



**HAL**  
open science

## Structural basis of bacteriophage T5 infection trigger and *E. coli* cell wall perforation

Romain Linares, Charles-Adrien Arnaud, Grégory Effantin, Claudine Darnault, Nathan Hugo Epalle, Elisabetta Boeri Erba, Guy Schoehn, Cécile Breyton

► **To cite this version:**

Romain Linares, Charles-Adrien Arnaud, Grégory Effantin, Claudine Darnault, Nathan Hugo Epalle, et al.. Structural basis of bacteriophage T5 infection trigger and *E. coli* cell wall perforation. *Science Advances* , 2023, 9 (12), pp.eade9674. 10.1126/sciadv.ade9674 . hal-04055766

**HAL Id: hal-04055766**

**<https://hal.science/hal-04055766v1>**

Submitted on 22 Nov 2023

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

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

## STRUCTURAL BIOLOGY

Structural basis of bacteriophage T5 infection trigger and *E. coli* cell wall perforation

Romain Linares, Charles-Adrien Arnaud, Grégory Effantin, Claudine Darnault, Nathan Hugo Epalle, Elisabetta Boeri Erba, Guy Schoehn, Cécile Breyton\*

Most bacteriophages present a tail allowing host recognition, cell wall perforation, and viral DNA channeling from the capsid to the infected bacterium cytoplasm. The majority of tailed phages bear a long flexible tail (*Siphoviridae*) at the tip of which receptor binding proteins (RBPs) specifically interact with their host, triggering infection. In siphophage T5, the unique RBP is located at the extremity of a central fiber. We present the structures of T5 tail tip, determined by cryo-electron microscopy before and after interaction with its *E. coli* receptor, FhuA, reconstituted into nanodisc. These structures bring out the important conformational changes undergone by T5 tail tip upon infection, which include bending of T5 central fiber on the side of the tail tip, tail anchoring to the membrane, tail tube opening, and formation of a transmembrane channel. The data allow to detail the first steps of an otherwise undescribed infection mechanism.

## INTRODUCTION

Bacteriophages or phages, viruses that infect bacteria, represent the most abundant biological entity on our planet. They are present in all ecosystems where bacteria develop and outnumber their hosts by at least an order of magnitude, being instrumental in the development and evolution of microbial populations (1). Moreover, with the increasing number of pathogenic strains resistant to antibiotics, virulent phages are considered as a serious alternative or complement to classical treatments (2). The vast majority of known phages bear a tail whose tip serves to recognize the host, perforate the bacterial cell wall, and deliver the viral genome into the host cytoplasm. Tails can be long and contractile in *Myoviridae*, long and flexible in *Siphoviridae*, or short in *Podoviridae*. Bacterial tail-like machines also serve as a means to inject various macromolecules in neighboring prokaryotic and/or eukaryotic cells: All these systems derive from a common ancestor that share high structural similarities (3–6). The contracting tails and bacterial tail-like systems have seen their mechanism of cell wall perforation described in exquisite details: They literally drill a hole through the cell wall by contracting an outer sheath that propels a needled inner tube in a syringe-like manner (5, 7–9). On the contrary, very little is known for siphophages, which represent more than 60% of all phages (10). Structures of the tail tube before and after interaction with the receptor suggest that the tail tube does not play a direct role in infection (11, 12). Structures of isolated purified baseplate (13–15) and of a tail tip (16) of Gram-positive infecting siphophages are available, and lactophage p2 purified baseplate was solved in a closed and open state, suggesting activation for this sugar-binding siphophage (15). However, the mechanism of transmission of host recognition, tail tube opening, and cell wall perforation remains completely unknown in siphophages.

Phage T5 (17), a *Siphoviridae* infecting *Escherichia coli*, is a model phage belonging to the T series introduced by Delbrück and co-workers in the 1940s (18). It presents a 90-nm icosahedral capsid (19) to which is attached a 160-nm tail tube (Fig. 1A), formed

by the polymerization of 40 ring-shaped trimers of the tail tube protein pb6 (TTP<sub>pb6</sub>) (11) around the tape measure protein pb2 (TMP<sub>pb2</sub>) (20). At its distal end, the tail harbors the tip complex, also called baseplate: three dispensable L-shaped side tail fibers (LTF<sub>pb1</sub>) reversibly bind to the sugar moiety of the host lipopolysaccharide (21). They are linked, by the collar, to a conical structure formed by the distal tail protein pb9 (DTP<sub>pb9</sub>) (22) and the baseplate hub protein pb3 (BHP<sub>pb3</sub>, also called tail-associated lysin/lysozyme) (17). A central fiber protein (pb4), at the extremity of which is found the receptor binding protein pb5 (RBP<sub>pb5</sub>) (23, 24), completes the tip complex, with p140 and p132 of unknown location (Fig. 1B) (17). FhuA, an outer membrane *E. coli* iron-ferrichrome transporter, is the bacterial receptor recognized by T5 (25). The mere interaction of T5 with its purified receptor FhuA triggers the release of viral DNA in vitro (26), making this phage an excellent model for studying host recognition, DNA ejection (11, 27), and cell wall perforation mechanisms. We thus embarked on solving the structure of T5 tail tip before (Tip) and after (Tip-FhuA) interaction with its receptor; our results unravel the conformational changes underwent by T5 tail tip and allow to detail the molecular mechanism of tail opening and outer membrane perforation.

## RESULTS

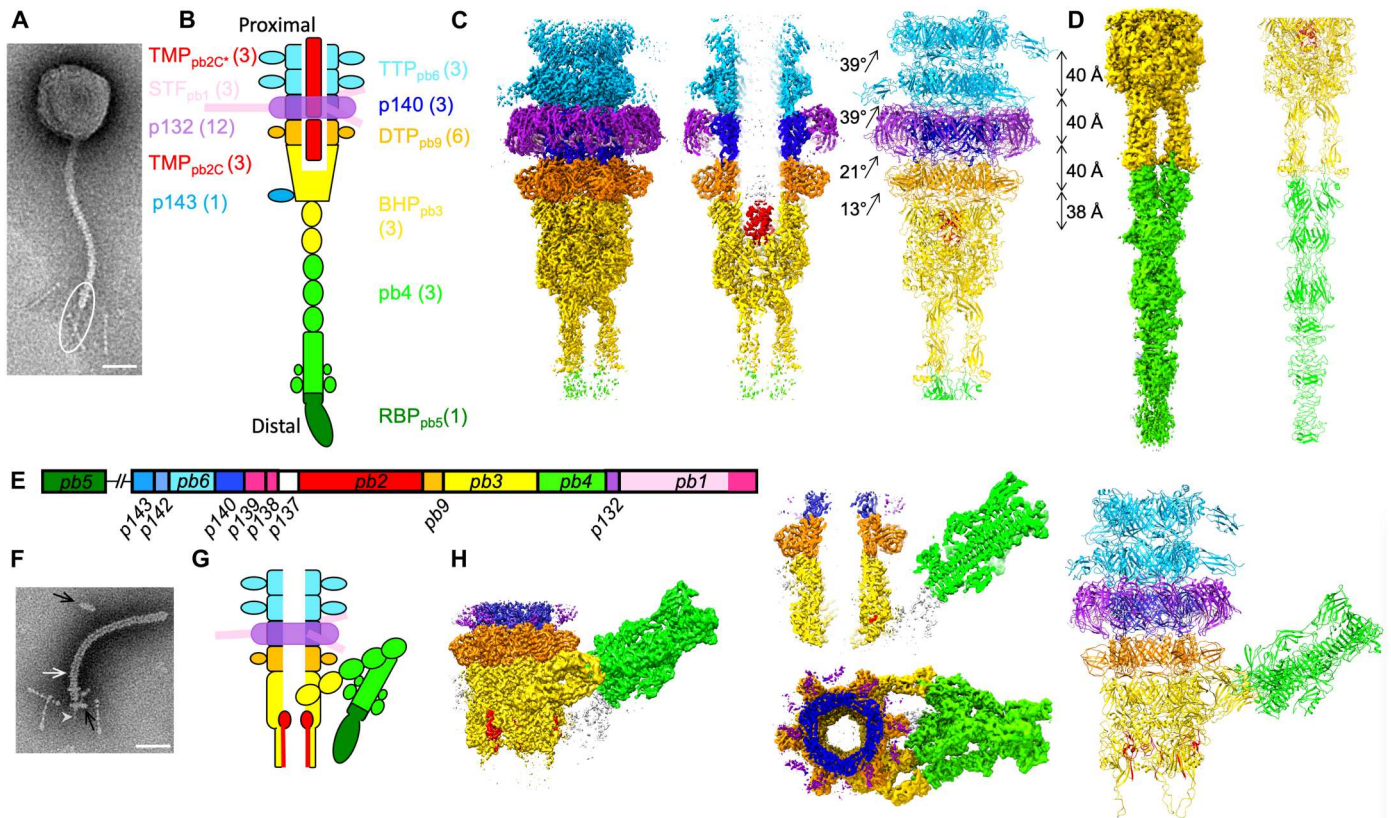
## General architecture

T5 tails (11, 17) were preferred over whole phages for cryo-electron microscopy (cryo-EM) as they allow better quality imaging (fig. S1A). Purified tails, obtained from a mutant bearing an amber mutation in the major capsid protein, behave as capsid-attached tails: They interact with their receptor and perforate outer membrane vesicles (17). Micrograph acquisition and extensive image processing (fig. S2 and table S1) yielded three maps of different tip subcomponents (Fig. 1, C and D, and fig. S1C), whose resolution allowed tracing all the proteins from the tail tube to the distal end of the central fiber, except for LTF<sub>pb1</sub> and RBP<sub>pb5</sub> (table S1). After two TTP<sub>pb6</sub> trimeric rings, the tube continues with a p140 trimer and then a DTP<sub>pb9</sub> hexamer. A BHP<sub>pb3</sub> trimer closes the tube and forms the beginning of the central fiber that continues with a pb4

Copyright © 2023  
The Authors, some  
rights reserved;  
exclusive licensee  
American Association  
for the Advancement  
of Science. No claim to  
original U.S. Government  
Works. Distributed  
under a Creative  
Commons Attribution  
License 4.0 (CC BY).

Univ. Grenoble Alpes, CNRS, CEA, IBS, Grenoble F-38000, France.

\*Corresponding author. Email: cecile.breyton@ibs.fr



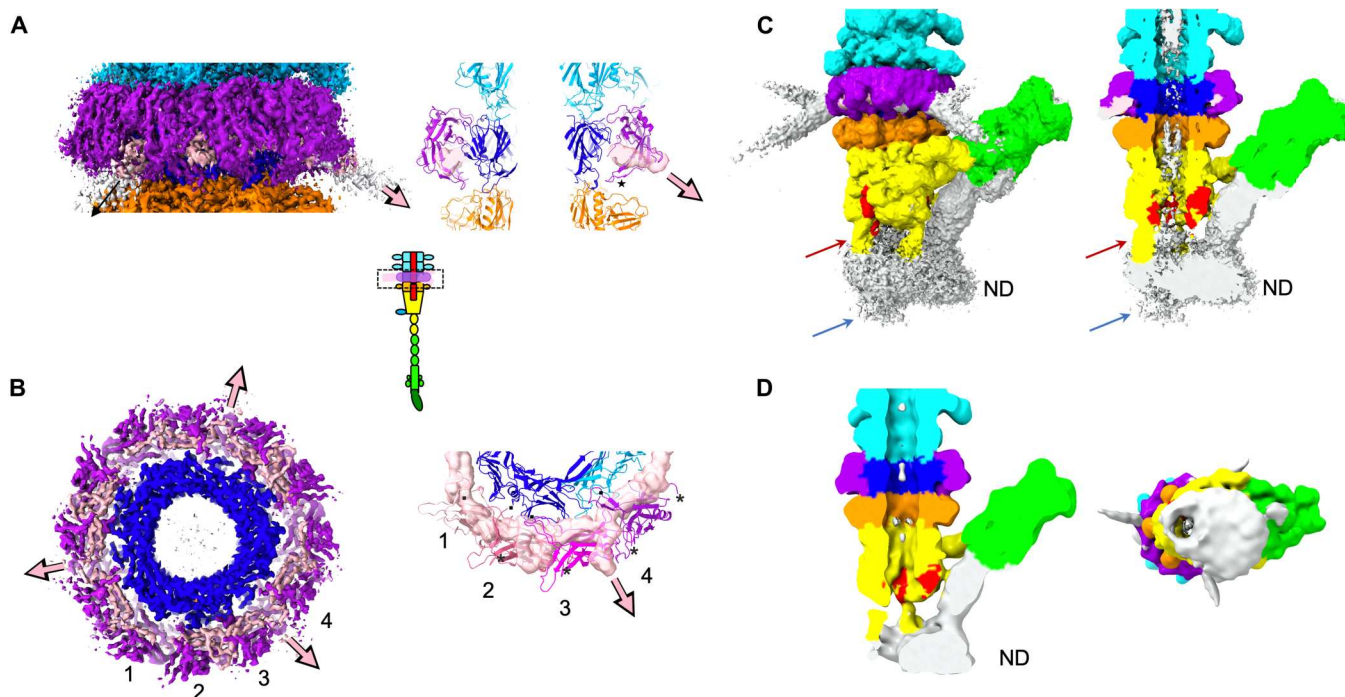
**Fig. 1. Structure of T5 tail tip before and after interaction with FhuA in nanodisc.** (A) Negative stain EM image of phage T5, with the tail tip circled. (B) Scheme of T5 tail tip with the assignment of the different proteins and their copy number. (C) Cryo-EM structure of T5 upper tail tip at 3.5-Å resolution. Left: Isosurface view of the map seen from the side; middle: central slice side view; right: ribbon representation of the modeled proteins. The twist and rise between each ring are noted. (D) Cryo-EM structure of the central fiber at 4.2-Å resolution. Left: Isosurface view of the map seen from the side; right: ribbon representation of the modeled proteins. (E) Map of T5 tail structural proteins and genes. (F) Negative stain EM image of a T5 tail interacting with a FhuA nanodisc (black arrows). The white arrow points to the empty-filled limit of the tail (see also extended data Fig. 1F). The gray arrowhead points to a density going through the nanodisc. (G) Scheme of T5 tip after interaction with FhuA (Tip-FhuA). (H) Cryo-EM structure of Tip-FhuA at resolutions ranging from 3.6- to 4.3-Å resolution. Isosurface view of a Tip-FhuA composite map seen from the side (left), a central slice side view of it (middle top), and a top view (middle bottom). This composite map is formed by the addition of Tip-FhuA C3 open tube and C1 bent fiber maps and is only for visualization purposes; right: ribbon representation of the modeled proteins. The color code in (C), (D), (G), and (H) is the same as in (B). Unattributed densities are in white. Scale bars, 50 nm.

trimer. The p140 trimeric ring is surrounded by a p132 dodecamer that forms the collar, onto which are grafted three LTF<sub>pb1</sub>.

Upon T5 tail incubation with detergent-solubilized FhuA, BHP<sub>pb3</sub> opens, TMP<sub>pb2</sub> is expelled from the tube lumen, and the central fiber disappears (fig. S1F, inset) (11). As the presence of a lipid bilayer might stabilize a cell wall perforation intermediate, we used instead FhuA reconstituted into nanodiscs. Nanodiscs are little patches of lipid bilayers stabilized by a membrane scaffold protein (MSP) (28). We used the largest MSP available to provide a lipid bilayer in addition to reconstituted FhuA (see also Material and Methods). Images of FhuA nanodisc-incubated tails clearly show (i) the presence of a nanodisc perpendicular to the tail tube at the rim of the open BHP<sub>pb3</sub>, (ii) a protrusion going through the nanodisc, (iii) TMP<sub>pb2</sub> partial ejection from the tail tube lumen, and (iv) the bending of the central fiber with a very acute angle on one side of the tip (Fig. 1F and fig. S1F), suggesting that we trapped an ejection intermediate. Extensive cryo-EM processing yielded three other maps (Fig. 1H, figs. S1H and S2, and table S1) allowing tracing all T5 tip proteins, except, again, LTF<sub>pb1</sub>

and RBP<sub>pb5</sub>. Densities belonging to RBP<sub>pb5</sub> are visible, but resolution is insufficient to build a model. However, a small-angle neutron scattering envelop of the FhuA-RBP<sub>pb5</sub> complex (24) and our recently determined FhuA-RBP<sub>pb5</sub> structure (29) could be very well fitted into the densities (fig. S3D).

The density corresponding to the nanodisc is clearly visible (Fig. 2, C and D), although nanodiscs are heterogeneous in size and in position relative to the tail (fig. S1F). Nanodisc density is not centered with respect to the tail tube axis: Its center of mass is shifted toward the bent fiber, below the density attributed to RBP<sub>pb5</sub>, under which the structure of FhuA could be fitted (Fig. 2, C and D, and fig. S3D). At low contour level, aligned with the tail tube lumen, a hole in the nanodisc is observed (Fig. 2D), strongly suggesting the presence of a channel at this position. At higher contour level, the tail tube lumen and the nanodisc are filled, and protrusions are visible above and below the nanodisc (Fig. 2C), as if a channel had perforated it. This channel is, however, poorly resolved, probably because of high heterogeneity in that region (fig. S1, F and H).



**Fig. 2. Structure of T5 Tip collar and Tip-FhuA complex.** (A) Left: Isosurface side view of the tip common core cryo-EM map at high contour level, centered on the p132 collar (boxed in the inset scheme of the tip). Right: Ribbon representation of a central slice of the collar. The star points to loop 52-60, and the density attributed to LTF<sub>pb1</sub> is in transparent isosurface representation. (B) Left: Isosurface bottom view of the map in (A), slice at the p132 collar level. Right: Bottom view of the four p132 monomers that are not related by the C<sub>3</sub> symmetry (colored from light pink to violet and numbered). They interact with two p140 monomers (cyan and blue) and with LTF<sub>pb1</sub> (transparent densities). The pink/black arrows point to the direction of the LTF<sub>pb1</sub>, and the N and C termini of the proteins are, respectively, indicated by black dots and asterisks. (C) Isosurface view at high contour level of Tip-FhuA unmasked and unfiltered cryo-EM map, side view (left) and slice (right). The red arrow points to one of the  $\beta$ -hairpin “leg.” The blue arrow points to the protrusion going through the nanodisc (ND). (D) Isosurface view at a lower contour level of Tip-FhuA unmasked cryo-EM map, after a 15-Å low-pass filtering, slice (left) and view from beneath the nanodisc (right). The color code is the same as in Fig. 1. Unattributed densities in (C) and (D) are in white.

Limited proteolysis experiments were performed on T5 tails and phages using subtilisin, and resulting particles analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and observed by negative stain EM. Within tail assembly, TMP<sub>pb2</sub> is among the first to be digested (fig. S4, A and D), as it has a tendency to be expelled from the tail by its proximal extremity. This suggests that the protein is contained in the tube in a metastable state. On the contrary, proteins of the tail tip complex are extremely stable. TTP<sub>pb6</sub> is also very stable with the exception of its decoration immunoglobulin (Ig)-like domain. Upon analysis with negative stain EM, particles appear intact, with the exception of TMP<sub>pb2</sub> in purified tails (fig. S4, B and E). Infectivity of proteolyzed phages is also only mildly affected, decreasing by only an order of magnitude (fig. S4C). Incubation of tails with FhuA expels TMP<sub>pb2</sub>, making it even more susceptible to the subtilisin action, whereas elements of the tail tip remain resistant to proteolysis despite the vast conformational changes they underwent.

### Core of the tip complex

When comparing the structures of the tip before and after interaction with FhuA, we could observe that they share a common core from the tail tube down to BHP<sub>pb3</sub> and only start to diverge from BHP<sub>pb3</sub> distal part. We previously determined a pseudo-atomic resolution structure of T5 tail tube. It is formed of a stack of TTP<sub>pb6</sub>

trimeric rings and exhibits an unusual pseudo sixfold symmetry (hexameric rings being the common tail tube architecture), the TTP<sub>pb6</sub> gene, resulting in a duplication/fusion of the canonical tail tube domain gene (table S2) (11). The tail tube domain is formed by a  $\beta$ -sandwich flanked by an  $\alpha$  helix and a long loop (fig. S5A) (11). At TTP<sub>pb6</sub> C terminus, an Ig-like domain decorates the tube, as in many *Siphoviridae* TTPs. Our tail tip structure includes two TTP<sub>pb6</sub> rings that could be modeled (Fig. 1C). This higher-resolution data confirm our previous modeling of the inter-ring interactions, mostly mediated by the long loops, the N terminus, the linker between the two tail tube domains of TTP<sub>pb6</sub>, and loops of the  $\beta$ -sandwiches (11), and have complementary charge surfaces (fig. S6B). The root mean square deviation (RMSD) between TTP<sub>pb6</sub> structures of the proximal and distal ring is only 0.5 Å over all 464 residues (fig. S5A), suggesting that the interface with the next p140 ring is very similar to that of two TTP<sub>pb6</sub> rings. The densities of the Ig-like domains are poorly defined (fig. S1C), witnessing a flexibility of this domain with respect to the tube scaffold (30).

The tube extends through the tip: After the last TTP<sub>pb6</sub> ring, it continues with a pseudo-hexameric p140 ring, a hexameric DTP<sub>pb9</sub> ring, the proximal domains of the BHP<sub>pb3</sub> trimer forming the last pseudo-hexameric ring of the tube. The structure of these proteins is also based on the canonical tail tube domain (6) but differently

decorated. Thus, the tube diameter is conserved, although the pitch and the twist between the different rings are different (Fig. 1C). As for TTP<sub>pb6</sub>, the interaction between the rings is mediated mostly by the long loops, the N terminus, and loops of the  $\beta$ -sandwiches (11); they also have complementary charged surfaces (fig. S6B). The inner surface of the tube is highly electronegative until DTP<sub>pb9</sub> (fig. S6A) allows DNA to slide along it (6).

p140 pairwise comparison with TTP<sub>pb6</sub> results in a very high DALI (31) Z score (fig. S5B), pointing to TTP<sub>pb6</sub> gene duplication to form p140, despite sequence identity between the two proteins being only 9% (fig. S5C). The main difference between the two proteins is the absence of the Ig-like domain in p140. The p140 ring is surrounded by a larger p132 dodecameric ring, p140 C terminus making direct contacts with a p132 monomer (Figs. 1C and 2, A and B), explaining the need for a decoration-less ring at this position. As suggested by p140 gene position in T5 genome (Fig. 1E), this protein is a bona fide component of its baseplate. p140 and p132 genes are a landmark of the large T5-like phages family only; thus, LTF anchoring would occur differently in other *Siphoviridae*, in particular, in the lambdoid phages. In the *Myoviridae* phage T4, the presence of an additional ring of the TTP-like protein gp54 between the “ring initiator” DTP<sub>gp48</sub> and the first bona fide TTP<sub>gp19</sub> rings (32) is also observed, and a role in sheath assembly initiation has been proposed (31). This additional ring is also not systematically present within the *Myoviridae* family.

We previously determined the crystal structure of DTP<sub>pb9</sub>, showing that the DTP was a common feature to both Gram-positive and Gram-negative infecting siphophages (22). In all other phages and tail-like machines, this protein ensures the six-to-threefold symmetry transition between the TTP hexamer and the BHP trimer. Here, however, the DTP<sub>pb9</sub> ring is sandwiched between two threefold symmetric rings (TTP<sub>pb6</sub> and BHP<sub>pb3</sub>), which both have a pseudo-sixfold symmetry. This explains the low RMSD between the two DTP<sub>pb9</sub> monomers that are not related by the imposed threefold symmetry of the map. DTP<sub>pb9</sub> tail tube domain is decorated with an OB (oligonucleotide-oligosaccharide binding) domain (fig. S5D). The OB domains, as the Ig-like domains, are proposed to interact with carbohydrates at the cell surface and serve to increase infectivity (4). Unlike DTPs of siphophages infecting Gram-positive bacteria that serve as a platform to anchor the RBPs (14–16), DTP<sub>pb9</sub> does not bind any other protein than those forming the tail tube.

### The collar, p132, and LTF<sub>pb1</sub>

The collar is made of a dodecamer of p132, whose position and structure were previously unknown. The p132 fold belongs to the Ig superfamily (fig. S5E), and a DALI search links it to the N-terminal domain of the baseplate protein upper (BppU; ORF48) of phage TP901-1 (14). Ig-like domains in phages are usually decoration domains; here, however, as for BppU, it is a bona fide structural protein that serves to anchor LTF<sub>pb1</sub>. The dodecameric p132 ring completely surrounds the trimeric p140 ring, with p132-p132 and p132-p140 contacts in both p132 and p140 proximal regions and, to a much lesser extent, p132-TTP<sub>pb6</sub> contacts, mainly mediated through the loops and termini of the three proteins (Figs. 1C and 2, A and B). There are no interactions between p132 and DTP<sub>pb9</sub> or the Ig-like domain of TTP<sub>pb6</sub>, as determined by PISA (33).

Unattributed densities in the lower part of the collar, intertwined between the p132 monomers, most probably belong to LTF<sub>pb1</sub>.

However, map quality/connectivity did not allow to unambiguously model it, but it could correspond to its ~50 N-terminal residues. These densities point out of the collar to form the start of the three LTF. The intricate LTF<sub>pb1</sub>-p132 interaction (Fig. 2, A and B) explains that in the fiberless T5-hd1 mutant, which bears a mutation in pb1 gene leading to a truncated protein, the collar is absent, and the p132 protein is not detected by Western blotting. Also, p132 could not be localized within T5 tails using anti-p132 IgGs (17). In solution, p132 bears a folded core with very flexible loops and termini (34), which could be the target of the IgGs. These loops are mainly unavailable in the phage context, as they are involved in protein-protein interactions (Fig. 2B).

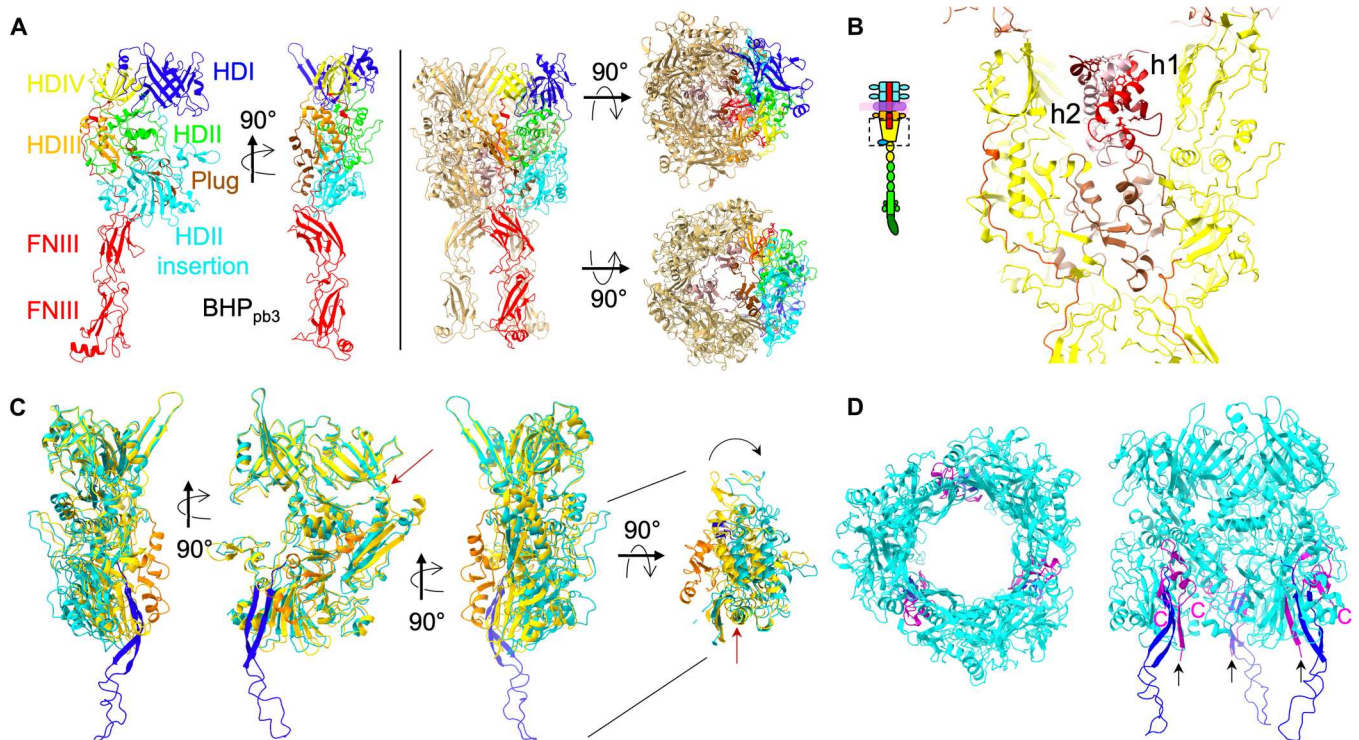
Four consecutive p132 monomers, not related by the threefold symmetry, show high structural similarity (Fig. 2B, and fig. S5, E and F), although their environment is different. There is a symmetry mismatch between LTF<sub>pb1</sub> 3-fold, p140/TTP<sub>pb6</sub> pseudo-6-fold, and p132 pseudo-12-fold symmetries. This symmetry mismatch is absorbed by the loops and N termini of p132 monomers. The C terminus and loop 52-60, involved in p132-p132 interactions, are less variable (fig. S5E).

### Closing the tube: BHP<sub>pb3</sub> and TMP<sub>pb2</sub>

BHP<sub>pb3</sub> trimer forms the most distal tail tube domain ring of the tube [hub domains (hd) I and IV; Fig. 3A and table S2]. The linker between the two tail tube domains and the long loop of the second tail tube domain have evolved into larger domains (hdII and insertion, and hdIII, respectively; fig. S7, A and B), the first of which is large enough to close the tube via a plug domain (Fig. 3A). A long linker runs along the protein down to the tip of the cone, inserted between two neighboring BHP<sub>pb3</sub> subunits and contributing to the stability of the closed tube (Fig. 3A). BHP<sub>pb3</sub> C terminus forms the beginning of the central fiber with two fibronectin domains (FNIII), further stabilizing the closed tube (Fig. 3B) and giving BHP<sub>pb3</sub> the shape of a trophy cup (Figs. 1, C and D, and 3A). BHP<sub>pb3</sub> results from a tail tube domain duplication/fusion, as TTP<sub>pb6</sub>. These two duplication/fusion events are clearly independent, however, as the fusion did not occur in the same way in the two proteins (fig. S7B).

Unexpectedly, on unsymmetrized EM reconstructions, an extra density is observed at the base of only one BHP<sub>pb3</sub> monomer (fig. S3A). Resolution was insufficient to build an atomic model de novo, but secondary structure features could be identified. From the 11 proteins that form the tail and that have been identified by mass spectrometry (MS) (table S3), only p143 has not been located. Its gene position in the tail structural module (Fig. 1E) suggests that it is the tail completion protein, located in the head-to-tail joining region (4, 6). However, we have solved T5 proximal tail region where it could not be attributed to any density. A flexible fit (see the “Protein model building” section) of an AlphaFold2 (35) structure prediction of p143 into this extra density was convincing (fig. S3B). Thus, we propose that this unattributed density corresponds to p143. This unusual position for a “Tail Completion Protein” suggests, for p143, a different role than that of completing the tail at its proximal extremity, a role which needs to be elucidated. This density is absent from reconstructions of the tip after interaction with PhuA, suggesting that the protein could detach during the infection process.

The tube lumen is filled with TMP<sub>pb2</sub>. TMP<sub>pb2</sub> (20, 36), as  $\lambda$  TMP (37), undergoes proteolytic cleavage during tail morphogenesis, but it was supposed that the cleaved peptide, TMP<sub>pb2C</sub>, was



**Fig. 3. BHP<sub>pb3</sub> closing and opening of the tube.** (A) Left: Two side views of a BHP<sub>pb3</sub> monomer in ribbon representation, with hdl to hdIV domains colored blue, green, orange, and yellow, respectively; the hdII insertion in cyan; the plug in brown; and the C terminus extension, comprising the hdIV-FNIII linker and the two FNIIIs, in red. Right: Side, top, and bottom views of the BHP<sub>pb3</sub> trimer. One monomer is colored as on the left, and the three plug domains are colored brown. In the bottom view, the FNIIIs have been removed for clarity. (B) Central slice through BHP<sub>pb3</sub> cup, boxed in the inset scheme of the tip (yellow; hdIV-FNIII linker, red; plug, brown) highlighting the 35 resolved residues of a TMP<sub>pb2C</sub> trimer in different shades of red. Hydrophobic residues of TMP<sub>pb2C</sub>, pointing to the center of the coil are represented in sticks. (C) Overlay of BHP<sub>pb3</sub> before (yellow; plug, orange) and after (cyan; plug, blue) opening of the BHP<sub>pb3</sub> cone, after superimposition of the whole tip. Three side views 90° apart and a top view are shown. In the top view, hdl and hdIV have been removed to highlight the pivotal movement of the hdII insertion domain. A red arrow points to the long helix of hdIII that acts as a hinge (see also movie S4). The long linker and the FNIIIs have been removed for clarity. (D) Top and side views of the open BHP<sub>pb3</sub> trimer with the same color code as in (C), and TMP<sub>pb2\*</sub> 42 C-terminal residues in magenta. TMP<sub>pb2\*</sub> C termini are indicated (C) as well as the last built residue in N-terminal (T1085, black arrow).

removed from the final tail assembly, leaving the rest of the protein, TMP<sub>pb2\*</sub>, in the tail tube lumen. TMP<sub>pb2</sub> density is very ill-defined along the tail tube, probably due to a poor interaction network between the predicted coiled-coil (20) of TMP<sub>pb2\*</sub> and the wall of the tube, except in the BHP<sub>pb3</sub> cup (Figs. 1C and 3B). There, we could model TMP<sub>pb2C</sub> 35 C-terminal residues. Three TMP<sub>pb2C</sub> copies are coiled in a superhelix, burying hydrophobic residues in its center, interacting closely with BHP<sub>pb3</sub> plug (Fig. 3B). The TMP<sub>pb2C</sub> presence in the tail is confirmed by proteomics and liquid chromatography/electrospray ionization/time-of-flight MS (LC-ESI-TOF-MS) (table S3 and fig. S8C) and allowed us to localize TMP<sub>pb2</sub> cleavage site after R1127 (fig. S5G). This latter is located between the TMP<sub>pb2\*</sub> hydrophobic stretch and a metallopeptidase motif that was shown to have muralytic activity (20), separating this enzymatic domain from the rest of the protein. This is reminiscent of the situation in T4, where the cell-puncturing protein gp5 is cleaved during tail assembly between its lysozyme domain and the  $\beta$  helix spike (38). It was previously suggested that TMP<sub>pb2</sub> (36), as other phages TMPs [e.g.,  $\lambda$  (39)], is a hexamer, but our data clearly indicate that only a trimer is present. A trimer of the 20 C-terminal residues of phage 80 $\alpha$  TMP was also modeled (16), suggesting that this might be a general feature of siphophages.

Interactions between TMP<sub>pb2</sub> and BHP<sub>pb3</sub> are mainly electrostatic (fig. S6, B and C), unlike in 80 $\alpha$ .

### The central fiber and its rearrangement upon receptor binding

BHP<sub>pb3</sub> C-terminal ~210 residues and pb4 N-terminal ~350 residues were modeled de novo (Fig. 1, C and D). As the resolution of the fiber map drops toward the fiber tip because of its flexibility, the rest of pb4 protein was modeled using flexible fitting of the better resolved equivalent domains built in the Tip-FhuA maps (see below and fig. S1H).

The proximal region of the central fiber is made of three strings of five consecutive FNIIIs, two at BHP<sub>pb3</sub> C terminus and three at pb4 N terminus. It starts as three independent strings; the repulsion between them could be caused by an important negative patch at the surface of the BHP<sub>pb3</sub> FNIIIs (fig. S6D). BHP<sub>pb3</sub>-pb4 interaction is ensured by two distal loops of the second FNIII of BHP<sub>pb3</sub> and the N terminus and two proximal loops of the first pb4 FNIII (fig. S8A). After a hinge region, the three pb4 monomers merge to form a 110-Å-long  $\beta$  helix spike, formed of a 24  $\beta$  strand longitudinal mixed  $\beta$  sheet prism. It has a triangular section with a mean diameter of 20 Å, delineating a very dense and hydrophobic interior (Fig. 4B).  $\beta$

Helix spikes/fibers are very common in phage host recognition or/and perforation apparatus, and a DALI search indeed relates pb4 spike to different phage and tail-like machines spikes/fibers (fig. 8B). Here, however, pb4 does not have a direct role in perforation or host recognition.

The spike is decorated with two small globular domains inserted between  $\beta$  strands 15 and 16 of each pb4 subunit and which are connected by relatively long linkers (Fig. 4B). These domains are not visible in the central fiber map, probably because of the degrees of freedom offered by the linkers. A DALI search links the second small domain to phage spike decoration domains [Protein Data Bank (PDB) codes 7CHU-A, 6TGF-D, 6E1R-A, 5M9F-A, 6NW9-C, 6EU4-A, and 5W6H-A] with a Z score of 3.9 to 2.3, RMSD from 2.4 to 3.0 Å over ~45 residues and identity ranging from 4 to 14%. This further illustrates the ability of phages to reuse and exchange functional modules (40).

The central fiber ends with RBP<sub>pb5</sub>, but this part of the map is poorly resolved and did not allow building an atomic model of this protein. This 310-Å-long central fiber bears two hinges, one between pb4 last FNIII and the spike and the other at the spike-RBP<sub>pb5</sub> interface: It introduces some controlled flexibility to this otherwise rigid assembly and may ease RBP<sub>pb5</sub> exploring space and encountering its bacterial receptor.

Upon RBP<sub>pb5</sub> binding to FhuA, the central fiber reorganizes: The three FNIII strings dissociate, two strings relocate on one side of BHP<sub>pb3</sub>, and the third one on the other side (Fig. 4A and movies S1 and S2). This reorganization of the central fiber is allowed by the long linker that connects BHP<sub>pb3</sub> hdIV to the first FNIII and which now runs perpendicular to the tail axis along BHP<sub>pb3</sub>. The central fiber bends by ~160° at the level of the hinge between the third pb4 FNIII and pb4 spike: The latter is now surrounded and stabilized by pb4 FNIIIs, which interact with and stabilize pb4 spike small decoration domains (Fig. 4A). pb4 spike also undergoes structural rearrangement, with a different twist of the spike (Fig. 4C). This bending and stabilization of the central fiber bring the tail tube closer to the membrane, orient it correctly, and anchor the tail to the membrane.

### Tube opening and anchoring of the tail to the membrane

As mentioned above, the structures of the tip before and after interaction with FhuA start to diverge from BHP<sub>pb3</sub> distal domains. More precisely, hdI and hIV overlay remarkably (Fig. 3C and movie S3), with an RMSD of 0.85 Å over 244 residues. However, hdII and hdIII rotate around the long helix of hdIII as a rigid body (the RMSD before and after opening of hdII-III is 0.75 Å over 367 residues), and the plug unwinds in a long  $\beta$ -hairpin (Figs. 3, C and D, and 4A; and movies S3 and S4). These conformational changes result in BHP<sub>pb3</sub> trimer opening, creating a channel with a constant ~40 Å in diameter from the hdI-hdIV ring to the hdII insertion tip (Fig. 3, C and D, and movie S4). The three  $\beta$ -hairpin legs connect BHP<sub>pb3</sub> to the nanodisc (Fig. 2, C and D): Their tip is composed of 233-Leu-Phe-Gly-Leu-236, which would insert into the outer leaflet of the lipid bilayer hydrophobic core. Above these hydrophobic residues stand 230-Lys-Lys-Lys-232 and Arg<sup>238</sup>, conferring a strong positive charge to the  $\beta$ -hairpin (fig. S6E). They could interact with the negatively charged phosphate groups of the lipopolysaccharides, further stabilizing the anchoring of the  $\beta$ -hairpin to the membrane.

In the crevice opened at the interface between two BHP<sub>pb3</sub> subunits, extra densities were identified, in which the 43 C-terminal residues of TMP<sub>pb2\*</sub> could be modeled (Figs. 3D and 4A). These densities merge with the ill-defined densities of BHP<sub>pb3</sub>  $\beta$ -hairpin, strongly suggesting that TMP<sub>pb2\*</sub> continues toward the nanodisc along BHP<sub>pb3</sub>  $\beta$ -hairpin, forming with the latter a three-stranded  $\beta$  sheet. TMP<sub>pb2\*</sub> continues with a stretch of nine residues, long enough to reach the nanodisc, followed by a stretch of 46 hydrophobic residues, compatible with two transmembrane helices (fig. S5G), which would insert into the outer membrane and form a channel (fig. S3C). Thus, T5 tail tube is anchored to the outer membrane by both BHP<sub>pb3</sub> and TMP<sub>pb2\*</sub>, in addition to FhuA-RBP<sub>pb5</sub>: It ensures that the tail tube is locked in register with the channel formed by TMP<sub>pb2\*</sub> in the outer membrane.

## DISCUSSION

### Baseplate comparison

As expected, BHP<sub>pb3</sub> structure partially aligns with other phages and tail-like machine BHPs, with high DALI Z scores (fig. S7C): The four hd of the canonical T4 BHP<sub>gp27</sub> (38) are also present in BHP<sub>pb3</sub>, but there is a large insertion in hdII, which closes the tube (Fig. 3A and fig. S7D). In *Myoviridae* and tail-like machines, the tail tube is closed by an OB domain followed by a spike (fig. S7D) (7–9, 38). In *Siphoviridae*, there is more diversity for closing the tail tube. The four baseplate structures available to date [phages T5, 80 $\alpha$  (16), p2 (15), and gene transfer agent (GTA) (41)] exhibit three different closing modes (fig. S7D): In p2 BHP, two hdII loops pointing toward the lumen of the tube are longer than in *Myoviridae* and close the tube. There, tube opening is induced by an iris-like movement, triggered by Ca<sup>2+</sup> binding, of hdII-III (15). In 80 $\alpha$ , the tube is closed by a helix in the C-terminal extension of the BHP that forms a twisted tripod in the trimer lumen (16). Last, T5 and GTA tubes are closed by the large hdII insertion (41). The two proteins composing GTA BHP, the Hub and the Megatron, align extremely well with BHP<sub>pb3</sub> (fig. S7, C and E).

Superimposing the *Siphoviridae*-related baseplate structures available, it was notable to observe the remarkable structural superposition of helix 2 of TMP<sub>pb2C</sub>, the resolved helix of TMP<sub>80 $\alpha$</sub> , and helix  $\alpha$ 1 of the iris/penetration domain of GTA Megatron (fig. S5H). Sequence alignment showed no detectable sequence similarity, and in the case of T5 TMP<sub>pb2C</sub> and TMP<sub>80 $\alpha$</sub>  C terminus, the interaction with the BHP is via the bottom of the BHP cup (Figs. 1C and 3B). In GTA, helix  $\alpha$ 1 of the Megatron is proposed to be a pore-lining helix that could insert in the outer membrane to allow the DNA across it, which however cannot be a general feature in *Siphoviridae*.

### T5 trigger for infection and formation of a channel

Comparing the structures of T5 tail tip before and after interaction with FhuA, we can propose a mechanism for T5 trigger for infection (Fig. 5). Upon RBP<sub>pb5</sub> binding to FhuA, a constraint at the RBP<sub>pb5</sub>-pb4 interface occurs (29), resulting in the different twist of pb4 spike observed in its proximal part. This twisting of the spike would pull on pb4 FNIII-spike linker, leading to the disruption of the FNIII-spike interaction network (Fig. 5B). The association between this FNIII and the spike is thus loosened; the pb4 FNIII-spike linker reorganizes and stabilizes a new interaction network between the three FNIII strings and the spike (Fig. 5C). This series of events results in the bending of the central fiber, at the level of the

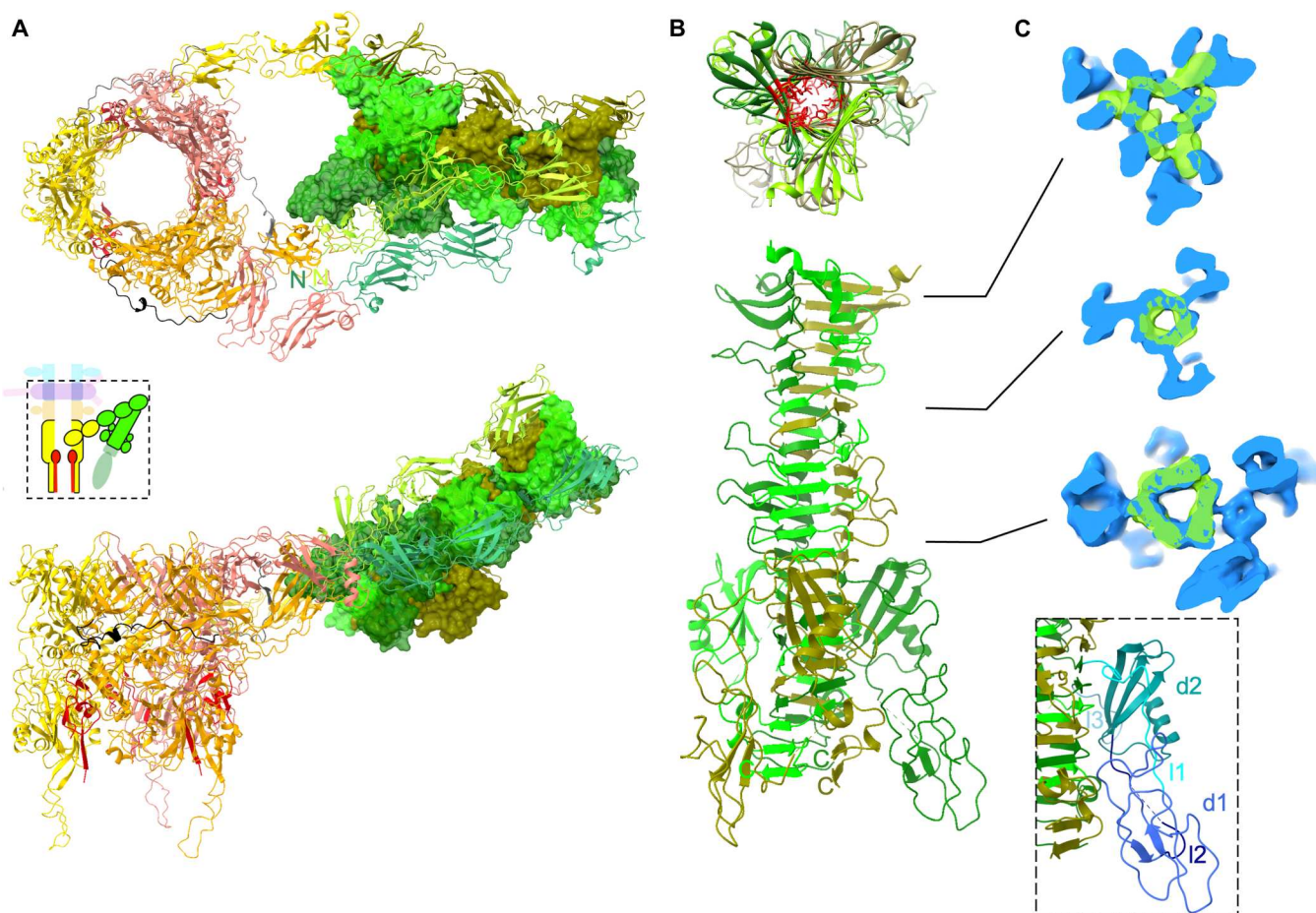
FNIII-spike hinge, on the side of the tube, pulling the tube toward the host membrane (Fig. 5D).

To accommodate for the drastic conformational changes of the FNIII strings induced by the bending of the central fiber, BHP<sub>pb3</sub> hdIV-FNIII linker is pulled and, like a zip, disrupts the interaction network between neighboring BHP<sub>pb3</sub> monomers and with BHP<sub>pb3</sub> plug. This then allows the rotation of BHP<sub>pb3</sub> hdII-III, plug unfolding/refolding, opening of the tube, and anchoring of BHP<sub>pb3</sub> to the membrane via the  $\beta$ -hairpin legs (Fig. 5D). BHP<sub>pb3</sub> closed conformation would be stabilized in a metastable state by the assembly process and interaction with its tip partners, TMP<sub>pb2C</sub> in particular. Open BHP<sub>pb3</sub> would be of lower energy and would drive the conformational changes underwent by the tip complex upon infection. T5 tip structure was proposed to the CASP14 competition: BHP<sub>pb3</sub> open structure only was correctly predicted (42).

BHP<sub>pb3</sub> opening disrupts the interaction with TMP<sub>pb2C</sub>, which is expelled from the tube and translocated to the host periplasm where it would locally digest the peptidoglycan; a refolding step could be

necessary. TMP<sub>pb2\*</sub> is in turn expelled from the tail tube and anchor its C terminus in the crevice created between BHP<sub>pb3</sub> monomers upon opening (Fig. 5E). Its hydrophobic segments would then insert in the outer membrane to form a transmembrane channel. These events are thermodynamically favored by hydrophobic segment insertion in the membrane and TMP<sub>pb2\*</sub> alleged metastable state within the tail tube. TMP<sub>pb2</sub> was shown to be involved in contact points between outer and inner membranes (43). Whether TMP<sub>pb2\*</sub> inserts into both the outer and the inner membrane remains to be determined: To form a channel wide enough to allow DNA through would require the six TMP<sub>pb2\*</sub> transmembrane helices. Insertion into the inner membrane could thus occur via another part of the protein and/or the recruitment of host proteins.

The mechanism presented here, by which receptor binding triggers the opening of its tail tube, its anchoring to the host membrane, and formation of a transmembrane channel is, to our knowledge, the first one described for *Siphoviridae*, the most prevalent family of phages. It is furthermore entirely original compared to what



**Fig. 4. Bending of T5 straight fiber.** (A) Structure of Tip-FhuA BHP<sub>pb3</sub>, 42 C-terminal TMP<sub>pb2\*</sub> residues, and pb4 (boxed in the inset scheme of the tip). BHP<sub>pb3</sub> monomers are in gold, orange, and salmon with the hdIV-FNIII linker colored in different shades of gray, pb4 in different shades of green, and TMP<sub>pb2\*</sub> in red. All proteins are in ribbon representation, except for pb4 spike that is in surface representation. Top: Top view. pb4 N termini (N) and BHP<sub>pb3</sub> C termini (C) are indicated; bottom: side view. (B) Top: Top view of pb4 spike. pb4 monomers are in different shades of green. The hydrophobic residues pointing toward the interior of the spike are depicted red and in sticks in one subunit only. Bottom: Side view of pb4 spike. C termini are indicated (C). (C) Superimposition of pb4 spike in isosurface view of the tip (green) and Tip-FhuA (blue) maps (unsharpened). Three slices are shown, and their position along the spike is indicated in (B). The map after interaction with the receptor contains the spike decoration domains and the FNIIIIs while that before interaction contains only pb4 spike. Inset: pb4 spike decoration domains and linkers are colored in different shades of blue on one pb4 subunit [linker 1 (l1, residues 484 to 547), domain 1 (d1, 548 to 566), linker 2 (l2, 548 to 566), domain 2 (d2, 567 to 618), and linker 3 (l3, 519 to 626)].



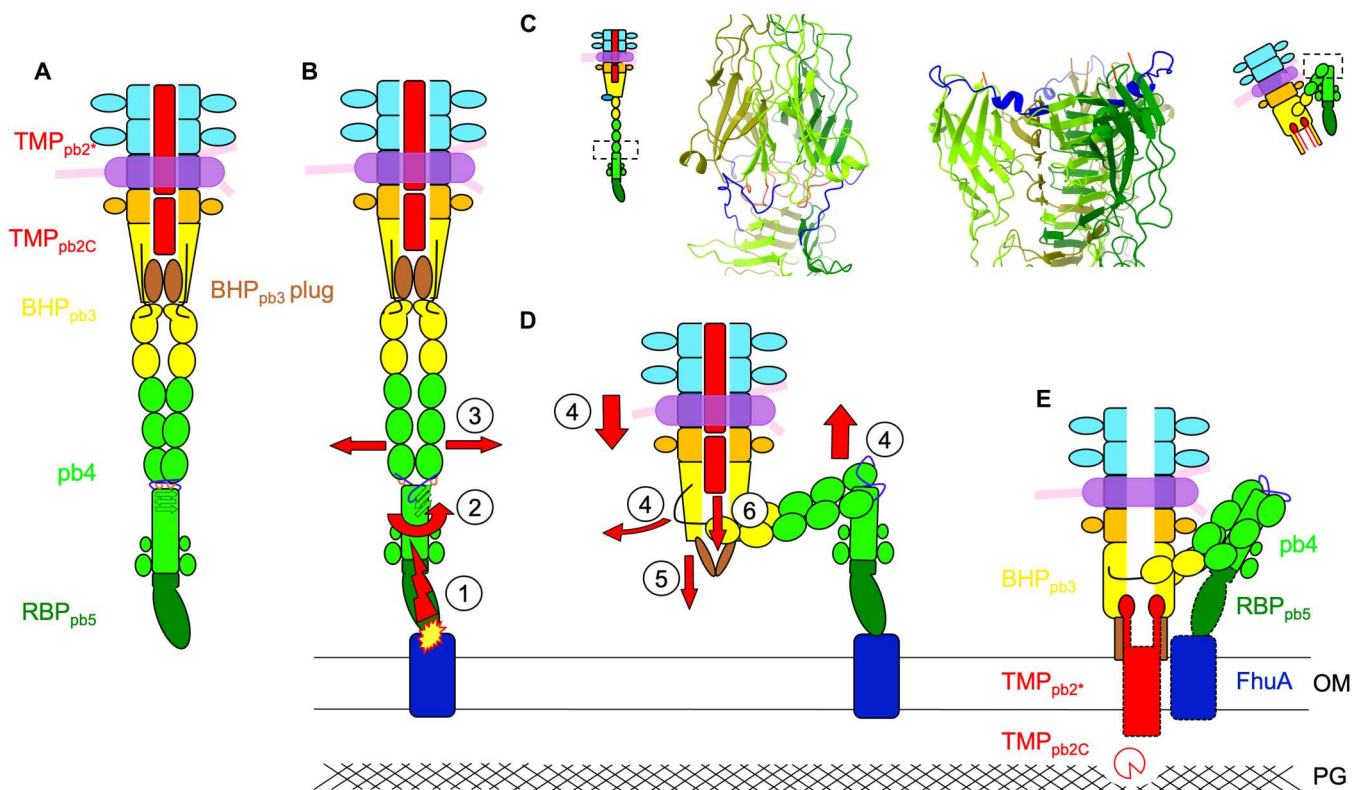
was known and described so far for the more complex *Myoviridae* and related tail-like bacterial machines. Our study provides a solid structural basis to further explore the diversity of viral entry mechanisms and their properties. This mechanism was recently complemented by the determination of FhuA-RBP<sub>pb5</sub> structure, which gives insights into how host recognition (RBP<sub>pb5</sub> binding to FhuA) triggers infection (29).

## MATERIALS AND METHODS

### T5 tail purification

T5 tails were preferred over whole phages for cryo-EM as the former allow thinner ice and no DNA background and, thus, better quality images. Purified tails are able to join filled heads to make infective particles (44). *E. coli* strain F cultures at 37°C were infected during the exponential growth phase with the amber mutant phage T5D20*am30d*, which bears an amber mutation in the major capsid protein gene, at a multiplicity of infection of 8. After complete cell lysis [optical density at 600 nm (OD<sub>600</sub>) < 0.15], the cell lysate was incubated with ribonuclease (1 µg/ml), 0.2% chloroform, and 0.5 M NaCl at 37°C for an hour and centrifuged for 20 min at

6000g to remove cell debris and unlysed cells. T5 tails were then precipitated from the culture medium by incubation with 8% (w/w) polyethylene glycol (PEG) 6000 overnight at 4°C. The pellet of a 1.5-hour 6000g centrifugation was resolubilized in 10 mM tris (pH 7.5), 100 mM NaCl, 1 mM CaCl<sub>2</sub>, and 1 mM MgCl<sub>2</sub> and purified on a glycerol step gradient (10 to 40%) in the same buffer centrifuged for 2 hours at 20,000 rpm (SW41 rotor). The gradient fractions containing the tails (usually ~10% glycerol), diluted five times in 10 mM tris (pH 7.5), 1 mM CaCl<sub>2</sub>, and 1 mM MgCl<sub>2</sub>, were loaded onto an ion exchange column (HiTrap Q HP 1 ml, GE HealthCare), equilibrated, and washed in the same buffer. The tails were eluted by a 0 to 0.5 M NaCl linear gradient. Purified tails were incubated 30 min with FhuA-loaded nanodiscs at room temperature before preparation of the cryo grids as a longer incubation time leads to a heterogeneous and aggregated sample. The tail/FhuA nanodisc ratio was first screened in negative stain. The chosen ratio included a large majority of tails having interacted with a nanodisc without too high a background in nanodisc.



**Fig. 5. Proposed mechanism of trigger for infection.** (A) Scheme of T5 tail tip. The hdIV-FNIII linker (black) and the plug (brown) are highlighted in BHP<sub>pb3</sub>, as well as the FNIII-spike linker (blue), loop 224-232 of the third FNIII (salmon), and the orientation of the proximal three  $\beta$  strands of the spike (black arrows) in pb4. (B) Following RBP<sub>pb5</sub>-FhuA interaction, a constraint (1) would induce a twisting of the proximal pb4 spike (2), pulling on pb4 FNIII-spike linker. This in turn would destabilize the FNIII string network (3). (C) Blow up on pb4 FNIII-spike interface before (left) and after (right) interaction with FhuA. The two spikes are aligned on the middle sheet of the spike (residues 413 to 465). pb4 subunits are in different shades of green, the FNIII-spike linker in blue, and FNIII loop 224-232 in salmon. (D) The FNIII string reorganization around pb4 spike induces pb4 bending, brings the tube close to the membrane, and disengages BHP<sub>pb3</sub> hdIV-FNIII linker (4). This latter event liberates the plug, opening the tube (5) and destabilizing TMP<sub>pb2C</sub>, which is expelled (6). (E) BHP<sub>pb3</sub> plugs refold as a  $\beta$ -hairpin legs and anchor in the outer membrane (OM), TMP<sub>pb2\*</sub> is also expelled, its C termini inserting in the crevice between BHP<sub>pb3</sub> subunits, its hydrophobic segment inserting in the OM to form a channel. TMP<sub>pb2C</sub>, released in the periplasm, would digest the peptidoglycan (PG). In (E), colored boxes depict proteins that could be modeled (full line) or for which densities are visible (dotted line). TMP<sub>pb2C</sub> for which no densities is visible but for which we propose a location, is represented as an empty Pacman.

### FhuA-containing nanodiscs

The gene coding for the MSP1E3D1 (28) was cloned in a pET28a plasmid used to transform BL21(DE3) *E. coli*. Protein expression was induced by the addition of 1 mM isopropyl- $\beta$ -D-thiogalactopyranoside to a 37°C growing Terrific Broth/kanamycin (50  $\mu\text{g ml}^{-1}$ ) culture when it reached an OD<sub>600</sub> of 1.2. Cells were harvested 4 hours later, resuspended in a lysis buffer [20 mM NaPO<sub>4</sub> (pH 7.4), 1% (w/v) Triton X-100, lysozyme (0.2 mg.ml<sup>-1</sup>), and deoxyribonuclease I (0.2 mg.ml<sup>-1</sup>)] and broken through three to four passages in a microfluidizer (13 kpsi). After clarification of the cell lysate, the supernatant was loaded on a nickel affinity column (HiTrap Chelating HP 5 ml, GE HealthCare) previously equilibrated with 40 mM tris-HCl (pH 8.0), 300 mM NaCl, and 1% (w/v) Triton X-100. The column was then washed with the same buffer, then with 40 mM tris-HCl (pH 8.0), 300 mM NaCl, and 50 mM sodium cholate, then with 40 mM tris-HCl (pH 8.0) and 300 mM NaCl, and then with 40 mM tris-HCl (pH 8.0), 300 mM NaCl, and 10 mM imidazol. The protein was eluted with 40 mM tris-HCl (pH 8.0), 300 mM NaCl, and 300 mM imidazol. The histidine tag was cleaved by incubating the MSP with the Tobacco Etch Virus (TEV) protease overnight in a MSP/TSV 1:10 (w/w) ratio at room temperature in 50 mM tris-HCl (pH 8.0), 0.5 mM EDTA, and 1 mM dithiothreitol and then dialyzed for 2 hours against 20 mM tris-HCl (pH 8.0) and 150 mM NaCl. The cleaved protein was loaded onto the same nickel affinity column equilibrated with 20 mM tris-HCl (pH 8.0) and 150 mM NaCl. Cleaved MSP was recovered in the flow through, concentrated, and loaded onto a size exclusion column (SD75 10/300 GL, GE HealthCare) equilibrated in 20 mM tris-HCl (pH 7.4), 150 mM NaCl, and 0.5 mM EDTA. Fractions of pure protein were pooled, and the protein was concentrated to 2 mg/ml on Amicon concentrator [molecular weight cutoff (MWCO) 100 kDa].

FhuA was produced and purified as described (23): *E. coli* AW740 (FhuA31  $\Delta$ ompF *zcb::Tn10*  $\Delta$ ompC) transformed with the pHX405 plasmid, in which the *fhuA*, gene under control of its natural promoter, was grown at 37°C in LB medium supplemented with ampicillin (125  $\mu\text{g ml}^{-1}$ ), tetracyclin (10  $\mu\text{g ml}^{-1}$ ), and 2,2'-bipyridyl (100  $\mu\text{M}$ ), an iron chelator used to induce FhuA production. After clarification of the cell lysate, total membranes were recovered by ultracentrifugation and solubilized using 50 mM tris (pH 8.0) and 2% (w/w) *N*-octylpolyoxyethylene (Bachem) at 37°C for half an hour. The insoluble material was recovered by ultracentrifugation and solubilized for 1 hour at 37°C using 50 mM tris (pH 8.0), 1 mM EDTA, and 1% (w/w) LDAO (*N,N*-dimethyl dodecylamine-*N*-oxide, Anatrace). The solubilized fraction, recovered after ultracentrifugation, was supplemented with 4 mM MgCl<sub>2</sub> and 5 mM imidazole and loaded on a nickel affinity column (HiTrap Chelating HP 5 ml, GE HealthCare) previously equilibrated with 0.1% LDAO, 20 mM tris (pH 8.0), and 200 mM NaCl and washed with the same buffer. The protein was eluted from the column with 0.1% LDAO, 20 mM tris (pH 8.0), and 200 mM imidazole and loaded onto an ion exchange column (HiTrap Q HP 1 ml, GE HealthCare) equilibrated with 0.05% LDAO and 20 mM tris (pH 8.0). The protein was eluted by a 0 to 1 M NaCl linear gradient.

To produce FhuA-loaded nanodiscs, purified FhuA was incubated with 1,2-dioleoyl-*sn*-glycero-3-phosphocholine solubilized in 100 mM sodium cholate and MSP1E3D1 in a 1:6:360 FhuA:MSP:lipid molar ratio. After 1 hour of incubation at 4°C, detergent was removed by the addition of 0.5 g ml<sup>-1</sup> BioBeads (Bio-Rad) and

incubation on a stirring wheel at room temperature for 2 hours. FhuA-loaded nanodiscs were diluted 6.5 times in 20 mM tris (pH 7.5), 150 mM NaCl, and 5 mM imidazol and further purified on a nickel affinity column equilibrated in the same buffer; eluted with the same buffer containing 200 mM imidazol; desalted on a PD10 (GE HealthCare) desalting column equilibrated in 10 mM tris (pH 7.5), 100 mM NaCl, 1 mM CaCl<sub>2</sub>, and 1 mM MgCl<sub>2</sub>; and concentrated ~40 times on a 50-kDa MWCO Amicon device.

### Cryo-EM sample preparation

Typically, 3.5  $\mu\text{l}$  of T5 tail sample (with or without FhuA nanodisc) was deposited on a freshly glow discharged (25 mA, 30 s) Cu/Rh 300 mesh Quantifoil R 2/1 EM grids and plunge-frozen in nitrogen-cooled liquid ethane using a Thermo Fisher Scientific Mark IV Vitrobot device (100% humidity, 20°C, 5-s blotting time, blot force 0).

### EM data acquisition

Respectively, 3208 and 9608 micrographs (split over two data collections for the latter) were collected for tails alone, and tails were incubated with FhuA nanodisc. Forty-frame movies were acquired on a Thermo Fisher Scientific Titan Krios G3 transmission EM (European Synchrotron Radiation Facility, Grenoble, France) (45) operated at 300 kV and equipped with a Gatan Quantum energy filter coupled to a Gatan K2 summit direct electron detector. Automated data collection was performed using Thermo Fisher Scientific EPU software, with a typical defocus range of  $-1.0$  to  $-3.0$   $\mu\text{m}$  and a total dose of 40 e<sup>-</sup>/Å<sup>2</sup> per movie. A nominal magnification of  $\times 105,000$  was used, resulting in a calibrated pixel size at the specimen level of 1.351 Å.

### EM image processing

Frame alignment was performed using Motioncor2 (46) keeping, respectively, frames 3 to 30 and 1 to 40 for Tip and Tip-FhuA and applying dose weighting. Contrast transfer function parameters were then determined using Gctf (47); manual particle picking was performed with EMAN2/e2helixboxer (48). The first picking coordinate was consistently centered on T5 collar and the second one a few hundred angstrom toward BHP<sub>pb3</sub>, along the central fiber (Tip) or the tail axis (Tip-FhuA) (extended data Fig. 2). This "vectorial" picking allowed us to choose and adapt the position of the box along that axis before extraction and proved to be very efficient. All subsequent image processing was performed using Relion (versions 3.0 and 3.1) (49). Flowchart of the EM processing pipeline is presented in extended data Fig. 2.

### Tip

After particle extraction (box size of 340 pixels by 340 pixels) centered 80 Å under the collar and two-dimensional (2D) classification, a homogeneous dataset of 9290 particles was obtained. No 3D classification was performed. Using a 15-Å resolution map determined from a previous cryo-EM data collection (50) as an initial model, a C3 reconstruction of the tip was calculated, from the second TTP<sub>pb6</sub> ring to the beginning of the central fiber. After masking and sharpening, the overall estimated resolution of the map reached 3.53 Å [Fourier Shell Correlation (FSC)<sub>0.143</sub>]. A new set of particles (box size of 400 pixels by 400 pixels) was extracted after a 150-pixel coordinate shift on the *z* axis, toward RBP<sub>pb5</sub>. A 15-Å low-pass-filtered initial model was generated from the newly extracted particles using relion\_reconstruct tool and determined a C3 reconstruction of the central fiber from BHP<sub>pb3</sub> to the beginning of

RBP<sub>pb5</sub>. After masking and sharpening, the overall estimated resolution of the map reached 4.2 Å (FSC<sub>0.143</sub>). An additional map of the full tip, of overall lower resolution (FSC<sub>0.143</sub> 3.88 Å), was also calculated to be able to fit the entire tip model but was not used for model building. Further image processing was necessary to obtain a map including the monomeric p143 protein; refined particles from the tip reconstruction were reextracted (box size of 200 pixels by 200 pixels) and recentered on the lower part of BHP<sub>pb3</sub>, on the side of which p143 is located. After reclassification/selection, symmetry relaxation, and a new run of 3D refinement using suitable masking, a nonsymmetrized map of the central part of the tip was obtained, where densities for the monomeric p143 are visible.

#### Tip-FhuA

After particle extraction (box size of 340 pixels by 340 pixels) and extensive 2D and 3D classifications, a homogeneous dataset of 20,349 particles was obtained. As an initial model, an 8-Å resolution map (low-pass-filtered at 15 Å) obtained from a previous cryo-EM data collection was used, and a reconstruction of the full nonsymmetrized tip after interaction with FhuA nanodisc (full tip FhuA) was calculated. After masking and sharpening, the overall estimated resolution of the map reached 4.3 Å (FSC<sub>0.143</sub>). Signal subtraction was then performed to enhance the resolution of specific parts of the structure. On the basis of the previously determined reconstruction, two soft masks were created, for the bent fiber only and for the C3 open tube. After reextraction and 3D reconstruction, the overall quality of these two areas greatly improved, with overall estimated resolution of 4.3 and 3.60 Å, respectively, for the bent fiber and the open tube.

#### Tip/Tip-FhuA common core

To improve the resolution of the tip common core (TTP<sub>pb6</sub>, p132, p140, and DTP<sub>pb9</sub>), and because we observed that it was invariant whether the tails were incubated with FhuA nanodisc, particles from all three datasets were merged. A soft mask was created using a 20-Å resolution model generated with Chimera tool molmap, using a previously built atomic model (see the "Protein model building" section below) containing two TTP<sub>pb6</sub> trimers, a p140 trimer, a DTP<sub>pb9</sub> hexamer, and a p132 dodecamer and used to perform signal subtraction on the merged particles. Subtracted particles were then refined to obtain a better C3 reconstruction for the tip common core, whose resolution reached an overall estimated value of 3.4 Å, allowing us to build slightly more accurate atomic models for TTP<sub>pb6</sub>, p140, DTP<sub>pb9</sub>, and p132 proteins. Efforts to specifically isolate and align the ill-defined RBP<sub>pb5</sub>/FhuA-RBP<sub>pb5</sub> parts of the maps did not result in any improvement, probably due to the small size of the protein and/or the low number of particles. For every reconstruction, a local resolution map was calculated using Relion built-in local resolution tool (extended data Fig. 1, C and H).

#### Protein model building

Atomic protein models were built into the different cryo-EM maps (table S4) using the Coot software (51) by tracing the protein sequence into the densities and were then iteratively refined alternating Coot manual refinement and PHENIX (52) real space refine tool until convergence. p140, p132, BHP<sub>pb3</sub>, and TMP<sub>pb2C</sub> models were built ab initio. For TTP<sub>pb6</sub> and DTP<sub>pb9</sub> models, existing x-ray models [5NGJ (11)/4JMQ (22)] were used as a starting point and were refined into the EM maps. MolProbity (53) was used for model quality assessment. The densities corresponding to the

BHP<sub>pb3</sub> β-hairpin legs (residues ~225 to 245) in the Tip-FhuA map are poorly resolved with regard to the rest of the protein. As a consequence, we only propose a likely model for BHP<sub>pb3</sub> β-hairpins, which should be considered with caution.

For p143, we used an AlphaFold2 (35)-predicted model as a starting point, which was fitted into the corresponding densities. Notably, AlphaFold2's level of confidence was not optimal throughout the whole sequence, which could explain the partial fit of the initial model. It was though coherent regarding the global shape and size. We then used a combination of Flex-EM (54)/Namdinator (55) (flexible fitting) and PHENIX (52) (real space refine), in an iterative way, to obtain a better model for this protein, with a convincing fit of most of its secondary structures.

#### Proteomics based on high-performance LC/ESI orbitrap

T5 tail proteins were stacked on the top of a 4 to 12% NuPAGE gel (Invitrogen) and stained with R-250 Coomassie blue. Gel bands were subjected to digestion using modified trypsin (Promega, sequencing grade) as previously described (56). Peptides were analyzed by online nano-LC tandem MS (LC-MS/MS) (UltiMate 3000 RSLCnano and QExact Plus, Thermo Fisher Scientific) with two replicates per sample. Peptides were sampled on a 300 μm by 5 mm PepMap C18 precolumn and separated on a 75 μm by 250 mm C18 column (PepMap, Dionex). The nanoLC method consisted of a 120-min gradient at a flow rate of 300 nl/min, ranging from 5 to 37% acetonitrile in 0.1% formic acid for 114 min, before reaching 72% acetonitrile in 0.1% formic acid for the last 6 min.

Spray voltage was set at 1.6 kV; heated capillary was adjusted to 270°C. Survey full-scan MS spectra [mass/charge ratio (*m/z*) = 400 to 1600] were acquired with a resolution of 70,000 after accumulation of 10<sup>6</sup> ions (maximum fill time, 200 ms). The 10 most intense ions were fragmented by higher-energy collisional dissociation after accumulation of 10<sup>5</sup> ions (maximum fill time, 50 ms). LC-MS/MS data files were processed using MaxQuant, version 1.5.1.2 (57). Spectra were searched against the UniProt database and frequently observed contaminants database. The minimum number of unique peptides is 1. Matching between runs option was activated. Proteins identified in the reverse and contaminant databases, or with less than two razor + unique peptides, or exhibiting less than six intensity based absolute quantification (iBAQ) values in one condition were discarded. After log<sub>2</sub> transformation, iBAQ values for the remaining proteins were normalized by condition-wise centering, missing values were imputed for each injected sample as the 2.5 percentile value, and statistical testing was conducted using Welch's *t* test.

#### Liquid chromatography/electrospray ionization mass spectrometry

To measure the accurate mass of the tail proteins, the T5 tails were diluted 2:3 in 0.1% trifluoroacetic acid (TFA) to obtain a final tail concentration of 0.20 to 0.22 μM and were analyzed using ESI-TOF mass spectrometer (6210 instrument, Agilent Technologies) coupled to an LC system (1100 series, Agilent Technologies). The instrument was calibrated using tuning mix (ESI-L, Agilent Technologies). The following instrumental settings were used: gas (nitrogen) temperature, 300°C; drying gas (nitrogen), 7 liters min<sup>-1</sup>; nebulizer gas (nitrogen), 10 psig; Capillary voltage, 4 kV; fragmentor, 250 V; skimmer, 60 V; and peak-to-peak voltage (octopole radio frequency), 250 V. The high-performance LC (HPLC) mobile

phases were prepared with HPLC-grade solvents. Mobile phase A composition was 5% acetonitrile (ACN) and 0.03% TFA. Mobile phase B composition was 95% ACN and 0.03% TFA.

Eight microliters of each sample (1.6 to 1.8 pmol) was first desalted on-line for 3 min with 100% of mobile phase A (flow rate of 50  $\mu\text{l}/\text{min}$ ) using a C8 reverse phase micro-column (Zorbax 300SB-C8, 5  $\mu\text{m}$ , 5 mm by 0.3 mm, Agilent Technologies). The sample was then eluted with 70% of mobile phase B (flow rate of 50  $\mu\text{l}/\text{min}$ ), and MS spectra were acquired in the positive ion mode in the 300 to 3000  $m/z$  range. Data were processed with MassHunter software (v. B.02.00, Agilent Technologies) and GPMW software (v. 7.00b2, Lighthouse Data, Denmark).

## Supplementary Materials

### This PDF file includes:

Figs. S1 to S8

Tables S1 to S4

Legends for movies S1 to S4

References

### Other Supplementary Material for this manuscript includes the following:

Movies S1 to S4

## REFERENCES AND NOTES

- C. A. Suttle, Marine viruses—major players in the global ecosystem. *Nat. Rev. Microbiol.* **5**, 801–812 (2007).
- S. Uytendaele, B. Chen, J. Onsea, F. Ruythooren, Y. Debaveye, D. Devolder, I. Spriet, M. Depytere, J. Wagemans, R. Lavigne, J.-P. Pirnay, M. Merabishvili, P. De Munter, W. E. Peetermans, L. Dupont, L. Van Gerven, W.-J. Metsemakers, Safety and efficacy of phage therapy in difficult-to-treat infections: A systematic review. *Lancet Infect. Dis.* **22**, e208–e220 (2022).
- D. Veesler, C. Cambillau, A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. *Microbiol. Mol. Biol. Rev.* **75**, 423–433 (2011).
- A. R. Davidson, L. Cardarelli, L. G. Pell, D. R. Radford, K. L. Maxwell, Long noncontractile tail machines of bacteriophages. *Adv. Exp. Med. Biol.* **726**, 115–142 (2012).
- M. Brackmann, S. Nazarov, J. Wang, M. Basler, Using force to punch holes: Mechanics of contractile nanomachines. *Trends Cell Biol.* **27**, 623–632 (2017).
- R. Linares, C.-A. Arnaud, S. Degroux, G. Schoehn, C. Breyton, Structure, function and assembly of the long, flexible tail of siphophages. *Curr. Opin. Virol.* **45**, 34–42 (2020).
- A. Desfosses, H. Venugopal, T. Joshi, J. Felix, M. Jessop, H. Jeong, J. Hyun, J. B. Heymann, M. R. H. Hurst, I. Gutsche, A. K. Mitra, Atomic structures of an entire contractile injection system in both the extended and contracted states. *Nat. Microbiol.* **4**, 1885–1894 (2019).
- F. Jiang, N. Li, X. Wang, J. Cheng, Y. Huang, Y. Yang, J. Yang, B. Cai, Y.-P. Wang, Q. Jin, N. Gao, Cryo-EM structure and assembly of an extracellular contractile injection system. *Cell* **177**, 370–383.e15 (2019).
- P. Ge, D. Scholl, N. S. Prokhorov, J. Avaylon, M. M. Shneider, C. Browning, S. A. Buth, M. Plattner, U. Chakraborty, K. Ding, P. G. Leiman, J. F. Miller, Z. H. Zhou, Action of a minimal contractile bactericidal nanomachine. *Nature* **580**, 658–662 (2020).
- H.-W. Ackermann, Phage classification and characterization. *Methods Mol. Biol.* **501**, 127–140 (2009).
- C.-A. Arnaud, G. Effantin, C. Vivès, S. Engilbergh, M. Bacia, P. Boulanger, E. Girard, G. Schoehn, C. Breyton, Bacteriophage T5 tail tube structure suggests a trigger mechanism for *Siphoviridae* DNA ejection. *Nat. Commun.* **8**, 1953 (2017).
- P. L. Campbell, R. L. Duda, J. Nassur, J. F. Conway, A. Huet, Mobile loops and electrostatic interactions maintain the flexible tail tube of bacteriophage lambda. *J. Mol. Biol.* **432**, 384–395 (2020).
- A. Goulet, J. Lai-Kee-Him, D. Veesler, I. Auzat, G. Robin, D. A. Shepherd, A. E. Ashcroft, E. Richard, J. Lichière, P. Tavares, C. Cambillau, P. Bron, The opening of the SPP1 bacteriophage tail, a prevalent mechanism in gram-positive-infecting Siphophages. *J. Biol. Chem.* **286**, 25397–25405 (2011).
- D. Veesler, S. Spinelli, J. Mahony, J. Lichière, S. Blangy, G. Bricogne, P. Legrand, M. Ortiz-Lombardia, V. Campanacci, D. van Sinderen, C. Cambillau, Structure of the phage TP901-1 1.8 MDa baseplate suggests an alternative host adhesion mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8954–8958 (2012).
- G. Sciarra, C. Bebeacua, P. Bron, D. Tremblay, M. Ortiz-Lombardia, J. Lichière, M. van Heel, V. Campanacci, S. Moineau, C. Cambillau, Structure of lactococcal phage p2 baseplate and its mechanism of activation. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6852–6857 (2010).
- J. L. Kizziah, K. A. Manning, A. D. Dearborn, T. Dokland, Structure of the host cell recognition and penetration machinery of a *Staphylococcus aureus* bacteriophage. *PLOS Pathog.* **16**, e1008314 (2020).
- Y. Zivanovic, F. Confalonieri, L. Ponchon, R. Lurz, M. Chami, A. Flayhan, M. Renouard, A. Huet, P. Decottignies, A. R. Davidson, C. Breyton, P. Boulanger, Insights into bacteriophage T5 structure from analysis of its morphogenesis genes and protein components. *J. Virol.* **88**, 1162–1174 (2014).
- M. Demerec, U. Fano, Bacteriophage-resistant mutants in *Escherichia coli*. *Genetics* **30**, 119–136 (1945).
- A. Huet, R. L. Duda, P. Boulanger, J. F. Conway, Capsid expansion of bacteriophage T5 revealed by high resolution cryoelectron microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 21037–21046 (2019).
- P. Boulanger, P. Jacquot, L. Plançon, M. Chami, A. Engel, C. Parquet, C. Herbeuval, L. Letellier, Phage T5 straight tail fiber is a multifunctional protein acting as a tape measure and carrying fusogenic and muralytic activities. *J. Biol. Chem.* **283**, 13556–13564 (2008).
- C. Garcia-Doval, J. R. Castón, D. Luque, M. Granell, J. M. Otero, A. L. Llamas-Saiz, M. Renouard, P. Boulanger, M. J. van Raaij, Structure of the receptor-binding carboxy-terminal domain of the bacteriophage T5 L-shaped tail fiber with and without its intramolecular chaperone. *Viruses* **7**, 6424–6440 (2015).
- A. Flayhan, F. M. D. Vellieux, R. Lurz, O. Maury, C. Contreras-Martel, E. Girard, P. Boulanger, C. Breyton, Crystal Structure of pb9, the Distal tail protein of bacteriophage T5: A conserved structural motif among all siphophages. *J. Virol.* **88**, 820–828 (2014).
- A. Flayhan, F. Wien, M. Paternostre, P. Boulanger, C. Breyton, New insights into pb5, the receptor binding protein of bacteriophage T5, and its interaction with its *Escherichia coli* receptor FhuA. *Biochimie* **94**, 1982–1989 (2012).
- C. Breyton, A. Flayhan, F. Gabel, M. Lethier, G. Durand, P. Boulanger, M. Chami, C. Ebel, Assessing the conformational changes of pb5, the receptor-binding protein of phage T5, upon binding to its *Escherichia coli* receptor FhuA. *J. Biol. Chem.* **288**, 30763–30772 (2013).
- V. Braun, FhuA (TonA), the career of a protein. *J. Bacteriol.* **191**, 3431–3436 (2009).
- M. Bonhivers, A. Ghazi, P. Boulanger, L. Letellier, FhuA, a transporter of the *Escherichia coli* outer membrane, is converted into a channel upon binding of bacteriophage T5. *EMBO J.* **15**, 1850–1856 (1996).
- N. Chiaruttini, M. de Frutos, E. Augarde, P. Boulanger, L. Letellier, V. Viasnoff, Is the in vitro ejection of bacteriophage DNA quasistatic? A bulk to single virus study. *Biophys. J.* **99**, 447–455 (2010).
- I. G. Denisov, S. G. Sligar, Nanodiscs in membrane biochemistry and biophysics. *Chem. Rev.* **117**, 4669–4713 (2017).
- S. Degroux, G. Effantin, R. Linares, G. Schoehn, C. Breyton, Deciphering bacteriophage T5 host recognition mechanism and infection trigger. *J. Virol.* **2023**, e01584–22 (2023).
- H. Fraga, C.-A. Arnaud, D. F. Gauto, M. Audin, V. Kurauskas, P. Macek, C. Krichel, J.-Y. Guan, J. Boisbouvier, R. Sprangers, C. Breyton, P. Schanda, Solid-state NMR H-N-(C)-H and H-N-C-C 3D/4D correlation experiments for resonance assignment of large proteins. *ChemPhysChem* **18**, 2697–2703 (2017).
- L. Holm, DALI and the persistence of protein shape. *Protein Sci.* **29**, 128–140 (2020).
- N. M. I. Taylor, N. S. Prokhorov, R. C. Guerrero-Ferreira, M. M. Shneider, C. Browning, K. N. Goldie, H. Stahlberg, P. G. Leiman, Structure of the T4 baseplate and its function in triggering sheath contraction. *Nature* **533**, 346–352 (2016).
- E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
- M. Noirdlerc-Savoie, A. Flayhan, C. Pereira, B. Gallet, P. Gans, C. Ebel, C. Breyton, Tail proteins of phage T5: Investigation of the effect of the His6-tag position, from expression to crystallisation. *Protein Expr. Purif.* **109**, 70–78 (2015).
- J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Židek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
- M. Zweig, D. J. Cummings, Cleavage of head and tail proteins during bacteriophage T5 assembly: Selective host involvement in the cleavage of a tail protein. *J. Mol. Biol.* **80**, 505–518 (1973).
- L. C. Tsui, R. W. Hendrix, Proteolytic processing of phage lambda tail protein gpH: Timing of the cleavage. *Virology* **125**, 257–264 (1983).

38. S. Kanamaru, P. G. Leiman, V. A. Kostyuchenko, P. R. Chipman, V. V. Mesyanzhinov, F. Arisaka, M. G. Rossmann, Structure of the cell-puncturing device of bacteriophage T4. *Nature* **415**, 553–557 (2002).
39. S. R. Casjens, R. W. Hendrix, Locations and amounts of major structural proteins in bacteriophage lambda. *J. Mol. Biol.* **88**, 535–545 (1974).
40. L. Cardarelli, L. G. Pell, P. Neudecker, N. Pirani, A. Liu, L. A. Baker, J. L. Rubinstein, K. L. Maxwell, A. R. Davidson, Phages have adapted the same protein fold to fulfill multiple functions in virion assembly. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 14384–14389 (2010).
41. P. Bárdy, T. Füzik, D. Hrebík, R. Pantůček, J. Thomas Beatty, P. Plevka, Structure and mechanism of DNA delivery of a gene transfer agent. *Nat. Commun.* **11**, 3034 (2020).
42. L. T. Alexander, R. Lepore, A. Kryshatafovych, A. Adamopoulos, M. Alahuhta, A. M. Arvin, Y. J. Bomble, B. Böttcher, C. Breyton, V. Chiarini, N. B. Chinnam, W. Chiu, K. Fidelis, R. Grinter, G. D. Gupta, M. D. Hartmann, C. S. Hayes, T. Heidebrecht, A. Ilari, A. Joachimiak, Y. Kim, R. Linares, A. L. Lovering, V. V. Lunin, A. N. Lupas, C. Makbul, K. Michalska, J. Moulit, P. K. Mukherjee, W. S. Nutt, S. L. Oliver, A. Perrakis, L. Stols, J. A. Tainer, M. Topf, S. E. Tsutakawa, M. Valdivia-Delgado, T. Schwede, Target highlights in CASP14: Analysis of models by structure providers. *Proteins* **89**, 1647–1672 (2021).
43. G. Guihard, P. Boulanger, L. Letellier, Involvement of phage T5 tail proteins and contact sites between the outer and inner membrane of Escherichia coli in phage T5 DNA injection. *J. Biol. Chem.* **267**, 3173–3178 (1992).
44. M. Zweig, D. J. Cummings, Structural proteins of bacteriophage T5. *Virology* **51**, 443–453 (1973).
45. E. Kandiah, T. Giraud, A. de Maria Antolinos, F. Dobias, G. Effantin, D. Flot, M. Hons, G. Schoehn, J. Susini, O. Svensson, G. A. Leonard, C. Mueller-Dieckmann, CM01: A facility for cryo-electron microscopy at the European Synchrotron. *Acta Crystallogr. D Struct. Biol.* **75**, 528–535 (2019).
46. X. Li, P. Mooney, S. Zheng, C. R. Booth, M. B. Braunfeld, S. Gubbens, D. A. Agard, Y. Cheng, Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584–590 (2013).
47. K. Zhang, Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
48. G. Tang, L. Peng, P. R. Baldwin, D. S. Mann, W. Jiang, I. Rees, S. J. Ludtke, EMAN2: An extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
49. J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. Hagen, E. Lindahl, S. H. Scheres, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
50. C.-A. Arnaud, thesis, University Grenoble-Alpes (2017).
51. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
52. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
53. C. J. Williams, J. J. Headd, N. W. Moriarty, M. G. Prisant, L. L. Videau, L. N. Deis, V. Verma, D. A. Keedy, B. J. Hintze, V. B. Chen, S. Jain, S. M. Lewis, W. B. Arendall III, J. Snoeyink, P. D. Adams, S. C. Lovell, J. S. Richardson, D. C. Richardson, MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
54. A. P. Pandurangan, M. Topf, RIBFIND: A web server for identifying rigid bodies in protein structures and to aid flexible fitting into cryo EM maps. *Bioinformatics* **28**, 2391–2393 (2012).
55. R. T. Kidmose, J. Juhl, P. Nissen, T. Boesen, J. L. Karlsen, B. P. Pedersen, Namdinator - automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *IUCr* **6**, 526–531 (2019).
56. M. G. Casabona, Y. Vandenbrouck, I. Attree, Y. Couté, Proteomic characterization of Pseudomonas aeruginosa PAO1 inner membrane. *Proteomics* **13**, 2419–2423 (2013).
57. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
58. H. Delattre, O. Souiai, K. Fagoonee, R. Guerois, M.-A. Petit, Phagonaut: A web-based interface for phage synteny browsing and protein function prediction. *Virology* **496**, 42–50 (2016).

**Acknowledgments:** We acknowledge the European Synchrotron Radiation Facility for provision of beam time on CM01. This work used the platforms of the Grenoble Instruct-ERIC center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI and GRAL, financed within the University Grenoble Alpes Graduate School (Ecoles Universitaires de Recherche) CBH-EUR-GS. The IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA). Proteomic experiments were partially supported by the French National Research Agency under projects ProFI (Proteomics French Infrastructure, ANR-10-INBS-08). We thank L. Signor, E. Neumann, and D. Fenel for technical assistance; Y. H. de Kermadec for help with nanodisc optimization; L. Estrozi, A. Desfosses, and B. Arragain for help and discussion on image analysis; A. Peuch for help with the EM computing cluster; C. Masselon, K. Clement, and the EdyP MS platform for MS analysis; and C. Ebel for useful discussions. **Funding:** This work was funded by the Agence Nationale pour la Recherche ANR-16-CE11-0027 (C.B. and G.S.), Agence Nationale pour la Recherche ANR-21-CE11-0023 (C.B. and G.S.), Labex Grenoble Alliance GRAL ANR-10-LABX-49-01 (C.B.), CEA grant DRF-Impulsion T5-MS (C.B. and E.B.E.), FRISBI ANR-10-INBS-05-02, University Grenoble Alpes Ecoles Universitaires de Recherche CBH-EUR-GS ANR-17-EURE-0003, Auvergne-Rhône-Alpes Region, Fontation pour la Recherche Médicale FRM DGE2012112563, Fonds FEDER (46475), and GIS-Infrastructures en Biologie Santé et Agronomie (IBISA). **Author contributions:** Conceptualization: C.B. T5 tails and FhuA nanodisc preparation: C.-A.A., R.L., C.D., and C.B. Limited proteolysis experiments: C.D., R.L., and C.-A.A. MS data generation and analysis: N.H.E. and E.B.E. Cryo-grid optimization: G.E. and R.L. Cryo-EM data recording: G.E. Cryo-EM data processing: R.L., C.-A.A., and G.E. Atomic model building: R.L. Data analysis: C.B. with help of C.-A.A. Funding acquisition: C.B. and G.S. Writing—original draft: C.B. Writing—review and editing: C.B., R.L., G.E., C.-A.A., E.B.E., and G.S. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Cryo-EM density maps of T5 tip resolved in this study and the associated atomic coordinates have been, respectively, deposited in the Electron Microscopy Data Bank (EMDB) and in the PDB under the following accession codes: EMD-13953/PDB 7QG9 (Tip/Tip-FhuA common core), EMD-14733/PDB 7ZHJ (Tip without fiber), EMD-14790/PDB 7ZLV (Tip fiber), EMD-14869/PDB 7ZQB (Tip full), EMD-14799/PDB 7ZN2 (Tip-FhuA full), EMD-14800/PDB 7ZN4 (Tip-FhuA bent fiber), and EMD-14873/PDB 7ZQP (Tip-FhuA open tube). See also table S1. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 20 September 2022

Accepted 21 February 2023

Published 24 March 2023

10.1126/sciadv.ade9674

Supplementary Materials for  
**Structural basis of bacteriophage T5 infection trigger and *E. coli* cell wall perforation**

Romain Linares *et al.*

Corresponding author: Cécile Breyton, [cecile.breyton@ibs.fr](mailto:cecile.breyton@ibs.fr)

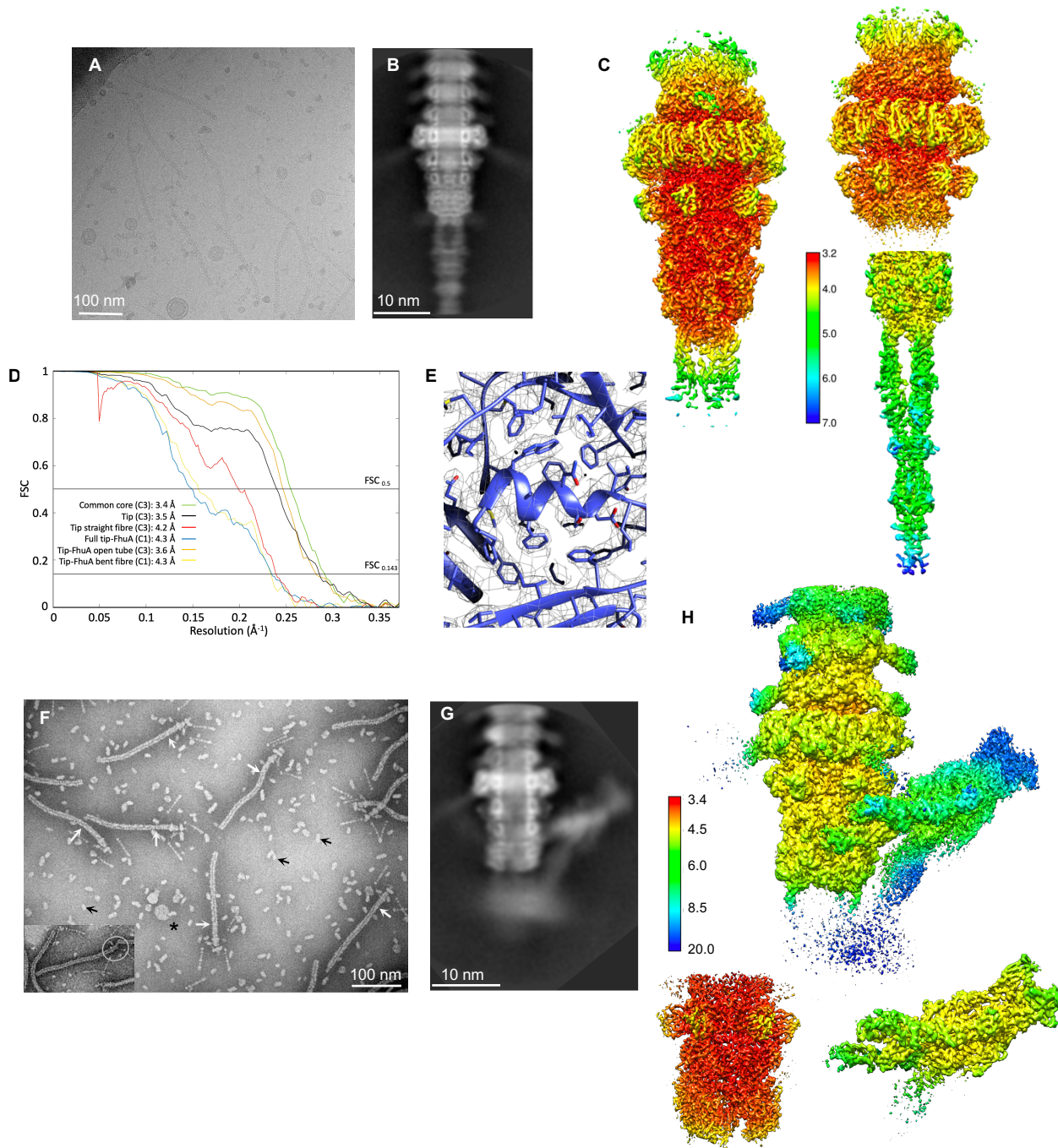
*Sci. Adv.* **9**, eade9674 (2023)  
DOI: 10.1126/sciadv.ade9674

**The PDF file includes:**

Figs. S1 to S8  
Tables S1 to S4  
Legends for movies S1 to S4  
References

**Other Supplementary Material for this manuscript includes the following:**

Movies S1 to S4



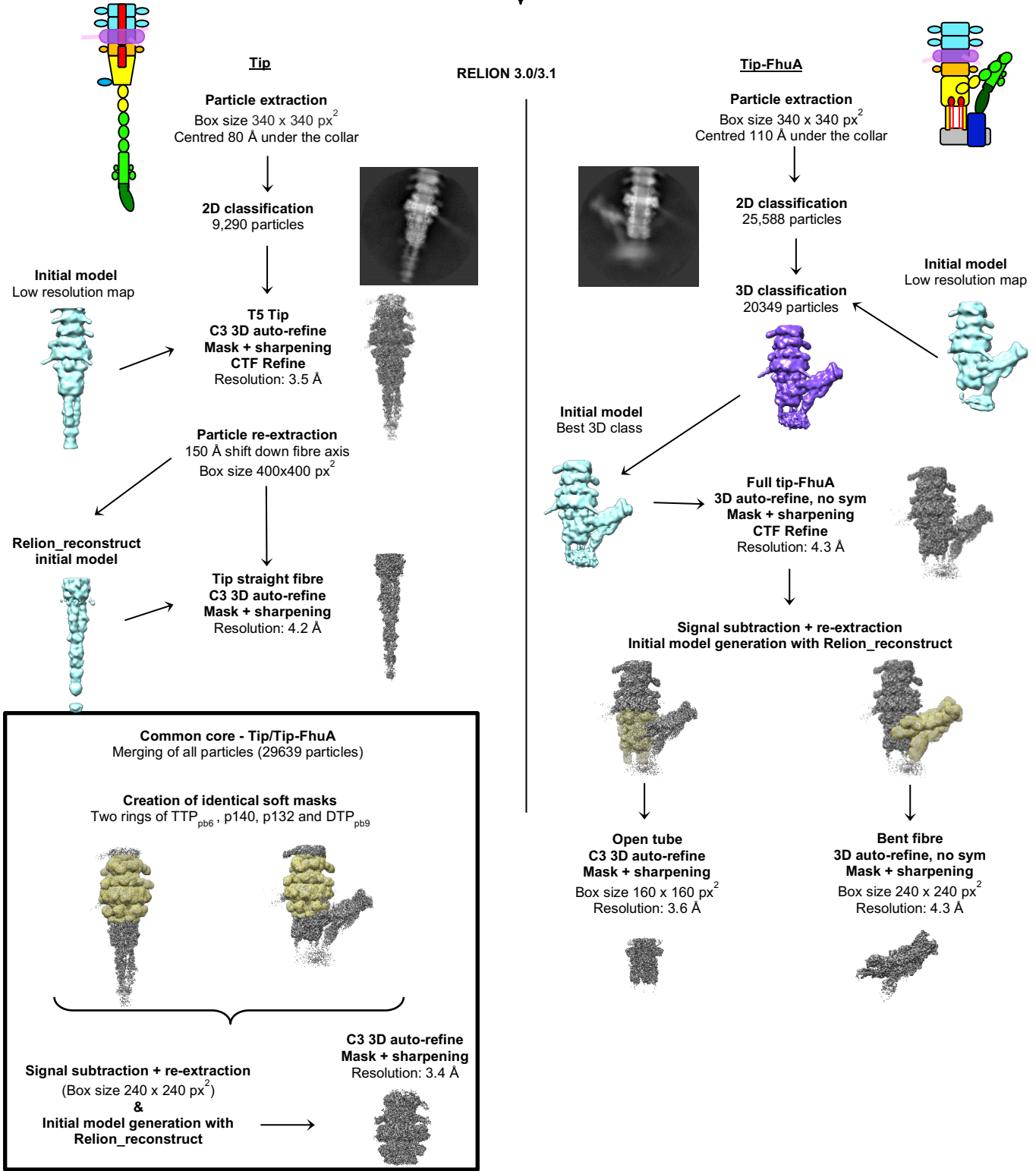
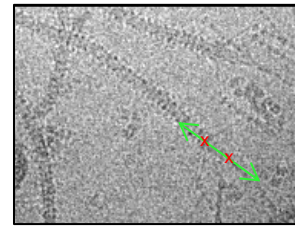
**Figure S1: Structure of T5 tail tip before and after interaction with FhuA in nanodisc.** (A) Cryo-EM image of T5 tails. (B) A 2D class average of T5 tail tip. (C) Local resolution maps, as determined by Relion of the tip (left), the Tip/Tip-FhuA common core (top right, from TTP<sub>pb6</sub> to the BHP<sub>pb3</sub> top part, map obtained gathering particles from both Tip and Tip-FhuA datasets) and of the tip straight central fibre (bottom right). All maps were calculated with a C3 symmetry and the key (in Å) is the same for the three maps. Resolution is the highest at the centre (3.2 Å) and falls off radially to ~7 Å at the tip of TTP<sub>pb6</sub> Ig-like domain. Because of the flexibility of the central fibre and of the tube, resolution also drops rapidly along the tube, above the first TTP<sub>pb6</sub> ring and downwards along the central fibre. (D) Fourier shell correlation plot for the six maps presented in C and H. FSC<sub>0.5</sub> and FSC<sub>0.143</sub> cutoffs are indicated, as well as the estimated resolution (FSC<sub>0.143</sub>) for each map. (E) Close up view of an  $\alpha$ -helix of p140 monomer model, built using Tip/Tip-FhuA common core map. (F) Large field, negative stain EM image of T5 tails incubated with FhuA-nanodiscs. The background is filled with FhuA-nanodiscs, mainly seen lying on the side (black arrows) but also from the top (\*). T5 tails are partially emptied from TMP<sub>pb2</sub>. White arrows point to the empty/filled limit in the tails. Inset: T5 tails incubated with detergent-solubilised FhuA. In those conditions, TMP<sub>pb2</sub> has been completely ejected and the central fibre is not visible. (G) A 2D class average of Tip-FhuA. (H) Local resolution maps, as determined by Relion, of the C1 full Tip-FhuA (top), C1 bent fibre (BHP<sub>pb3</sub> FNIII, pb4, bottom right) and C3 open tube (DTP<sub>pb9</sub>, BHP<sub>pb3</sub> ~700 N-terminal residues, TMP<sub>pb2</sub>\* 43 C-terminal residues, bottom left). The key (in Å) is the same for the three maps.

40-frame movies acquired on a Titan Krios with Gatan K2 Summit direct electron detector + Quantum energy filter  
Magnification 105,000x => Pixel size : 1.351 Å  
3,208/9,608 movies for Tip/Tip-FhuA

Frame alignment with Motioncor2  
Frames 3-30/1-40 for Tip/Tip-FhuA  
Dose weighting

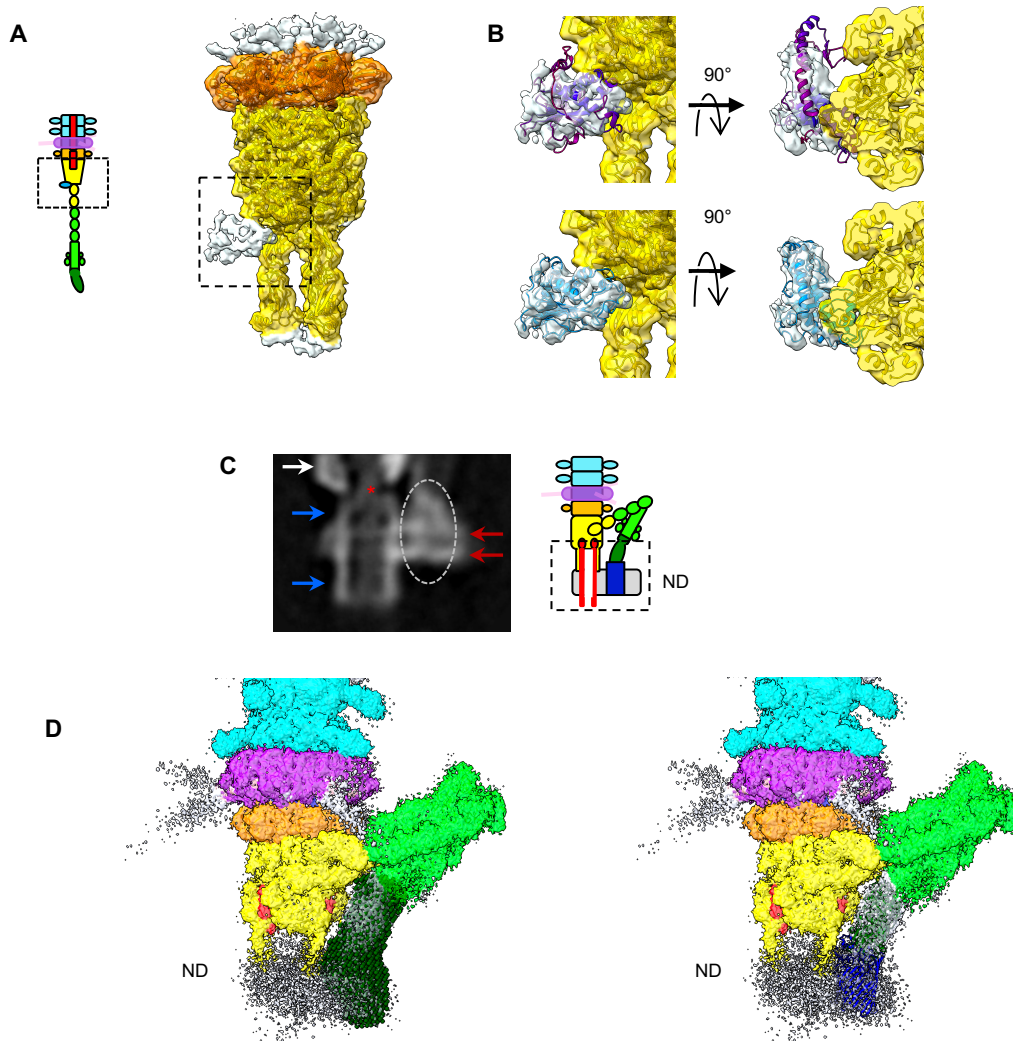
CTF estimation with Gctf

Manual picking with EMAN2/e2helixboxer  
Coordinate 1 : on the collar  
Coordinate 2 : lower on the fibre

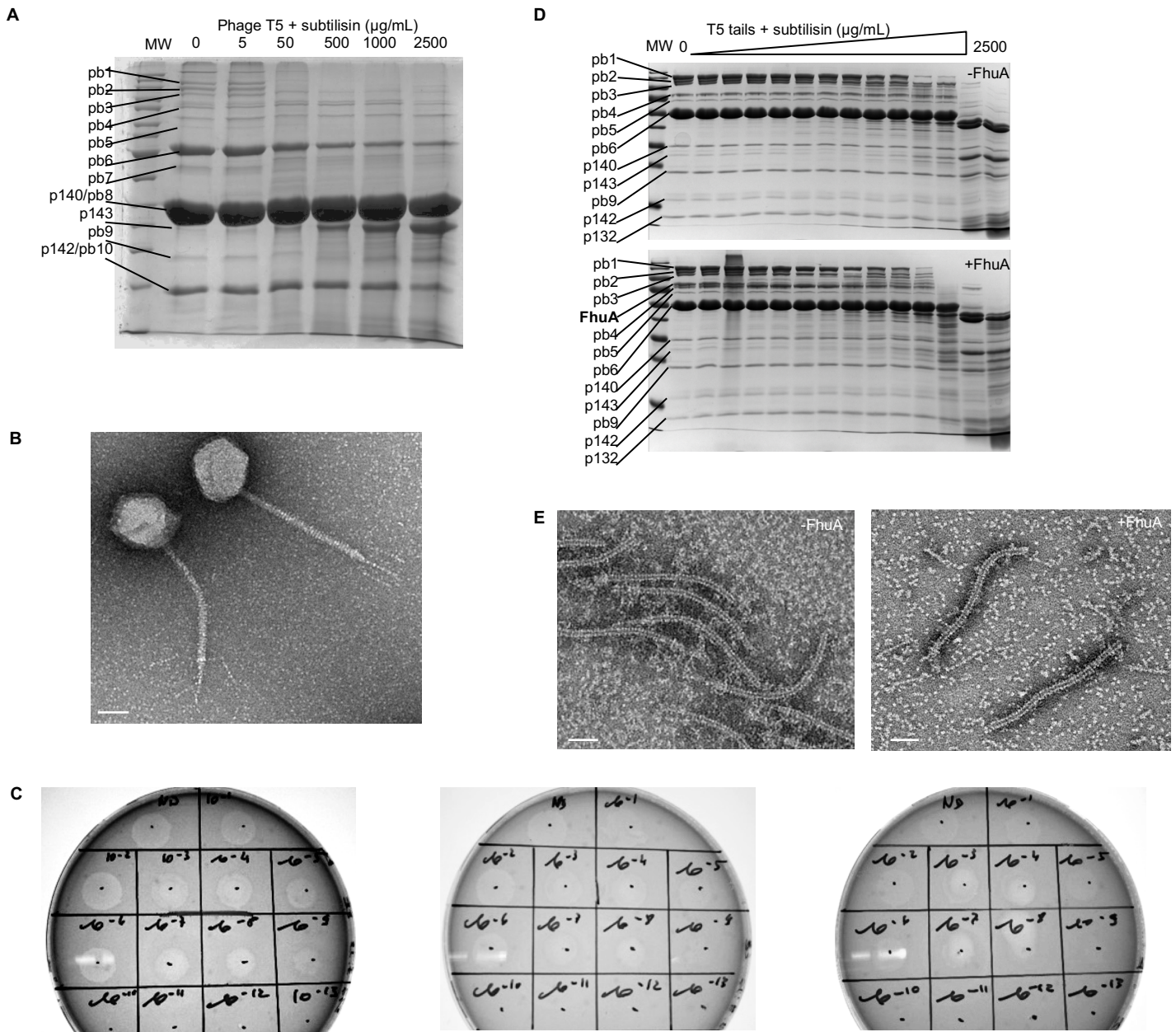


**Figure S2: Flowchart of the EM processing pipeline for T5 Tip and Tip-FhuA.** Common steps for the Tip/Tip-FhuA datasets are framed in blue. Initial models are in blue, 3D classes in purple, soft masks in yellow and 3D refine in grey. Top right: vectorial picking of the tip particles on a cryo-EM image of a T5 tail tip.





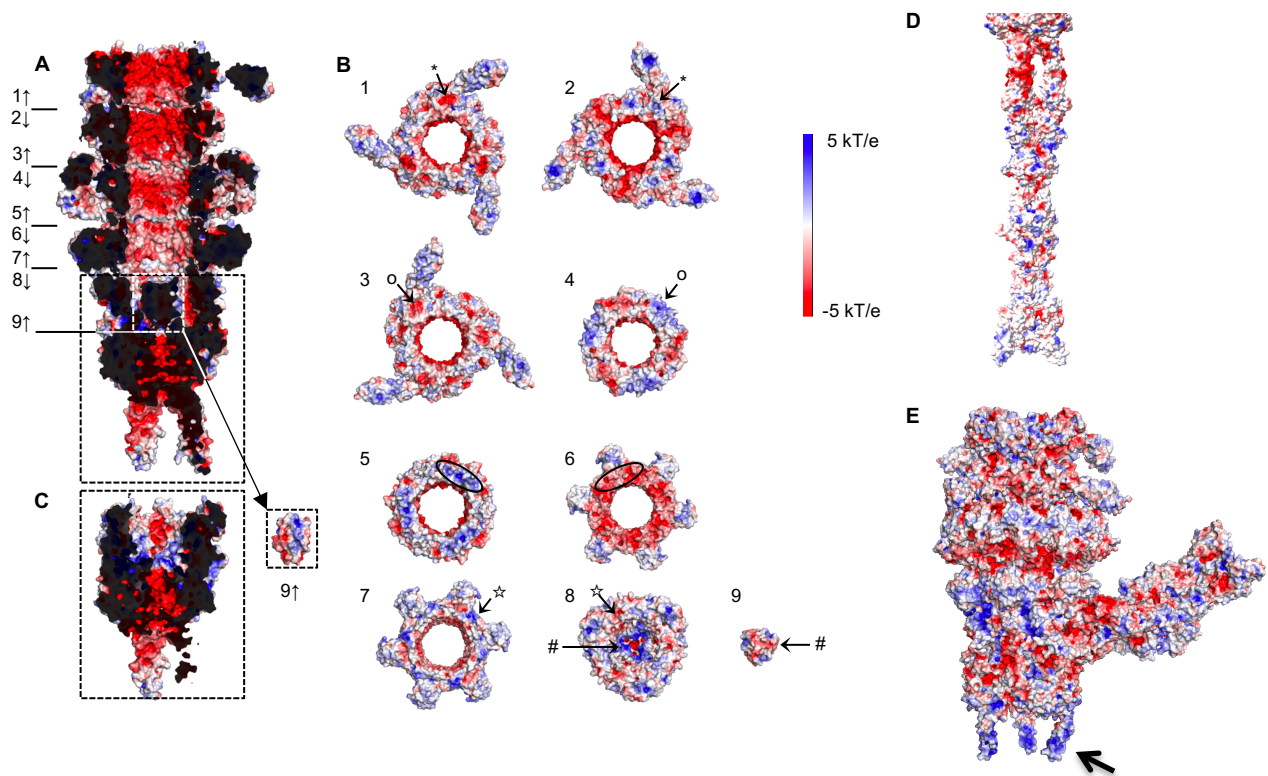
**Figure S3: Fit of p143, FhuA and RBP<sub>pb5</sub> in unattributed densities.** (A) Isosurface view of an unsymmetrised cryo-EM map of the tip (50% transparency), with models for the BHP<sub>pb3</sub> trimer (yellow) and the DTP<sub>pb9</sub> hexamer (orange) fitted in it. Additional densities at the base of BHP<sub>pb3</sub>, corresponding to a monomeric protein, are clearly visible (dotted box). (B) Enlargement of the dotted box in A, with a fit of the AlphaFold2 predicted structure for p143 before (top, coloured according to the pLDDT (high: blue, low: magenta)) and after (bottom, light blue model) flexible fitting. Left: side view, right: top view from a slice. (C) 2D Slice of a low resolution EM reconstruction from Tip-FhuA showing the bottom of BHP<sub>pb3</sub> (white arrow), the nanodisc bilayer (red arrows) and what appears to be a cylindrical structure connected to BHP<sub>pb3</sub> and spanning the nanodisc (blue arrows), probably the channel or part of it. Densities corresponding to TMP<sub>pb2\*</sub> can also be seen inside the tube/channel (red asterisk). The position of FhuA is marked by a white dotted oval. (D) Left: Fit of a SANS envelop of the FhuA-pb5 complex (24)(dark green beads) into the densities prolonging pb4 spike and merging into the nanodisc. Right: Fit of FhuA-RBP<sub>pb5</sub> (PDB 8B14) in the same map (29)(FhuA: blue and RBP<sub>pb5</sub> green ribbon). The map is that of Tip-FhuA C1 filtered and masked reconstruction, coloured as in figure 1 and 50% transparent. Unattributed densities are white. ND indicates the position of the nanodisc.



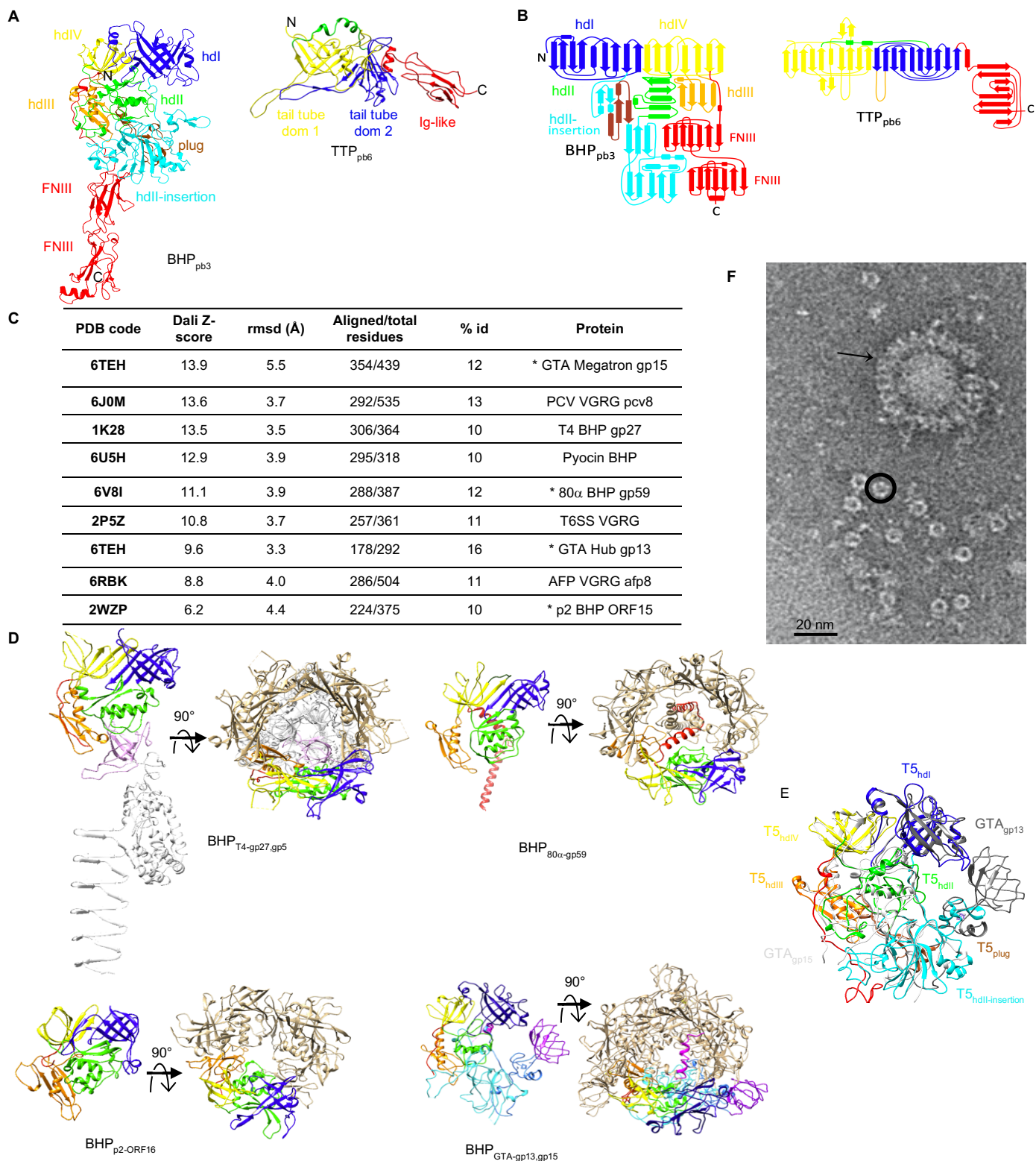
**Figure S4: Limited proteolysis of phage T5 (A, B, C) and T5 tails (D, E).** (A) Phage T5 was incubated with subtilisin (0, 5, 50, 500, 1000, 2500  $\mu\text{g/mL}$ ) for 1h at room temperature. The reaction was stopped by the addition of 10 mM PMSF, and phage ghosts (17) were prepared for migration on 12% SDS-PAGE. Molecular weight markers: 200, 150, 120, 100, 85, 70, 60, 50, 40, 30, 25, 20, 15, 10 kDa. (B) Negative stain EM image of phage T5 after 1h incubation with 2.5 mg/mL subtilisin. (C) Titration of unproteolysed T5 (left), T5 incubated 1h with 50  $\mu\text{g/mL}$  (middle) and 2.5 mg/mL (right) subtilisin on a lawn of *E. coli* strain F in soft Agar. (D) Upper panel: Purified T5 tails were incubated with subtilisin (0, 0.1, 0.3, 0.6, 1, 2, 3, 5, 7.5, 10, 20, 30, 1000, 2500  $\mu\text{g/mL}$ ). Lower panel: T5 tails were pre-incubated with LDAO-solubilised FhuA for 1h at room temperature (final LDAO concentration 0.05%), before being incubated with subtilisin. Molecular weight markers: 180, 130, 100, 70, 55, 40, 35, 25, 15, 10 kDa. (E) Negative stain EM images of T5 tails (right) and T5 tails incubated with FhuA (left) incubated with 2.5 mg/mL subtilisin. Subtilisin is clearly visible in the background. Bar: 50 nm.



summarising the pairwise superpositions of the four p132 subunits of the asymmetric unit. In the diagonal in grey is the number of residues built for each subunit, above the diagonal the rmsd in Å and below the DALI Z-score. DALI search links the p132 fold to the N-terminal domain of the Baseplate Protein Upper (BppU, ORF48) of phage TP901-1 with a Z-score of 5.9 and a rmsd of 4.0 Å over 99 residues and 9% identity. **(G)** Sequence analysis of the TMP<sub>pb2</sub> 259 C-terminal residues. Bold: predicted coiled-coil region from COILS; underlined: predicted transmembrane domain from PSI-PRED and TMHMM; green: peptides identified in proteomics; blue: Zinc carboxypeptidase motif, with conserved residues from Prosite analysis indicated with a star; purple arrow: cleavage site, separating TMP<sub>pb2\*</sub> and TMP<sub>pb2C</sub>. The red box corresponds to TMP<sub>pb2C</sub> 35 modelled residues from the Tip map, those pointing towards the centre of the coil are underlined. The magenta box correspond to TMP<sub>pb2\*</sub> C-terminus 42 modelled residues in Tip-FhuA map. **(H)** Overlay of T5 TMP<sub>pb2C</sub> (red), TMP<sub>80a</sub> (green) and the N-terminal helix of GTA Megatron (yellow) when superimposing the BHP trimers of each baseplate. Sequence alignment of the aligned helices, with the hydrophobic residues pointing towards the centre of the coil underlined.

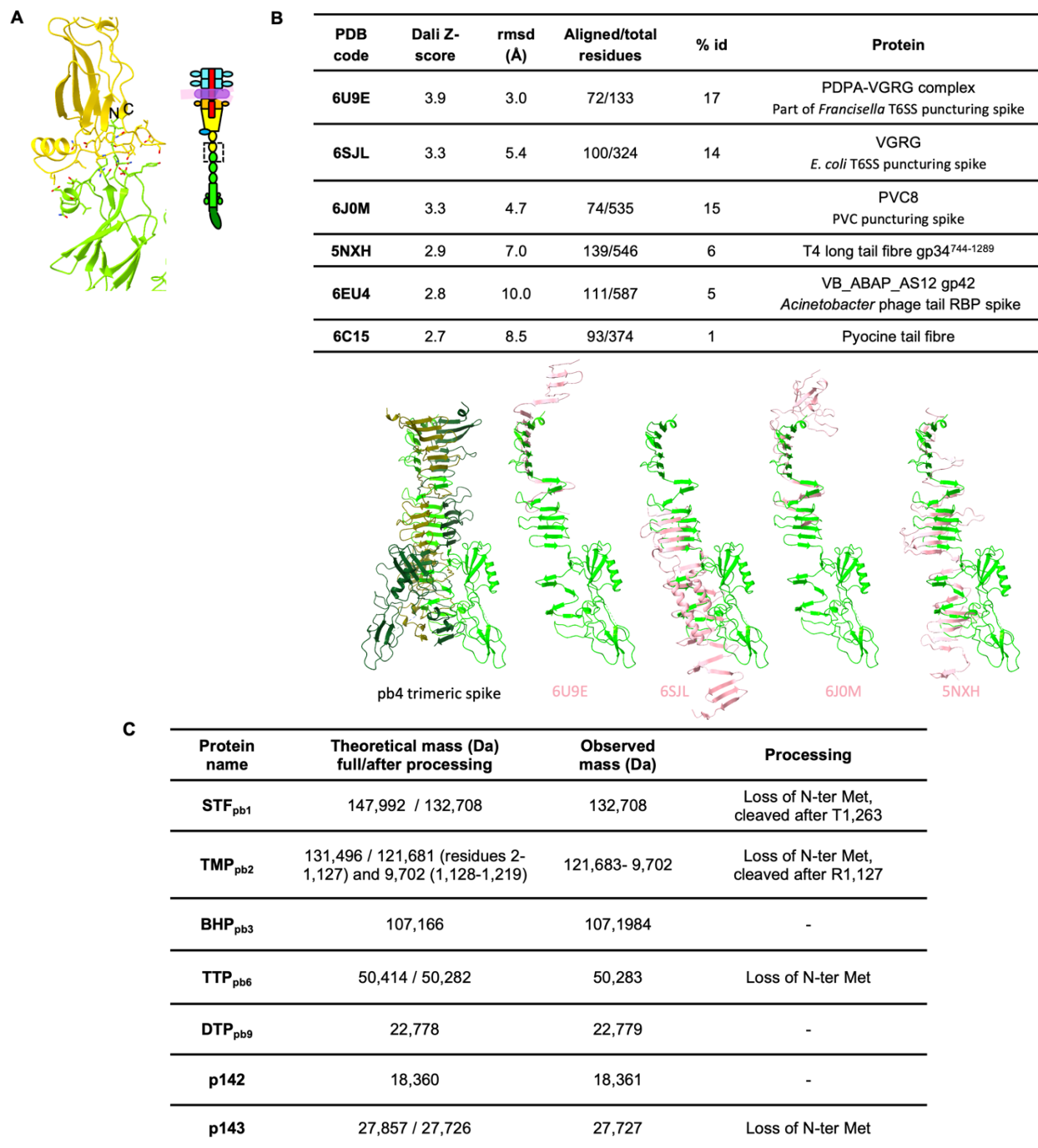


**Figure S6: Electrostatic charge distribution.** (A) Interior of T5 tail tube, (B) complementary interfaces of the different rings, (C) closed BHP<sub>pb3</sub> in which TMP<sub>pb2c</sub> has been removed *in silico* and is shown on the right side, (D) tail tip before and after interaction with FhuA. The arrow points to one of the highly charged three  $\beta$ -hairpins 'legs'. The position of the views in B is indicated by a number in A, and complementary charge patches are noted by symbols. Electrostatic charge distribution was calculated from the APBS plugin of either PyMol or ChimeraX.



**Figure S7: Structural analysis of BHP<sub>pb3</sub>.** (A) Ribbon representation in the same side view orientation of TTP<sub>pb6</sub> (left) and BHP<sub>pb3</sub> (right) and (B) topology diagram of the two proteins. BHP<sub>pb3</sub> hdl-VI are coloured blue, green, orange and yellow, respectively, the hdlII insertion in cyan and the two FNIII in red. The same colour code was used for structurally homologous domain in TTP<sub>pb6</sub>. The link between the two tail tube domains in BHP<sub>pb3</sub> and TTP<sub>pb6</sub> is topologically different in the two proteins: it connects the C-terminus of the first domain to the N-terminus of the

second one in TTP<sub>pb6</sub>, and the N-terminus the first domain to the C-terminus of the second one in BHP<sub>pb3</sub>. N- and C-terminus are indicated. **(C)** DALI search using BHP<sub>pb3</sub> without its FNIIIs (716 residues) as a bait. Asterix point to *Siphoviridae* BHPs. **(D)** Side views of monomeric BHP proteins of T4-gp27, 80 $\alpha$ -gp59, p2-ORF16 and GTA-gp13-15 with the same colour code as in A, and top views of the trimeric complex, with one monomer coloured. In T4, gp5 was added, with the OB domain that closes the tube coloured in pink and the needle in white. Note that GTA BHP is composed of two proteins, the Hub (gp13) and the Megatron (gp15). The Hub protein comprises hdl (darker blue), part of the insertion domain (darker cyan) in which is inserted an OB domain (purple), while the Megatron comprises hdII (green), part of the insertion domain (cyan), hdIII (orange), hdIV (yellow) and an N-terminal “iris helix” (magenta). **(E)** Side view superimposition of BHP<sub>pb3</sub> (coloured) with GTA-Megatron (light grey) and GTA-Hub (dark grey) in ribbon representation. **(F)** Negative stain EM images of purified BHP<sub>pb3</sub> showing isolated monomers coexisting with either free trimers seen in top views (circled), or that aggregate around impurities, seen in side views (arrow). The dimensions of the trimer (height: 45 nm, diameter: 9 nm) correspond to those of the BHP<sub>pb3</sub> trophy cup. In the top views, six subdomains can be distinguished (see circled particle).



**Figure S8: Tip central fiber structure analysis (A,B) and MS analysis of T5 tails(C).** (A) interface between BHP<sub>pb3</sub> last FNIII (yellow) and pb4 first FNIII (green) in ribbon representation. BHP<sub>pb3</sub> C-terminus (C) and pb4 N-terminus (N) are highlighted. Residues at the interface between the two proteins are shown in sticks. (B) Table of relevant DALI hits for pb4 spike (top panel), and alignment of a pb4 monomer with a monomer of the four first proteins of the table (bottom panel). (C) LC-ESI-TOF-MS of purified tails. The experimental intact mass of seven T5 tail proteins were determined. Even though proteomics detected their presence (Table S3), masses of pb4, RBP<sub>pb5</sub>, p140 and p132 as intact proteins could not be measured. This may be explained by a difficulty in their ionisation.



**Table S1: Cryo-EM data collection, refinement and validation statistics**

	Common core (EMD-13953) (PDB 7QG9)	Tip (without pb4) (EMD-14733) (PDB 7ZHJ)	Tip (fibre) (EMD-14790) (PDB 7ZLV)	Tip (full) (EMD-14869) (PDB 7ZQB)
<b>Data collection and processing</b>				
Magnification	105.000 x	105.000 x	105.000 x	105.000 x
Voltage (kV)	300	300	300	300
Electron exposure (e-/Å <sup>2</sup> )	40	40	40	40
Defocus range (µm)	-1.0 to -3.0	-1.0 to -3.0	-1.0 to -3.0	-1.0 to -3.0
Pixel size (Å)	1.351	1.351	1.351	1.351
Symmetry imposed	C3	C3	C3	C3
Micrographs (no.)	12816	3208	3208	3208
Final particle images (no.)	29639	9290	9290	9290
Map resolution (Å) 0.143 FSC threshold	3.44	3.53	4.22	3.88
Map resolution range (Å)	3.2 - 4.5	3.3 - 5	4 - 6	3.5 - 7
<b>Refinement</b>				
Model resolution (Å) 0.5 FSC threshold	3.51	3.5	4.6	4
Map sharpening B factor (Å <sup>2</sup> )	-112	-75	-123	-40
Model composition				
Chain count	27	33	6	36
Non-hydrogen atoms	50829	74346	12237	86583
Protein residues	6576	9528	1581	11109
Ligands	0	0	0	0
B factors (Å <sup>2</sup> )				
Protein (min/max/mean)	17.76/180.35/61.99	17.76/180.5/55.05	83.25/292.09/136.25	17.76/292.09/66.53
Ligand	n/a	n/a	n/a	n/a
R.m.s. deviations				
Bond lengths (Å)	0.004	0.004	0.004	0.004
Bond angles (°)	0.955	0.945	0.993	0.952
Validation				
MolProbity score	1.69	1.7	2.46	1.87
Clashscore	8.97	8.47	25.02	10.81
Poor rotamers (%)	0	0.04	0	0.03
Ramachandran plot				
Favoured (%)	96.67	96.41	89.29	95.4
Allowed (%)	3.33	3.59	10.52	4.57
Disallowed (%)	0	0	0.19	0.03
	Tip-FhuA (full) (EMD-14799) (PDB 7ZN2)	Tip-FhuA (bent fibre) (EMD-14800) (PDB 7ZN4)	Tip-FhuA (open tube) (EMD-14873) (PDB 7ZQP)	
<b>Data collection and processing</b>				
Magnification	105.000 x	105.000 x	105.000 x	
Voltage (kV)	300	300	300	
Electron exposure (e-/Å <sup>2</sup> )	40	40	40	
Defocus range (µm)	-1.0 to -3.0	-1.0 to -3.0	-1.0 to -3.0	
Pixel size (Å)	1.351	1.351	1.351	
Symmetry imposed	C1	C1	C3	
Micrographs (no.)	9608	9608	9608	
Final particle images (no.)	20349	20349	20349	
Map resolution (Å) 0.143 FSC threshold	4.25	4.32	3.60	
Map resolution range	4 - 20	4 - 6	3.4 - 4.5	
<b>Refinement</b>				
Model resolution (Å) 0.5 FSC threshold	4.3	4.3	3.6	
Map sharpening B factor (Å <sup>2</sup> )	-90	-135	-141	
Model composition				
Chain count	36	6	6	
Non-hydrogen atoms	89393	19989	18315	
Protein residues	11481	2581	2289	
Ligands	0	0	0	
B factors (Å <sup>2</sup> )				
Protein (min/max/mean)	17.76/200.25/67.84	54.9/144.16/94.69	21.56/200.25/53.75	
Ligand	n/a	n/a	n/a	
R.m.s. deviations				
Bond lengths (Å)	0.004	0.004	0.004	
Bond angles (°)	0.946	0.934	0.931	
Validation				
MolProbity score	1.66	1.62	1.52	
Clashscore	7.87	6.89	5	
Poor rotamers (%)	0	0	0	
Ramachandran plot				
Favoured (%)	96.49	96.36	96.18	
Allowed (%)	3.51	3.64	3.82	
Disallowed (%)	0	0	0	

**Table S2: Domain definition of T5 tip proteins.** *STF<sub>pb1</sub>*, *RBP<sub>pb5</sub>* and *p143*, in italics, are partially or not resolved in our EM maps.

<b>T5 protein name / function</b>	<b>Number of residues</b>	<b>Domains</b>	<b>Sequence</b>	<b>Colour</b>
TPP <sub>pb6</sub> Tail Tube Protein	464	Tail tube domain1 Tail tube domain 2 Ig-Like	1-203 204-355 356-464	cyan
p140 Baseplate Tube Protein	298	Tail tube domain 1 Tail tube domain 2	1-167 168-298	blue
p132 Collar protein	140	Ig-like	1-140	purple
<i>STF<sub>pb1</sub></i> <i>Side Tail Fibre</i>	<i>1,263</i>			pink
DTP <sub>pb9</sub> Distal Tail Protein	204	Tail tube domain OB-Fold	1-83; 172-204 84-171	orange
BHP <sub>pb3</sub> Baseplate Hub Protein	949	hdI hdII-Insertion plug/ $\beta$ -hairpin hdII hdIII hdIV hdIV-FNIII linker FNIII-1 FNIII-2	1-159 160-437 210-262 438-570 586-658 571-585; 659-709 711-742 743-835 836-949	gold
TMP <sub>pb2</sub> Tape Measure Protein	1,219	TMP <sub>pb2</sub> * TMP <sub>pb2C</sub>	2-1127, 1128-1219	red
pb4 Central Fibre Protein	688	FNIII-1 FNIII-2 FNIII-3 FNIII-spike linker Spike Small domains	1-105 106-208 209-316 317-332 333-465; 627-688 466-626	lime
<i>RBP<sub>pb5</sub></i> <i>Receptor Binding Protein</i>	<i>640</i>			dark green
<i>p143</i> <i>Tail Completion Protein</i>	<i>262</i>			cornflower blue

**Table S3: Proteomics of T5 tail proteins.**

Protein	Accession number	Gene name	Protein set score	Theoretical mass (Da) of full protein	#observable peptides	Coverage %
<b>STF<sub>pb1</sub></b>	<b>FIBL1_BPT5</b>	<b>ltf</b>	<b>8529.02</b>	<b>147992</b>	<b>81</b>	<b>73.14</b>
<b>TMP<sub>pb2</sub></b>	<b>TMP_BPT5</b>	<b>D18-19</b>	<b>7609.99</b>	<b>131496</b>	<b>75</b>	<b>62.84</b>
<b>BHP<sub>pb3</sub></b>	<b>BPPB3_BPT5</b>	<b>D16</b>	<b>4554.46</b>	<b>107166</b>	<b>53</b>	<b>75.66</b>
<b>pb4</b>	<b>FIBC_BPT5</b>	<b>D17</b>	<b>3347.97</b>	<b>74788</b>	<b>31</b>	<b>68.02</b>
<b>RBP<sub>pb5</sub></b>	<b>RBP5_BPT5</b>	<b>oad</b>	<b>2063.73</b>	<b>68726</b>	<b>25</b>	<b>54.38</b>
<b>TTP<sub>pb6</sub></b>	<b>TUBE_BPT5</b>	<b>N4</b>	<b>5054.41</b>	<b>50414</b>	<b>22</b>	<b>95.69</b>
<b>p143</b>	<b>COMPL_BPT5</b>	<b>ORF136</b>	<b>970.79</b>	<b>27857</b>	<b>18</b>	<b>63.53</b>
<b>p140</b>	<b>TAIL1_BPT5</b>	<b>ORF133</b>	<b>1221.08</b>	<b>34334</b>	<b>12</b>	<b>62.75</b>
<b>DTP<sub>pb9</sub></b>	<b>DIT_BPT5</b>	<b>D16</b>	<b>1315.54</b>	<b>22778</b>	<b>11</b>	<b>85.29</b>
<b>p132</b>	<b>FIBL2_BPT5</b>	<b>ORF125</b>	<b>375.13</b>	<b>15067</b>	<b>5</b>	<b>57.86</b>
<b>p142</b>	<b>TTTP_BPT5</b>	<b>ORF135</b>	<b>650.65</b>	<b>18360</b>	<b>5</b>	<b>40.37</b>
MCP <sub>pb8</sub>	CAPSD_BPT5	D20	537.11	50885	31	24.45
pb10	DECO_BPT5	N5	84	17247	11	9.15
putative metalloproteinase/ribonuclease	Q5DML2_BPT5	ORF082	66.64	25001	12	10.36
Uncharacterized protein	Q66M03_BPT5	T5p046	256.58	27425	16	26.29
Glucose-1-phosphatase	AGP_ECOLI	agp	114.46	45683	24	6.05
Elongation factor Tu	EFTU1_ECOLI	tufA	86.7	43284	26	4.31
Uncharacterized protein	YAGL_ECOLI	yagL	29.97	27274	16	3.88
Glyceraldehyde-3-phosphate dehydrogenase A	G3P1_ECOLI	gapA	27.76	35532	24	3.32
Uncharacterized protein	YDHW_ECOLI	ydhW	26.29	24421	13	4.19

Bold proteins are tail proteins (17). MCP<sub>pb8</sub> and pb10 are capsid proteins, identified because of a small proportion of revertant ( $10^{-7}$ ) in the T5D20am30d mutant, leading to the production of full T5 particles.

Q66M03\_BPT5: uncharacterised T5 protein, gene located in the early gene region. It is surrounded by genes coding for lysins, holins and endolysins.

Q5DML2\_BPT5: T5 protein, annotated as a putative metalloproteinase/carboxypeptidase. HHPred aligns it with nucleases or a subunit of DNA polymerase III. It is located in the early gene portion of the genome, surrounded by DNA and RNA interacting proteins. Phagonaute (58) aligns it with similar proteins in T5-like phages, but also other siphon- and myo-phages.

**Table S4: Validation statistics and model building for T5 tip individual proteins**

**Tip / Tip-FhuA common core proteins**

Protein Chain ID	TTP <sub>pb6-1</sub> D/E/F	TTP <sub>pb6-2</sub> A/B/C	p132-1 F/M/Q	p132-2 J/N/R	p132-3 K/O/S	p132-4 L/P/T	p140 G/H/I	DTP <sub>pb9-1</sub> V/X/Z	DTP <sub>pb9-2</sub> W/Y/a
Map(s) used for building	Tip / Tip-FhuA common core								
EMDB entry	EMD-13953								
Bond outliers (%)	0	0	0	0	0	0	0	0	0
Lengths	0	0	0	0	0.07	0	0	0	0
Angles	0	0	0	0	0	0	0	0	0
MolProbity score	1.63	4.77	1.56	1.56	1.68	1.83	1.47	1.31	1.4
Clashscore	7.37	10.07	7.19	7.19	8.07	9.01	4.82	5.63	5.63
Rotamer outliers (%)	0	0	0	0	0	0	0	0	0
Ramachandran plot (%)									
Outliers	0	0	0	0	0	0	0	0	0
Allowed	3.46	3.68	2.92	2.92	3.62	5.07	3.38	1.98	2.48
Favored	96.54	96.32	97.08	97.08	96.38	94.93	96.62	98.02	97.52
Residues built/total	464/464	464/464	139/140	139/140	140/140	140/140	298/298	204/204	204/204

**Tip proteins**

Protein Chain ID	BHP <sub>pb3</sub> b/c/d	TMP <sub>pb2c</sub> e/f/g	pb4 h/i/j
Map(s) used for building	Tip (without fibre)		First built into Tip-FhuA (bent fibre), then fitted and refined into Tip (fibre)
EMDB entry	EMD-14733		EMD-14790
Bond outliers (%)	0 (0)	0 (0)	0 (0)
Lengths	0.01 (1)	0 (0)	0 (0)
Angles			
MolProbity score	1.67	1.36	2.1
Clashscore	6.81	3.96	10.36
Rotamer outliers (%)	0.12	0	0
Ramachandran plot (%)			
Outliers	0	0	0.19
Allowed	4.22	3.03	10.52
Favored	95.78	96.97	89.29
Residues built/total	949/949	35/1219	527/688

**Tip-FhuA proteins**

Protein Chain ID	BHP <sub>pb3-1</sub> b	BHP <sub>pb3-2</sub> c	BHP <sub>pb3-3</sub> d	pb4-1 h	pb4-2 l	pb4-3 j	TMP <sub>pb2</sub> e/f/g
Map(s) used for building	Tip-FhuA (open tube): residues 1-710 Tip-FhuA (full): hdIV-FNIII linker, residues 711-729 Tip-FhuA (bent fibre): FNIII, residues 730-949			Tip-FhuA (bent fibre)			Tip-FhuA (open tube) Tip-FhuA (full)
EMDB entry	EMD-14799 EMD-14800 EMD-14873			EMD-14800			EMD-14799 EMD-14873
Bond outliers (%)	0	0	0	0	0	0	0
Lengths	0	0	0	0	0	0	0
Angles							
MolProbity score	1.56	1.48	1.62	1.58	1.38	1.38	1.23
Clashscore	4.6	5.61	5.01	5.19	5.89	5	4.62
Rotamer outliers (%)	0	0	0	0	0	0	0
Ramachandran plot (%)							
Outliers	0	0	0	0	0	0	0
Allowed	4.65	3.06	5.17	4.36	2.26	90.57	0
Favored	95.35	96.94	94.83	95.64	97.74	9.43	100
Residues built/total	949/949	949/949	949/949	671/688	669/688	589/688	43/1219

**Supplementary Movie S1 and S2 | Morphs of the BHP<sub>pb3</sub> trimer between the conformation before and after interaction with the receptor.** Side (S1) and top (S2) views. Monomers are coloured yellow, orange and salmon, with the hdIV-FNIII linker and the two FNIII domains in darker shades of the same colour.

**Supplementary Movie S3 and S4 | Morphs of the BHP<sub>pb3</sub> trimer without its FNIII, between the conformation before and after interaction with the receptor.** Side (S3) and top (S4) views. Monomers are coloured yellow, orange and salmon, with the hdl and hdIV domains in lighter colours, the hdIV-FNIII linker in dark orange and the plugs in brown.

## REFERENCES AND NOTES

1. C. A. Suttle, Marine viruses--major players in the global ecosystem. *Nat. Rev. Microbiol.* **5**, 801–812 (2007).
2. S. Uytendaele, B. Chen, J. Onsea, F. Ruythooren, Y. Debaveye, D. Devolder, I. Spriet, M. Depypere, J. Wagemans, R. Lavigne, J.-P. Pirnay, M. Merabishvili, P. De Munter, W. E. Peetermans, L. Dupont, L. Van Gerven, W.-J. Metsemakers, Safety and efficacy of phage therapy in difficult-to-treat infections: A systematic review. *Lancet Infect. Dis.* **22**, e208–e220 (2022).
3. D. Veesler, C. Cambillau, A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. *Microbiol. Mol. Biol. Rev.* **75**, 423–433 (2011).
4. A. R. Davidson, L. Cardarelli, L. G. Pell, D. R. Radford, K. L. Maxwell, Long noncontractile tail machines of bacteriophages. *Adv. Exp. Med. Biol.* **726**, 115–142 (2012).
5. M. Brackmann, S. Nazarov, J. Wang, M. Basler, Using force to punch holes: Mechanics of contractile nanomachines. *Trends Cell Biol.* **27**, 623–632 (2017).
6. R. Linares, C.-A. Arnaud, S. Degroux, G. Schoehn, C. Breyton, Structure, function and assembly of the long, flexible tail of siphophages. *Curr. Opin. Virol.* **45**, 34–42 (2020).
7. A. Desfosses, H. Venugopal, T. Joshi, J. Felix, M. Jessop, H. Jeong, J. Hyun, J. B. Heymann, M. R. H. Hurst, I. Gutsche, A. K. Mitra, Atomic structures of an entire contractile injection system in both the extended and contracted states. *Nat. Microbiol.* **4**, 1885–1894 (2019).
8. F. Jiang, N. Li, X. Wang, J. Cheng, Y. Huang, Y. Yang, J. Yang, B. Cai, Y.-P. Wang, Q. Jin, N. Gao, Cryo-EM structure and assembly of an extracellular contractile injection system. *Cell* **177**, 370–383.e15 (2019).
9. P. Ge, D. Scholl, N. S. Prokhorov, J. Avaylon, M. M. Shneider, C. Browning, S. A. Buth, M. Plattner, U. Chakraborty, K. Ding, P. G. Leiman, J. F. Miller, Z. H. Zhou, Action of a minimal contractile bactericidal nanomachine. *Nature* **580**, 658–662 (2020).

10. H.-W. Ackermann, Phage classification and characterization. *Methods Mol. Biol.* **501**, 127–140 (2009).
11. C.-A. Arnaud, G. Effantin, C. Vivès, S. Engilberge, M. Bacia, P. Boulanger, E. Girard, G. Schoehn, C. Breyton, Bacteriophage T5 tail tube structure suggests a trigger mechanism for *Siphoviridae* DNA ejection. *Nat. Commun.* **8**, 1953 (2017).
12. P. L. Campbell, R. L. Duda, J. Nassur, J. F. Conway, A. Huet, Mobile loops and electrostatic interactions maintain the flexible tail tube of bacteriophage lambda. *J. Mol. Biol.* **432**, 384–395 (2020).
13. A. Goulet, J. Lai-Kee-Him, D. Veessler, I. Auzat, G. Robin, D. A. Shepherd, A. E. Ashcroft, E. Richard, J. Lichière, P. Tavares, C. Cambillau, P. Bron, The opening of the SPP1 bacteriophage tail, a prevalent mechanism in gram-positive-infecting Siphophages. *J. Biol. Chem.* **286**, 25397–25405 (2011).
14. D. Veessler, S. Spinelli, J. Mahony, J. Lichière, S. Blangy, G. Bricogne, P. Legrand, M. Ortiz-Lombardia, V. Campanacci, D. van Sinderen, C. Cambillau, Structure of the phage TP901-1 1.8 MDa baseplate suggests an alternative host adhesion mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8954–8958 (2012).
15. G. Sciara, C. Bebeacua, P. Bron, D. Tremblay, M. Ortiz-Lombardia, J. Lichière, M. van Heel, V. Campanacci, S. Moineau, C. Cambillau, Structure of lactococcal phage p2 baseplate and its mechanism of activation. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6852–6857 (2010).
16. J. L. Kizziah, K. A. Manning, A. D. Dearborn, T. Dokland, Structure of the host cell recognition and penetration machinery of a *Staphylococcus aureus* bacteriophage. *PLOS Pathog.* **16**, e1008314 (2020).
17. Y. Zivanovic, F. Confalonieri, L. Ponchon, R. Lurz, M. Chami, A. Flayhan, M. Renouard, A. Huet, P. Decottignies, A. R. Davidson, C. Breyton, P. Boulanger, Insights into bacteriophage T5 structure from analysis of its morphogenesis genes and protein components. *J. Virol.* **88**, 1162–1174 (2014).

18. M. Demerec, U. Fano, Bacteriophage-resistant mutants in *Escherichia coli*. *Genetics* **30**, 119–136 (1945).
19. A. Huet, R. L. Duda, P. Boulanger, J. F. Conway, Capsid expansion of bacteriophage T5 revealed by high resolution cryoelectron microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 21037–21046 (2019).
20. P. Boulanger, P. Jacquot, L. Plançon, M. Chami, A. Engel, C. Parquet, C. Herbeuval, L. Letellier, Phage T5 straight tail fiber is a multifunctional protein acting as a tape measure and carrying fusogenic and muralytic activities. *J. Biol. Chem.* **283**, 13556–13564 (2008).
21. C. Garcia-Doval, J. R. Castón, D. Luque, M. Granell, J. M. Otero, A. L. Llamas-Saiz, M. Renouard, P. Boulanger, M. J. van Raaij, Structure of the receptor-binding carboxy-terminal domain of the bacteriophage T5 L-shaped tail fiber with and without its intra-molecular chaperone. *Viruses* **7**, 6424–6440 (2015).
22. A. Flayhan, F. M. D. Vellieux, R. Lurz, O. Maury, C. Contreras-Martel, E. Girard, P. Boulanger, C. Breyton, Crystal Structure of pb9, the Distal tail protein of bacteriophage T5: A conserved structural motif among all siphophages. *J. Virol.* **88**, 820–828 (2014).
23. A. Flayhan, F. Wien, M. Paternostre, P. Boulanger, C. Breyton, New insights into pb5, the receptor binding protein of bacteriophage T5, and its interaction with its *Escherichia coli* receptor FhuA. *Biochimie* **94**, 1982–1989 (2012).
24. C. Breyton, A. Flayhan, F. Gabel, M. Lethier, G. Durand, P. Boulanger, M. Chami, C. Ebel, Assessing the conformational changes of pb5, the receptor-binding protein of phage T5, upon binding to its *Escherichia coli* receptor FhuA. *J. Biol. Chem.* **288**, 30763–30772 (2013).
25. V. Braun, FhuA (TonA), the career of a protein. *J. Bacteriol.* **191**, 3431–3436 (2009).
26. M. Bonhivers, A. Ghazi, P. Boulanger, L. Letellier, FhuA, a transporter of the *Escherichia coli* outer membrane, is converted into a channel upon binding of bacteriophage T5. *EMBO J.* **15**, 1850–1856 (1996).



27. N. Chiaruttini, M. de Frutos, E. Augarde, P. Boulanger, L. Letellier, V. Viasnoff, Is the in vitro ejection of bacteriophage DNA quasistatic? A bulk to single virus study. *Biophys. J.* **99**, 447–455 (2010).
28. I. G. Denisov, S. G. Sligar, Nanodiscs in membrane biochemistry and biophysics. *Chem. Rev.* **117**, 4669–4713 (2017).
29. S. Degroux, G. Effantin, R. Linares, G. Schoehn, C. Breyton, Deciphering bacteriophage T5 host recognition mechanism and infection trigger. *J. Virol.* **2023**, e01584–22 (2023).
30. H. Fraga, C.-A. Arnaud, D. F. Gauto, M. Audin, V. Kurauskas, P. Macek, C. Krichel, J.-Y. Guan, J. Boisbouvier, R. Sprangers, C. Breyton, P. Schanda, Solid-state NMR H-N-(C)-H and H-N-C-C 3D/4D correlation experiments for resonance assignment of large proteins. *ChemPhysChem* **18**, 2697–2703 (2017).
31. L. Holm, DALI and the persistence of protein shape. *Protein Sci.* **29**, 128–140 (2020).
32. N. M. I. Taylor, N. S. Prokhorov, R. C. Guerrero-Ferreira, M. M. Shneider, C. Browning, K. N. Goldie, H. Stahlberg, P. G. Leiman, Structure of the T4 baseplate and its function in triggering sheath contraction. *Nature* **533**, 346–352 (2016).
33. E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
34. M. Noirclerc-Savoye, A. Flayhan, C. Pereira, B. Gallet, P. Gans, C. Ebel, C. Breyton, Tail proteins of phage T5: Investigation of the effect of the His6-tag position, from expression to crystallisation. *Protein Expr. Purif.* **109**, 70–78 (2015).
35. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

36. M. Zweig, D. J. Cummings, Cleavage of head and tail proteins during bacteriophage T5 assembly: Selective host involvement in the cleavage of a tail protein. *J. Mol. Biol.* **80**, 505–518 (1973).
37. L. C. Tsui, R. W. Hendrix, Proteolytic processing of phage lambda tail protein gpH: Timing of the cleavage. *Virology* **125**, 257–264 (1983).
38. S. Kanamaru, P. G. Leiman, V. A. Kostyuchenko, P. R. Chipman, V. V. Mesyanzhinov, F. Arisaka, M. G. Rossmann, Structure of the cell-puncturing device of bacteriophage T4. *Nature* **415**, 553–557 (2002).
39. S. R. Casjens, R. W. Hendrix, Locations and amounts of major structural proteins in bacteriophage lambda. *J. Mol. Biol.* **88**, 535–545 (1974).
40. L. Cardarelli, L. G. Pell, P. Neudecker, N. Pirani, A. Liu, L. A. Baker, J. L. Rubinstein, K. L. Maxwell, A. R. Davidson, Phages have adapted the same protein fold to fulfill multiple functions in virion assembly. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 14384–14389 (2010).
41. P. Bárdy, T. Füzik, D. Hřebík, R. Pantůček, J. Thomas Beatty, P. Plevka, Structure and mechanism of DNA delivery of a gene transfer agent. *Nat. Commun.* **11**, 3034 (2020).
42. L. T. Alexander, R. Lepore, A. Kryshatovych, A. Adamopoulos, M. Alahuhta, A. M. Arvin, Y. J. Bomble, B. Böttcher, C. Breyton, V. Chiarini, N. B. Chinnam, W. Chiu, K. Fidelis, R. Grinter, G. D. Gupta, M. D. Hartmann, C. S. Hayes, T. Heidebrecht, A. Ilari, A. Joachimiak, Y. Kim, R. Linares, A. L. Lovering, V. V. Lunin, A. N. Lupas, C. Makbul, K. Michalska, J. Moul, P. K. Mukherjee, W. S. Nutt, S. L. Oliver, A. Perrakis, L. Stols, J. A. Tainer, M. Topf, S. E. Tsutakawa, M. Valdivia-Delgado, T. Schwede, Target highlights in CASP14: Analysis of models by structure providers. *Proteins* **89**, 1647–1672 (2021).
43. G. Guihard, P. Boulanger, L. Letellier, Involvement of phage T5 tail proteins and contact sites between the outer and inner membrane of Escherichia coli in phage T5 DNA injection. *J. Biol. Chem.* **267**, 3173–3178 (1992).

44. M. Zweig, D. J. Cummings, Structural proteins of bacteriophage T5. *Virology* **51**, 443–453 (1973).
45. E. Kandiah, T. Giraud, A. de Maria Antolinos, F. Dobias, G. Effantin, D. Flot, M. Hons, G. Schoehn, J. Susini, O. Svensson, G. A. Leonard, C. Mueller-Dieckmann, CM01: A facility for cryo-electron microscopy at the European Synchrotron. *Acta Crystallogr. D Struct. Biol.* **75**, 528–535 (2019).
46. X. Li, P. Mooney, S. Zheng, C. R. Booth, M. B. Braunfeld, S. Gubbens, D. A. Agard, Y. Cheng, Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584–590 (2013).
47. K. Zhang, Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
48. G. Tang, L. Peng, P. R. Baldwin, D. S. Mann, W. Jiang, I. Rees, S. J. Ludtke, EMAN2: An extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
49. J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. Hagen, E. Lindahl, S. H. Scheres, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
50. C.-A. Arnaud, thesis, University Grenoble-Alpes (2017).
51. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
52. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

53. C. J. Williams, J. J. Headd, N. W. Moriarty, M. G. Prisant, L. L. Videau, L. N. Deis, V. Verma, D. A. Keedy, B. J. Hintze, V. B. Chen, S. Jain, S. M. Lewis, W. B. Arendall III, J. Snoeyink, P. D. Adams, S. C. Lovell, J. S. Richardson, D. C. Richardson, MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
54. A. P. Pandurangan, M. Topf, RIBFIND: A web server for identifying rigid bodies in protein structures and to aid flexible fitting into cryo EM maps. *Bioinformatics* **28**, 2391–2393 (2012).
55. R. T. Kidmose, J. Juhl, P. Nissen, T. Boesen, J. L. Karlsen, B. P. Pedersen, Namdinator - automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *IUCrJ* **6**, 526–531 (2019).
56. M. G. Casabona, Y. Vandenbrouck, I. Attree, Y. Couté, Proteomic characterization of *Pseudomonas aeruginosa* PAO1 inner membrane. *Proteomics* **13**, 2419–2423 (2013).
57. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
58. H. Delattre, O. Souiai, K. Fagoonee, R. Guerois, M.-A. Petit, Phagonaute: A web-based interface for phage synteny browsing and protein function prediction. *Virology* **496**, 42–50 (2016).