Virology 464-465 (2014) 55–66

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Three-dimensional reconstructions of the bacteriophage CUS-3 virion reveal a conserved coat protein I-domain but a distinct tailspike receptor-binding domain

Kristin N. Parent [a,*], Jinghua Tang [a], Giovanni Cardone [a], Eddie B. Gilcrease [b], Mandy E. Janssen [a], Norman H. Olson [a], Sherwood R. Casjens [b,**], Timothy S. Baker [a,c,***]

[a] Department of Chemistry & Biochemistry, University of California, San Diego, La Jolla, CA 92093-0378, United States
[b] University of Utah School of Medicine, Division of Microbiology and Immunology, Department of Pathology, Salt Lake City, UT 84112, United States
[c] University of California, San Diego, Division of Biological Sciences, La Jolla, CA, 92093, United States

ARTICLE INFO

Keywords:
Bacteriophage
CUS-3
Virion structure
Cryo-electron microscopy

ABSTRACT

CUS-3 is a short-tailed, dsDNA bacteriophage that infects serotype K1 *Escherichia coli*. We report icosahedrally averaged and asymmetric, three-dimensional, cryo-electron microscopic reconstructions of the CUS-3 virion. Its coat protein structure adopts the "HK97-fold" shared by other tailed phages and is quite similar to that in phages P22 and Sf6 despite only weak amino acid sequence similarity. In addition, these coat proteins share a unique extra external domain ("I-domain"), suggesting that the group of P22-like phages has evolved over a very long time period without acquiring a new coat protein gene from another phage group. On the other hand, the morphology of the CUS-3 tailspike differs significantly from that of P22 or Sf6, but is similar to the tailspike of phage K1F, a member of the extremely distantly related T7 group of phages. We conclude that CUS-3 obtained its tailspike gene from a distantly related phage quite recently.

Introduction

Tailed bacteriophages are the most abundant "organisms" in Earth's biosphere (Whitman et al., 1998; Wommack and Colwell, 2000). In addition to the immense numbers of these virions, they exhibit vast genetic diversity. Indeed, thousands of tailed phage and prophage genomes have been sequenced, but it is not uncommon for more than half of the genes in a newly sequenced tailed phage genome to have no recognizable homolog in the sequence database (Rohwer, 2003; Hatfull and Hendrix, 2011). No doubt, this enormous diversity reflects at least in part their early appearance on Earth and the resulting long time available for their divergence. Tailed phages have host cell adsorption organelles (tails) that have one of three different morphologies, and nearly complete icosahedral heads (interrupted at one vertex by the presence of the portal and associated tail components) that encapsidate the dsDNA chromosome (Hendrix and Casjens, 2005). The only three tailed phage genes for which homologs are known to be universally present encode (1) the coat protein (CP, or major capsid protein) that assembles into the icosahedral head shell, (2) the portal protein that forms the conduit or portal in that shell through which DNA is packaged and released, and (3) the large terminase subunit, TerL, which is the ATPase molecular motor that pumps DNA into the preformed CP shell or procapsid (Casjens, 2005). However, the extant diversity of these proteins is such that some members of each of these three protein families have amino acid sequences that are not recognizably similar to other members of their group.

X-ray crystal structures of three moderately, distantly related phage portal proteins show nearly identical folds in their core domains (Simpson et al., 2000; Lebedev et al., 2007; Olia et al., 2011), and the same is true for the four TerL protein structures that have been determined (Sun et al., 2008; Smits et al., 2009; Roy and Cingolani, 2012; Zhao et al., 2013). Analysis of these structures strongly supports the idea that all portal proteins are homologs, as are all TerL proteins. Although the CPs are less

* Corresponding author. Present address: Department of Biochemistry & Molecular Biology, Michigan State University, 603 Wilson Road, 519 Biochemistry, East Lansing, MI 48824, United States. Tel.: +1 517 432 8434; fax: +1 517 353 9334.
** Corresponding author at: University of Utah School of Medicine, Division of Microbiology and Immunology, Department of Pathology, Rm 2200 EEJMRB Salt Lake City, UT 84112, United States. Tel.: +1 801 581 5980; fax: +1 801 585 2417.
*** Corresponding author at: Department of Chemistry & Biochemistry, University of California, San Diego, 9500 Gilman Drive MC-0378, La Jolla, CA 92093-0378, United States. Tel.: +1 858 534 5845; fax: +1 858 534 5846.
E-mail addresses: kparent@msu.edu (K.N. Parent), sherwood.casjens@path.utah.edu (S.R. Casjens), tsb@ucsd.edu (T.S. Baker).

http://dx.doi.org/10.1016/j.virol.2014.06.017
0042-6822/

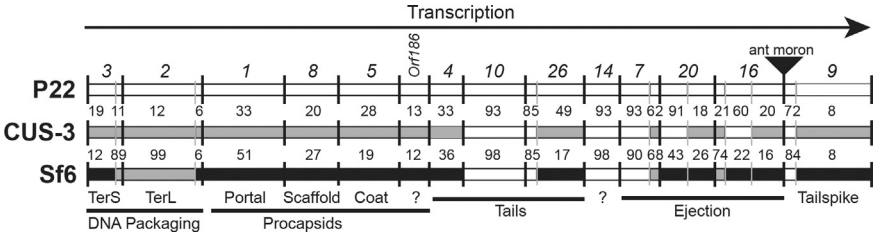


Fig. 1. Relationships among the CUS-3, P22 and Sf6 virion assembly genes. The fourteen genes in the virion assembly gene cluster are shown (not precisely to scale) for the three phages indicated on the left. Top, the P22 gene names are indicated; the "ant moron" is a variable insert of several non-virion assembly genes that is present in P22 and CUS-3 but not Sf6 (Pedulla et al., 2003; Casjens et al., 2004; King et al., 2007). Bottom, the functions of the genes are indicated; the roles of gene *14* and orf186 are not well understood, and neither gene is essential under normal laboratory conditions. In each map black vertical lines separate the genes and thin gray lines mark intragenic mosaic boundaries that are evident from the comparison of these three phages (see Casjens and Thuman-Commike, 2011), and the same shading in vertically aligned sections indicates $>60\%$ amino acid sequence identity. The non-italicized numbers between the gene maps indicate the percentage of amino acid sequence identity of the proteins encoded by the genes immediately above and below as determined by DNA Strider (Douglas, 1994).

well conserved than portal and TerL, the crystal structure of *Escherichia coli* phage HK97 (Wikoff et al., 2000; Helgstrand et al., 2003) is known and sub-nanometer cryo-electron microscopic (cryoEM) three-dimensional (3D) reconstructions of several tailed-phages show a similar protein fold, suggesting that they are distant relatives of one another. These include P22 (Parent et al., 2010; Chen et al., 2011), Sf6 (Parent et al., 2012b), T7 (Agirrezabala et al., 2007), T4 (Fokine et al., 2005b), λ(Lander et al., 2008), ϕ29 (Morais et al., 2005), T5 (Effantin et al., 2006), SIO-2 (Lander et al., 2012), ε15 (Jiang et al., 2008), 80α (Spilman et al., 2011), P-SSP7 (Liu et al., 2010), C1 (Aksyuk et al., 2012) and ϕRSL1 (Effantin et al., 2013). In each of these reconstructions the basic polypeptide fold of the major CP closely resembles that of HK97. The recent 3.5-Å structure of the phage BPP-1 CP indicates that its structure has the essentially the same fold as HK97, but the polypeptide chain connectivity (topology) in the structure is a permutation of that of HK97 (Zhang et al., 2013). The number of CP molecules in the heads of known tailed phages varies from 235 in ϕ29 (Morais et al., 2005) to 1615 in phage ϕKZ and others (Fokine et al., 2005a; Skurnik et al., 2012; Effantin et al., 2013), but despite the large distribution of head sizes that this leads to, it appears quite certain that all tailed phage CPs adopt an HK97-like fold or the related BPP-1-like fold. We note that the Herpes simplex I virus coat and large terminase proteins have folds similar to their tailed phage counterparts (Baker et al., 2005; Selvarajan Sigamani et al., 2013), which strongly suggests that tailed phage and herpesvirus CPs share some common ancestry. The CPs of these viruses have diverged to the point where many of them are no longer convincingly related by amino acid sequence similarity.

The short tailed phage CUS-3, the subject of this study, is a member of a group of phages that is typified by the well-studied phage P22. The virion assembly genes of 151 members of this phage group have been compared (Casjens and Thuman-Commike, 2011 S. Casjens, unpublished), and although these virions are all built from largely homologous sets of twelve proteins, their CPs fall into three distinct classes that exhibit only 10–30% amino acid sequence identity. These three coat protein types are represented by *Salmonella* phage P22, *Shigella* phage Sf6, and *E. coli* phage CUS-3. P22, Sf6, and CUS-3 have similar genome sizes (41.7, 39.0 and 40.2 kbp, respectively), genome organization, and gene regulation during lytic growth and lysogeny (Pedulla et al., 2003; Casjens et al., 2004; King et al., 2007). Their virions are virtually indistinguishable by negative stain electron microscopy (Botstein et al., 1973; Lindberg et al., 1978; King et al., 2007). Although the corresponding genes of the three phages have, for historical reasons, been assigned different names (Botstein et al., 1973; Casjens et al., 2004; King et al., 2007), to simplify discussion, we hereafter refer to the virion assembly genes and proteins of Sf6 and CUS-3 using the names of the better-studied P22 homologs (Fig. 1).

CUS-3, P22, and Sf6 have diverged to a point where their virion assembly gene clusters exhibit only a few regions of recognizable nucleotide sequence similarity; however, comparisons of their virion assembly proteins show greater sequence similarity than can be detected in nucleotide sequence comparisons. This similarity in protein sequences is irregular. For example, the CUS-3 and Sf6 TerL proteins are very similar, whereas other encoded proteins such as the C-terminal receptor-binding domains of these phages' tailspikes are not recognizably related in amino acid sequence (Fig. 1). Diversity of tailed phages is increased beyond simple linear divergence by means of the extensive horizontal exchange of genetic information, which produces "mosaic genomes" (Casjens et al., 1992; Hendrix, 2002; Casjens, 2005). The large P22-like phage group that includes these three coat protein subtypes has proven to be fertile ground for the examination of past horizontal exchanges of virion assembly genes (Casjens and Thuman-Commike, 2011; Leavitt et al., 2013a; Leavitt et al., 2013b). However, many questions remain about whether all phage genes are free to participate in random horizontal transfer or whether functionally-related sets of genes that encode interacting proteins co-evolve to a point where they can no longer be separated by such exchange events. The relationships among the virion assembly proteins encoded by CUS-3, P22, and Sf6 clearly show that the assembly gene clusters of these phages are mosaically related.

As discussed above, current evidence suggests that the CPs of all tailed phages have the same fold, but they nonetheless vary widely in amino acid sequence. The CPs of different phages employ different "accessory" domains that embellish the core structure of the HK97-like fold. For example, P22 and Sf6 carry an external domain that is not present in the other tailed phage coat proteins whose structures are known (Parent et al., 2012a; Parent et al., 2012b). This domain was previously called "telokin-like" based on structural similarity, but recent high-resolution NMR data has led to a change in nomenclature (Rizzo et al., 2014), and as such we will refer to this domain as the "I-domain". We sought to ascertain whether the CUS-3 CP also includes an I-domain and thus determine whether it is a defining characteristic that is common among all three branches of the P22-like phage group. We report here 6.8-Å resolution, icosahedrally-averaged and 20-Å resolution asymmetric structures of phage CUS-3 virions and compare them to previously determined P22 and Sf6 virion structures (Tang et al., 2011; Parent et al., 2012b).

Results and discussion

Icosahedral cryoEM reconstruction of the CUS-3 virion

CUS-3 virions were purified from phage-infected *E. coli* (see Materials and methods) and were subsequently vitrified and imaged by standard cryoEM methods (Baker et al., 1999). Image

preprocessing and icosahedral reconstruction methods were performed as previously described (Parent et al., 2012b). In brief, 419 transmission electron micrographs were recorded with a Direct Electron DE-12 camera (Fig. 2), and 7766 particle images were included in the final, icosahedrally-averaged 3D reconstruction that was estimated to have a resolution of 6.8 Å (Table 1). Density for components in the virion that do not strictly follow icosahedral symmetry, such as the portal protein complex, the ejection proteins, and the tail machine, was greatly reduced and is thus essentially invisible in this cryo-reconstruction as a result of the icosahedral averaging that was imposed during image processing.

The gross morphology of the CUS-3 virion CP shell is quite similar to those of P22 (EMDB ID 5232) and Sf6 (EMDB ID 5728) (Fig. 3A). All three exhibit a $T=7$ arrangement of coat subunits (the handedness of the CUS-3 capsid structure was not determined in this study, but it is likely to be *laevo,* the same as P22), and the shells have similar maximum outer diameters (~690 Å). The CPs in all three virions are organized as oligomers ("capsomers"), 11 of which are pentamers ("pentons" at the 5-fold vertices; one of the 5-fold vertices in the unaveraged structure is occupied by the portal ring and tail machine, see below) and 60 of which are hexamers ("hexons"). The dsDNA genome is tightly packed inside the CUS-3 capsid within a series of discrete, concentric shells (Fig. 3C) in a manner similar to that seen in P22, Sf6, and other tailed dsDNA phages and herpesviruses (reviewed by Johnson and Chiu, 2007; Casjens and Molineux, 2012).

The protein portion of the CUS-3 virion cryo-reconstruction contains many tube-shaped density features, consistent with the presence of α-helical secondary structural elements. These and other clearly defined features indicate the coat subunits of CUS-3, like those of Sf6 and P22, have an HK97-like core fold (Fig. 4A, See Supplementary material Movie M1). In addition, all three have a distinct protrusion at the same location on the outer surface. The NMR structure of the P22 I-domain (residues 222–345, PDB ID 2M5S) and a P22 model of the HK97-like spine helix were fit into a CP subunit extracted from each of the three density maps using Chimera (Fig. 4).

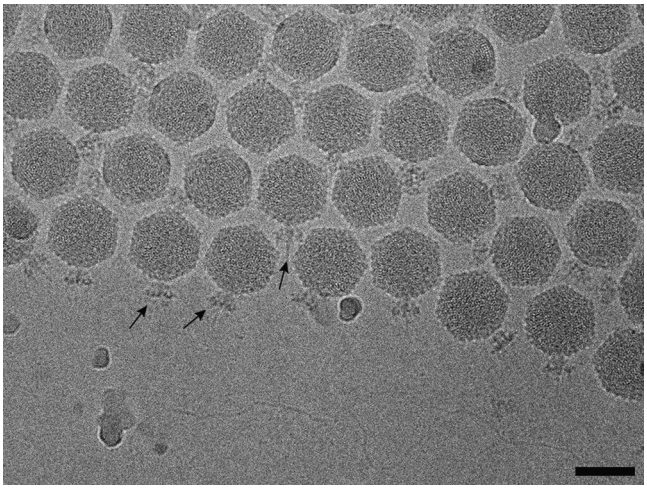


Fig. 2. A representative cryoelectron micrograph of an unstained, vitrified sample of CUS-3 virions recorded on a DE-12 detector. Black arrows highlight a few phage tails. Scale bar is 500 Å.

Recent evidence points to a clear role for the I-domain during proper coat protein folding prior to assembly (Suhanovsky and Teschke, 2013) and also an electrostatic interface that might be important for procapsid stability (Rizzo et al., 2014). The I-domain has a 6-stranded, antiparallel Greek key, β-barrel with a large "D loop" (residues 239–254) radiating perpendicularly downwards. Density corresponding to the barrel domain is evident in all three maps, and although the D loop is clearly defined in both CUS-3 and P22, it is not well defined in the Sf6 structure. Overall, comparison of the density maps shows that the I-domain is present in all three classes of P22-like phages, and is most similar between P22 and CUS-3. The fitting into the cryo-EM density maps indicates the I-domain occurs in a similar position in all three CPs. A BLAST search comparison of I-domain sequences between CUS-3 and P22 revealed that there is 40% similarity between these domains. There was no convincing hit for the CUS-3 I-domain within the Sf6 coat protein. In summation, our data implies that the I-domain was likely inserted into the HK97 fold in an ancestral phage that was common to P22, Sf6, and CUS-3.

Although the cryo-reconstructions of the CUS-3, P22, and Sf6 virion coat protein shells share several overall common structural features mentioned above, closer inspection reveals some notable differences. These are evident in radial density projections of the virions (Fig. 5). Similarities are most obvious in regions of intercapsomer contacts and all three capsids appear quite comparable at lower radii where the HK97-like core fold resides. However, at higher radii, differences in these phage structures are evident. The I-domains of CUS-3 and P22 adopt similar morphologies, but differ somewhat from that in Sf6. To emphasize this, at $r=290$ Å, a hexon is encircled in red for each density map (Fig. 5). This view shows that the I-domains do not contact neighboring subunits in CUS-3 at this radius, similar to what is seen in P22 (Parent et al., 2010). However, *within* each Sf6 hexon there is a distinct bridging among the I-domains of the three CPs that are nearest to the penton (Parent et al., 2012b). At $r=315$ Å, an individual subunit of each penton is highlighted to show that the protruding I-domain has very similar shape and five-fold extension in CUS-3 and P22 pentons that differ somewhat from those in Sf6 pentons.

Asymmetric cryoEM reconstruction of the CUS-3 virion

The structure of the CUS-3 virion was also reconstructed without imposing any symmetry during the image processing from a set of image data collected on film. This asymmetric reconstruction (Fig. 6A) was computed from 4819 particle images and reached an estimated 20-Å resolution (Table 1). It includes reliable structural details for the dodecameric portal complex (gp1) and the tail machine, comprised of twelve gp4 subunits, six gp10 subunits, six tailspikes (each a trimer of gp9), and the tail

Table 1
3D reconstruction statistics.

Symmetry[1]	Image media	Micrographs	Particles[2]	Defocus (μm)[3]	Pixel size (Å)	Resolution (Å) [4]
532	DE-12 camera	419	7766	0.60–4.08	1.35	6.8
1	Kodak SO-163 film	153	4819	0.15–3.66	3.28	20

1. Point group symmetry applied during image processing procedures and 3D reconstruction.
2. Number of boxed particles used in each image reconstruction.
3. Range of objective lens underfocus settings for micrographs.
4. Estimation based on $FSC_{0.5}$ criterion (van Heel and Schatz, 2005).

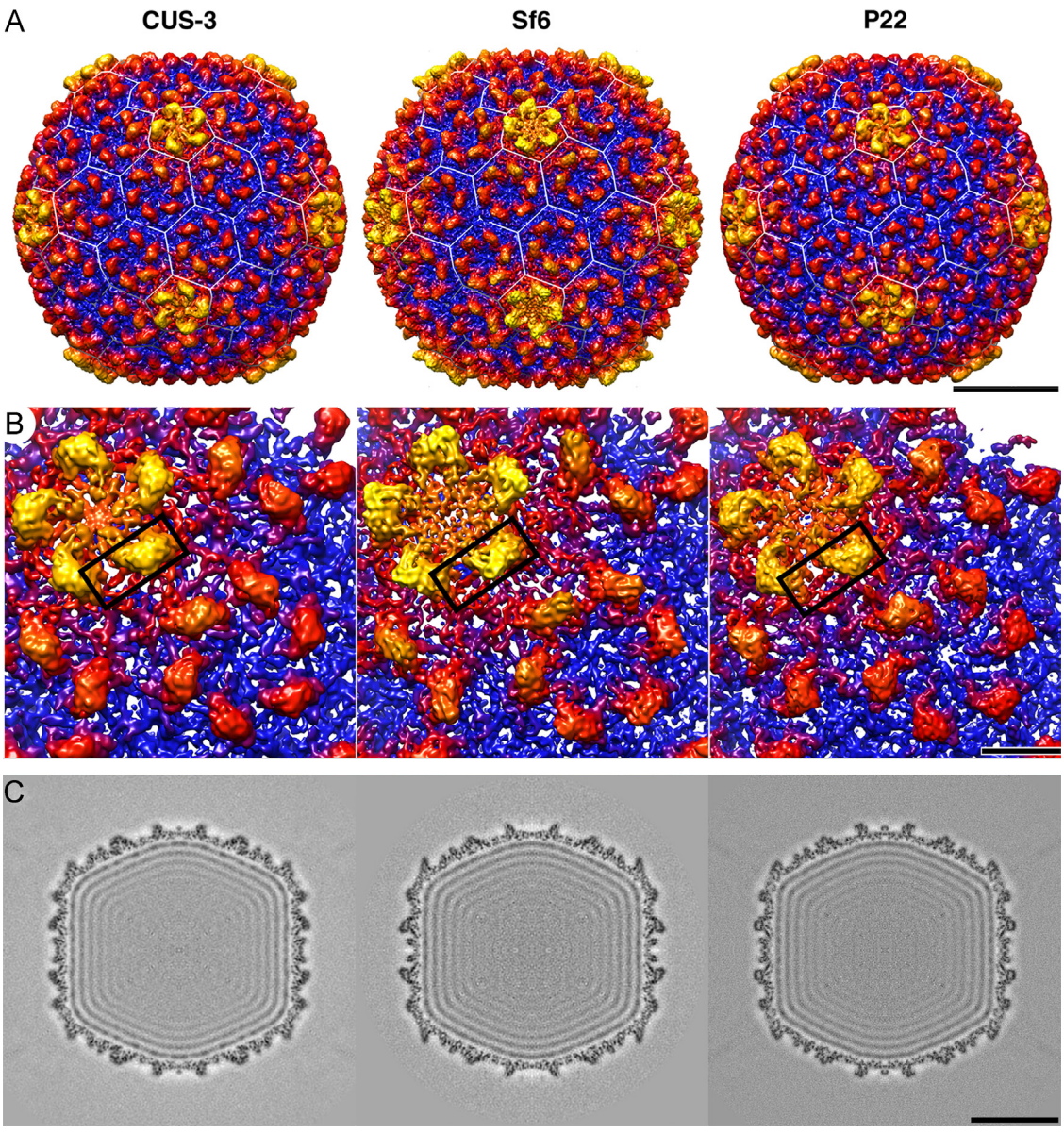


Fig. 3. Comparison of icosahedrally-averaged, 3D reconstructions of CUS-3, Sf6, and P22 virions. (A) Radially, color-coded surface views along a two-fold axis of symmetry of the capsids of CUS-3, Sf6, and P22 (color scheme transitions from dark blue at low radii to red and orange and finally yellow at highest radii). (B) Same as (A) showing close-up views of one penton and an adjacent hexon from each virion cryo-reconstruction; the spine helix of the CP enclosed within a black box. (C) Equatorial density sections (1-pixel thick) of the three reconstructions show the densely packed, concentric shells of dsDNA genome on the inside of the capsids. The scale bars in A & C denote 200 Å, while the scale bar in the enlarged view (B) denotes 50 Å.

needle (a trimer of gp26) (Tang et al., 2011). A central, planar density section through the CUS-3 virion reconstruction shows the organization of virion components (Fig. 6B) as deduced from the locations of the P22 homologs. These include the CP shell (415 copies of gp5), the concentric layers of densely packed dsDNA genome, the portal complex at the unique tail vertex, and the tail machine.

We fit known atomic structures of phage P22 orthologs into the CUS-3 virion structure (Fig. 6C). These included the crystal structures of the complete dodecameric portal complex (PDB ID 3LJ4; (Olia et al., 2011), the trimeric tail needle, (PDB ID 2C91; (Olia et al., 2007), and the P22 gp4 tail protein (PDB ID 1VT0; (Olia et al., 2011). All of these virion protein components, which have similar masses in CUS-3, Sf6, and P22, were each fit without remodeling into the CUS-3 density map. The only tail machine protein omitted from the fitting is the P22 gp10 homolog, for which no atomic-resolution structure is currently available. The CUS-3 portal structure closely mimics the one observed in Sf6 and P22 virions, and it includes the C-terminal, long barrel domain that was previously hypothesized to be present in CUS-3 as well as many other tailed phages (Tang et al., 2011). In addition, a single, protruding tail needle (postulated to make initial contact with the host outer membrane during infection as observed by cryo-tomograms of Sf6 interacting with membranes (Parent et al., 2014)) was apparent in each density map. The needles of CUS-3 and P22 have similar, long and slender morphologies. The obvious difference is the distal "knob" domain that is only present in Sf6. For the P22-like phages, this knob domain is highly conserved when present, and it is encoded by about 40% of the members of this group of phages whose genomes have been sequenced (Leavitt et al., 2013b).

The sequence of the receptor-binding domain of the CUS-3 tailspike is unrelated to that in P22 or Sf6 (8.0% and 8.4% identity, respectively), and furthermore, the CUS-3 domain is about 50%

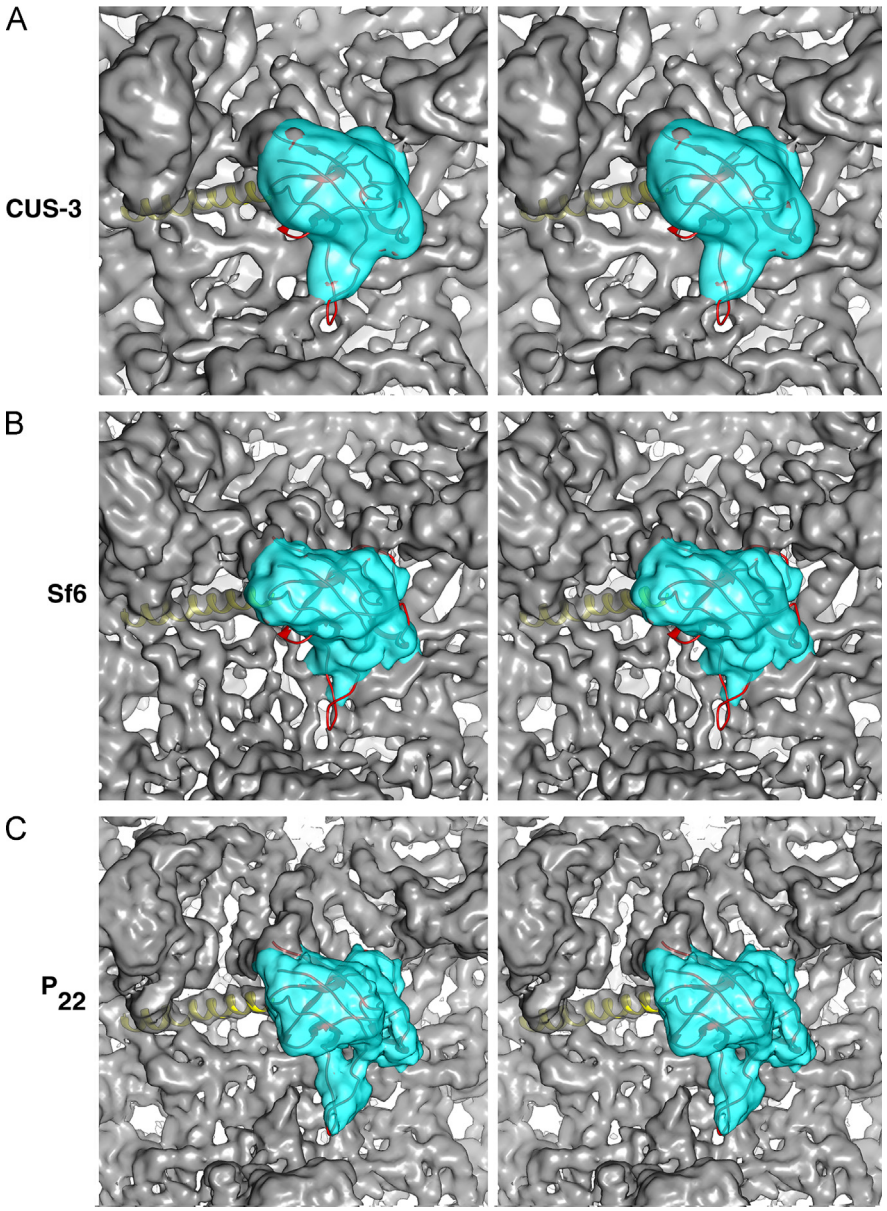
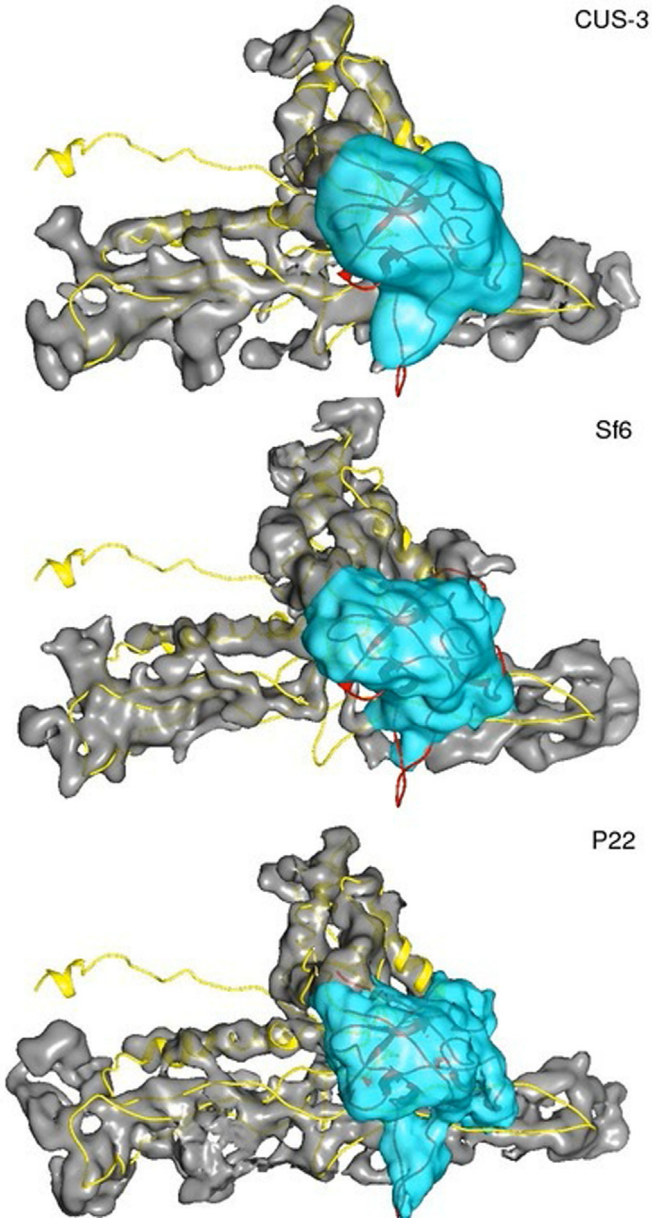


Fig. 4. I-domain structure in the P22-like CPs. Stereo views of the docked I-domain (PDB ID 2M5S) and spine helix of P22 in each of the three virion density maps (yellow ribbon=spine helix; red ribbon=P22 I domain). The cyan density envelop corresponds to the I domain density in each structure.

larger than the parallel P22 and Sf6 domains (King et al., 2007; Casjens and Thuman-Commike, 2011). The CUS-3 tailspikes have a significantly different shape than those of P22 or Sf6 (Fig. 7). The CUS-3 tailspike has a more massive "midsection" and its distal, putative receptor-binding domain has a narrower tip than Sf6 or P22. The host receptor utilized by CUS-3 is the capsular, poly-sialic acid polysaccharide that defines the *E. coli* serotype K1 (King et al., 2007), and its tailspike receptor-binding domain is 63% identical to the tailspike of phage K1F. K1F belongs to the T7 phage group, which is only extremely distantly related to the P22 phage group. It also infects K1 *E. coli* and utilizes its poly-sialic acid surface chain as its receptor (Scholl and Merril, 2005). The X-ray crystal structure of the K1F tailspike receptor-binding domain has been determined (Stummeyer et al., 2005), and it has a very different tertiary structure from that of the P22 and Sf6 tailspike receptor-binding domains. Given the high sequence similarity that exists between the CUS-3 and phage K1F tailspikes, we docked the structure of the receptor-binding domain of a K1F tailspike trimer (PDB ID 3GVK) as a rigid body into the cryoEM density of the CUS-3 virion tailspike. The K1F tailspike trimer fits unambiguously and neatly within the density envelope of the CUS-3 tailspike, which strongly supports the notion that these two tailspikes have closely similar structures (Fig. 7A). The X-ray crystal structures of P22 (PDB 2XC1) and Sf6 (PDB ID 2VBK) were also docked into their cognate virions (Fig. 7B and C) to highlight the different tailspike morphologies and to show that the accuracy of each fit is comparable in all three cases.

Horizontal exchange of tailspike protein genes

Tailspike and tail fiber genes in tailed phage are the most frequent of the virion assembly genes to be horizontally transferred among different phages (Haggard-Ljungquist et al., 1992; Sandmeier et al., 1992; Casjens and Thuman-Commike, 2011). Among 151 P22-like phage genomes that we have examined, 34 have CUS-3 type CPs and eleven of these encode a closely related homolog of the CUS-3 tailspike (these 151 contain all the characterized P22-like phages and 140 P22-like prophages that



Movie M1. Comparison of subunits for P22, Sf6, and CUS-3 reveal the I-domain is present in each phage. Individual coat protein subunits were extracted from each density map. Ribbon models of the HK97-like core and I-domain (yellow and red, respectively) were fitted into each subunit using Chimera. The density envelope corresponding to the I-domain in each subunit is colored in cyan. Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.virol.2014.06.017.

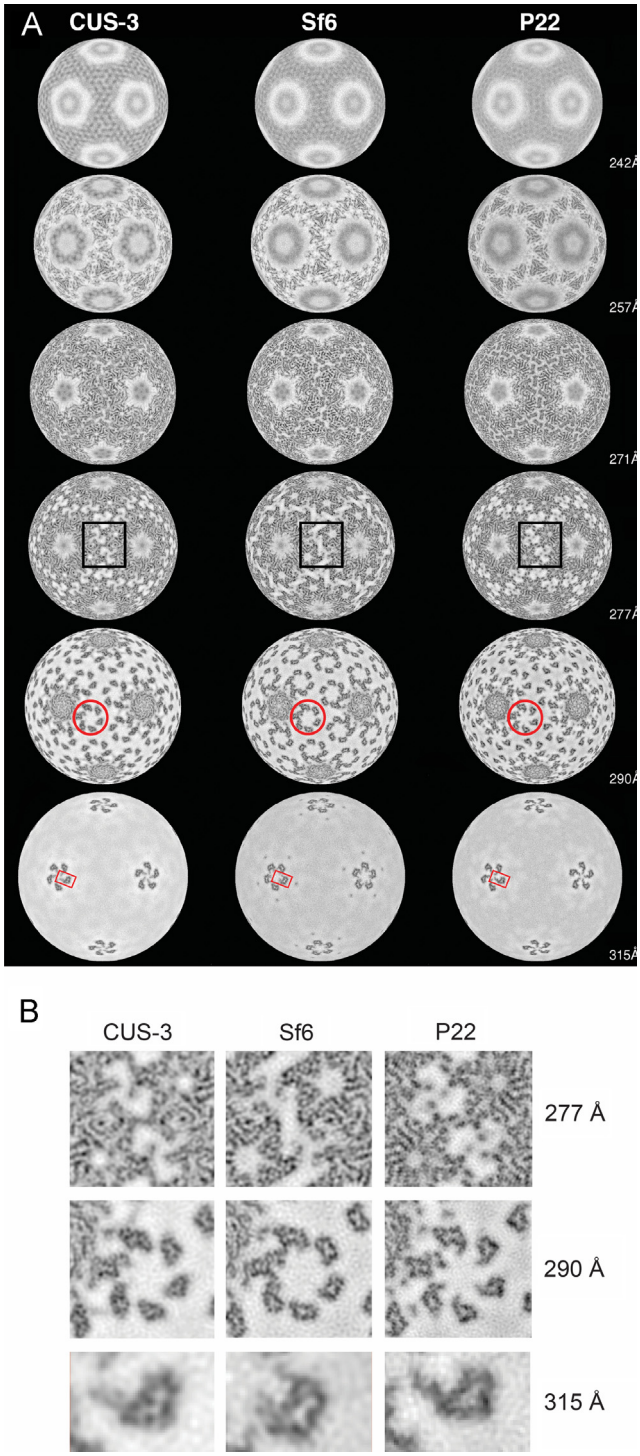


Fig. 5. Comparisons based on radial density projections of the 3D reconstructions of the phage CUS-3, Sf6, and P22 virions. A) At $r=$ 277 Å the contacts between neighboring hexons (highlighted by black boxes) show a larger separation of the rings of five hexons at this level of the Sf6 capsid. At $r=$ 290 Å individual hexons (highlighted with red circles) exhibit bridging between neighboring I-domains in Sf6, which does not occur in CUS-3 or P22. At $r=$ 315 Å, the I-domain of the penton subunit assumes different morphologies in the different phages. B) Enlarged views of the boxed and circled areas shown in A. Note, the enlarged view at $r=$ 277 and $r=$ 299 are at the same scale, while the enlarged view at Å $r=$ 315 Å is larger.

were chosen at random from those available) (Casjens and Thuman-Commike, 2011; S. Casjens, unpublished). The other 23 each carry one of nine different types of tailspike sequence whose receptor-binding domain sequences are not recognizably related to each other or to those of P22 or Sf6. Among the 151 phages, 117 have P22 or Sf6 type CPs and none of these encode a CUS-3 type tailspike. The eleven P22-like phages with CUS-3 tailspike homologs, were all discovered as prophages in *E. coli* isolates (see legend to Fig. 8). CUS-3, a prophage in the genome of serotype K1 *E. coli* strain RS218, is the only one known to encode a fully functional phage genome (King et al., 2007); the other ten are all very closely related to CUS-3 and are likely to be functional phages (Casjens and Thuman-Commike, 2011). These eleven *E. coli* isolates form a tightly clustered group in the *Enterobacteriaceae* family clade

analysis of Gillespie et al. (2011) on the PATRIC World-wide web site http://patricbrc.vbi.vt.edu/portal/portal/patric/Home.

Phage proteins of similar function but with no recognizable homology can be either (i) unrelated proteins in which one

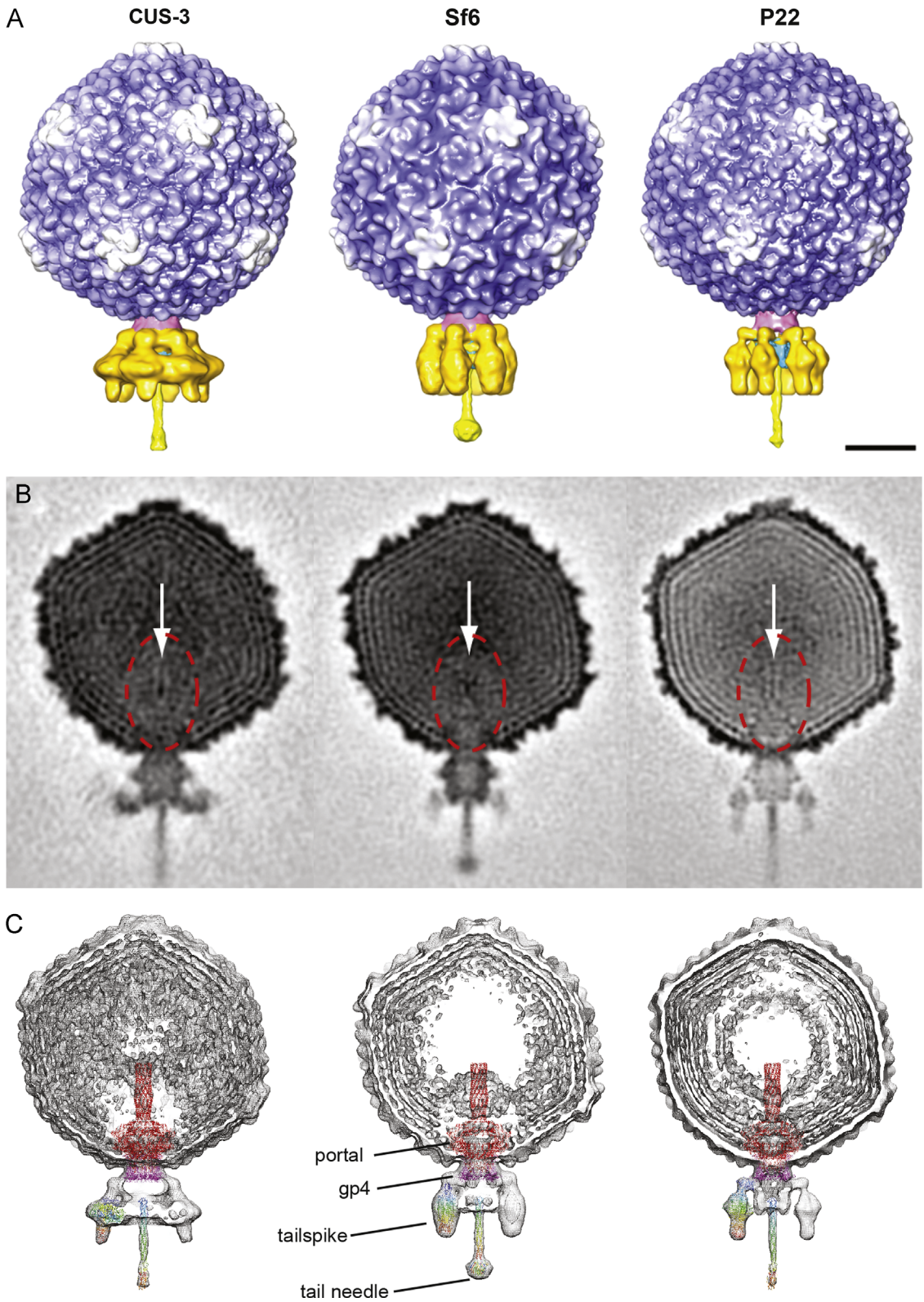


Fig. 6. Comparison of asymmetric 3D reconstructions of CUS-3, Sf6, and P22 virions. (A) Segmented, surface-rendered views of each virion with the head protein (P22, gp5) colored blue and the tail proteins colored gold (tailspike, P22 gp9), yellow (needle, gp26), and pink (gp4). (B) Central slice through the density map of each virion is shown in (A) with the portal structures highlighted with dashed red ovals and white arrows pointing towards linear stretches of dsDNA filling the distal end of the portal barrel. (C) Known crystal structures docked into the CUS-3 tail machine (see text). The scale bar denotes 200 Å.

was obtained from a distant source by horizontal gene transfer or (ii) homologous proteins that have diverged and no longer show any convincing sequence similarity. The tailspikes of these phages exemplify both of these situations. The O-antigen, polysaccharide-receptor-binding domains of P22 and Sf6 are known to have very similar polypeptide folds but no convincing sequence similarity (Steinbacher et al., 1994; Freiberg et al., 2003; Seul et al., 2014). The receptor-binding domain of the tailspike protein of CUS-3 is

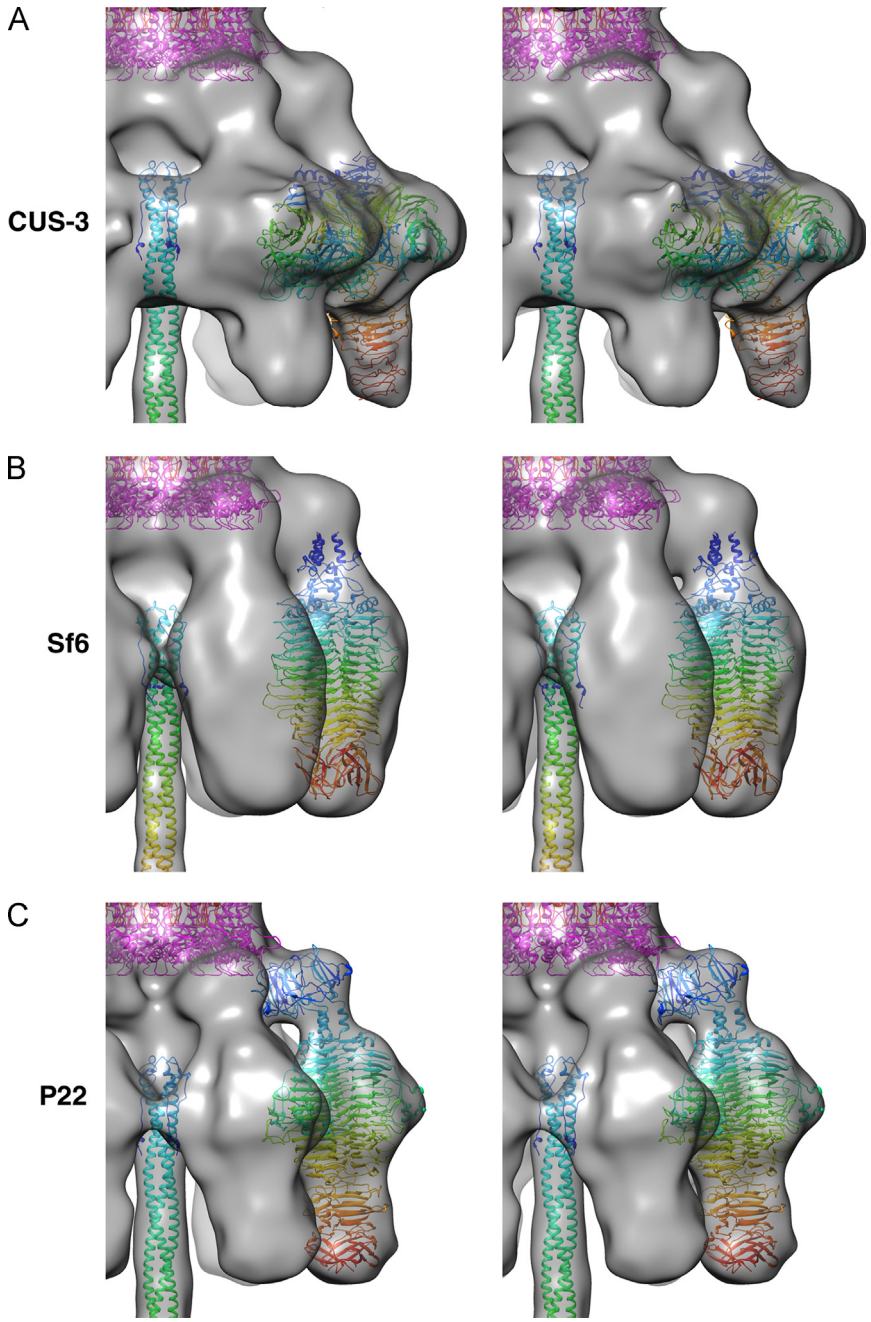


Fig. 7. Stereo views of the tailspikes in phages CUS-3, Sf6 and P22 virions. The portal and gp4 atomic structures from P22 are depicted in red and magenta at the top. Each protein chain in the trimeric tail needle and tail spike is colored in a rainbow scheme, transitioning from blue at the N-termini to red at the C-termini.

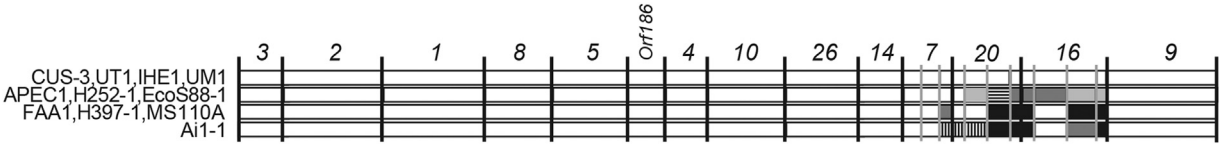


Fig. 8. Relationships among eleven CUS-3 type phages that encode a K1F-type tailspike. The virion assembly gene clusters are shown as in Fig. 1 (Haggard-Ljungquist et al., 1992; Sandmeier et al., 1992; Casjens and Thuman-Commike, 2011); mosaic sections with different patterns are >30% different from the others in encoded protein sequence. Phage P22 gene names are indicated at the top. The prophages, tailspike protein accession numbers, and host *E. coli* strains are as follows: prophage APEC1, tailspike protein accession number is YP_853451, bacterial strain APEC01 (Johnson et al., 2007)); UT1, ABE08102, UT189 (Chen et al., 2006); EcoS88-1, YP_002392174, S88 (Touchon et al., 2009); H252-1, EGB44572, H252; IHE1, YP_006101778, IHE3034 (Moriel et al., 2010); UM1, ADN70420, UM146 (Krause et al., 2011); FAA1, EEH87454, 3_2_53FAA; Ai1-1, EHG00336, A_i1; H397-1, EHN99038, H397; and MS110A, EFU47675, MS110-3. These strains are all serotype K1 except the last three, whose serotypes are not known. Since the sequence reports did not name the prophages, we gave them these tentative names to facilitate discussion.

similarly unrelated in sequence to those of P22 and Sf6, but it also has a different fold and therefore almost certainly was acquired by horizontal transfer.

All tailspike proteins in the P22-like phages, including that of CUS-3, have an ~113 a.a. N-terminal, head binding region, with about 50% sequence identity across these phages (Casjens and

Thuman-Commike, 2011). In P22, this region is a folded protein domain that is separate from the C-terminal portion (which includes the receptor-binding domain (Seul et al., 2014)) and contacts the virion via tail machine proteins gp10 and gp4 (Chang et al., 2006; Lander et al., 2006; Tang et al., 2011). Sequence comparisons show 62% identity between C-terminal receptor-binding domains of CUS-3 (868 a.a.) and K1F (851 a.a.). In contrast, their N-terminal head-binding domains are not recognizably related to one another. On the other hand, this N-terminal domain of the CUS-3 tailspike is 96% identical to *E. coli* phage HK620, whose C-terminal domain is unrelated to that of CUS-3 but is similar to Sf6 and P22 (Barbirz et al., 2008). Thus, the K1F-like tailspike C-terminal domain was clearly obtained by a rather precisely positioned fusion of that domain to the resident CUS-3 N-terminal domain.

Evidence for the recent nature of this latter acquisition is supported by two observations. First, the receptor-binding domains of the tailspikes of the eleven phages with CUS-3-like tailspikes are all $\geq$99.3% identical to one another. Second, this cluster of phage genomes shows considerably less overall mosaicism relative to one another than the CUS-3 coat-carrying subgroup as a whole (Fig. 8 and Casjens and Thuman-Commike, 2011). Most of the proteins encoded by the virion assembly gene cluster (homologs of P22 genes *3* through *16* and gene *9*) of these eleven phages are nearly identical. For example, their coat proteins are all $\geq$99.4% identical to one another. Thus, the CUS-3-type tailspike appears to have entered the CUS-3 lineage so recently that there has been little time for sequence divergence within the virion assembly gene cluster.

There is some genome mosaicism present within the virion assembly gene clusters of the eleven CUS-3-like phages that have K1F-related tailspikes, but only for the ejection protein genes *7, 16* and *20*. These proteins are not required for virion assembly but they are required for successful ejection of DNA into susceptible cells, and they are released from the virion with the DNA during injection (Israel, 1977). Our analysis has previously identified three, four, and four mosaic sections (exchangeable genetic regions) within these three genes, respectively (Casjens and Thuman-Commike, 2011). The rightmost section of gene *7* and all four sections of genes *20* and *16* show variability within the eleven phages that carry the CUS-3 type tailspike (P22 gene *9*) (Fig. 8). Not shown in Fig. 8 is the fact that all of these different gene *7, 16,* and *20* mosaic sectional sequences have very closely related homologs that are present in the ejection protein genes of one or more of the P22-like phages that have the P22 or Sf6 type CPs. These mosaic sections, present as four different patterns in the eleven phages in Fig. 8, have clearly not simply diverged in place within these CUS-3-like phages. The ejection proteins have no known homologs outside this phage group, so the mosaic variations are not likely to have been obtained by exchange with other phage types. Thus, if we make the plausible assumption that this tailspike entered this phage group from an outside source only once, the ejection protein gene variants of at least three of the four patterns must have been acquired by horizontal exchange from other members of the larger P22-like group after acquisition of the K1F type tailspike. We have previously noted that, in addition to tailspike, the ejection genes exhibit exceptionally high levels of diversity and mosaicism (Casjens and Thuman-Commike, 2011), and the above reasoning indicates that their rates of horizontal exchange are, like the tailspike genes, more rapid than the other virion assembly genes. It is not known why the ejection proteins are so diverse or why they exchange rapidly. However, since the ejection proteins and tailspikes are the most diverse proteins encoded by the P22-like virion assembly gene clusters, and both almost certainly interact with the host during DNA delivery into bacteria, evolutionary sparring between host and virus is likely to have driven up the level of diversity of these proteins and perhaps forced rapid generation of new combinations of mosaic sections.

We also note that homologs of the K1F/CUS-3 type tailspike receptor-binding domain are encoded by a number of very different phage types. Phage K1F is a lytic phage with short tail, and CUS-3 is a very different temperate phage with a short tail; although both are members of the *Podoviridae,* they are extremely distantly related outside their tailspikes. In addition to some other members of these two phage groups (ACG-C91, K1-5, and K1E, which are closely related to K1F (Stummeyer et al., 2006; Chibeu et al., 2012) and the CUS-3-like prophages, above), this type of tailspike is encoded by the genomes of phages K1-dep(1), K1-dep(4), ϕ92, and phAPEC8 (Bull et al., 2010; Schwarzer et al., 2012; Tsonos et al., 2012). Of these, the first two belong to the SETP3 type of lytic phages, which have long, noncontractile tails (*i.e.,* are members of the family *Siphoviridae*) (De Lappe et al., 2009), and the last two are members of the rv5-like type of large lytic phages with contractile tails (*Myoviridae*) (Kropinski et al., 2013). In addition, we have identified this type of tailspike in ε15-like prophages in the genome sequences of a number of *E. coli* strains (*e.g.,* tailspike accession no. CAR18780 in the K1 serotype *E. coli* strain IAI39 (Touchon et al., 2009)). Phage ε15 is representative of a disparate type of temperate phages. Thus these tailspikes are encoded by at least five, very different types of tailed phages that include phages with all three tail morphologies. In all of these cases there are group members that have other kinds of tailspikes, so horizontal exchange of this K1F/CUS-3 tailspike appears to be widespread among the tailed phages. Acquisition of these particular tailspikes was likely advantageous since it allowed these phages access to bacterial hosts with polysialic acid surface polysaccharides.

Conclusions

Our symmetric and asymmetric 3D reconstructions of the phage CUS-3 virion show that, despite substantial divergence among the proteins involved in the assembly of phage virions in the P22-like phage group, the structures of P22, Sf6, and CUS-3 virions and the shapes of their CPs are remarkably alike. They also clearly show that the CUS-3 CP, like those of P22 and Sf6, contains a similar shape and located I-domain. The fact that this domain is uniquely and universally present in virions of the large, P22-like phage group, strongly suggests that the CPs of all P22-like phages have diverged within this phage group. For this group of phages, the I-domain likely plays a strong role in directing CP folding, eliminating a requirement for extraneous chaperones like GroEL/ES (Suhanovsky and Teschke, 2013; Rizzo et al., 2014). The observed CP divergences are not the consequence of horizontal acquisition of distantly related CPs from other types of tailed phages. Instead, our asymmetric 3D cryo-reconstruction of the CUS-3 virion strongly supports the notion that a progenitor of CUS-3 and its close relatives obtained their K1 capsule-specific tailspike genes by horizontal exchange from a different phage type.

Materials and methods

Phage preparation

Phage CUS-3 was obtained from Dr. Eric Vimr, and it was propagated on *E. coli* EV36 (King et al., 2007). Virions were purified through two rounds of CsCl step gradient centrifugation (Earnshaw et al., 1976; Parent et al., 2012b).

Cryo-transmission electron microscopy

Small (3.5 μL) aliquots of purified CUS-3 (at $\sim 1 \times 10^{14}$ phage/mL or ~10 mg/mL) were vitrified and examined using established procedures (Baker et al., 1999). Samples were applied to holey Quantifoil grids that had been glow-discharged for ~15 s in an Emitech K350 evaporation unit. Grids were then blotted with Whatman filter paper for ~5 s, plunged into liquid ethane, and transferred into a precooled, FEI Polara, multi-specimen holder, which maintained the specimen at liquid nitrogen temperature. Micrographs were recorded either using Kodak SO-163 electron-image film or a Direct Electron DE-12 camera in an FEI Polara microscope operated at 200 keV and under minimal-dose conditions (~22 $e/Å^2$) at nominal magnifications of 39,000× (1.69 Å/pixel on film) and 31,000× (1.359 Å/pixel on the DE-12 camera). The DE-12 camera was configured as follows: 40 msec readout delay with 40 frames per second capture rate for a total exposure of 700 msec totaling 28 frames. This corresponds to a final dose of 35 $e^-/Å^2$, or 1.25 $e^-/Å^2$ per frame. The data collected on the DE-12 camera was under control of the automated software, Leginon (Potter et al., 1999). The programs RobEM (http://cryoEM.ucsd.edu/programs.shtm) and ctffind3 (Mindell and Grigorieff, 2003) were used to extract individual particles and estimate the level of defocus and astigmatism for each micrograph. To reconstruct the final map 15 frames from the 28 per image (the first two frames and the last nine were omitted); frames 3–17 were selected for processing, corresponding to a dose of 19 $e^-/Å^2$.

Icosahedral cryo-reconstructions of CUS-3 virion

Micrographs that exhibited minimal astigmatism and specimen drift were selected for further processing. A subset of 150 particle images was used as input to the random-model computation procedure to generate an initial 3D density map at ~25 Å resolution (Yan et al., 2007a). This map was then used to initiate determination and refinement of particle orientations and origins for the complete set of images using AUTO3DEM (Yan et al., 2007b). Phases and amplitudes of the particle structure factor data were corrected to compensate for the effects caused by the microscope contrast-transfer function (Bowman et al., 2002). The Fourier Shell Correlation criterion ($FSC_{0.5}$) was used to estimate the resolution of each reconstruction (Table 1) (van Heel and Schatz, 2005). Graphical representations were generated with the RobEM and Chimera visualization software package (Goddard et al., 2007).

Asymmetric cryo-reconstruction of CUS-3 virion

The strategy used to process the CUS-3 virion film images and compute an asymmetric reconstruction was similar to that used to derive an asymmetric reconstruction of the P22 (Lander et al., 2006) and Sf6 (Parent et al., 2012a) virions. First, all digitized micrographs of the Sf6 virion samples were binned 2×, which generated a pixel size of 3.28 Å. Then, we re-boxed all virion images within a box window that was large enough to assure that the tail of each particle was not excluded as was the case during the icosahedral processing strategy described above. Also, the origin and orientation parameters of each virion image were set to match those determined during the icosahedral processing. Next, we used the "ticos_equiv" option in AUTO3DEM (v4.01.07) to determine an initial estimate of the orientation that properly aligns the tail in each virion image to a common reference, which was the 3D cryo-reconstruction of P22 (EMDB ID 1220 (Lander et al., 2006), after being low-pass filtered to 50 Å). Next, iterative refinement of the origin and orientation of each virion image was performed in AUTO3DEM set to operate in an "asymmetric" mode, in which no symmetry is assumed or imposed on any of the data.

Accession numbers

Asymmetrically-averaged and icosahedrally-averaged CUS-3 virion density maps have been deposited to the Electron Microscopy Data Bank (IDs, EMD-5946 and EMD-5947, respectively).

Acknowledgments

We thank K. and J. Pogliano (UCSD) for access to equipment and laboratory space, R. Khayat and J. E. Johnson (Scripps Research Institute) for helpful discussions, and E. Vimr for the kind gift of phage CUS-3 and its host *E. coli* strain. We also thank L. Jin and Direct Electron, LP (San Diego, CA) for support and access to their demonstration DE-12 camera. Thanks to Carol Teschke (UConn) for helpful discussion. This work was supported in part by NIH, USA grants R37 GM-033050 and 1S10 RR-020016 to TSB, AI074825 to SRC, NIH fellowship F32A1078624 to KNP, and funds from UCSD and the Agouron Foundation to TSB to establish and support the UCSD cryoEM facilities.

References

Agirrezabala, X., Velazquez-Muriel, J.A., Gomez-Puertas, P., Scheres, S.H., Carazo, J.M., Carrascosa, J.L., 2007. Quasi-atomic model of bacteriophage T7 procapsid shell: insights into the structure and evolution of a basic fold. Structure 15, 461–472.

Aksyuk, A.A., Bowman, V.D., Kaufmann, B., Fields, C., Klose, T., Holdaway, H.A., Fischetti, V.A., Rossmann, M.G., 2012. Structural investigations of a *Podoviridae streptococcus* phage C1, implications for the mechanism of viral entry. Proc. Natl. Acad. Sci. USA 109, 14001–14006.

Baker, M.L., Jiang, W., Rixon, F.J., Chiu, W., 2005. Common ancestry of herpesviruses and tailed DNA bacteriophages. J. Virol. 79, 14967–14970.

Baker, T.S., Olson, N.H., Fuller, S.D., 1999. Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. Microbiol. Mol. Biol. Rev. 63, 862–922.

Barbirz, S., Muller, J.J., Uetrecht, C., Clark, A.J., Heinemann, U., Seckler, R., 2008. Crystal structure of *Escherichia coli* phage HK620 tailspike: podoviral tailspike endoglycosidase modules are evolutionarily related. Mol. Microbiol. 69, 303–316.

Botstein, D., Waddell, C.H., King, J., 1973. Mechanism of head assembly and DNA encapsulation in *Salmonella* phage P22. I. Genes, proteins, structures and DNA maturation. J. Mol. Biol. 80, 669–695.

Bowman, V.D., Chase, E.S., Franz, A.W., Chipman, P.R., Zhang, X., Perry, K.L., Baker, T. S., Smith, T.J., 2002. An antibody to the putative aphid recognition site on cucumber mosaic virus recognizes pentons but not hexons. J. Virol. 76, 12250–12258.

Bull, J.J., Vimr, E.R., Molineux, I.J., 2010. A tale of tails: Sialidase is key to success in a model of phage therapy against K1-capsulated *Escherichia coli*. Virology 398, 79–86.

Casjens, S., Hatfull, G., Hendrix, R., 1992. Evolution of dsDNA tailed-bacteriophage genomes. Semin. Virol. 3, 383–397.

Casjens, S., Winn-Stapley, D., Gilcrease, E., Moreno, R., Kühlewein, C., Chua, J.E., Manning, P.A., Inwood, W., Clark, A.J., 2004. The chromosome of *Shigella flexneri* bacteriophage Sf6: complete nucleotide sequence, genetic mosaicism, and DNA packaging. J. Mol. Biol. 339, 379–394.

Casjens, S.R., 2005. Comparative genomics and evolution of the tailed-bacteriophages. Curr. Opin. Microbiol. 8, 451–458.

Casjens, S.R., Molineux, I.J., 2012. Short noncontractile tail machines: adsorption and DNA delivery by podoviruses. Adv. Exp. Med. Biol. 726, 143–179.

Casjens, S.R., Thuman-Commike, P.A., 2011. Evolution of mosaically related tailed bacteriophage genomes seen through the lens of phage P22 virion assembly. Virology 411, 393–415.

Chang, J., Weigele, P., King, J., Chiu, W., Jiang, W., 2006. Cryo-EM asymmetric reconstruction of bacteriophage P22 reveals organization of its DNA packaging and infecting machinery. Structure 14, 1073–1082.

Chen, D.H., Baker, M.L., Hryc, C.F., DiMaio, F., Jakana, J., Wu, W., Dougherty, M., Haase-Pettingell, C., Schmid, M.F., Jiang, W., Baker, D., King, J.A., Chiu, W., 2011. Structural basis for scaffolding-mediated assembly and maturation of a dsDNA virus. Proc. Natl. Acad. Sci. USA 108, 1355–1360.

Chen, S.L., Hung, C.S., Xu, J., Reigstad, C.S., Magrini, V., Sabo, A., Blasiar, D., Bieri, T., Meyer, R.R., Ozersky, P., Armstrong, J.R., Fulton, R.S., Latreille, J.P., Spieth, J., Hooton, T.M., Mardis, E.R., Hultgren, S.J., Gordon, J.I., 2006. Identification of

genes subject to positive selection in uropathogenic strains of *Escherichia coli*: a comparative genomics approach. Proc. Natl. Acad. Sci. USA 103, 5977–5982.
Chibeu, A., Lingohr, E.J., Masson, L., Manges, A., Harel, J., Ackermann, H.W., Kropinski, A.M., Boerlin, P., 2012. Bacteriophages with the ability to degrade uropathogenic *Escherichia coli* biofilms. Viruses 4, 471–487.
De Lappe, N., Doran, G., O'Connor, J., O'Hare, C., Cormican, M., 2009. Characterization of bacteriophages used in the *Salmonella enterica* serovar Enteritidis phage-typing scheme. J. Med. Microbiol. 58, 86–93.
Douglas, S.E., 1994. DNA Strider. A Macintosh program for handling protein and nucleic acid sequences. Methods Mol. Biol. 25, 181–194.
Earnshaw, W., Casjens, S., Harrison, S., 1976. Assembly of the head of bacteriophage P22, X-ray diffraction from heads, proheads and related structures. J. Mol. Biol. 104, 387–410.
Effantin, G., Boulanger, P., Neumann, E., Letellier, L., Conway, J.F., 2006. Bacteriophage T5 structure reveals similarities with HK97 and T4 suggesting evolutionary relationships. J. Mol. Biol. 361, 993–1002.
Effantin, G., Hamasaki, R., Kawasaki, T., Bacia, M., Moriscot, C., Weissenhorn, W., Yamada, T., Schoehn, G., 2013. Cryo-electron microscopy three-dimensional structure of the jumbo phage ØRSL1 infecting the phytopathogen *Ralstonia solanacearum*. Structure 21, 298–305.
Fokine, A., Kostyuchenko, V.A., Efimov, A.V., Kurochkina, L.P., Sykilinda, N.N., Robben, J., Volckaert, G., Hoenger, A., Chipman, P.R., Battisti, A.J., Rossmann, M.G., Mesyanzhinov, V.V., 2005a. A three-dimensional cryo-electron microscopy structure of the bacteriophage øKZ head. J. Mol. Biol. 352, 117–124.
Fokine, A., Leiman, P.G., Shneider, M.M., Ahvazi, B., Boeshans, K.M., Steven, A.C., Black, L.W., Mesyanzhinov, V.V., Rossmann, M.G., 2005b. Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry. Proc. Natl. Acad. Sci. USA 102, 7163–7168.
Freiberg, A., Morona, R., Van den Bosch, L., Jung, C., Behlke, J., Carlin, N., Seckler, R., Baxa, U., 2003. The tailspike protein of *Shigella* phage Sf6. A structural homolog of *Salmonella* phage P22 tailspike protein without sequence similarity in the beta-helix domain. J. Biol. Chem. 278, 1542–1548.
Gillespie, J.J., Wattam, A.R., Cammer, S.A., Gabbard, J.L., Shukla, M.P., Dalay, O., Driscoll, T., Hix, D., Mane, S.P., Mao, C., Nordberg, E.K., Scott, M., Schulman, J.R., Snyder, E.E., Sullivan, D.E., Wang, C., Warren, A., Williams, K.P., Xue, T., Yoo, H.S., Zhang, C., Zhang, Y., Will, R., Kenyon, R.W., Sobral, B.W., 2011. PATRIC: the comprehensive bacterial bioinformatics resource with a focus on human pathogenic species. Infect. Immun. 79, 4286–4298.
Goddard, T.D., Huang, C.C., Ferrin, T.E., 2007. Visualizing density maps with UCSF Chimera. J. Struct. Biol. 157, 281–287.
Haggard-Ljungquist, E., Halling, C., Calendar, R., 1992. DNA sequences of the tail fiber genes of bacteriophage P2: evidence for horizontal transfer of tail fiber genes among unrelated bacteriophages. J. Bacteriol. 174, 1462–1477.
Hatfull, G.F., Hendrix, R.W., 2011. Bacteriophages and their genomes. Curr. Opin. Virol 1, 298–303.
Helgstrand, C., Wikoff, W.R., Duda, R.L., Hendrix, R.W., Johnson, J.E., Liljas, L., 2003. The refined structure of a protein catenane: the HK97 bacteriophage capsid at 3.44 Å resolution. J. Mol. Biol. 334, 885–899.
Hendrix, R., Casjens, S., 2005. Caudovirales. In: Fauquet, C., Mayo, M., Maniloff, J., Desselberger, L., Ball, L. (Eds.), Virus Taxonomy. Elsevier Academic Press, London, pp. 35–42.
Hendrix, R.W., 2002. Bacteriophages: evolution of the majority. Theor. Popul. Biol. 61, 471–480.
Israel, V., 1977. E proteins of bacteriophage P22. I. Identification and ejection from wild-type and defective particles. J. Virol. 23, 91–97.
Jiang, W., Baker, M.L., Jakana, J., Weigele, P.R., King, J., Chiu, W., 2008. Backbone structure of the infectious epsilon15 virus capsid revealed by electron cryomicroscopy. Nature 451, 1130–1134.
Johnson, J.E., Chiu, W., 2007. DNA packaging and delivery machines in tailed bacteriophages. Curr. Opin. Struct. Biol. 17, 237–243.
Johnson, T.J., Kariyawasam, S., Wannemuehler, Y., Mangiamele, P., Johnson, S.J., Doetkott, C., Skyberg, J.A., Lynne, A.M., Johnson, J.R., Nolan, L.K., 2007. The genome sequence of avian pathogenic *Escherichia coli* strain O1:K1:H7 shares strong similarities with human extraintestinal pathogenic *E. coli* genomes. J. Bacteriol. 189, 3228–3236.
King, M.R., Vimr, R.P., Steenbergen, S.M., Spanjaard, L., Plunkett 3rd, G., Blattner, F.R., Vimr, E.R., 2007. *Escherichia coli* K1-specific bacteriophage CUS-3 distribution and function in phase-variable capsular polysialic acid O acetylation. J. Bacteriol. 189, 6447–6456.
Krause, D.O., Little, A.C., Dowd, S.E., Bernstein, C.N., 2011. Complete genome sequence of adherent invasive *Escherichia coli* UM146 isolated from Ileal Crohn's disease biopsy tissue. J. Bacteriol. 193, 583.
Kropinski, A.M., Waddell, T., Meng, J., Franklin, K., Ackermann, H.W., Ahmed, R., Mazzocco, A., Yates 3rd, J., Lingohr, E.J., Johnson, R.P., 2013. The host-range, genomics and proteomics of *Escherichia coli* O157:H7 bacteriophage rV5. Virol. J. 10, 76.
Lander, G.C., Baudoux, A.C., Azam, F., Potter, C.S., Carragher, B., Johnson, J.E., 2012. Capsomer Dynamics and Stabilization in the T=12 marine bacteriophage SIO-2 and its procapsid studied by CryoEM. Structure 20, 498–503.
Lander, G.C., Evilevitch, A., Jeembaeva, M., Potter, C.S., Carragher, B., Johnson, J.E., 2008. Bacteriophage lambda stabilization by auxiliary protein gpD: timing, location, and mechanism of attachment determined by cryo-EM. Structure 16, 1399–1406.
Lander, G.C., Tang, L., Casjens, S.R., Gilcrease, E.B., Prevelige, P., Poliakov, A., Potter, C.S., Carragher, B., Johnson, J.E., 2006. The structure of an infectious P22 virion shows the signal for headful DNA packaging. Science 312, 1791–1795.
Leavitt, J.C., Gilcrease, E.B., Wilson, K., Casjens, S.R., 2013a. Function and horizontal transfer of the small terminase subunit of the tailed bacteriophage Sf6 DNA packaging nanomotor. Virology 440, 117–133.
Leavitt, J.C., Gogkhia, L., Gilcrease, E., Bhardwaj, A., Cingolani, G., Casjens, S., 2013b. The tip of the tail needle affects the rate of DNA delivery by bacteriophage P22. PLoS One 8, e70936.
Lebedev, A.A., Krause, M.H., Isidro, A.L., Vagin, A.A., Orlova, E.V., Turner, J., Dodson, E.J., Tavares, P., Antson, A.A., 2007. Structural framework for DNA translocation via the viral portal protein. EMBO J. 26, 1984–1994.
Lindberg, A.A., Wollin, R., Gemski, P., Wohlhieter, J.A., 1978. Interaction between bacteriophage Sf6 and *Shigella flexneri*. J. Virol. 27, 38–44.
Liu, X., Zhang, Q., Murata, K., Baker, M.L., Sullivan, M.B., Fu, C., Dougherty, M.T., Schmid, M.F., Osburne, M.S., Chisholm, S.W., Chiu, W., 2010. Structural changes in a marine podovirus associated with release of its genome into *Prochlorococcus*. Nat. Struct. Mol. Biol. 17, 830–836.
Mindell, J.A., Grigorieff, N., 2003. Accurate determination of local defocus and specimen tilt in electron microscopy. J. Struct. Biol. 142, 334–347.
Morais, M.C., Choi, K.H., Koti, J.S., Chipman, P.R., Anderson, D.L., Rossmann, M.G., 2005. Conservation of the capsid structure in tailed dsDNA bacteriophages: the pseudoatomic structure of ø29. Mol. Cell 18, 149–159.
Moriel, D.G., Bertoldi, I., Spagnuolo, A., Marchi, S., Rosini, R., Nesta, B., Pastorello, I., Corea, V.A., Torricelli, G., Cartocci, E., Savino, S., Scarselli, M., Dobrindt, U., Hacker, J., Tettelin, H., Tallon, L.J., Sullivan, S., Wieler, L.H., Ewers, C., Pickard, D., Dougan, G., Fontana, M.R., Rappuoli, R., Pizza, M., Serino, L., 2010. Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic *Escherichia coli*. Proc. Natl. Acad. Sci. USA 107, 9072–9077.
Olia, A.S., Casjens, S., Cingolani, G., 2007. Structure of phage P22 cell envelope-penetrating needle. Nat. Struct. Mol. Biol. 14, 1221–1226.
Olia, A.S., Prevelige Jr., P.E., Johnson, J.E., Cingolani, G., 2011. Three-dimensional structure of a viral genome-delivery portal vertex. Nat. Struct. Mol. Biol. 18, 597–603.
Parent, K.N., Deedas, C.T., Egelman, E.H., Casjens, S.R., Baker, T.S., Teschke, C.M., 2012a. Stepwise molecular display utilizing icosahedral and helical complexes of phage coat and decoration proteins in the development of robust nanoscale display vehicles. Biomaterials 33, 5628–5637.
Parent, K.N., Erb, M.L., Cardone, G., Nguyen, K., Gilcrease, E.B., Porcek, N.B., Pogliano, J., Baker, T.S., Casjens, S.R., 2014. OmpA and OmpC are critical host factors for bacteriophage Sf6 entry in *Shigella*. Mol. Microbiol. 92, 47–60.
Parent, K.N., Gilcrease, E.B., Casjens, S.R., Baker, T.S., 2012b. Structural evolution of the P22-like phages: comparison of Sf6 and P22 procapsid and virion architectures. Virology 427, 177–188.
Parent, K.N., Khayat, R., Tu, L.H., Suhanovsky, M.M., Cortines, J.R., Teschke, C.M., Johnson, J.E., Baker, T.S., 2010. P22 coat protein structures reveal a novel mechanism for capsid maturation: stability without auxiliary proteins or chemical crosslinks. Structure 18, 390–401.
Pedulla, M.L., Ford, M.E., Karthikeyan, T., Houtz, J.M., Hendrix, R.W., Hatfull, G.F., Poteete, A.R., Gilcrease, E.B., Winn-Stapley, D.A., Casjens, S.R., 2003. Corrected sequence of the bacteriophage P22 genome. J. Bacteriol. 185, 1475–1477.
Potter, C.S., Chu, H., Frey, B., Green, C., Kisseberth, N., Madden, T.J., Miller, K.L., Nahrstedt, K., Pulokas, J., Reilein, A., Tcheng, D., Weber, D., Carragher, B., 1999. Leginon: a system for fully automated acquisition of 1000 electron micrographs a day. Ultramicroscopy 77, 153–161.
Rizzo, A.A., Suhanovsky, M.M., Baker, M.L., Fraser, L.C.R., Jones, L.M., Rempel, D.L., Gross, M.L., Chiu, W., Alexandrescu, A.T., Teschke, C.M., 2014. Multiple functional roles of the accessory I-domain of bacteriophage P22 coat protein revealed by NMR structure and cryoEM modeling. Structure 22, 830–841.
Rohwer, F., 2003. Global phage diversity. Cell 113, 141.
Roy, A., Cingolani, G., 2012. Structure of P22 headful packaging nuclease. J. Biol. Chem. 287, 28196–28205.
Sandmeier, H., Iida, S., Arber, W., 1992. DNA inversion regions Min of plasmid p15B and Cin of bacteriophage P1: evolution of bacteriophage tail fiber genes. J. Bacteriol. 174, 3936–3944.
Scholl, D., Merril, C., 2005. The genome of bacteriophage K1F, a T7-like phage that has acquired the ability to replicate on K1 strains of *Escherichia coli*. J. Bacteriol. 187, 8499–8503.
Schwarzer, D., Buettner, F.F., Browning, C., Nazarov, S., Rabsch, W., Bethe, A., Oberbeck, A., Bowman, V.D., Stummeyer, K., Muhlenhoff, M., Leiman, P.G., Gerardy-Schahn, R., 2012. A multivalent adsorption apparatus explains the broad host range of phage ø92: a comprehensive genomic and structural analysis. J. Virol. 86, 10384–10398.
Selvarajan, S. Sigamani, Zhao, H., Kamau, Y.N., Baines, J.D., Tang, L., 2013. The structure of the herpes simplex virus DNA-packaging terminase pUL15 nuclease domain suggests an evolutionary lineage among eukaryotic and prokaryotic viruses. J. Virol. 87, 7140–7148.
Seul, A., Muller, J.J., Andres, D., Stettner, E., Heinemann, U., Seckler, R., 2014. Bacteriophage P22 tailspike: structure of the complete protein and function of the interdomain linker. Acta Crystallogr. 70, 1336–1345.
Simpson, A.A., Tao, Y., Leiman, P.G., Badasso, M.O., He, Y., Jardine, P.J., Olson, N.H., Morais, M.C., Grimes, S., Anderson, D.L., Baker, T.S., Rossmann, M.G., 2000. Structure of the bacteriophage ø29 DNA packaging motor. Nature 408, 745–750.
Skurnik, M., Hyytiainen, H.J., Happonen, L.J., Kiljunen, S., Datta, N., Mattinen, L., Williamson, K., Kristo, P., Szeliga, M., Kalin-Manttari, L., Ahola-Iivarinen, E., Kalkkinen, N., Butcher, S.J., 2012. Characterization of the genome, proteome, and structure of yersiniophage øR1-37. J. Virol. 86, 12625–12642.

Smits, C., Chechik, M., Kovalevskiy, O.V., Shevtsov, M.B., Foster, A.W., Alonso, J.C., Antson, A.A., 2009. Structural basis for the nuclease activity of a bacteriophage large terminase. EMBO Rep. 10, 592–598.

Spilman, M.S., Dearborn, A.D., Chang, J.R., Damle, P.K., Christie, G.E., Dokland, T., 2011. A conformational switch involved in maturation of Staphylococcus aureus bacteriophage 80α capsids. J. Mol. Biol. 405, 863–876.

Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R., Reinemer, P., 1994. Crystal structure of P22 tailspike protein: interdigitated subunits in a thermostable trimer. Science 265, 383–386.

Stummeyer, K., Dickmanns, A., Muhlenhoff, M., Gerardy-Schahn, R., Ficner, R., 2005. Crystal structure of the polysialic acid-degrading endosialidase of bacteriophage K1F. Nat. Struct. Mol. Biol. 12, 90–96.

Stummeyer, K., Schwarzer, D., Claus, H., Vogel, U., Gerardy-Schahn, R., Muhlenhoff, M., 2006. Evolution of bacteriophages infecting encapsulated bacteria: lessons from *Escherichia coli* K1-specific phages. Mol. Microbiol. 60, 1123–1135.

Suhanovsky, M.M., Teschke, C.M., 2013. An intramolecular chaperone inserted in bacteriophage P22 coat protein mediates its chaperonin-independent folding. J. Biol. Chem. 288, 33772–33783.

Sun, S., Kondabagil, K., Draper, B., Alam, T.I., Bowman, V.D., Zhang, Z., Hegde, S., Fokine, A., Rossmann, M.G., Rao, V.B., 2008. The structure of the phage T4 DNA packaging motor suggests a mechanism dependent on electrostatic forces. Cell 135, 1251–1262.

Tang, J., Lander, G.C., Olia, A.S., Li, R., Casjens, S., Prevelige Jr., P., Cingolani, G., Baker, T.S., Johnson, J.E., 2011. Peering down the barrel of a bacteriophage portal: the genome packaging and release valve in P22. Structure 19, 496–502.

Touchon, M., Hoede, C., Tenaillon, O., Barbe, V., Baeriswyl, S., Bidet, P., Bingen, E., Bonacorsi, S., Bouchier, C., Bouvet, O., Calteau, A., Chiapello, H., Clermont, O., Cruveiller, S., Danchin, A., Diard, M., Dossat, C., Karoui, M.E., Frapy, E., Garry, L., Ghigo, J.M., Gilles, A.M., Johnson, J., Le Bouguenec, C., Lescat, M., Mangenot, S., Martinez-Jehanne, C.S., Matic, I., Nassif, X., Oztas, S., Petit, M.A., Pichon, C., Rouy, Z., Ruf, C.S., Schneider, D., Tourret, J., Vacherie, B., Vallenet, D., Medigue, C., Rocha, E.P., Denamur, E., 2009. Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. PLoS Genet. 5, e1000344.

Tsonos, J., Adriaenssens, E.M., Klumpp, J., Hernalsteens, J.P., Lavigne, R., De Greve, H., 2012. Complete genome sequence of the novel *Escherichia coli* phage phAPEC8. J. Virol. 86, 13117–13118.

van Heel, M., Schatz, M., 2005. Fourier shell correlation threshold criteria. J. Struct. Biol. 151, 250–262.

Whitman, W.B., Coleman, D.C., Wiebe, W.J., 1998. Prokaryotes: the unseen majority. Proc. Natl. Acad. Sci. USA 95, 6578–6583.

Wikoff, W.R., Liljas, L., Duda, R.L., Tsuruta, H., Hendrix, R.W., Johnson, J.E., 2000. Topologically linked protein rings in the bacteriophage HK97 capsid. Science 289, 2129–2133.

Wommack, K.E., Colwell, R.R., 2000. Virioplankton: viruses in aquatic ecosystems. Microbiol. Mol. Biol. Rev. 64, 69–114.

Yan, X., Dryden, K.A., Tang, J., Baker, T.S., 2007a. *Ab initio* random model method facilitates 3D reconstruction of icosahedral particles. J. Struct. Biol. 157, 211–225.

Yan, X., Sinkovits, R.S., Baker, T.S., 2007b. AUTO3DEM-an automated and high throughput program for image reconstruction of icosahedral particles. J. Struct. Biol. 157, 73–82.

Zhang, X., Guo, H., Jin, L., Czornyj, E., Hodes, A., Hui, W.H., Nieh, A.W., Miller, J.F., Zhou, Z.H., 2013. A new topology of the HK97-like fold revealed in *Bordetella* bacteriophage by cryoEM at 3.5 Å resolution. Elife 2, e01299.

Zhao, H., Christensen, T.E., Kamau, Y.N., Tang, L., 2013. Structures of the phage Sf6 large terminase provide new insights into DNA translocation and cleavage. Proc. Natl. Acad. Sci. USA 110, 8075–8080.