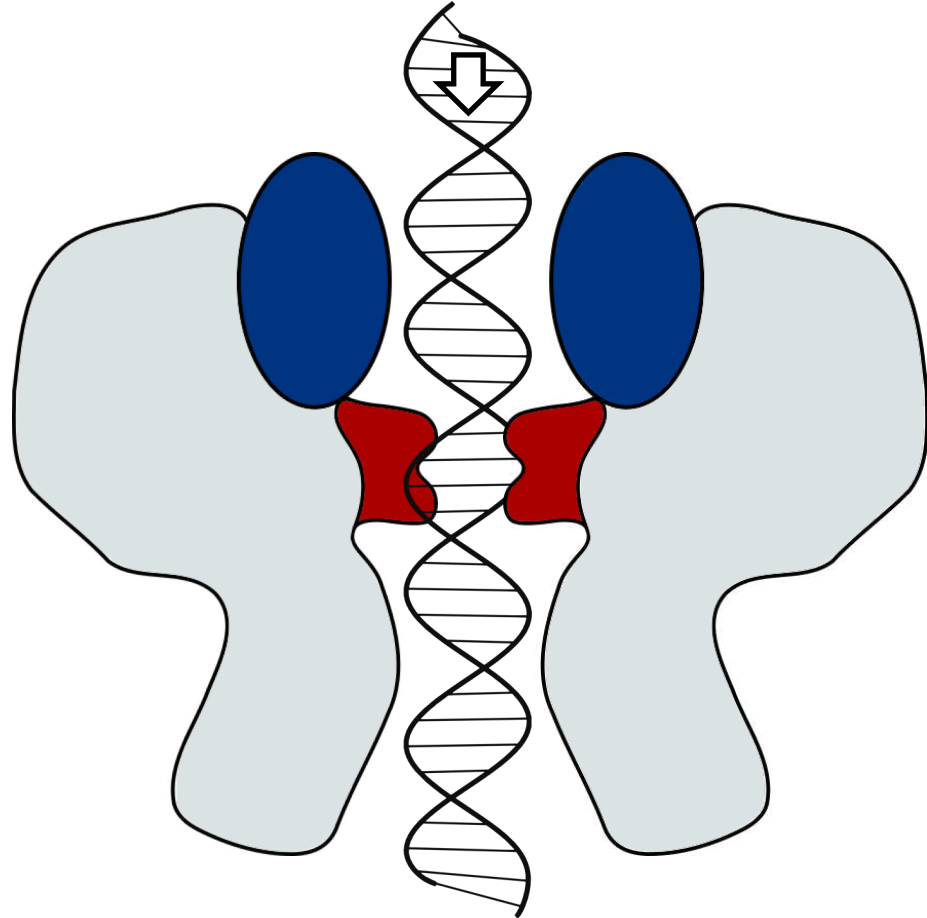
hydrophobic

hydrophilic





Crown **up**, Loops **out**



Crown **down**, Loops **in**

