Supporting Information

Chaban et al. 10.1073/pnas.1504039112

SI Materials and Methods

Microbiology and Genetic Methods. Bacteriophages wild-type SPP1, *sus31* (defective in gene *13*), *sus117* (defective in gene *16*), and *sus128* (defective in gene *15*) were previously described (1–3). SPP1*sus31sus117* was constructed by mixed infection of the permissive host *B. subtilis* HA101B. The parental phages were added at an input multiplicity of 10 infective particles of each phage strain per bacterium. Individual progeny phage clones then were screened using appropriate plasmid constructs that complement or marker-rescue the phage mutations.

Gene *16* alleles coding for $gp16_{Q51C}$, $gp16_{Q43CQ51C}$, or $gp16_{Q43CY47CQ51C}$ were transferred from plasmids to the SPP1 genome by double cross-over during SPP1*sus31sus117* infection of the nonpermissive host *B. subtilis* YB886 bearing constructs pCC48, pCC42, or pCC50 (Table S1), respectively. Individual phage clones multiplying in *B. subtilis* YB886 were screened by PCR and DNA sequencing to confirm the presence of the desired mutations. These phages showed a more stable phenotype than phage particles carrying gp16 cysteine substitutions produced by complementation in which strains bearing plasmids coding for the same gp16 mutant proteins were infected with SPP1*sus117* (4).

Multiplication and purification of phage particles were carried out as described (5, 6). Virions carrying cysteine mutations in gp16 were preincubated with 1 mM (liquid cultures) or 4 mM (plating in medium with 0.7% agar) DTT for 15 min at room temperature before infection. Semisolid medium used for SPP1 titrations was supplemented with 4 mM DTT and 10 mM $CaCl_2$.

Plasmid Construction. Plasmid pCC6 was constructed by cloning a PCR fragment spanning genes *14–16* [coordinates 8787–9664 of the SPP1 genome sequence; GenBank accession no. X97918 (7)] downstream of the inducible promoter $P_{N25/0}$ present in a pHP13-derived plasmid (8). The PCR product was cleaved at PstI and HindIII sites engineered in the sequence of the primers designed for PCR amplification (14Fw-PstI and 16rev-HindIII; Table S2) and ligated to the vector cut with the same restriction endonucleases. The resulting pCC6 plasmid was used as template for site-directed mutagenesis with the QuikChange Site-Directed Mutagenesis Kit (Stratagene) to engineer mutations in genes *15* and *16* (Table S1). The replacement of $R_{98}KMAR_{102}$ by MAG at the gp15 C terminus was engineered in gene *15* by overlap PCR (9): First, two PCR fragments were obtained using the oligonucleotide pairs 14fw-PstI together with 15 MAG-reverse and 15 MAG-forward together with 16rev-HindIII. The DNA fragments purified with the QIAquick PCR Purification Kit (Qiagen) then were used as templates for amplification with 14fw-PstI and 16rev-HindIII followed by cloning of the resulting PCR product cleaved with PstI-HindIII into pCC6 cut with the same enzymes.

A fragment of SPP1 DNA was amplified by PCR with primers 13CS0PstI and 16revHindIII (Table S2), cleaved with PstI-*Aat*II, and ligated to pCC6 cut with the same enzymes to generate pCC40. The PstI-*Aat*II fragment spans genes *13* (coordinate 7782 in GenBank X97918) to the beginning of gene *16*. This gene was fused in frame to the *Aat*II site of gene *16* present in pCC6, restoring the complete gene sequence. Cloning of the same PstI-*Aat*II fragment into plasmids pCC36 and pCC39 (Table S1) yielded pCC48 and pCC42 that code for $gp16_{Q51C}$ and $gp16_{Q43CQ51C}$, respectively. Site-directed mutagenesis of pCC42 was used to obtain pCC50 encoding $gp16_{Q43CY47CQ51C}$. Constructs pCC40, pCC42, pCC48, and pCC50 thus are isogenic coding for genes *13*, *14*, *15*, and different alleles of gene *16*.

All plasmid genetic manipulations were carried out in *Escherichia coli* XLI Blue (Stratagene) or DH5α. Selected clones were transformed into *B. subtilis* YB886 for genetic and functional infection experiments with SPP1.

Purification of Capsid Structures Carrying gp15 Mutant Forms. *B. subtilis* YB886 bearing pCC6, pCC45, pCC46, or pPT322 (Table S1) was grown (150 mL) at 37 °C in LB medium with 5 μg/mL chloramphenicol to a density of 10^8 cfu/mL and supplemented with 10 mM $CaCl_2$. The culture was infected with SPP1*sus128* at an input multiplicity of five infective particles per bacterium. Purification of capsid structures was carried out as described (2). Gradient fractions were analyzed by Coomassie Blue-stained gels, Ponceau staining after protein transfer to membranes by Western blot, and immunodetection with rabbit polyclonal antibodies anti-SPP1 (for detection of the major capsid protein gp13), anti-gp15, and anti-gp16. The assay shown in Fig. 3*B* was reproduced at least twice for every mutation using different capsid preparations in independent experiments.

Assay for Stable DNA Packaging in SPP1 Particles in Vivo. Cultures of YB886 bearing the constructs under analysis (2 mL) were grown at 37 °C in LB medium with 5 μg/mL chloramphenicol to a density of 10^8 cfu/mL and were supplemented with 10 mM $CaCl_2$. Infection with wild-type SPP1 or with SPP1*sus128*, a phage strain defective in gp15 production (2), was carried out at an input multiplicity of 5. At 25 min postinfection the infected bacteria were sedimented by centrifugation at 8,000 × *g* for 5 min. The supernatant was discarded. After a short spin the remaining traces of supernatant were carefully pipetted away to eliminate all free input phages used for infection. The bacterial pellet was resuspended in 100 μL of 50 mM glucose, 1 mg/mL lysozyme, protease inhibitors at the concentration recommended by the manufacturer (Roche), and 50 mM Tris-Cl (pH 8) and was incubated for 5 min at room temperature. An identical volume of 2× lysis buffer [500 mM NaCl, 1% Nonidet P-40, 50 mM Tris-Cl (pH 8), 5 mM $MgCl_2$, 50 U Benzonase (Novagen), 100 μg/mL RNase] was added followed by incubation for 30 min at 4 °C. The sample was deproteinized by treatment with 10 mM EDTA-Na (pH 8), 0.5% SDS, and 50 μg/mL proteinase K for 1 h at 65 °C (10). DNA was analyzed by conventional 0.8% gel agarose electrophoresis or pulse-field gel electrophoresis (PFGE) (11). The complete assay presented in Fig. 3*D* was reproduced in two independent experiments, and the phenotype of $gp15_{R5ER8E}$ was observed in five independent assays.

DNA Ejection Assays in Vitro. SPP1 virions (3×10^9 pfu or an equivalent number of phage particles for each experimental condition assayed) were incubated with 0.6 U/μL Benzonase and 0.25 mg/mL RNase for 30 min at 37 °C. Samples then were split, mixed with DTT for a final concentration of 4 mM or with an identical volume of water (control), and maintained at 37 °C for 15 min. After the NaCl final concentration was increased to 300 mM, YueB780 was added to 12 μg/mL for DNA ejection. Plasmid DNA (180 ng) was mixed with each reaction to assure that free DNA was digested in all conditions assayed. After 1 h of incubation at 37 °C the sample was deproteinized followed by agarose gel electrophoresis and/or PFGE analysis as described above. The assay was reproduced in at least eight independent experiments, and several different phage preparations were tested.

Structures in DNA ejection experiments were imaged by EM after negative staining with uranyl acetate (12). DNA was observed

after adsorption to mica and rotary shadowing (13) in samples not treated with Benzonase.

Image Processing and 3D Reconstruction.

Ab initio reconstruction. An initial dataset of ~8,000 particles was selected manually with BOXER [EMAN (14)]. The images were low-pass filtered to remove details in the images that could be affected by noise. The individual particle images were centered, subjected to reference-free MSA (15), and classified with ~50 images per class. The particle images were aligned and classified [multi-reference alignment (MRA) (15)]). The best classes representing characteristic views were used as new references, followed by MSA. The process was iterated several times to obtain the consistent classes. The alignment parameters found during this step were used to shift and rotate nonfiltered particles. These images of the particles were again subjected to MSA and classification. The best class averages, characterized by the smallest variance between members of the class averages, were selected semiautomatically. Their angular orientations were determined using the angular reconstitution (AR) technique, assuming C12 symmetry for the connector region (16). The classes with the lowest AR error were used for 3D reconstruction using a filtered back-projections algorithm (15).

Structure refinement. Refinement of both HTIfull and HTIEmpty complexes was carried out by the iterative procedure of MRA and classification reducing the number of images per class to between three and four followed by refinement of angular orientation using the reconstruction from the previous iteration. C12 symmetry was applied during the first steps of analysis; later it was reduced to C6 symmetry. During each round, AR error and 3D reconstruction error were used for automatic discrimination of poor classes: Classes with the highest errors in the 3D reconstruction were removed, and a new reconstruction was calculated using only the best 70% of classes. The final reconstructions included ~14,000 particle images for HTIFull and ~18,000 particle images for HTIEmpty complexes.

Assessment of the components symmetry. After several rounds of refinement, parts of the images corresponding to the tail region were extended to include longer segments of the bacteriophage tail. The SPP1 tail tube is formed by rings of six subunits piled up with a helical symmetry (17); for that reason, the symmetry of the reconstruction during the refinement was relaxed to C6. Slices of the 3D map orthogonal to the tail axis were tested for symmetry using the rotational cross-correlation function (Fig. S3*C*). It revealed that a part of the connector adjacent to the tail segments has C6 symmetry, whereas the part of the connector containing gp6 and gp15 has clear C12 symmetry. Gp16 has prevailing C6 symmetry in the external part of the structure that faces tail proteins and C12 symmetry in the part adjacent to gp15 (Fig. S3*C*). The final 3D map was corrected for amplitudes using a rotationally averaged amplitude spectrum of the atomic models of gp16, gp15, gp16, and gp17. The resolution of the 3D maps was estimated with 0.5 FSC criteria (Fig. S3 *A* and *B*).

Fitting of atomic models into EM 3D maps. To improve the fit and to localize protein–protein contacts, flexible fitting was done using Modeler FlexEM software (18). The procedure comprised three steps that were iterated until no further improvement was achieved:

Step 1. Map segmentation. 3D EM densities corresponding to subunits of different proteins were extracted using Chimera from the 3D maps using rigid body fit of known or predicted atomic models with sufficient space around the fitted models (6–7 Å). Extracted segments of EM map then were used to refine the fit of atomic models using Flex-EM molecular dynamics (MD) optimization for each subunit. For MD optimization, secondary structure elements of the atomic model were defined as rigid body elements, and connecting loops were allowed to be flexible.

Step 2. The extracted area of the map density was expanded to three adjacent subunits. This step to refine the atomic model fitting allowed the reduction of clashes between subunits and definition of their boundaries. Each Flex-EM round was followed by symmetrization of the refined atomic model to C12 for gp6 and gp15. Transition from C12 to C6 takes place in gp16 (Fig. S3*C*). Rotational symmetry C6 thus was used for gp16 and for components of the sixfold symmetric tail (gp17 and gp17.1).

Step 3. Tests of pseudoatomic models for clashes followed by optimization were carried out using Flex-EM Conjugate Gradient optimization (CG; 18). The program was run on atomic coordinates of fitted structures to remove clashes between amino acids of subunits composing the oligomeric rings and between the adjacent rings of different proteins. This procedure did not change the previous MD fit significantly. For proteins gp15, gp16, and gp17 atomic structures were obtained by NMR (4, 19) providing several conformations for the protein flexible regions. We tested all of them and used the models that had the highest correlation with the EM density map (Table S3) and minimal clashes between adjacent subunits. To analyze and optimize the contacts between the rings of different proteins, we used columns of protein subunits throughout the full height of the HTI (as highlighted in colors in Fig. 1*C*), consisting of three subunits from each protein ring.

1. Riva S, Polsinelli M, Falaschi A (1968) A new phage of *Bacillus subtilis* with infectious DNA having separable strands. *J Mol Biol* 35(2):347–356.
2. Becker B, et al. (1997) Head morphogenesis genes of the *Bacillus subtilis* bacteriophage SPP1. *J Mol Biol* 268(5):822–839.
3. Orlova EV, et al. (2003) Structure of a viral DNA gatekeeper at 10 A resolution by cryo-electron microscopy. *EMBO J* 22(6):1255–1262.
4. Lhuillier S, et al. (2009) Structure of bacteriophage SPP1 head-to-tail connection reveals mechanism for viral DNA gating. *Proc Natl Acad Sci USA* 106(21):8507–8512.
5. Dröge A, Tavares P (2000) *In vitro* packaging of DNA of the *Bacillus subtilis* bacteriophage SPP1. *J Mol Biol* 296(1):103–115.
6. Jakutytė L, et al. (2011) Bacteriophage infection in rod-shaped gram-positive bacteria: Evidence for a preferential polar route for phage SPP1 entry in *Bacillus subtilis*. *J Bacteriol* 193(18):4893–4903.
7. Alonso JC, et al. (1997) The complete nucleotide sequence and functional organization of *Bacillus subtilis* bacteriophage SPP1. *Gene* 204(1-2):201–212.
8. Isidro A, Henriques AO, Tavares P (2004) The portal protein plays essential roles at different steps of the SPP1 DNA packaging process. *Virology* 322(2):253–263.
9. Heckman KL, Pease LR (2007) Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat Protoc* 2(4):924–932.
10. de Frutos M, Brasilès S, Tavares P, Raspaud E (2005) Effect of spermine and DNase on DNA release from bacteriophage T5. *Eur Phys J E Soft Matter* 17(4):429–434.
11. São-José C, de Frutos M, Raspaud E, Santos MA, Tavares P (2007) Pressure built by DNA packing inside virions: Enough to drive DNA ejection *in vitro*, largely insufficient for delivery into the bacterial cytoplasm. *J Mol Biol* 374(2):346–355.
12. Steven AC, et al. (1988) Molecular substructure of a viral receptor-recognition protein. The gp17 tail-fiber of bacteriophage T7. *J Mol Biol* 200(2):351–365.
13. Portmann R, Sogo JM, Koller T, Zillig W (1974) Binding sites of *E. coli* RNA polymerase on T7 DNA as determined by electron microscopy. *FEBS Lett* 45(1):64–67.
14. Ludtke SJ, Baldwin PR, Chiu W (1999) EMAN: Semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* 128(1):82–97.
15. van Heel M, et al. (2000) Single-particle electron cryo-microscopy: Towards atomic resolution. *Q Rev Biophys* 33(4):307–369.
16. Lurz R, et al. (2001) Structural organisation of the head-to-tail interface of a bacterial virus. *J Mol Biol* 310(5):1027–1037.
17. Plisson C, et al. (2007) Structure of bacteriophage SPP1 tail reveals trigger for DNA ejection. *EMBO J* 26(15):3720–3728.
18. Topf M, et al. (2008) Protein structure fitting and refinement guided by cryo-EM density. *Structure* 16(2):295–307.
19. Chagot B, et al. (2012) Solution structure of gp17 from the Siphoviridae bacteriophage SPP1: Insights into its role in virion assembly. *Proteins* 80(1):319–326.

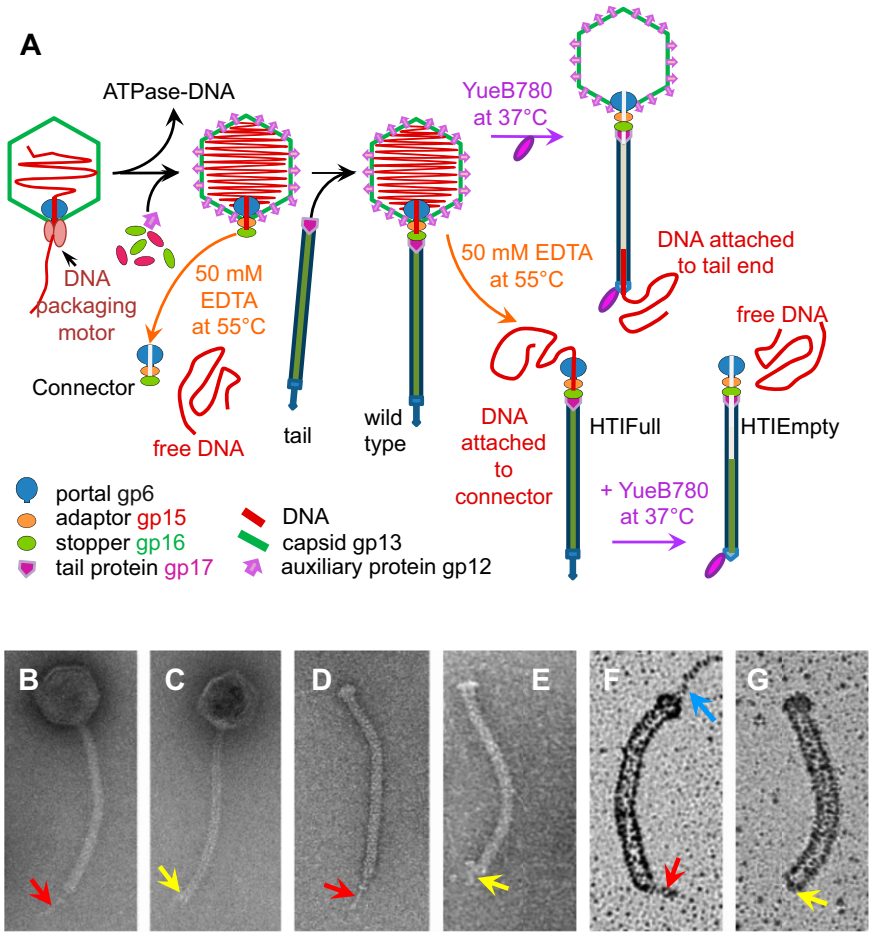

Fig. S1. Bacteriophage assembly, DNA ejection, and disruption of viral particles for cryoEM structure determination. (*A*) Schematics of the experimental strategy. SPP1 DNA-filled capsids and intact virions were disrupted by treatment with EDTA at 55 °C to provide connectors and HTI structures (HTIFull and HTIEmpty, the latter obtained after incubation with YueB780), respectively, for cryoEM imaging. Capsid and HTI molecular components are identified at the bottom left of the figure. (*B* and *C*) EM of negatively stained wild-type SPP1 particles (*B*) and after incubation with the receptor ectodomain YueB780 (*C*) in the presence of Benzonase. (*D* and *E*) EM of negatively stained connector–tail complexes produced by disruption of wild-type SPP1 virions with EDTA and incubated with receptor buffer (*D*) or with YueB780 (*E*) in presence of Benzonase. (*F* and *G*) EM observation of DNA–connector–tail complexes produced as in *D* and *E*, respectively, after adsorption to mica. These samples were not treated with Benzonase. In *F*, DNA emerging from HTIFull is indicated by a blue arrow. In *B*–*G* the tail spikes before and after interaction with YueB780 are labeled with red and yellow arrows, respectively.

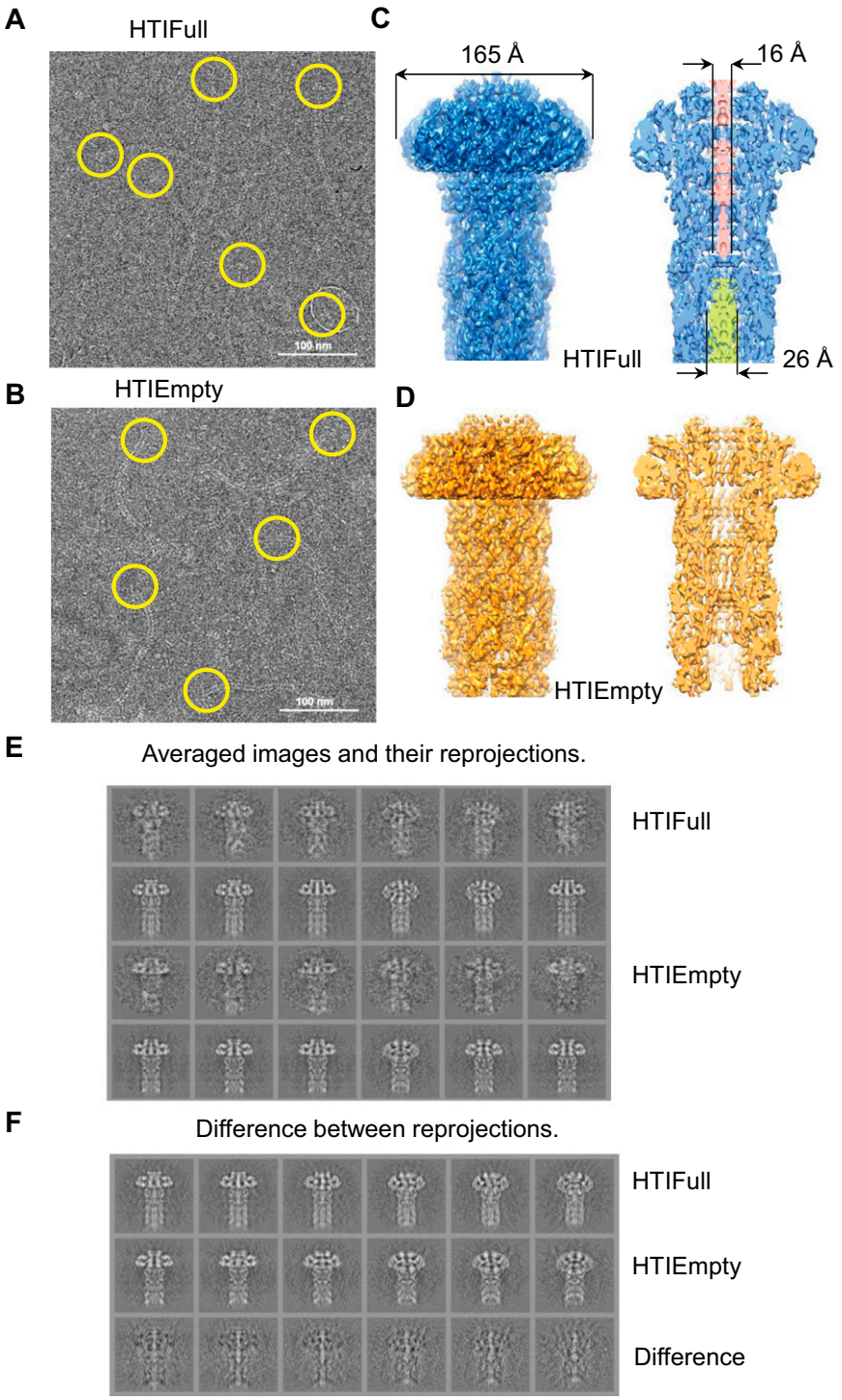

Fig. S2. CryoEM and reconstructions of the HTIFull and HTIEmpty structures. (*A* and *B*) Wild-type SPP1 particles disrupted with EDTA yielded tail structures bound to the connector to image the HTIFull (yellow circles in *A*), and their subsequent incubation with YueB780 provided HTIEmpty (*B*). (*C* and *D*) Surface (*Left*) and cut-away (*Right*) views of HTIFull (*C*) and HTIEmpty (*D*) reconstructions. Densities corresponding to DNA (salmon) and tape measure (light green) are highlighted in HTIFull (*C*). (*E*) Averaged images and their reprojections. (*F*) Differences between reprojections.

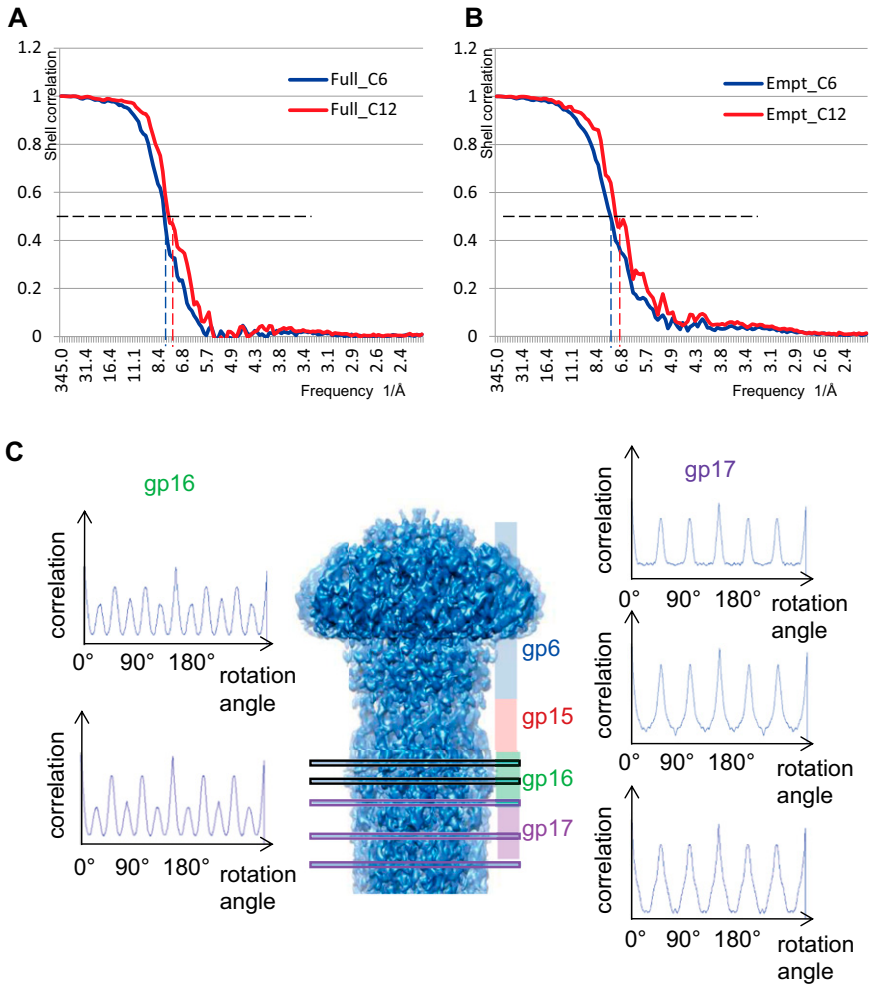

Fig. S3. Reconstructions resolution and symmetry. (*A* and *B*) Resolution of the HTIFull and HTIEmpty reconstructions assessed by the Fourier shell correlation at the threshold 0.5. (*C*) Symmetry assessment of HTIFull slices along the reconstruction height.

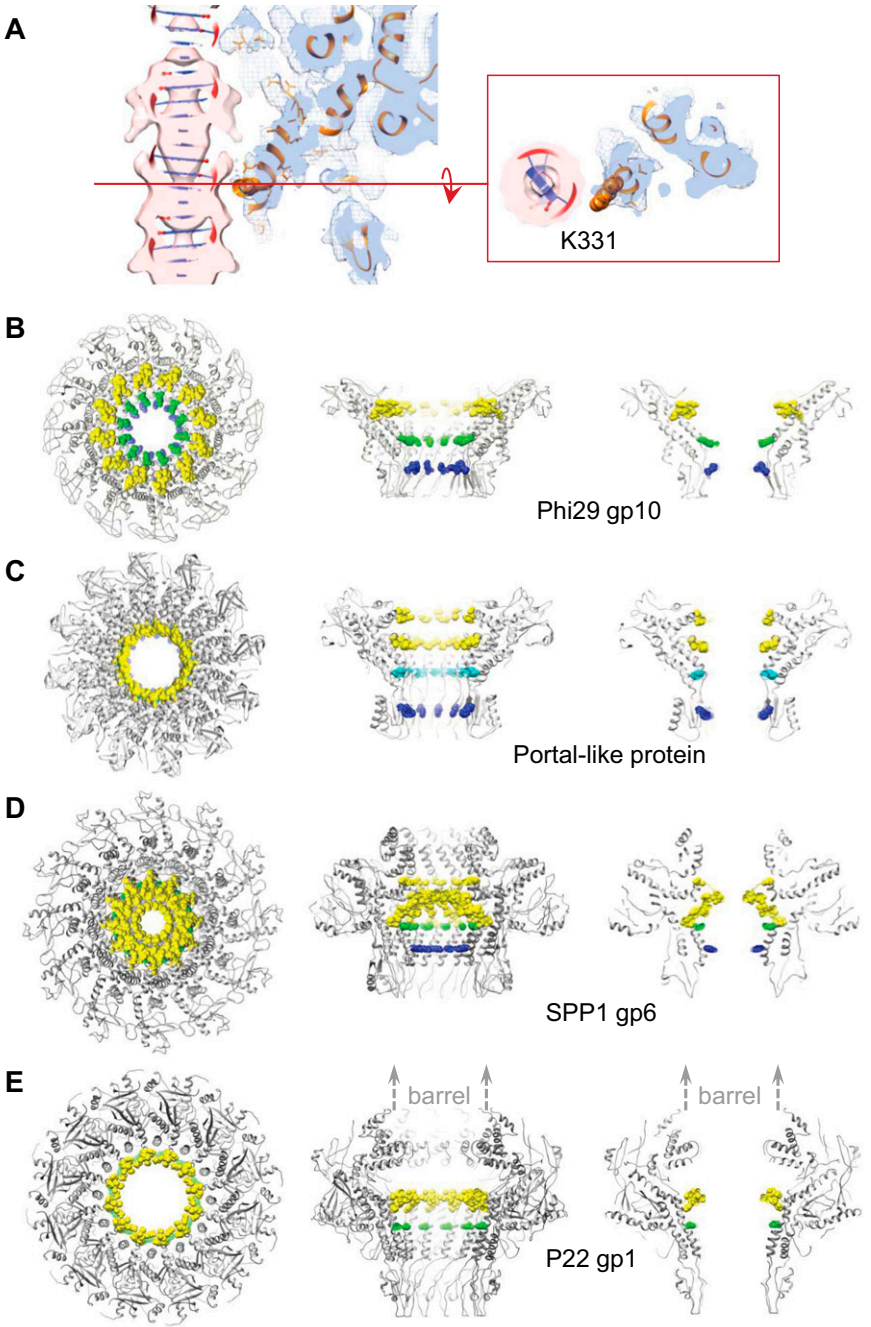

Fig. S4. Portal protein–DNA interaction. (*A*) Densities of DNA and of the gp6 stem region highlighting the contacts with DNA. The density slice reveals the close proximity of K331 to the double helix. (*B*–*E*) Structures of portal cyclical oligomers are shown in gray with colored positively charged side chains (spheres in green and blue) that are oriented toward the portal channel interior. Tunnel loops are shown as yellow spheres. Cross-sections of structures in the areas of the tunnel loops are shown at the left of each panel, cut-away views of atomic models are shown in the center, and central sections along the symmetry axes are shown at the right. (*B*) Phage phi29 gp10 (PDB ID code 1FOU). (*C*) Protein encoded by the *Corynebacterium diphtheriae* genome (PDB ID code 3KDR). (*D*) Phage SPP1 gp6 (HTIFull; this work). (*E*) Phage P22 gp1 (PDB ID code 3LJ4) without the barrel that extends the portal channel ~200 Å to the capsid interior (1). Note that in the structures in *B* and *C* most of the tunnel loops are not defined. Although the topologically equivalent loops (yellow) are present in *E*, they are far from the central channel axis and thus cannot interact with DNA.

1. Olia AS, Prevelige PE, Jr, Johnson JE, Cingolani G (2011) Three-dimensional structure of a viral genome-delivery portal vertex. *Nat Struct Mol Biol* 18(5):597–603.

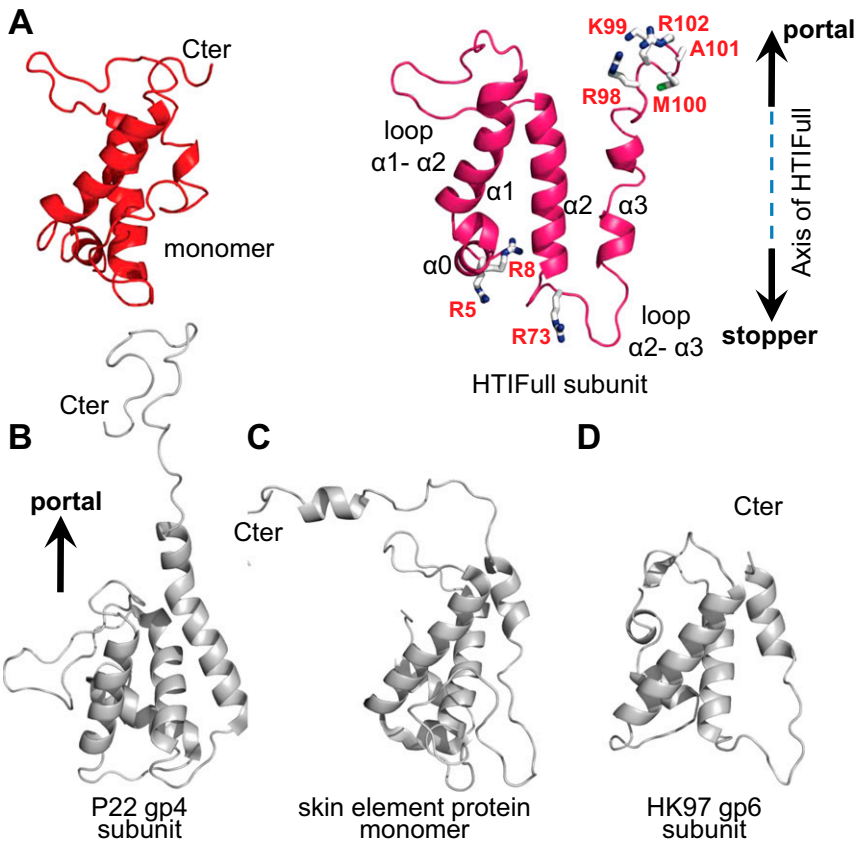

Fig. S5. Comparison of gp15 structures with other phage adaptor proteins. (*A*) The gp15 monomer NMR structure (red) (1) and the gp15 subunit fitted in the HTIFull (magenta) reconstruction. Secondary structure elements of gp15 are labeled in the subunit of HTIFull and residues mutagenized in this work are shown as sticks. (*B–D*) Structures of a gp4 subunit of phage P22 (PDB ID code 3LJ4) (*B*), of the monomer of YqbG encoded by the skin element of the *B. subtilis* genome (PDB ID code 1XN8) (*C*), and of a subunit of gp6 from phage HK97 (PDB ID code 3JVO) (*D*) presented in gray. The subunit of gp15 in HTIFull (*A*, *Right*) is shown as a side view relative to the HTIFull plane of symmetry, and the other structures (*B–D*) are displayed in similar orientations.

1. Lhuillier S, et al. (2009) Structure of bacteriophage SPP1 head-to-tail connection reveals mechanism for viral DNA gating. *Proc Natl Acad Sci USA* 106(21):8507–8512.

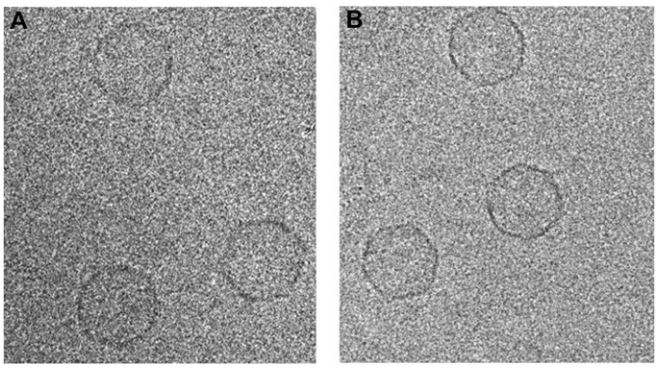

Fig. S6. Cryomicrographs of purified capsids carrying $gp15_{R5ER8E}$ (*A*) and $gp15_{R73E}$ (*B*). Their composition is shown in Fig. 3*B*.

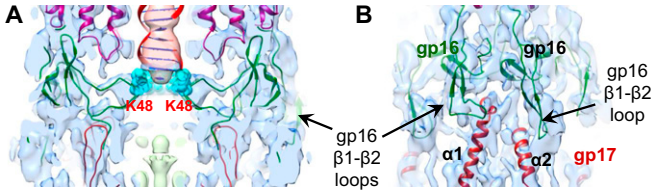

Fig. S7. Gp16 interactions with DNA and with gp17. (*A*) Region of the HTIFull reconstruction superimposed on the gp16 fit with the side chains of K48 residues forming a basic ring positioned underneath the DNA density end (cyan spheres). (*B*) Two subunits of gp16 interact with helices α1 and α2 of one gp17 subunit (dark red) as viewed from outside the HTIFull complex.

Table S1. Plasmids used in this work

Plasmid*	Parental plasmid	Mutation gp15*	Mutation gp16*
pCC6 (genes *14,15,16*)	pHP13 vector engineered with promoter $P_{N25/0}$ (1)	None	None
pCC40 (genes *13*, *14*, *15*, *16*)	pCC6	None	None
pCC43 (genes *14,15,16*)	pCC6	R5E	
pCC45 (genes *14,15,16*)	pCC6	R5E R8E	
pCC46 (genes *14,15,16*)	pCC6	R73E	
pCC51 (genes *14,15,16*)	pCC6	R8E	
pPT322 (genes *14,15,16*)	pCC6	Replacement of gp15 C-ter $R_{98}KMAR_{102}$ by MAG	
pCC36 (genes *14,15,16*)	pCC6		Q51C
pCC39 (genes *14,15,16*)	pCC6		Q43C Q51C
pCC42 (genes *13*, *14*, *15*, *16*)	pCC40		Q43C Q51C
pCC48 (genes *13*, *14*, *15*, *16*)	pCC40		Q51C
pCC49 (genes *14,15,16*)	pCC6		Q43C Y47C Q51C
pCC50 (genes *13*, *14*, *15*, *16*)	pCC40		Q43C Y47C Q51C

*The SPP1 genes cloned in each plasmid and mutations engineered in genes *15* or *16* are listed.

1. Isidro A, Henriques AO, Tavares P (2004) The portal protein plays essential roles at different steps of the SPP1 DNA packaging process. *Virology* 322(2):253–263.

Table S2. Oligonucleotides used in this study

Name	Sequence	Restriction sites	Mutations in gp15
13CS0Pst	AGGT*CTGCAG*CCCaaaggagAAATTAACATGGC	PstI	
14fw-PstI	CAA*CTGCAG*GCTAACGCCTCTCTTTTTTTGaaaagg	PstI	
16rev-HindIII	GCCC*AAGCTT*GCTAGGGTCGATTctaACCGACATAAGC	HindIII	
15 MAG forward	CGATCTTGAAAAAACTCAATCCATAC ATGGCG**G**GGtaaGATGTACG		Replacement of gp15 C-ter $R_{98}KMAR_{102}$ by MAG
15 MAG reverse	CGTACATCttaCC**C**CGCCATGTATGGATTGAGTTTTTTCAAGATCG		Replacement of gp15 C-ter $R_{98}KMAR_{102}$ by MAG

Sequences homologous to SPP1 DNA are underlined. Relevant translation signals are in lowercase letters. Mutations in the genes sequence are in bold. Relevant restriction sites are in italics. Oligonucleotides used for site-directed mutagenesis of genes *15* and *16* (not shown) were designed according to the instructions of the QuikChange Site-directed Mutagenesis Kit from Stratagene.

Table S3. Cross-correlation values for the gp6, gp15, gp16, and gp17 flexible fits in the electron densities of HTIEmpty and HTIFull reconstructions

Atomic coordinates	HTIEmpty	HTIFull
gp6 rigid body fit (X-ray)	0.797	0.809
gp6 empty	**0.840**	0.820
gp6 full	0.828	**0.830**
gp15 empty	**0.838**	0.813
gp15 full	0.806	**0.817**
gp16 empty dimer*	**0.844**	0.835
gp16 full dimer*	0.839	**0.848**
gp17 empty	**0.811**	0.760
gp17 full	0.781	**0.795**

The final flexible fittings were tested for the two maps. Fits with the best correlation are shown in bold.
*gp16 dimers were used for fitting because the gp16 dodecamer has sixfold symmetry in HTIEmpty and HTIFull.