
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

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]

[a]York Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5DD, United Kingdom.

[b]Laboratory 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

concerning these structures have yet to be rationalised in the context of the portal's many functional roles, including: the variable diameter of the central tunnel and flexibility of tunnel loops (Lebedev et al., 2007; Simpson et al., 2000; Sun et al., 2015); the symmetry mismatch between the portal and capsid vertex (12-fold versus 5-fold) (Simpson et al., 2000; Sun et al., 2008); and the portal's role in DNA translocation (Harvey, 2015; Ray et al., 2010). The influence of the properties of the internal tunnel, and how these can be modulated by external factors to coordinate DNA translocation, remains unclear. Cryo-EM studies on mesophilic herpesviruses characterised the shape of the portal protein tunnel in the mature virion and showed how DNA can be locked inside (Liu et al., 2019; McElwee et al., 2018). However, there are no detailed structural data on portal proteins *in situ* for unexpanded capsids, primed for DNA packaging. Moreover, it has proven difficult to derive accurate models for tunnel loops of the portal protein, such as in the case of tailed bacteriophage φ29, where the flexible nature of the tunnel loops prevented their observation in a crystal structure (Simpson et al., 2000) and also in cryo-EM structures of the procapsid and mature capsid (Xu et al., 2019).

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

Results

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

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

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

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

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

this symmetry-mismatched interface. In the asymmetric reconstruction of the procapsid, the portal protein appears 12-fold symmetric (Bayfield et al., 2019). Residues 185–189 (β-hairpin loop) of the portal Wing are involved in interactions with the capsid (Figure 3B). These loops may pivot downwards to make closer contact with the capsid inner wall, in select chains (Figure 3B,C). Such adjustments in specific subunits of the portal protein would not be resolved in a symmetrically averaged structure; however, the bridging regions observed between the capsid and the portal in the asymmetric procapsid reconstruction suggests local deviations from C12 symmetry are possible. The portal Wing (β-hairpin loop) is in close proximity to residues 24-30 and 337-340 of the major capsid protein (Figure 3B). Interactions of the portal Clip likely involve portal protein residues 263-275 (α-helix and adjacent loop within the Clip domain) interacting with the major capsid protein residues 119-127 and 357-358 (Figure 3D). In common with φ29 (Simpson et al., 2000), portal-capsid interactions are mediated by residues of both polar and hydrophobic character. The portal–capsid symmetry mismatch means that only select portal chains make contact with the capsid: these are chains A-C-(E/F)-H-J at the Wing (Figure 3E), and chains C-E-(G/H)-J-L at the Clip (Figure 3F).

Discussion

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

1. Assembly of capsid proteins around the portal stabilises a ~8° rotational adjustment in Wing domains (Figure 2B).
2. As the Wing domain pivots, the C-terminal helix of the Crown domain that interacts with the Wing, slides, facilitating a ~5 Å shift of the Crown towards the capsid centre (Figure 2B).
3. Movement of the Crown creates space between the Wing and Crown, which allows remodelling of tunnel loops, facilitated by an unfolding of a 5-residue segment of the long helix (residues 331-335, Figure 2D) within the Wing domain.
4. The loop remodelling “switches” the properties of the tunnel surface from hydrophobic to hydrophilic, causing the tunnel to “open” at its most constricted part (Figure2 E,F)
5. Reversal of the Crown and tunnel movements (steps 4 to 1 above) would cause the tunnel to revert to the “closed” state, as shown schematically on Figure 4.

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

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

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

Mechanism preventing DNA slippage during translocation. Based on structural observations, we propose the following mechanism (Figure 4). At the start of a packaging cycle, the Crown is protruding towards capsid and hence the tunnel is open for DNA translocation (Figure 4) and its internal surface is hydrophilic (Figure 2F). As shown for the φ29 system, DNA is expected to be translocated in bursts followed by dwell intervals, serving to reset the motor (Chistol et al., 2012; Moffitt et al., 2009). When the packaging driving force is removed, as when the motor is resetting to bind ATP (Feiss and Rao, 2012), or when the motor fully detaches in preparation for tail docking (Cuervo et al., 2019), the risk of the genome leaking from the capsid increases. This risk is highest when the pressure inside the head is at its greatest, when the head becomes fuller. In this instance, the portal tunnel can act to negate this risk, constricting to prevent DNA from slipping out (Figure 4). In this scenario, the tunnel loops engage with DNA to prevent its slippage, in a manner analogous to a ratchet. This would be caused by the downward movement of the Crown, pushing on the tunnel loops. Such a mechanism is consistent with variation in the length of packaging dwell periods, which become longer as the capsid fills, as observed for the φ29 system (Chistol et al., 2012; Liu et al., 2014; Moffitt et al., 2009), and with the arresting DNA slippage, as observed in "single molecule" experiments for T4 (Ordyan et al., 2018).

As a result of this synergy in the movement of portal Crown domain and tunnel loops, the closed state could be induced more readily by a higher internal pressure pushing on the Crown domain, which builds as the capsid fills with DNA, or by the occasional slippage of DNA which could interact transiently with the Crown ("snagging"). The role of the tunnel loops in engaging with the DNA, particularly during the late stages of packaging, is supported by the observations that tunnel loop deletions allow DNA to escape from the capsid in phages φ29 and T4 (Grimes et al., 2011; Padilla-Sanchez et al., 2014). Overall, this describes a packaging mechanism that is naturally safeguarded against genome loss by the portal protein. When fully packaged, DNA can be held in place by the constricted tunnel (Liu et al., 2019;

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

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

The high order of symmetry of the portal helps to mitigate these problems. Its 12-fold symmetry is advantageous in that regions of the portal which can interact with the capsid are repeated at a correspondingly high frequency, which reduces the distance between the mismatched interacting residues. The remaining distance can easily be closed by pivoting of flexible loops towards the capsid, such as at the β-hairpin loops of the portal Wing (residues 185–189). These loops are in equivalent positions in φ29 (Xu et al., 2019). As a result, only minimal, localised deviations from ideal C12 symmetry are needed to make interactions with the capsid. The portal can therefore utilise the same few residues to interact around its circumference, which contrasts with the situation that would exist if the portal possessed C6, C3, or other lower symmetries matching that of tail components. The symmetry mismatch of the interaction is a general feature amongst all tailed bacteriophages and related viruses, including herpesviruses (Liu et al., 2019; McElwee et al., 2018). In the case of bacteriophage φ29 prohead (Mao et al., 2016; Xu et al., 2019), one of the structural roles of the pRNA appears to be equivalent to that of the capsid protein P-domain, in interacting with the outside of the portal Clip, with the φ29 capsid protein P-domain, instead making contact with an N-terminal segment of the portal protein.

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

Conclusions

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

Methods

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

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

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

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

Acknowledgements

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

References

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

doi:10.1107/S0907444909052925
Baumann RG, Mullaney J, Black LW. 2006. Portal fusion protein constraints on function in DNA packaging of bacteriophage T4. *Mol Microbiol* **61**:16–32. doi:10.1111/j.1365-2958.2006.05203.x
Bayfield OW, Klimuk E, Winkler DC, Hesketh EL, Chechik M, Cheng N, Dykeman EC, Minakhin L, Ranson NA, Severinov K, Steven AC, Antson AA. 2019. Cryo-EM structure and in vitro DNA packaging of a thermophilic virus with supersized T=7 capsids. *Proc Natl Acad Sci* 201813204. doi:10.1073/PNAS.1813204116
Bergh Ø, Børsheim KY, Bratbak G, Heldal M. 1989. High abundance of viruses found in acquatic environments. *Nature* **340**:467–468. doi:10.1038/340467a0
Cerritelli ME, Trus BL, Smith CS, Cheng N, Conway JF, Steven AC. 2003. A second symmetry mismatch at the portal vertex of bacteriophage T7: 8-Fold symmetry in the procapsid core. *J Mol Biol* **327**:1–6. doi:10.1016/S0022-2836(03)00117-7
Chen D-H, Baker ML, Hryc CF, DiMaio F, Jakana J, Wu W, Dougherty M, Haase-Pettingell C, Schmid MF, Jiang W, Baker D, King J a, Chiu W. 2011. Structural basis for scaffolding-mediated assembly and maturation of a dsDNA virus. *Proc Natl Acad Sci U S A* **108**:1355–60. doi:10.1073/pnas.1015739108
Chistol G, Liu S, Hetherington CL, Moffitt JR, Grimes S, Jardine PJ, Bustamante C. 2012. High degree of coordination and division of labor among subunits in a homomeric ring ATPase. *Cell* **151**:1017–1028. doi:10.1016/j.cell.2012.10.031
Cuervo A, Fàbrega-Ferrer M, Machón C, Conesa JJ, Fernández FJ, Pérez-Luque R, Pérez-Ruiz M, Pous J, Vega MC, Carrascosa JL, Coll M. 2019. Structures of T7 bacteriophage portal and tail suggest a viral DNA retention and ejection mechanism. *Nat Commun* **10**:1–11. doi:10.1038/s41467-019-11705-9
Emsley P, Cowtan K. 2004. Coot: Model-building tools for molecular graphics. *Acta Crystallogr Sect D Biol Crystallogr* **60**:2126–2132. doi:10.1107/S0907444904019158
Feiss M, Rao VB. 2012. Viral Molecular Machines In: Rossmann MG, Rao VB, editors. Advances in Experimental Medicine and Biology, Advances in Experimental Medicine and Biology. Boston, MA: Springer US. pp. 489–509. doi:10.1007/978-1-4614-0980-9
Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE. 2018. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**:14–25. doi:10.1002/pro.3235
Grimes S, Ma S, Gao J, Atz R, Jardine PJ. 2011. Role of φ29 Connector Channel Loops in Late-Stage DNA Packaging. *J Mol Biol* **410**:50–59. doi:10.1016/j.jmb.2011.04.070.Role
Harvey SC. 2015. The Scrunchworm Hypothesis: Transitions Between A-DNA and B-DNA Provide the Driving Force for Genome Packaging in Double-Stranded DNA Bacteriophages. *J Struct Biol* **189**:1–8. doi:10.1109/EMBC.2016.7590696.Upper
Hendrix RW. 1978. Symmetry mismatch and DNA packaging in large bacteriophages. *Proc Natl Acad Sci U S A* **75**:4779–4783. doi:10.1073/pnas.75.10.4779
Ilca SL, Kotecha A, Sun X, Poranen MM, Stuart DI, Huiskonen JT. 2015. Localized reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes. *Nat Commun* **6**:1–8. doi:10.1038/ncomms9843
Lebedev AA, Krause MH, Isidro AL, Vagin AA, Orlova E V, Turner J, Dodson EJ, Tavares P, Antson AA. 2007. Structural framework for DNA translocation via the viral portal protein. *EMBO J* **26**:1984–1994. doi:10.1038/sj.emboj.7601643
Liu S, Chistol G, Hetherington CL, Tafoya S, Aathavan K, Schnitzbauer J, Grimes S, Jardine PJ, Bustamante C. 2014. A viral packaging motor varies its DNA rotation and step size to preserve subunit coordination as the capsid fills. *Cell* **157**:702–713. doi:10.1016/j.cell.2014.02.034
Liu Y-T, Jih J, Dai X, Bi G-Q, Zhou ZH. 2019. Cryo-EM structures of herpes simplex virus type 1 portal vertex and packaged genome. *Nature*. doi:10.1038/s41586-019-1248-6
Lokareddy RK, Sankhala RS, Roy A, Afonine P V, Motwani T, Teschke CM, Parent KN, Cingolani G. 2017. Portal protein functions akin to a DNA-sensor that couples genome-packaging to icosahedral capsid maturation. *Nat Commun* **8**:14310. doi:10.1038/ncomms14310
Mao H, Reyes-Aldrete E, Sherman MB, Woodson M, Atz R, Grimes S, Jardine PJ, Morais MC. 2016.

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

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

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

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

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

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

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

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

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

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

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

Simpson A, Tao Y, Leiman P, Badasso M, He Y, Jardine PJ, Olson N, Morais M, Grimes S, Anderson D, Baker T, Rossmann M. 2000. Structure of the bacteriophage ϕ29 DNA packaging motor. *Nature* **408**:745–750. doi:10.1038/35047129

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

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

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

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

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

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

Xu J, Wang D, Gui M, Xiang Y. 2019. Structural assembly of the tailed bacteriophage ϕ29. *Nat Commun* **10**:2366. doi:10.1038/s41467-019-10272-3

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

Figure Legends

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

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

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

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

Supplementary Legends

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

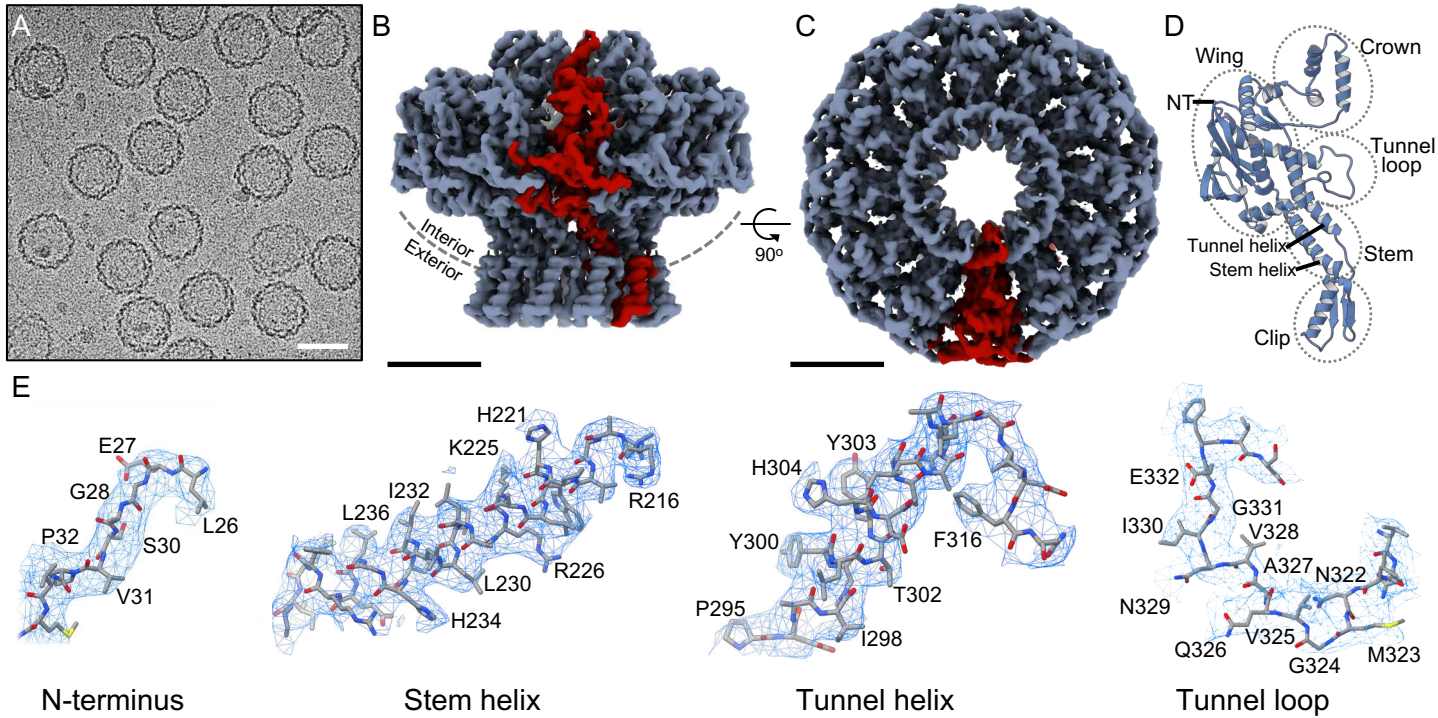
Figure 1-figure supplement 2. Mass spectrometry analysis of the portal protein from unexpanded capsids. Blue bars beneath the sequence denote regions for which peptides were detected.

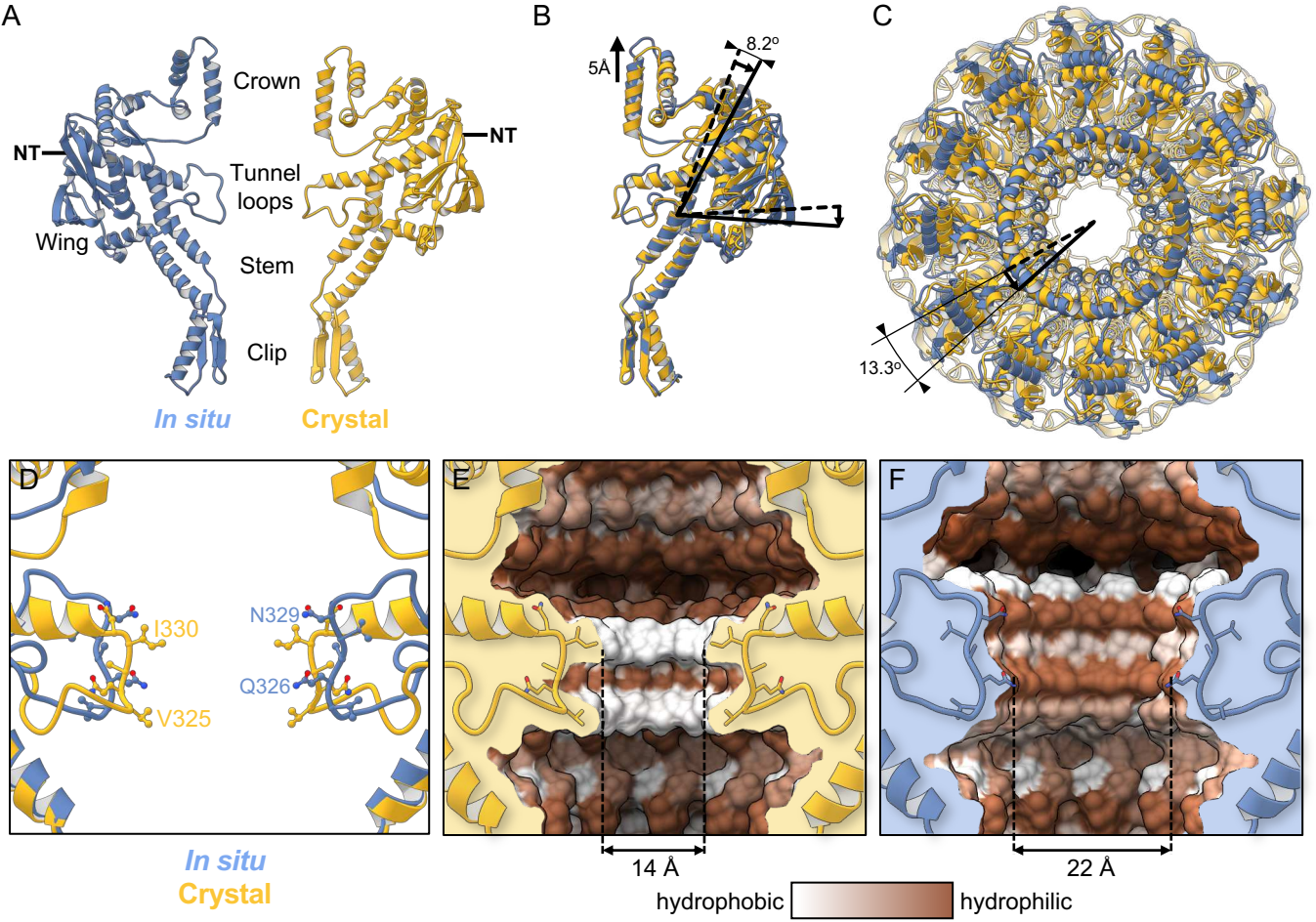
Video 1. Reconstruction of the *in situ* portal. Surface rendering, first viewed perpendicular to the tunnel axis, then viewed along the tunnel axis.

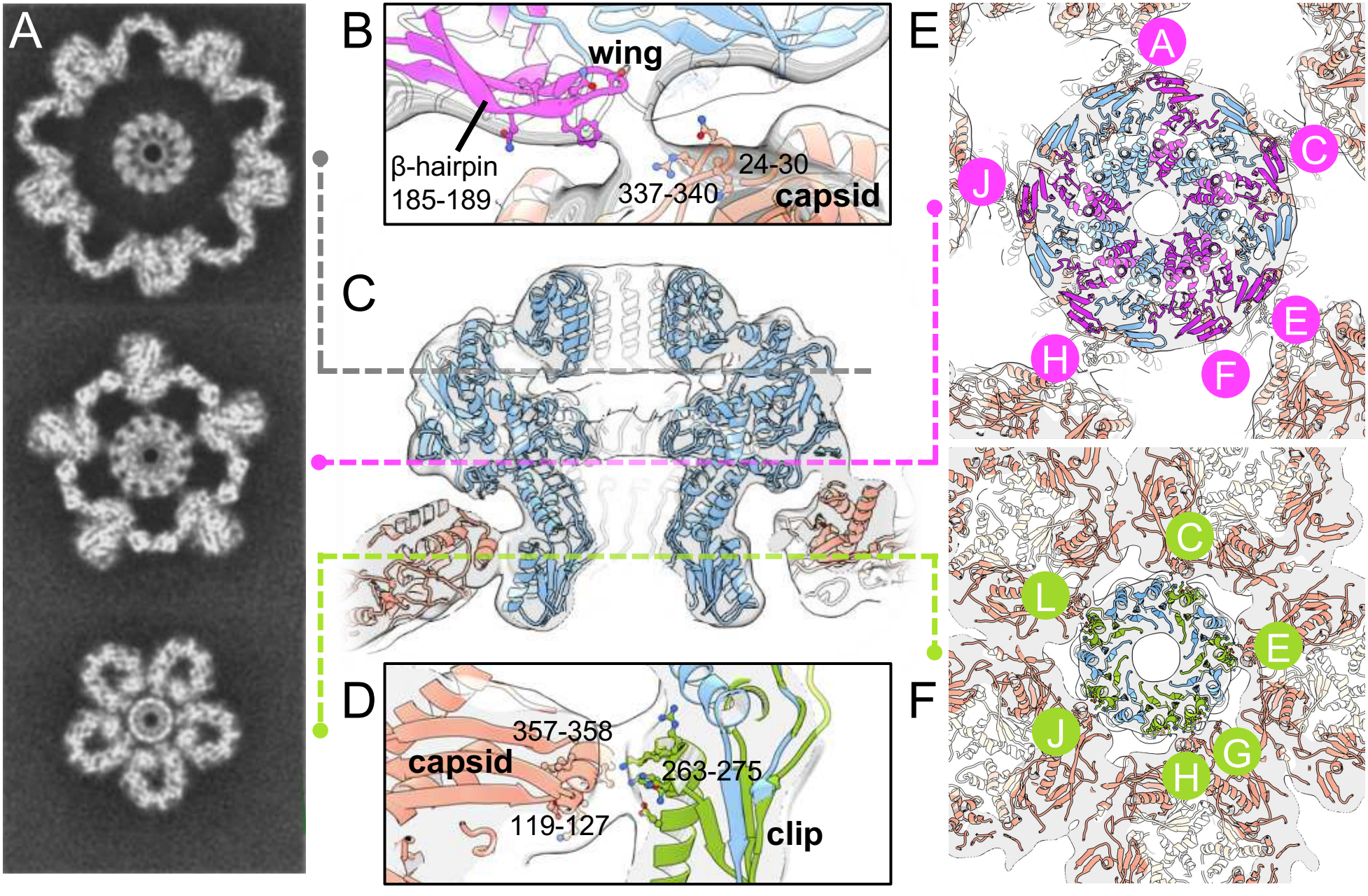
Video 2. Morph between the *in situ* structure (first) and crystal structure (second). Ribbon diagram, first viewed perpendicular to the tunnel axis, then viewed along the tunnel axis, then rotated back to initial view with two apposing chains displayed.

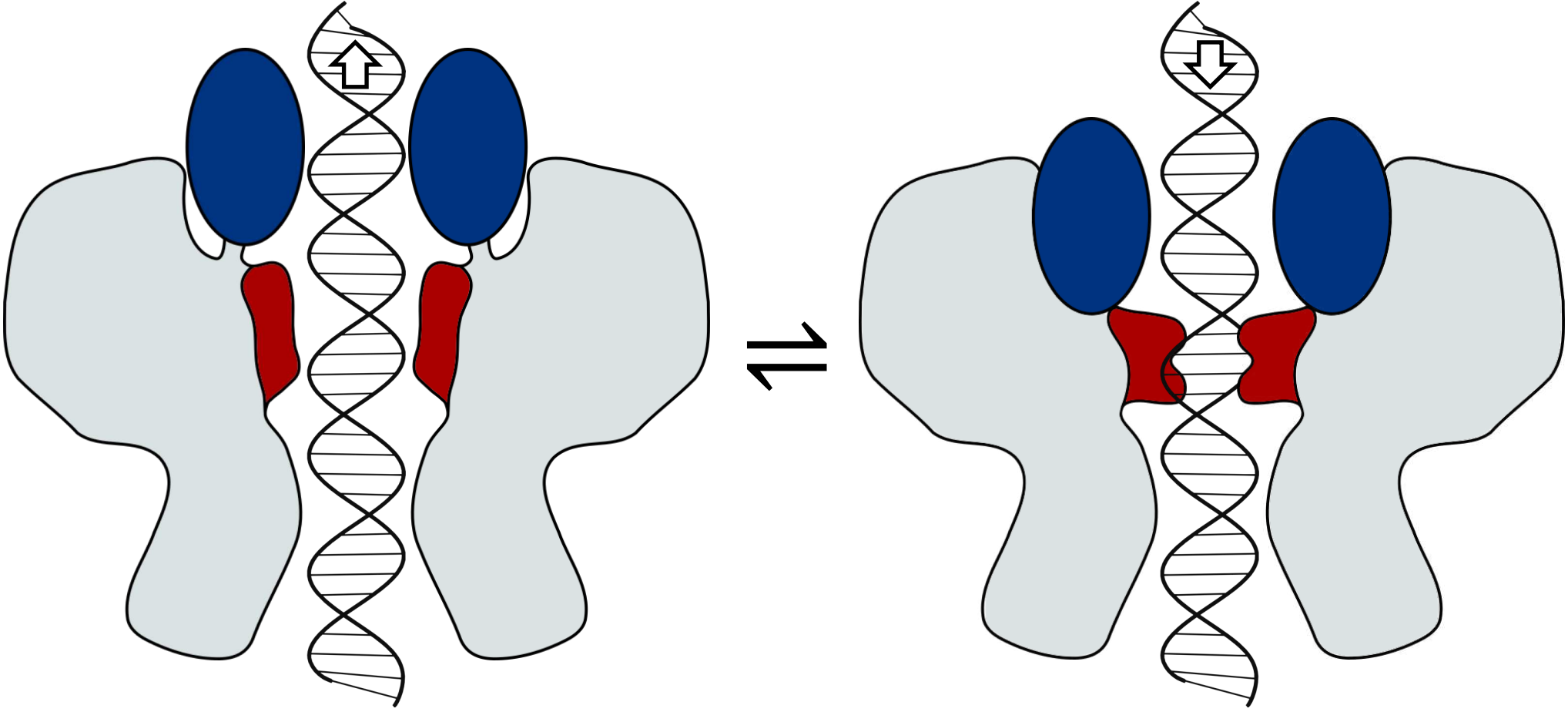
Video 3. Portal–capsid registers. One-degree step change in portal register (inner 12-fold circle) with respect to capsid vertex (outer 5-fold circle), beginning with "0^{o}". Portal register "6^{o}" is superposable on register "0^{o}" by 144^{o} rotation of the whole capsid (i.e. rotating both inner and outer circles together).

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


A
B
Interior
Exterior
90°
C
D
Wing
Crown
NT
Tunnel loop
Tunnel helix
Stem
Stem helix
Clip
E
E27
G28
L26
P32
S30
V31
H221
K225
I232
L236
R216
R226
L230
H234
Y303
H304
Y300
F316
T302
P295
I298
E332
G331
I330
V328
A327
N322
N329
Q326
V325
G324
M323
N-terminus
Stem helix
Tunnel helix
Tunnel loop

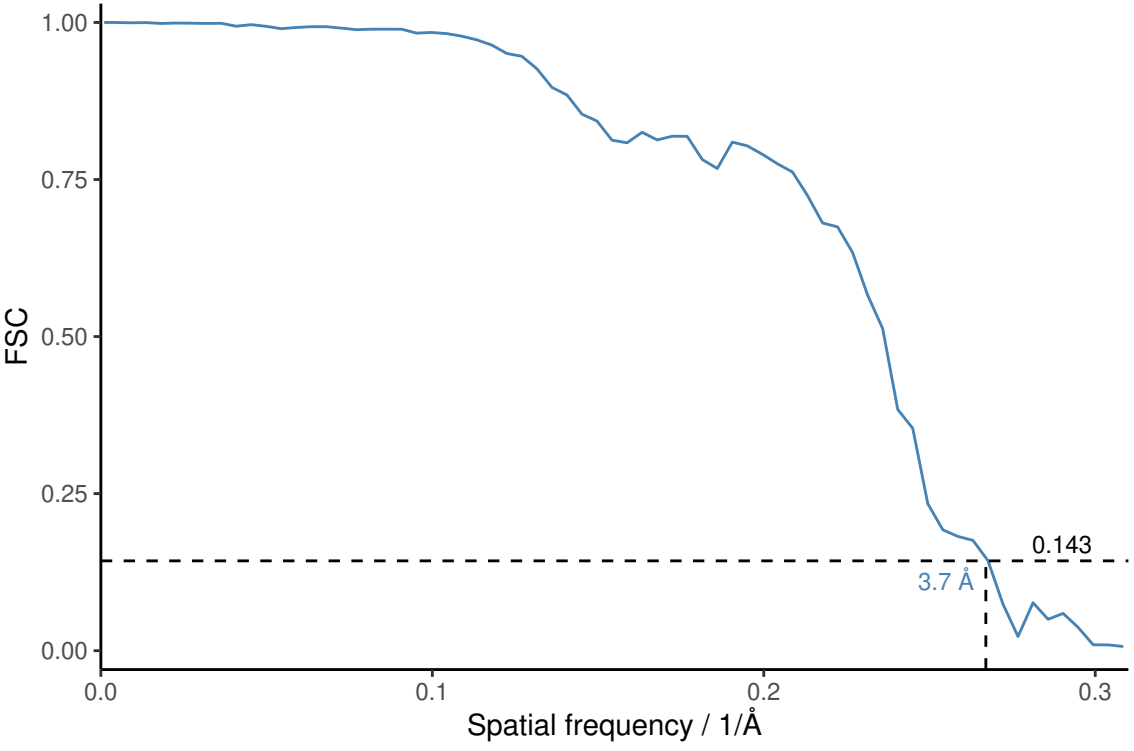

A
Crown
NT
Tunnel
loops
Wing
Stem
Clip
NT
In situ
Crystal
B
5Å
8.2°
C
13.3°
D
I330
N329
Q326
V325
In situ
Crystal
E
14 Å
F
22 Å
hydrophobic
hydrophilic


A
B
wing
β-hairpin
185-189
337-340
24-30
capsid
C
D
357-358
capsid
263-275
119-127
clip
E
A
C
J
E
H
F
F
C
L
E
J
G
H



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

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


1.00
0.75
0.50
0.25
0.00
FSC
0.143
3.7 Å
0.0
0.1
0.2
0.3
Spatial frequency / 1/Å

1 **MAKRGRKPKELVPGPGSIDPSDVPKLEGASVPVMSTSYDVVVDREFDELL**

51 **QGKDGLLVYHKML**SDGTVKNALNYIFGRIRS**AKWYVEPASTDPE**DIAIAA

101 FIHAQLGIDDASVGKYPFGRLFAIYENAYIYGMAAGEIVLTLGAD**GKLIL**

151 **DKIVPIHPFNIDE**VLYDEEGGPKALKLSGEVKGGSQFVNGLEIPIWKTVV

201 FLHNDDGSFTGQSALRAAVPHWLAKRAL**ILLINHGLERFMIGVPTLTIPK**

251 **SVRQGTKQWE**AAKEIVKNFVQKPRHGIILPDDWKFDTVDLKSAMPDAIPY

301 LTYHDAGIARALGIDFNTVQLNMGVQAVNIGE**FVSLTQQTIISLQRE**FAS

351 AVNLYLIPKLVLPNWPGATRFPRLTFEMEERNDFSAAANLM**GMLINAVKD**

401 **SEDIPTE**LKALIDALPSKMR**RALGVVDEVREAVRQPADSRYLYTR**RRR