

BIROn - Birkbeck Institutional Research Online

White, H.E. and Orlova, Elena (2019) Bacteriophages: their structural organisation and function. In: UNSPECIFIED (ed.) Bacteriophages - Biology and Applications [Working Title]. IntechOpen.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/29349/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>
contact lib-eprints@bbk.ac.uk.

or alternatively

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,300

Open access books available

116,000

International authors and editors

130M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Bacteriophages: Their Structural Organisation and Function

Helen E. White and Elena V. Orlova

Abstract

Viruses are infectious particles that exist in a huge variety of forms and infect practically all living systems: animals, plants, insects and bacteria. Viruses that infect and use bacterial resources are classified as bacteriophages (or phages) and represent the most abundant life form on Earth. A phage can be described as a specific type of nano-machine that is able to recognise its environment, find a host cell, start infection, self-assemble and safeguard its genome until the next cycle of replication is initiated. Remarkable results have been obtained by combining cryo-EM, X-ray analysis and bioinformatics in structural studies of these nano-machines. In this review we will describe results of structural studies of phages that uncover their organisation in different conformations, thus facilitating our understanding of the functional mechanisms in supramolecular assemblies and helping us understand the usage of phages in medical treatments. Currently, antibiotic resistance is an enormous challenge that we face. The tailed phages could be used in place of antibiotics due to their high specificity to host cells, but more knowledge of their organisation and function is required.

Keywords: viruses, bacteriophage, structural organisation, infectivity, function, structural methods, electron microscopy

1. Introduction

All living systems have many diseases that are often caused by small organisms such as bacteria or infectious particles consisting of proteins, nucleic acids and sometimes lipids. These particles are called viruses, use the resources of living cells for their own propagation and can be transmitted from one organism to another. Each type of particle infects its own host cells, and they can survive outside living organisms in very harsh conditions. Some of them continue to replicate with cells despite the host's defence mechanisms and remain dormant (latent) in their host cell, e.g. herpesviruses which reactivate at a later date to produce further attacks of the disease if the host's defence system weakens [1].

Bacteriophages (or phages) are viruses that infect and use bacterial resources for their own reproduction. They are characterised by a high specificity to bacteria at infection and are very common in all environments. Their number is directly related to the number of bacteria present. It is estimated that there are more than 10^{30} tailed phages in the biosphere [2]. Phages are common in soil and readily isolated from faeces and sewage, as well as being very abundant in freshwater and oceans with an estimate of more than 10 million virus-like particles in 1 mL of seawater [3, 4].

Why study the structure-function relationship of phages? Currently, there are substantial problems with diseases caused by bacteria, especially in hospitals.

Many pathogenic bacteria exist such as *Mycobacterium tuberculosis*, *Enterococcus faecalis*, *Staphylococcus aureus*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and methicillin-resistant *S. aureus* (MRSA) and have become modified in hospitals due to the overuse of antibiotics. Bacteria have become resistant to some of the most potent drugs used in modern medicine, and this causes treatment problems [5–7]. It appears that the pathogenic bacteria adapt quicker to antibiotics than the new ones that can be produced. The number of new antibiotics being introduced has decreased since their first introduction [8].

A powerful method to circumvent this resistance is the use of phages in the treatment of bacterial infections [9]. Most current studies of phage therapy have focussed on acute infections in animals [10]. In order to regulate the mechanisms of phage infection, we need to know not only the phage structure but also the phage-cell surface interaction mechanism and the process of switching the cell replication machinery for phage propagation. One important factor that has to be considered is how phages are reproduced. Phages have two ways of propagation: lytic and lysogenic [11]. In the first case, phages cause the complete lysis of a cell, where it breaks open and subsequently dies after phage replication. In the second type of replication, a phage integrates its genome into the host bacterium's genome or forms a circular replicon in the bacterial cytoplasm. The bacterium then continues to live and reproduce normally, but the phage genome is transmitted to progeny cells at each subsequent cell division. Changes in cell conditions such as radiation or certain chemicals can release the phage genome, causing proliferation of new phages via the lytic cycle. Therefore, for medical treatments we need to use only lytic phages, so they will exist in an organism, while the pathogenic bacteria are around but only infect those bacteria that have the appropriate receptors in the outer membrane. This is an important factor that can be used to affect specific bacteria without harming those ones that are essential for the health of humans and animals [10]. In this review we will focus on tailed phages as they are abundant and well studied and could be beneficial to medicine [12]. We will describe the general organisation and structural features of their components revealed by current structural methods.

2. Phages and their classification

Virus classification is based on characteristics such as morphology, type of nucleic acid, replication mode, host organism and type of disease. The International Committee on Taxonomy of Viruses (ICTV) has produced an ordered system for classifying viruses (<https://talk.ictvonline.org/taxonomy/>). Phages are found in a variety of morphologies: filamentous phages, phages with a lipid-containing envelope and phages with lipids in the particle shell (**Figure 1A**). They have a genome, either DNA or RNA, which can be single or double stranded, and contain information on the proteins that constitute the particles, additional proteins that are responsible for switching cell molecular metabolism in favour of viruses and, therefore, the information on the self-assembly process. The genome can be one or multipartite and is located inside the phage capsid. Nearly 5500 bacterial viruses have been characterised by electron microscopy (EM) [15]. The shape of viruses is closely related to their genome, and a large genome indicates a large capsid and therefore a more complex organisation. The most studied group of phages is the tailed phages (order *Caudovirales*) which are classified by the type of tail; *Siphoviridae* have a long non-contractile tail, *Podoviridae* have a short non-contractile tail and *Myoviridae* have a complex contractile tail (**Figure 1B**).

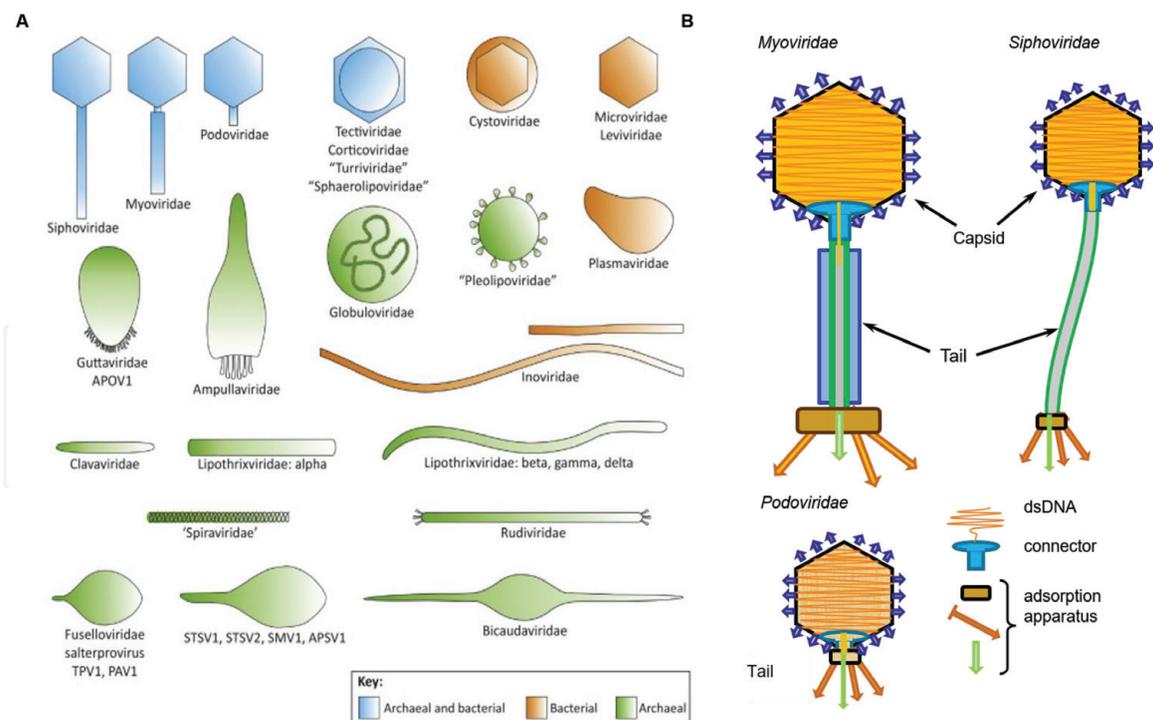


Figure 1. (A) Representation of prokaryote bacteriophage morphotypes [13]. (B) Members of the Caudovirales family [14].

3. Methods used for structural studies of viruses

The first ideas on how viruses infect cells were based on results obtained by microbiology and bacteriology during the last century. Understanding the function of viruses and how this can be regulated and modified requires knowledge of their structural organisation. However, investigation of structure-function relationships needs a combination of different techniques. Microbiology has identified viruses as infectious agents, while bacteriology and light microscopy enabled us to identify specificity between viruses and host cell interactions and to recognise a level of survival of bacteria in the presence of different phages. In order to understand interactions at the molecular level, one needs to know the structural features of the viruses and their components at an atomic level. Different structural techniques are often utilised for smaller components, and the results fitted into larger EM structures.

3.1 X-ray crystallography

X-ray crystallography was the first method used to study proteins at the atomic level, which is essential to reveal protein-ligand interactions that can boost or suppress protein activity. It is based on the principles of beam scattering within a crystal. By using specific software packages, a 3D electron density map of the protein that forms the crystal can be calculated [16]. However, to produce protein crystals, we need solutions of a protein at high concentration. The proteins have to be stable, and often mutations are made to remove their flexible parts, but this may produce different conformations to those that are required for their natural activity.

X-ray analysis is an efficient tool for analysis of protein complexes from a few kDa to hundreds of kDa in size. In order to study the structure of a large protein or a complex of several proteins, the process of crystallisation becomes a more challenging step. The development of cryoprotection in X-ray crystallography, where the crystals are flash frozen, has improved the quality of the data and often resulted in higher resolution. Nowadays, many structures of large protein complexes (up to

2–3 MDa) have been determined by X-ray analysis, but these projects have required decades to obtain high-quality crystals [17].

Viruses are much bigger particles and often have flexible components. The large size of the complexes results in significantly bigger unit cells, which results in technical challenges in obtaining fine structural details. Viruses with a rigid icosahedral lattice of the capsid have been studied successfully by X-ray crystallography at near-atomic resolution. The first viral structure was that of the *Blue tongue virus* (700 Å diameter) determined at a resolution of 3.5 Å which was the largest virus structure determined at that time [18]. The capsid of the *Siphoviridae* phage HK97 (without a portal protein) was determined at a resolution of 3.5 Å [19]. Later studies have shown that the fold of the HK97 phage capsid protein, which forms the envelope to protect viral genome from the harsh outer environment, represents a conservative fold found nowadays in nearly all dsDNA viruses so far studied.

3.2 Nuclear magnetic resonance

Nuclear magnetic resonance (NMR) is an important technique that resolves structures of small proteins that are not suitable for crystallisation due to their flexibility. This method is based on exploiting the electrical charges and spins of the nuclei in a molecule. If an external magnetic field is applied, energy is transferred to the nuclei changing their state from the level of base energy to a higher energy. This energy is emitted when the spin returns back to its base level at a frequency corresponding to radio frequencies¹. The signal that matches this transfer is measured and processed in order to yield a NMR spectrum [17, 20]. This technique is typically used for proteins of less than 200 amino acids and an upper weight limit of about 50 kDa, so it is unsuitable for the structural determination of complete viruses. However, it can be used to analyse flexibility of bigger complexes [21]. The NMR structures can be docked into low-resolution cryo-EM structures.

3.3 Electron microscopy

Light microscopy has been used for several centuries to study objects that are hardly visible to the naked eye. In conventional microscopy, resolution is mostly restricted according to the theoretical context of the Rayleigh criterion [22]. This limit is defined by the diffraction properties of light in lenses and has restricted our view to objects bigger than 250 nm. New developments in technology and advances in optical quality, electronics and software have delivered new options and extended the field of applications for electron microscopes allowing visualisation of single molecules. Electron microscopes use a beam of electrons (wavelength of less than 0.1 nm) instead of visible light (wavelength 400–700 nm). Due to their charge, the electrons can be focused using an electromagnetic field, which is why the optical system of the electron microscope (EM) is similar to the general optical system of light [23]. The short wavelength of the electron beam allows details of small objects less than 0.1 nm in size to be seen. However, biological samples are not stable in the vacuum necessary to create an image using electrons that would otherwise become absorbed by air, and, moreover, biological samples are sensitive to the electron irradiation. These factors reduce the level of achievable resolution.

At the very beginning of EM evolution, a method called negative staining was used for visualisation of biological complexes. In this case a drop of biocomplex solution is placed on a support grid and embedded in a heavy atom salt, usually uranyl acetate [24]. Since the specific density of the negative stain is much higher than

¹ chem.ch.huji.ac.il/nmr/whatisnmr/whatisnmr.html

the density of the biological molecules in the microscope, we can see the cast of the molecule merged into the surrounding stain. Where the stain did not penetrate into the molecule, one can see light spots in the image as the stain has blocked electrons. Sample preparation is fast and produces very high contrast. However, this technique does not allow fine details to be seen, and the particle becomes distorted due to the drying procedure required. The stain has a relatively large grain (up to 1.5 nm) that obscures details of the molecules under study.

Nearly four decades ago, a cryo-technique for sample preparation was introduced that allows biocomplexes to be kept at nearly native conditions. A thin layer of sample on a grid is flash frozen at liquid nitrogen temperatures, thus trapping molecules in a native, hydrated state within a thin layer of amorphous ice [25]. This technique is used to study the structural organisation of biocomplexes by cryo-electron microscopy (cryo-EM) or electron tomography (cryo-ET). Until two decades ago, all data in EM was collected on films that had to be developed and digitised, which was time-consuming. The advent of charge-coupled devices (CCDs) allowed direct digital acquisition of images and the collection of large numbers of particles giving rise to structures of higher resolution. Later, direct electron detectors were introduced into EM and are now used in all high-end electron microscopes [26]. Together with new approaches in microtechnology and the automation of data collection, the results from image analysis have improved tremendously. Cryo-EM is now approaching the near-atomic resolution that had only been achieved by X-ray crystallography. New maps obtained by cryo-EM provide information on the main polypeptide chains and often reveal the positions of side chains. The current highest resolution of structures currently deposited in the EMDB is 1.5 Å [27], with many others at a resolution between 3.5 and 4 Å. At this resolution atomic models can be built and refined using the crystallographic methods.

In cryo-ET the samples are also flash frozen, but data is collected by tilting the grid with the sample between -60 and 60° around the horizontal axis (perpendicular to the optical axis of the microscope) with an increment typically of 2° . The 2D images taken at each angle are combined to calculate a 3D map of the object. The limitation in the range of the tilt results in a cone of missing data [28]. The resolution in structures obtained by cryo-ET is lower than that in single-particle analysis. However, this approach allows visualisation of important organelles within cells. If there are multiple small structures such as ribosomes or viruses, then each structure can be extracted and averaged. This is called subtomogram averaging and will give higher-resolution structures [29].

4. Overall structural organisation of phages

Phages may have different shapes and sizes (**Figure 1A**). The most studied group is that of tailed phages with a dsDNA genome, and it also represents the largest group (**Figure 1B**). The tailed phages have three major components: a capsid where the genome is packed, a tail that serves as a pipe during infection to secure transfer of genome into host cell and a special adhesive system (adsorption apparatus) at the very end of the tail that will recognise the host cell and penetrate its wall. Cell resources are used for the phage reproduction.

The functional phage is a result of a multistep process that starts with all the necessary proteins produced by the host cell after infection: capsid, portal, tail, scaffolding, terminase, etc. (**Figure 2**). The capsids of the dsDNA phages often have fivefold or icosahedral symmetries [30], which are broken at one of the fivefold axes by the head-to-tail interface (HTI). The main component of the HTI is a dodecameric portal protein (PP) within the capsid. The PP represents the DNA-packaging motor,

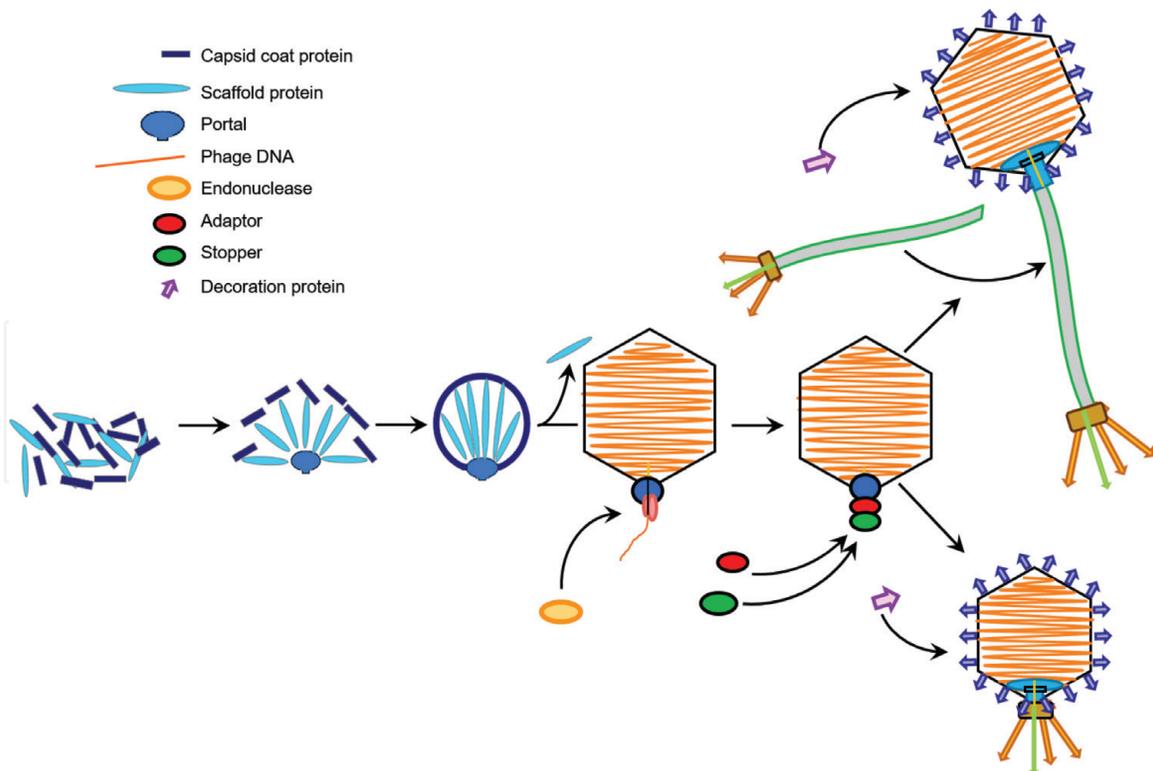


Figure 2. Self-assembly pathway of phages. Multiple copies of the capsid/scaffold complex bind the portal protein to form the procapsid; then, the scaffold proteins are ejected, and DNA is packaged into the procapsid, which expands to the size of the mature capsid. The head completion proteins (the stopper and the adaptor) are bound to the portal complex preventing DNA leakage. Next, decoration proteins bind to the capsid, and the tail, assembled separately or after DNA packaging, is attached; thus, the final infectious phage is produced. The preassembled tail attaches in *Myoviridae* and *Siphoviridae*, while in *Podoviridae* the tail assembles at the stopper.

which is the crucial part of these nano-machines. The HTI also includes oligomeric rings of head completion proteins that play dual roles: (1) making an additional interface to molecules of ATP which provide energy for DNA packaging and (2) then connecting the portal protein and the tail. Some HTIs also serve as valves that close the exit channel preventing leakage of genome from the capsid but opening as soon as the phage is attached to the host cell. However, symmetries other than dodeca-meric have been found for nearly all PPs in vitro if the PPs are assembled under naive conditions, without any other phage protein components [31–35]. Typically, the main phage proteins have conservative folds despite low sequence similarity, although they may have different additional domains [36, 37].

The phage tail is the structural component of the phage that is essential during infection. Its adsorption apparatus located on the distal end of the tail recognises a receptor, or the envelope chemistry, of the host cell and ensures genome delivery to the cell cytoplasm. In *Myoviridae* and *Siphoviridae*, the tail is composed of a series of stacked rings with the host recognition device being located at the end of the tail. In *Podoviridae* the adsorption apparatus is bound immediately to the HTI. The adsorption apparatus is surrounded in many phages by fibrils that ensure a tight connection to the host cell (**Figure 2**).

4.1 Procapsids

The capsid of a phage has a precursor formation, named the procapsid, during the assembly process (**Figure 2**). Scaffolding proteins (SPs) drive the assembly process by chaperoning major capsid protein (MCP) subunits to build an icosahedral procapsid that is later filled with dsDNA. The SPs are bound to the portal

Phage	Type of phage	Capsid protein	No. of residues	M. Mass (kDa)	Resolution (Å)	Structure analysis
HK97	<i>Sipho</i>	gp5	385 282 (AC)	42	3.44 (C) 12 (PC)	X-ray [42] EM [51]
T5	<i>Sipho</i>	pb8	458 299 (AC)	51	9 (C)	EM [52]
λ	<i>Sipho</i>	gpE	341	38	6.8 (C),13.3 (PC)	EM [47]
SPP1	<i>Sipho</i>	gp13	324	35	8.8 (C)	EM [53]
TP901-1	<i>Sipho</i>	ORF36	272	29	15	EM [54]
TW1	<i>Sipho</i>	gp57*	352	39	7	EM [55]
ϕ 29	<i>Podo</i>	gp8	448	50	8	EM [56]
T7	<i>Podo</i>	gp10	345	37	4.6 (PC) 3.6 (C)	EM [57]
P22	<i>Podo</i>	gp5	430	47	3.8 (PC) 4.0 (C) 3.3 (C)	EM [58–60]
ϵ 15	<i>Podo</i>	gp7	335	37	4.5	EM [50]
T4	<i>Myo</i>	gp24 gp24* gp23 gp23*	427 417 (AC) 521 456 (AC)	47 44 56 49	2.9 (monomer) 3.3 (EM)	X-ray [61] EM [62]
HSV-1	<i>virus</i>	VP5	1374	149	4.2 (C)	EM [63]

AC—after cleavage; C—capsid; PC—procapsid

Table 1.
Phage procapsids and capsids.

complex during formation of a procapsid with scaffolding inside. The sequence of conformational changes from a procapsid to the phage capsid where genome has been packed is named as the maturation process and goes through a series of intermediates [19, 38–40]. Some phages like HK97 and T5 do not have a separate SP; instead, the capsid protein is fused with a scaffolding domain at the N-terminus. As soon as the procapsid is assembled, the scaffolding domain is cleaved off and then like the separate SP will be removed from the capsid to make room for the genome [38, 39]. Structures of procapsids and mature virions have been determined for a number of phages (**Table 1**). The spherical capsid shell expands during maturation and becomes thinner due to alterations in the inter- and intra-subunit contacts.

4.2 Capsids

Most tailed phages have capsids of an icosahedral shape formed by multiple copies of one or more proteins. Icosahedral capsids are characterised by 12× five-fold, 20× threefold and 30× twofold axes, which give rise to 60 copies of the major independent parts [41]. A triangulation number (T number) describes the number of copies of the same protein within the independent part of the icosahedral lattice. The overall number of proteins in the virus corresponds to the T number multiplied by 60; for example, a T = 3 virus has 180 subunits [41]. Oligomers of the proteins that are located on the fivefold axes are referred to as pentons, while those complexes that are located on the faces of the icosahedron and form oligomers from six subunits are named as hexons.

HK97 phage. The first structure of a phage capsid was solved for the *Siphoviridae* HK97 phage (a mutant without the PP and tail, diameter ~ 650 Å, $T = 7$) at 3.6 Å resolution [19] by X-ray crystallography. This structure revealed a new type of protein fold, which has been found in many other phage capsid proteins (CP) despite low sequence identity (**Table 1**). Later, the structure was improved to 3.45 Å resolution [42]. This fold is also found in distantly related icosahedral tailed viruses that infect halophilic archaea [43] and human pathogens such as *Herpesviridae* [44–46]. The characteristic features of the HK97-fold are an N-arm, a peripheral P-domain with a long helix and a β -sheet, an axial A-domain and a long E-loop that fills the region between adjacent threefold axes (**Figure 3A**). The HK97 capsid is held together by molecular chain mail [19]. The two antiparallel β -strands in the E-loop are terminated by a loop, containing Lys169. This residue forms an isopeptide bond with the P-domain residue Asn356 of an adjacent subunit. The third residue involved is the catalytic residue Glu363 from a third subunit (**Figure 3A**). Conformational changes between HK97 procapsids and capsids were assessed by flexible fitting of atomic models [40]. The capsid subunits are skewed around a pseudo-twofold axis in the procapsid but form more symmetric pentons and hexons in the mature virion.

λ phage. The mature capsid of the *Siphoviridae* phage λ (diameter ~ 600 Å, $T = 7$) is stabilised with the help of a decoration protein gpD which is attached to the threefold vertices. A procapsid structure was determined with a resolution between 13.3 and 14.5 Å and a mature capsid at 6.8 Å [47]. The capsid protein gpE of the mature capsid shows clearly the HK97-fold so the HK97 model was used for rigid body fitting into the capsid [47]. However, the A-domain gave a poor fit so it was treated as a separate rigid body and was found to be rotated 15° clockwise compared with HK97. There was unassigned density, which could fit the extra 59 residues in

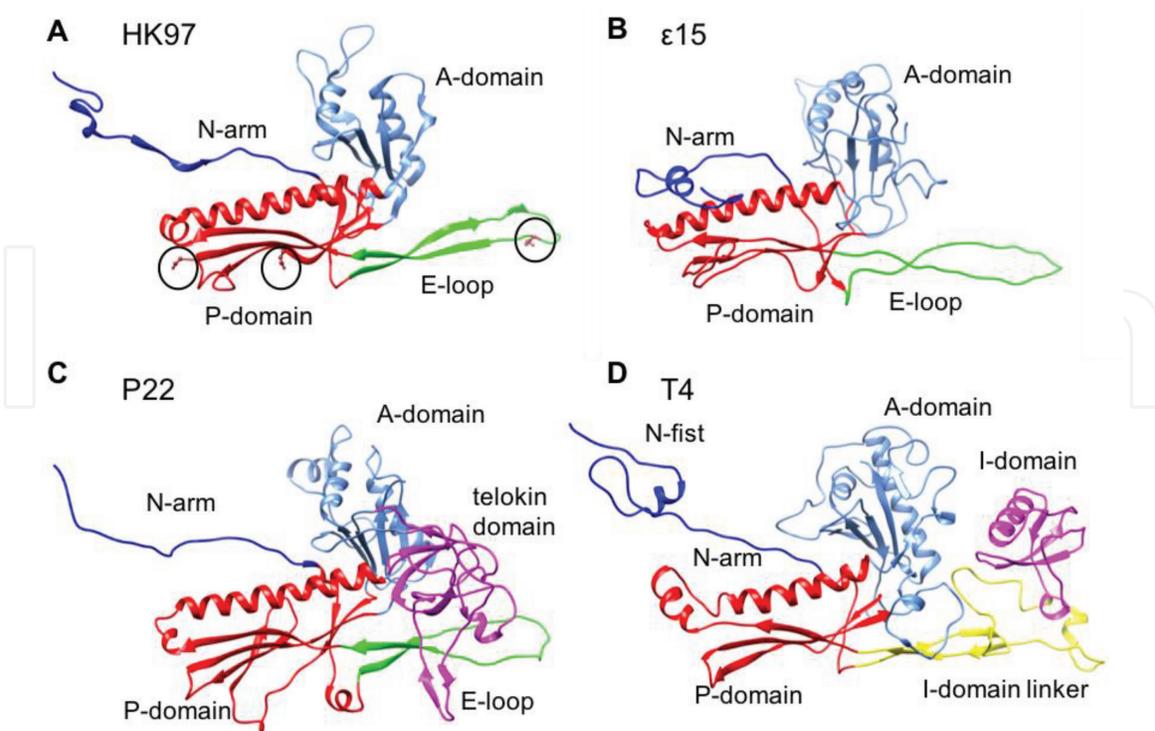


Figure 3. Structural organisation of the major capsid proteins. (A) Siphophage HK97 (1OHG). The catalytic residues are shown in brown and circled. (B) Podophage $\epsilon 15$ (3J40). (C) Podophage P22 (5UU5). (D) Myophage T4 gp23* (5VF3). The N-arm is dark blue, the P-domain is red, the A-domain is light blue and the E-loop is yellow. Extra inserted domains seen in P22 and T4 are magenta. The yellow linker in T4 is topologically equivalent to the E-loop seen in the other phages.

the λ phage. The crystal structure of gpD (1.1 Å) [48] was fitted into the cryo-EM capsid map and, when combined with the HK97 model, showed how gpD is held in place at the capsid threefold axis.

ε15 phage. The structure of gp7, the MCP of the *Salmonella* phage **ε15**, was determined by cryo-EM (4.5 Å) [49] and showed a fold similar to that in HK97 (**Figure 3B**). Later, the same group obtained a better resolution structure (~3.5 Å) [50] and, using their previous 4.5 Å structure for gp7, computationally analysed possible different gp7 conformations within EM density obtained from averaging the individual subunits in one asymmetric unit. They found that two models could be fitted in the structure: one that was consistent with the previous interpretation [49], and the other had a strand swap of the β -strands in the P-domain before the two helices in the A-domain. The strand order was 4-3-1-5-2, but in the swapped model, it was now 5-4-2-3-1. The model with the strand swap provided a better visual fit in the refined map [50].

P22 phage. Structures of a procapsid and a heat-expanded capsid of the *Podoviridae* **P22** phage (diameter ~620 Å, T = 7) were determined by cryo-EM at 9.1 and 8.2 Å, respectively [58]. This heat-expanded capsid was composed only of hexons with wide openings where the pentons would be. The MCP gp5 revealed the HK97-fold, but there was an extra domain (residues 223–349) with an immunoglobulin-like telokin domain [64]. Later, the procapsid and capsid structures were determined by cryo-EM at 3.8 and 4.0 Å resolution, respectively [59]. It was also found that gp5 has the HK97-fold with extra density above the E-loop corresponding to the telokin domain (**Figure 3C**). A model built from a cryo-EM map of P22 (3.3 Å) [60] identified the interactions that stabilise the capsid. In the mature virion, the N-arm forms an antiparallel β -strand pair between neighbouring subunits around the threefold axis. A second set of interactions involve hydrogen bonds and salt bridges between adjacent subunits involving the A-domain, the E-loop and the I-domain [60].

φ29 phage. The podovirus **φ29** is a relatively small phage (prolate icosahedra, height ~540 Å, diameter ~450 Å) that requires pRNA along with the ATPase gp16 to provide enough energy for the DNA translocation. The procapsid consists of the MCP gp8, the SP gp7, the head fibre protein gp8.5, the connector gp10 and a pRNA. After DNA packaging, the pRNA and ATPase come off the procapsid and are replaced by the gp11 and gp13 underneath the connector, gp29 at the end of the tail and the appendages (gp12*) which are attached to gp11 and gp13 [65]. During maturation an 18 kDa fragment of gp12 is cleaved off to give gp12* found in the appendages whose role is to adsorb the virion on the host cell. A fibreless, isometric φ29 variant capsid cryo-EM structure (7.9 Å) showed that gp8 has a structure with the HK97-fold but with extra density [66]. Domain profile-searching algorithms [67] showed that residues 348–429 at the C-terminus were found to be 32% identical to a BIG2 domain consensus sequence (group 2 bacterial immunoglobulin-like domains). An asymmetric reconstruction of fibreless full (7.8 Å) and empty (9.3 Å) capsids by cryo-EM revealed the interactions between the capsid shell and DNA [56]. All these interactions were in similar or identical locations in most of the gp8 subunits. The most prominent contact was at the end of a long tubular piece of density, which after fitting a HK97 homology model could be assigned to the N-terminal end of the HK97 long helix.

T4 phage. The protein elements of the *Myoviridae* phage **T4** (prolate icosahedra, height ~2000 Å, diameter ~900 Å) and its overall organisation have been extensively studied by X-ray and EM [62, 68–70]. The procapsid of T4 contains two proteins gp23 and gp24 which have 22% sequence identity. During maturation the gp21 protease cleaves off 65 N-terminal residues from the capsid protein gp23 and 10 N-terminal residues from the vertex protein gp24 to produce gp23* and gp24*,

respectively. In the mature capsid, gp23* forms 120 hexons, and 11 capsid vertices are formed by gp24* proteins, while the 12th is occupied by a dodecamer of gp20 PP. T4 has two decoration proteins Soc and Hoc [62]. A crystal structure of gp24 (2.9 Å) [61] showed a domain with the HK97-fold and a 60 residue insertion I-domain located on the outer capsid surface (**Figure 3D**). The 3.3 Å cryo-EM reconstruction of the isometric capsid of T4 [62] allowed the structure of gp23* and gp24* to be determined. The I-domain linker, missing from the crystal structure of gp24, could be seen and interacts with a neighbouring gp24* molecule to stabilise the capsid. The structure of gp23* is similar to gp24* but with an extra compact region formed by residues 66–93, termed the “N-fist” prior to the N-arm residues 94–110 [62].

Crystal structures were obtained for the Hoc protein from the T4-like phage RB49 with the capsid-binding C-terminal domain 4 missing [71] and Soc protein from the T4-like phage RB49 [72]. The Soc molecules, which are required for capsid stability, interact with three gp23* subunits [62] although not all binding sites were fully occupied possibly due to differences in the gp23* I-domain linkers. The immunogenic outer capsid Hoc protein was found in two different sites within the asymmetric unit: at the centre of the hexon near the icosahedral threefold axis and in the hexon close to the fivefold axis [62]. The density of Hoc near the threefold axis was less interpretable than that near the fivefold axis.

HSV-1 virus. Although it is not a phage, the human herpesvirus **HSV-1** capsid (1250 Å, T = 16) is a close relative and undergoes a pathway of self-assembly similar to that of dsDNA phages [73]. The virion is characterised by the following features: envelope, tegument, capsid and the viral genome. There are three types of HSV capsids: A-capsids have neither DNA nor the SP, B-capsids have the SP but no DNA and C-capsids contain DNA but no SP. The MCP VP5 forms pentons and hexons and VP26 binds to the VP5 hexons. A triplex of VP19C and VP23 found between capsomers [45]. The upper domain of residues 451–1054 was crystallised and the structure was determined at 2.9 Å. The structure of the whole virion of HSV-1 was determined at ~7 Å resolution [45]. The model of HK97 capsid protein was fitted to the lower domain of VP5 where the E-loop and N-arm were visible, the spline helix was longer and the central channel was wider. Unlike HK97, the E-loop does not form the covalent cross-links or reach an adjacent capsomer. Instead, it interacts with adjacent subunits, lower and middle domains of same VP5 subunit and a triplex molecule. A structure of HSV-1 with its capsid-associated tegument complex (CATC) has been obtained at 4.2 Å [63]. VP5 had the HK97-fold with six additional domains. The two β-barrels in HSV Tri1 (VP19c) and Tri2 (VP23) resemble the homotrimers found in proteins like gpD of phage λ.

4.3 Connectors

In phages and herpesviruses, one of the fivefold vertices of the capsid is replaced by a *head-to-tail interface* (HTI) [30], which is a multi-protein complex (connector). In all phages the HTI provides a platform for docking of preassembled tails in *Siphoviridae* or *Myoviridae* or initiates the assembly of a short tail in *Podoviridae* [30]. The HTI comprises a portal complex (PP) and head completion proteins (**Figure 2**) that serve as a valve for closing the channel and keeping the phage genome inside the capsid at high pressure and only opens to allow genome release from the capsid (under natural conditions) as soon as the phage becomes tightly attached to a host cell.

All currently known PPs are homo-dodecamers when extracted from the viral capsids, as that symmetry is imposed during self-assembly *in vivo*. However, naive assemblies *in vitro* of the PP complexes have some variations in their rotational symmetry with 13-mers being observed for SPP1, T7 and HK97 [31, 33, 74]. HSV has been shown to have 11-fold, 12-fold, 13-fold and sometimes even 14-fold symmetry [34].

While monomers of the different PPs vary in size, all of them share a common fold—shown by EM and X-ray structures that were obtained for the ϕ 29, SPP1 and P22 portals [75–78] and by cryo-EM for T7 and T4 (**Table 2**) [69, 79]. All known PP monomers are characterised by four domains: clip, stem, wing and crown (**Figure 4**) [77].

The clip domain is exposed to the capsid exterior and involved in binding to the terminase for DNA packaging [75, 80, 81] and later to a head completion protein during the HTI assembly [82]. The first high-resolution structure of a phage PP was obtained for the ϕ 29 phage (**Figure 4A**, [75]). The clip domain is linked to the wing region through a stem that comprises typically two α -helices and the outer loops (**Figure 4B, C**). X-ray structures of PP from ϕ 29 and SPP1 phages revealed major helical components that form the central channel through which DNA enters and exits the capsid. The structures of other PPs obtained later have confirmed that this is a conserved element characteristic for all known PPs. The wing domain radiates outwards from the central axis and has an α -helix, which is the longest one and serves as a spine of the wing. It has an α/β sub-fold at its periphery [77]. The crown domain consists of α -helices and is relatively small in SPP1 and surprisingly long (213 aa) in phage P22 (**Figure 4B, D, Table 2**).

SPP1 phage. The **SPP1 phage** connector has been studied for nearly two decades; the SPP1 PP structure was determined by X-ray crystallography, and all other portal complexes are compared with it (**Figures 4B, C and 5A, B**), but this structure was a 13-mer [31]. The HTI of SPP1 extracted from the capsid was a stable complex composed of the PP gp6, the adaptor protein (AP) gp15 and the stopper (SP) gp16, all organised as three stacked cyclical homo-oligomers [82–84] (**Figure 5A, B**). The cryo-EM structure of the SPP1 HTI before and after DNA release was obtained by cryo-EM at ~ 7 Å resolution, where the HTI is bound to the tail [82] with gp16 acting as a docking platform for the SPP1 preassembled tail [85, 86]. Binding of the tail induces changes in the position of the gp16 residues Ile9 to Thr33 that close the central channel of the connector.

HK97 phage. There is no structure of the PP gp3 of the **HK97 phage**, but structures of gp6 and gp7 that correspond to the AP gp15 and SP gp16 of SPP1, respectively, were determined by X-ray analysis [74]. The 2.1 Å crystal structure of the gp6 AP revealed that it forms a 13-mer during crystallisation. A model for a

Phage	PP	No. of residues	M. Mass (kDa)	Resolution (Å)	Structure analysis
HK97	gp3	424	47		none
T5	pb7	403	45		none
λ	gpB	533	59		none
SPP1	gp6	503	57	3.4 (X-ray), ~ 7 (EM)	X-ray [77], EM [82]
TP901-1	ORF32	452	52	20	EM [54]
TW1	gp24	459	51	21	EM [55]
ϕ 29	gp10	309	36	2.1	X-ray [76]
T7	gp8	536	59	8, 12	EM [87]
P22	gp1	725	83	10.5 (EM) 3.25 (X-ray)	EM [88] X-ray [78]
ϵ 15	gp4	556	61	20	EM [89]
T4	gp20	524	61	3.6	EM [69]
HSV-1	pUL6	676	74.2	8	EM [90]

Table 2.
Phage portal proteins.

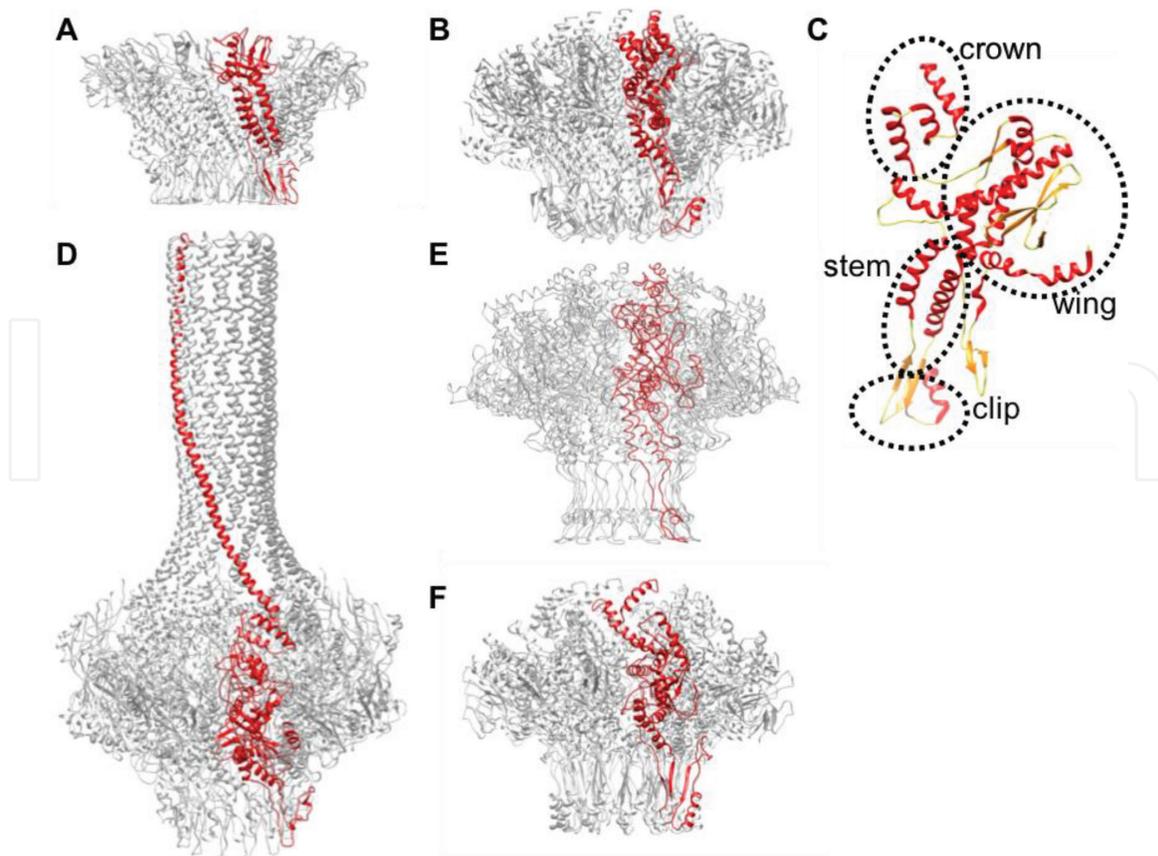


Figure 4.

Structures of portal proteins. One chain of the PP is highlighted in red. (A) gp10 of ϕ 29 (1FOU). (B) gp6 of SPP1 (2JES). (C) A gp6 SPP1 monomer with the crown, stem, wing and clip domains indicated. (D) gp1 of P22 (3LJ5). (E) gp8 of T7 (3JA4). (F) gp24 of T4 (3JA7).

dodecameric ring of gp6 was constructed from a monomer taken from the 13-mer and fitted into a cryo-EM map of SPP1 [83]. This fitted well in size and shape, but the helices of HK97 gp6 did not fit well into the densities of the SPP1 connector EM map which suggested that the assembly into a 13-mer in the absence of other phage components may produce a different conformation [74].

P22 phage. The HTI of the **P22 phage** consists of two proteins, the PP gp1 and the AP gp4. The first structural organisation of the P22 HTI complex was obtained by cryo-EM at a resolution of 9.4 Å [91]. A crystal structure (3.25 Å) was obtained for the PP by using low-resolution EM data for phasing (**Figure 5C, D** [78]). The complete polypeptide chain was traced apart from a loop between residues 464 and 492, and loop modelling was used to build this area from a 9.2 Å cryo-EM map [88]. The overall height is ~300 Å with the gp4 AP forming a dodecameric ring below the PP. The PP of P22 has the same fold in its central channel as SPP1 [90] and ϕ 29 [88, 89] (**Figures 4D, 5D**). The crystal structure of the full-length PP revealed that the C-terminal domain forms a ~200 Å long, α -helical barrel. At the same time, an asymmetric reconstruction of the entire P22 virion has been determined by cryo-EM at 7.8 Å resolution [92]. The 150 Å coiled-coil barrel structure extends from the PP to near the centre of the capsid. Fitting the crystal structure of the core-gp4 complex into the 7.8 Å virion density map revealed a stretch of about 21 gp4 C-terminal residues that lie wedged between the capsid and portal [92]. Overlap between gp4 and the MCP gp5 indicates that gp4 must undergo significant conformational change during phage assembly when the tail is added. A comparison of the portal position in the procapsid and the virion shows that the portal increases its contact with the capsid shell during maturation. It was proposed that a portion of the scaffold remains in place during dsDNA packaging to allow access of the gp4

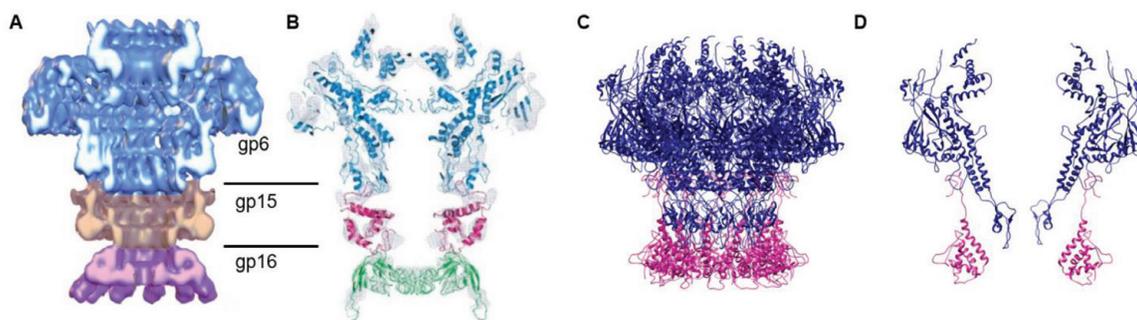


Figure 5.

Structures of the HTI. (A) The cryo-EM map of the SPP1 HTI coloured according to protein with the gp6 PP (blue), adaptor gp15 (brown) and stopper (purple) (EMD-1021 [83]). (B) Cutaway view of the SPP1 HTI with models gp6 (2JES), gp15 (2KBZ) and gp16 (2KCA) fitted into EMD-1021 (from [84]). (C) P22 connector complex determined by X-ray crystallography without barrel domain, PP pb1 (blue) and AP gp4 (red). (D) Central slice of C [91].

C-terminus to the bottom of the portal [92]. When gp4 binds, the SP is displaced allowing the final conformational change implied by the position gp4 C-terminal polypeptide that is wedged between the capsid and the HTI.

T7 phage. The HTI of T7 was determined at 8 Å resolution [79], and the dodecameric T7 PP was found to be structurally similar to the PPs from other phages (Figure 4E). A ~12 Å resolution structure of a recombinant HTI (gp8-gp11-gp12) complex of T7 was determined by cryo-EM [87], and pseudo-atomic models were obtained using gp1 and gp4 of P22 for fitting and analysis of the T7 gp8 (PP) and gp11 (AP), respectively [78]. The T7 gp8 model was superimposed on gp10 from ϕ 29 [75, 76], gp1 from P22 [78] and gp6 from SPP1 [77] which confirmed the presence of two stem helices, but the fold of the clip domain varies between the different PPs [87]. Previous structures of the head completion proteins have shown them to contain four helices, and when the T7 gp11 model was superimposed on gp4 from P22 and gp6 from HK97 [74, 78], the position of these four helices was conserved, but the C-termini were flexible.

T4 phage. The structure of the T4 HTI region was determined by cryo-EM for the fully assembled capsid (~17 Å) [93]. The neck region, which connects the tail to the dodecameric PP (gp20), comprises adaptor proteins gp3, gp15, gp13, gp14 and gp wac (fibrin). It was assumed that the T4 HTI would have a similar PP organisation to other phages; therefore, the crystal structure of the ϕ 29 PP was tentatively docked into the cryo-EM map of the T4 HTI region [75, 93]. Recently, the structure of the gp20 PP from T4 was determined by cryo-EM (3.6 Å) (Figure 4F, [69]). Interestingly, a gp20-N74 construct with the 73 N-terminal residues truncated contained 95% dodecamers and 5% 13-mers in solution [69]. The dodecameric PP was ~120 Å in height and varied in diameter from 90 Å (top) to 170 Å (middle) and to 80 Å close to the capsid, while the central channel goes from 90 to 28 Å near the middle of the channel. The connection to the capsid was by the PP wing domains, and the interactions are partially hydrophobic and partially polar [69].

ϕ 29 phage. The asymmetric 3D reconstructions of ϕ 29 with and without DNA obtained by cryo-EM at 7.8 and 9.3 Å, respectively, showed that the PP gp10 dodecamer has elongated densities lining the central channel and tilted at ~30° to the central axis of the virion [56]. These cylindrical densities correspond to the α -helices seen in the crystal structure of the ϕ 29 PP [76]. The density of the clip domain in the cryo-EM map lies further away from the PP axis than in the crystal structure possibly due to a different conformation of the PP within the fully assembled phage. Two cylindrical columns of high density were observed within the virions: one in the upper part and another at the bottom of the HTI. These densities were assigned to DNA based on the diameter, intensity and their location. The ϕ 29 DNA is visible as ringlike densities below the PP, and then DNA stretches into the tail to about ~100 Å

[56]. The DNA appears to contact with the PP crown domain and then with the density corresponding to the tunnel loops located between the PP crown and stem domains. These tunnel loops in the narrowest part of the connector channel were found in the SPP1 phage PP and are believed that they play a role in DNA translocation [77, 94]. The ϕ 29 DNA appears to be clamped at the top of the tail tube [56].

HSV-1 virus. Herpes simplex virus (**HSV-1**) has its DNA packaged into the capsid through a portal channel of the PP complex (pUL6) [95]. The structure of a dodecameric HSV-1 PP has been determined by cryo-EM at 16 Å resolution from purified portals [34]. The structure showed a close resemblance to the SPP1 PP [83]. The PP is about the same size as the pentons that occupy the other fivefold vertices, which explains the difficulty in localising the portal density in images of HSV capsids unlike in phages, which have a tail. The PP, pUL6, was well defined in the A-capsid structure located at one fivefold vertex on the outer surface of the capsid as shown by cryo-ET [96]. There is also a strong density within the portal channel and inside of the capsid, which is interpreted as the end of DNA as seen in ϕ 29 and SPP1 [56, 82].

4.4 Tails

The tail organisation in phages depends on their type: *Siphoviridae* have long flexible tails, and *Podoviridae* have very short tails that mostly consist of the adhesive

Phage	Tail proteins	No. of residues	M. Mass kDa	Resolution (Å)	Structure analysis
HK97	putative tail-component IPR010064, gp10	not defined	not known	n/a	None
T5	pb6	464	50	2.2 6	X-ray [97] EM [97]
λ	gpV (TP) gpH (TMP) gpU (terminator)	246 853 131	26 92 15	n/a 2.7	NMR [98, 102] X-ray [103]
SPP1	gp17 (TP) gp17* (TP) gp18 (TMP)	134 264 1032	15 28 111	n/a 14	NMR (gp17) [104] EM [85]
TW1	gp12 (TP) gp14 (TMP)	NF 675	18 72	23	EM [55]
ϕ 29	gp9 (knob) gp12 (tailspike)	599 854	68 92	2.04 1.8–2.05 7.8	X-ray [105, 106] EM [56]
T7	gp11 (TP) gp12 (TP) gp17 (fibres)	196 794 553	22 89 62	12.0 2.0 (X-ray)	EM (gp11,12,17) [87] X-ray (gp17) [107]
P22	gp10 (hub) gp9 (tailspike)	472 667	52 72	9.4 2.0	EM, tomography [88, 91], X-ray [108, 109]
ϵ 15	gp20 (tailspike)	1070	116	20n/r	EM [89, 110]
T4	gp19 (TP) gp15 (terminator)	163 272	18 32	4.11 3.4 15.0 3.2	EM [111, 112] X-ray [113]
HSV	does not have the tail	n/a	n/a	n/a	n/a

Table 3.
Phage tail structures.

device, while *Myoviridae* have rigid long contractile tails that consist of a number of different proteins forming the inner rigid tube and the outer contractile sheath. *Siphoviridae* and *Myoviridae* have an independent pathway for assembly of their tails and are attached to the capsid after it has been packed with genome. However, in *Podoviridae* the tails are assembled on the capsids after DNA packaging as the last step of self-assembly (**Figure 2, Table 3**). The long tails of *Siphoviridae* are composed of tail proteins (TPs) that form circular oligomeric rings with three- or sixfold rotational symmetry. The rings are assembled around a tape measure protein (TMP) that defines the length of the tail and are stacked on the top of each other with helical symmetry. A tail terminator protein (TrP) caps the tail when it reaches the length defined by the TMP; the TrP serves as an interface with the capsid. When the phage interacts with the host receptor, the HTI opens and the TMP is pushed out by DNA as a result of the inner pressure of the capsid. Most long tails have a smooth outer surface, but some have appendages that protrude outwards from the tail surface.

T5 phage. The T5 phage has a long tail (1600 Å) with a diameter of 90 Å. A crystal structure has been obtained for the TP pb6 of T5 (2.2 Å) [97]. The protein consists of two domains: an N-terminal domain with two subdomains both of which comprise a β -sandwich flanked by an α -helix and a long loop (**Figure 6A**) and a C-terminal domain with an immunoglobulin-like fold [98]. The overall tail structure has threefold symmetry (**Figure 6B**) [99, 100]. Cryo-EM was used to

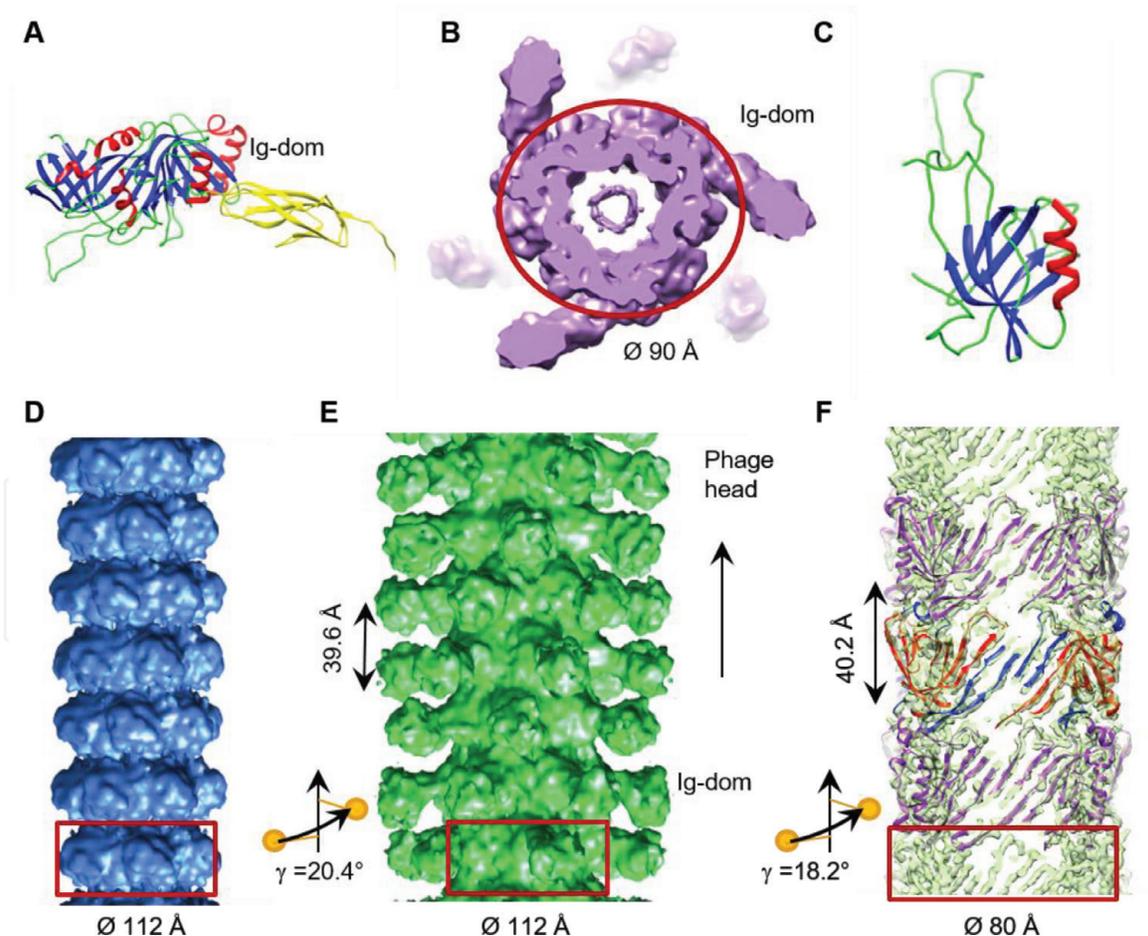


Figure 6. Bacteriophage tails. (A) The crystal structure of a monomer of T5 pb6 (5NGJ). The extra immunoglobulin domain is coloured yellow. (B) A slice of the combined EM map of T5 (EMD-3692) showing the fold symmetry of the tail. (C) The crystal structure of the N-terminal domain of the P22 TP gpV (2K4Q). (D) Cryo-EM map of SPP1 tails (gp17.1). (E) Cryo-EM map of SPP1 tails (gp17.1*). The protrusions are the size of an immunoglobulin domain. (F) The T4 cryo-EM map (EMD-8767) with fitted coordinates (5W5F). Alternate subunits in the central ring are coloured red and blue. The red circle in B and rectangles in D, E and F indicate the inner tail tube, γ —rotation between adjacent tail rings.

determine the tail structure of T5 at ~ 6 Å resolution before and after DNA ejection [97] and the atomic model of pb6 used to interpret the results. No differences were found between the two structures, apart from the absence of the tape measure protein pb2 after DNA ejection [97].

λ phage. In the phage λ , the tail has sixfold symmetry and is composed of gpV (TP) and gpH (TMP) [101]. The N-terminal domain of gpV (gpV_N) is required for assembly of the functional phage, and the structure was determined by solution NMR [102]. There are seven β -strands, arranged into two antiparallel sheets which fold into a twisted β -sandwich (**Figure 6C**). The single α -helix is located at the side of the sandwich. The C-terminus of gpV (gpV_C) was shown by solution NMR to have an Ig-like fold [98].

SPP1 phage. The structures of the tails from the *Siphoviridae* **SPP1** before and after DNA ejection were determined using negative stain EM at ~ 14.5 Å resolution [85]. This tail is ~ 1600 Å long and consists of three proteins gp17 and gp17* that form the tail tube and the TMP gp18. Even at low resolution, the structures revealed movements in the positions of inner domains gp17 and gp17* after DNA ejection [85]. The ratio of proteins gp17 and gp17* within the tail complex suggests that the tail has to have threefold symmetry. Reconstructions of mutants comprising either gp17 or gp17* were obtained and demonstrated sixfold symmetry. A comparison of structures indicated that protein gp17* has an immunoglobulin domain located on the outer surface of the tail (**Figure 6D, E**) (Orlova, Personal communication).

T4 phage. *Myoviridae* have contractile tails with a sheath that surrounds a central tail tube. On infection, the tail sheath contracts allowing the tail tube to penetrate the outer membrane of the host cell. The structure of the **T4** phage (a representative of *Myoviridae*) tail is well studied. It consists of a rigid tube, composed of multiple copies of gp19, surrounded by contractile sheath from gp18 subunits [93]. A structure of the central tube at 3.4 Å has been obtained by cryo-EM and showed sixfold symmetry [112] with an axial rise for the helical unit of ~ 40 Å. This resolution revealed the strands of the β -sheets indicating that four strands from one subunit become part of a continuous helical β -sheet lining the inner channel of the tail (**Figure 6F**). The structure of the T4 contracted tail was obtained at 17 Å resolution, and the tail sheath protein gp18 was found to be arranged into 23 hexameric rings [93]. A crystal structure of a gp18 mutant missing the C-terminal domain [114] was used as the basis for identifying domains in gp18. A homology model based on other *Myoviridae* tail sheath structures [115] was used to localise the C-terminal domain [113].

4.5 Adsorption apparatus

Most *Siphoviridae* phages have an oligomeric ring formed by distal tail proteins (DTPs), which is attached to the last ring of the tail tube [116, 117]. The DTP ring usually serves as an apparatus to recognise and connect to receptor-binding proteins; sometimes, this interaction is assisted by tail fibres found in T4, T5 and other phages. The DTP of SPP1 does not have the fibres [118]. Many phages that infect Gram-negative bacteria have lysozyme-like proteins on the ends of their tails, which enter the periplasm to digest the peptidoglycan barrier [119].

P22 phage. *Podoviridae* phage tails have a central needle with trimeric appendages (or spikes) around it, and the part of the tail attached to the capsid could be considered as the baseplate. The short tail of the podovirus **P22** has been studied by cryo-EM of the entire virion [120, 121]. The structure of the P22 tail was obtained at 9.4 Å and aided understanding of its organisation [91]. The 2.8 MDa multisubunit complex is formed by the dodecameric PP gp1, 6 trimeric tailspikes of gp9, 12 copies of the tail accessory protein gp4, a hexon of gp10 and a trimer of the tail needle gp26 (**Figure 7A**). The proteins gp4 and gp10 form a HTI on which the tail

assembles and are attached to the PP, while the N-terminal head-binding domain of the outer tailspikes (gp9) attaches to the interface of the gp4 and gp10 subunits. The second larger receptor-binding domain of gp9 contacts and destroys the cell surface lipopolysaccharide [121, 122]. Crystal structures have been obtained for both gp26 [123] and gp9 [108, 109], and these were fitted into the cryo-EM map to find how the different proteins hold the complex together. The tail needle gp26 is folded as triple-stranded coiled coil. The gp9 head-binding domain has a β -helical domain, and the receptor-binding domain is a triangular β -prism (**Figure 7B**) [91]. The asymmetric cryo-EM structure of P22 (7.8 Å) allowed a detailed analysis of the interactions between gp9 and the gp4–gp10 interface by fitting the gp9 and gp26 crystal structures and models of gp1, gp4 and gp10 [92].

SPP1 phage. The *Siphoviridae* SPP1 phage has the long flexible tail. A negative stain reconstruction of the adsorption apparatus revealed a hinge connection between the tip and the tail tube, allowing bending angles as high as 50° [85]. The gp24 protein keeps the tail closed before DNA ejection by forming a cap and is located between the tail and the tip. The first protein in the tip after the cap is gp21,

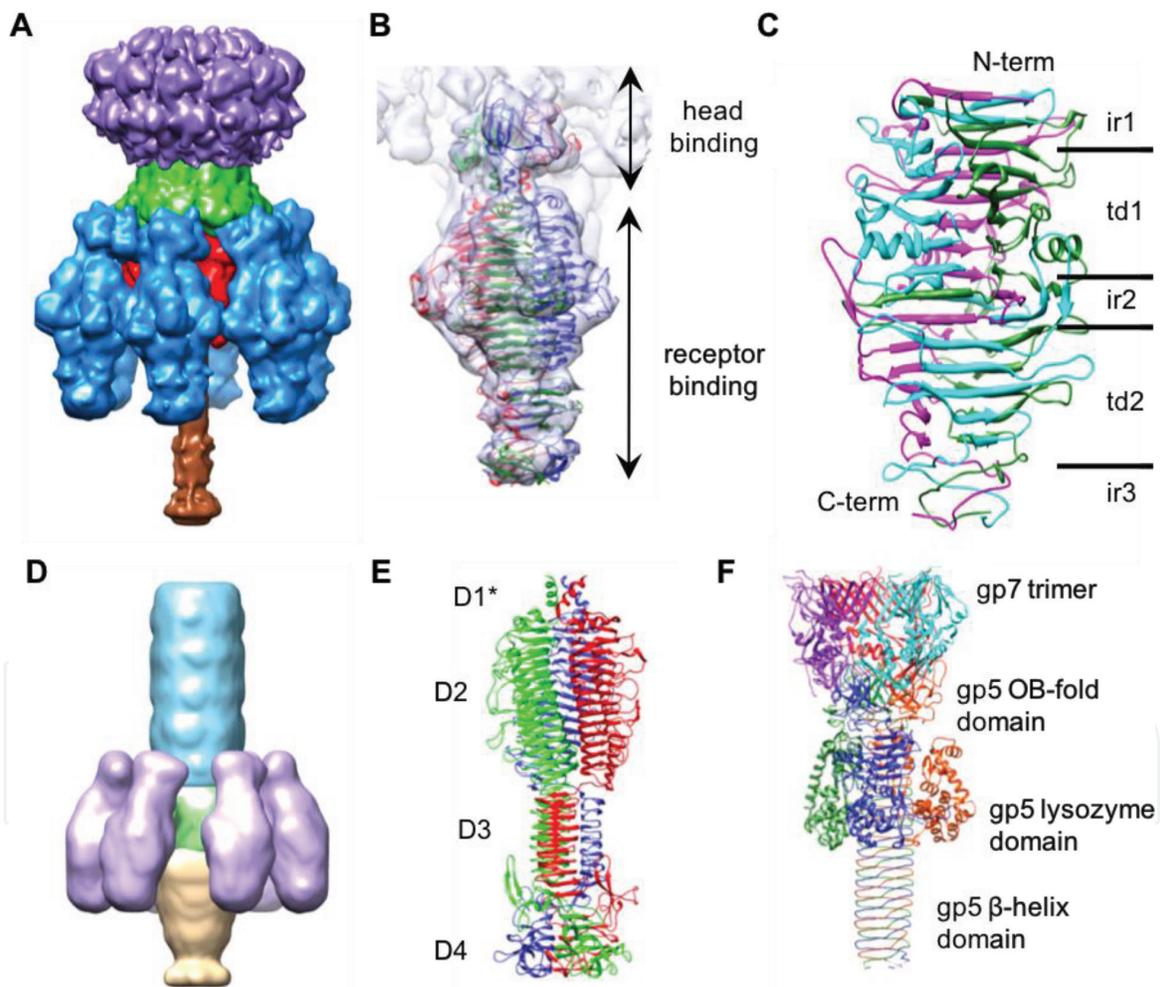


Figure 7. Adsorption apparatus. (A) The EM map (EMD-5051) of P22 coloured according to the constituent proteins. The PP (gp1), the gp4 12-mer, the gp10 6-mer, the gp9 tailspikes and the gp26 cell-puncturing needle are in purple, green, red, blue and brown, respectively. (B) The crystal structures for the N-terminal head-binding domain (1LKT) and the C-terminal receptor-binding domain (1TYU) of P22 gp9 docked into the cryo-EM density (EMDB-5051). (C) The receptor-binding carboxy-terminal domain of T5 tail fibre pb1 (5AQ5). The five different domain regions are labelled. The N- and C-termini are indicated. (D) The EM map (EMD-8868) of the TW1 tail showing the TP gp12 (in blue), gp15 (in green), gp16, gp18, gp27 (together in light brown) and the gp19 tailspikes (in purple). (E) The ϕ 29 tailspike protein gp12 (3SUC). The four domains D1*, D2, D3 and D4 are labelled. (F) The crystal structure (1K28) of the cell-puncturing device of T4 (gp27 ± gp5* ± gp5C)₃. Each monomer within in the gp5 trimer is coloured in blue, green and orange, and each one in gp27 is coloured in magenta, red and cyan.

and a structural similarity has suggested that SPP1 gp21 has a fold like the P22 tailspike [85]. When the P22 tailspike is fitted to the SPP1 density, the main domain of the P22 tailspike [124] occupies ~70% of the broad flattened area. The most remote protein in the tip is gp19.1, and the predicted secondary structure was structurally similar to the head-binding domain of the P22 tailspike [124]; and fitting three copies of the P22 trimer accounted for all the density in that region. During infection the tip is lost so DNA can pass into the cell and the cap remains in an open state.

T5 phage. The crystal structure of the T5 DTP pb9 has shown that it has two domains. The A-domain has a barrel-like fold with structural similarity to the N-domains of other phage DTPs [118, 125, 126]. In spite of low sequence identity, these proteins form a hexameric ring that occupies the central core of the baseplate. The peripheral B-domain has an oligosaccharide-/oligonucleotide-binding (OB) fold [127]. The attachment of phage T5 to the host cell is assisted by three side tail fibres attached to the distal end of the tail [107, 127, 128], and they are homo-trimers of the pb1 (1396 aa). The trimeric structure of the receptor-binding carboxy-terminal domain 970–1263(aa) was determined at 2.3 Å using X-ray crystallography [107] and could be divided into five different regions (**Figure 7C**) based on the structure of the P22 spike [124]. The N-terminal region (989-1009 aa) is shaped by the β -strands of the three monomers that wrap around each other to form a threefold beta-helix [124, 129]. The first “interdigitated” region (**ir1**) is followed by a triangular domain (1010–1129 aa) where three concave β sheets form a β -prism (**td1**). The second interdigitated region **ir2** (1130–1160 aa) also forms a short triple beta-helix. A second triangular domain **td2** (1161–1238 aa) is a β -prism like **td1**. At the distal end of the fibre, the third interdigitated region (aa1239–1263), **ir3**, forms a tapered triple-helical structure making the end of the structure pointed (**Figure 7C**). There is some similarity in the structure with the P22 tailspike [124] as both have a β -helical domain, an **ir** region, a triangular beta-prism domain and a second **ir** domain (called caudal fin). The triangular β -prism of P22 is the most similar to **td2** of pb1 and has the same topology.

TW1 phage. **TW1** has an unusual tail organisation for a siphophage, as a cryo-EM reconstruction of the tail (23.5 Å) [55] revealed six spikes on the distal end from the head. They are attached to the central tail tube, similar to the spikes seen in podophages P22 and Sf6 [120, 130] (**Figure 7D**). The TW1 gp19 tailspike (TS) protein is homologous to the TS protein of the podophage HK620 [131] so the crystal structure of the HK620 TS protein was fitted into the TW1 appendages. The TW1 gp19 TSs are thought to be attached to the phage via the DTP gp15 protein. However, the size of TW1 gp15 and the EM density suggest that this protein does not have a peripheral OB-fold domain as seen in the DTP of phage T5 [127]. Below gp15 are gp16 and gp18, which form the central tip of the phage tail (**Figure 7D**) and are similar to phage λ proteins gpL and gpJ, respectively [132]. At the tip of the tail is gp27 (**Figure 7D**) which is homologous to peptidoglycan-degrading enzymes. Many phages that infect Gram-negative bacteria have lysozyme-like proteins in their tails which enter the periplasm to digest the peptidoglycan barrier [119].

ϕ 29 phage. The tail of **ϕ 29** has 12 appendages, which are similar to the tailspikes of phage P22 and are attached to the bulge of densities close to the capsid [133]. Each appendage is a trimer of gp12* (the cleavage product of gp12 which during maturation loses an 18 kDa C-terminal fragment). Although there is no sequence similarity with the P22 tailspike, the P22 tailspike domain structure gave a good fit into the peripheral component of the **ϕ 29** appendages [65, 124]. A construct of gp12 residues 89–854 was cleaved in vivo to give an N-terminal fragment (up to Ser691) and a C-terminal fragment (from Asp692) and crystal structures obtained for each [105]. They are both trimers, and the N-terminal part attaches to the virion and has three domains: D1* is a coiled coil, D2 is mostly a β -helix and D3

is also a β -helix (**Figure 7E**). The C-terminal domain D4 acts as a chaperone for trimer assembly and is cleaved by autocatalysis. The ϕ 29 structure attached to the lipid bilayer has been obtained by cryo-ET (34 Å) [134]. The structure is comparable to cryo-EM structures of mature ϕ 29 [56, 133, 135]. Tomographic reconstructions demonstrated the different stages of infection [134]. In the adsorption stage, the phage is tilted to the cell wall, and both the appendages and the tail seem to contact the cell surface. The tail tip protein helps the phage penetrate the cell wall. When it contacts the cytoplasmic membrane, a pore is created which allows the genome to be injected into the cell.

T4 phage. The structure of the baseplate in *Myoviridae* is complex as illustrated by the T4 phage. The sixfold symmetric baseplate is 270 Å long and about 520 Å in diameter at the base and is connected to the distal end of the tail [136, 137]. It is composed at least by 16 different proteins [137]. A star-shaped baseplate is formed by sequential binding of four different proteins to form a wedge shape [137]. Six wedges are arranged around the independently assembled hub. Finally, other proteins are added to form the complete baseplate. Once gp48 and gp54 have bound to the top of the central hub, polymerisation of the tail tube is initiated, and after gp25 has attached to them, then polymerisation of the sheath is initiated [138]. Crystal structures of these constituent proteins were fitted into EM structure, and this showed the location of the proteins [137]. Six long fibres and six short fibres are attached to the baseplate. The long fibres reversibly interact with the cell surface receptors [139]. After recognition, the baseplate comes closer to the cell surface allowing the six short tail fibres to bind irreversibly to the cell outer membrane. This process is accompanied by a large conformational change in the baseplate from a “high-energy” to a “low-energy” structure [93, 140]. This induces contraction of the tail sheath and allows the inner tail tube to pierce the outer host cell membrane and penetrate the inner membrane so that the genome is transferred directly to the host's cytoplasm.

The structure of the T4 baseplate was assembled *in vitro* from gp10, gp7, gp8, gp6 and gp53, and the crystal structure was determined (4.2 Å) [141]. This indicated interesting differences compared to the structures when they are separately crystallised. However, about two-thirds of the structure was missing, but a cryo-EM structure of the same construct (3.8 Å) provided the positions of these missing parts [142]. The structures of T4 baseplate in its pre- and post-host attachment states were determined at 4.11 and 6.77 Å, respectively, by cryo-EM [111]. By combining high-resolution structures of the individual baseplate proteins, the authors were able to build a pseudo-atomic model for the baseplate proteins. The crystal structure at 2.9 Å of the gp5–gp27 cell-puncturing device was fitted into the EM structure (**Figure 7F**) [143]. Positions of gp27, gp5C (the C-terminal β -helix domain of gp5) and gp5* (the N-terminal OB-fold domain and the lysozyme middle domain) were identified. A monomeric protein gp5.4 caps the tip of the gp5 β -helix to sharpen the central spike [144]. During infection this spike punctures the cell membrane, and the lysozyme domain of gp5 digests the peptidoglycan in the *E. coli* periplasm.

ϵ 15 phage. A 20 Å cryo-EM map of ϵ 15 showed six gp20 tailspikes extending out from one of the fivefold capsid vertices. Each tailspike is composed of two domains [89] and has slightly different orientations with respect to the capsid. Cryo-ET has been used to show the interaction of ϵ 15 phage with the cell and to visualise the process of how ϵ 15 infects its host *Salmonella anatum* [110]. Initially, the tailspikes attach to the host cell followed by the tail hub attaching to a putative cell receptor. A bowl-shaped density was observed beneath the tail hub at the beginning of infection. It was proposed that phage indents the host outer membrane looking for a secondary receptor or for puncturing the membrane. A tunnel is established through the cell wall which allows the DNA to enter the cell.

5. Conclusions

Structural studies of the currently known tailed phages have shown a common organisation, which implies that they have a single ancestor and diversity has arisen through evolution [37]. All phages have a similar pathway of self-assembly: a procapsid formed with the help of a SP (or sometimes a scaffolding domain); conformational changes induced by release of the SP create a space for the DNA, and assisted by DNA terminases, the genome is packaged into the procapsid. This step is typically named as the maturation of the capsid. The tail is then attached or assembled on the capsid to form the infectious virion. The MCPs are characterised by the HK97 capsid protein fold. However, phages have a very low sequence similarity, which leads to differences in how the capsid stability is arranged to withstand the high inner pressure of the genome. In some phages like HK97 and SPP1, the interactions between capsid proteins are strong and hold the capsid intact. In many phages the process of capsid maturation is linked to attachment of additional proteins that are named as auxiliary or decoration proteins. They are often essential to enhance the capsid stability. The HK97 capsid is held together by chain mail covalent links between the MCPs; in SPP1 and T5, the decoration proteins enhance stability of the capsid, but in λ , T4 and ϵ 15 phages, these proteins are essential for keeping DNA inside the capsid [19, 52, 53].

The HTIs play an important role in all tailed phages as they provide a channel for DNA to enter and exit the capsid and at the same time provide a covalent connection to either the preassembled tails or tails assembled on the capsid. They all contain a dodecameric PP positioned within the capsid at one of the fivefold vertices and that acts as a gatekeeper holding the DNA within the capsid even in very harsh environments. Like the capsid proteins, the PPs have a common fold with the conserved elements being involved in interactions with DNA [145]. They have mostly α -helical domains in their central part and β -layers in the wing domains that interact with the capsid to fix the PP position. Head completion proteins below the PP also have similar folds to each other.

A much higher level of divergence is reflected in phage tail structures. The most common feature in all long-tailed phages is a central tube with a large number (30–40) of three- or sixfold circular rings of the major TPs. There is structural similarity between these major TPs: they have a similar fold of a β -sandwich flanked by alpha-helices and loops that provide links between adjacent rings. The helical tails have a typical rise of about 40 Å and rotation of around 20° between adjacent rings. Some tails also have appendages, which appear to have an immunoglobulin-like fold. Very little is known about the organisation of tail sheaths that have some similarities with type VI secretion systems, but sometimes they have extra appendages like immunoglobulin domains to help phages recognise their host cells. There is also some structural similarity of the TP with the tail terminator proteins and proteins in the T4 sheath.

Even higher diversity is found in the adsorption apparatus which are responsible for the recognition of the host cells and signalling the opening of the gate for the genome release. The tip of phage SPP1 recognises its receptor; induces the tail to be attached to the outer membrane of the host cell after disconnection of the tip. At the same time this interaction generates a signal that open the PP gate keeper. The T4 phage has a significantly more complex system of a baseplate which undergoes several steps of complex conformational changes.

Interestingly, the receptor-binding proteins also have similar organisation: they are all trimers, usually intertwined with β -helical regions, and use their N-terminal domain to bind to the phage. Spikes and fibres are also found in many phages. However, the number of spikes or fibres varies significantly between phages. Podophages have

trimeric tailspikes to recognise the specific host cell for infection. Like other phage components, they vary from six fibres in phage T7 to 12 in phage ϕ 29, but they all have a β -helical fold. The fibres can have different roles within a phage, for instance, T4 has six long fibres that serve as host recognition and six short fibres which then extend and bind to the cell.

Antibiotics (especially of the broad-spectrum type) are very effective at killing infectious bacteria; however, they kill typically multiple bacterial species indiscriminately, thus destroying beneficial bacteria of the host microbiome as well. Since phages are specific to one species of bacteria, they are unlikely to perturb microbiome bacterial species. Current problems with antibiotic resistance require new approaches, and here phages can be used [12]. For medicinal purposes it is necessary to design a phage that will recognise the specific bacteria we want to eliminate [146]. Phages can be modified for high specificity in the recognition of pathogens. The high level of phage specificity is based on recognition of receptor characteristic for a given type of bacteria which is where the differences in the adsorption systems of different phages play a crucial role. The important task in studying phages is to find those that are able to kill only antibiotic-resistant bacteria. Here, the lytic phages are of most interest, since rather than stopping bacteria from producing a certain type of protein that will slow down the bacterium proliferation, like in the case of antibiotics, these phages destroy the bacteria's cell wall and cell membrane completely. In addition, many bacteria develop biofilm—a thick layer of viscous materials that protect them from antibiotics. Some phages are equipped with tools that can digest this biofilm [147]. There are some problems with phages, since they are easy to use for topical applications, but often specific medications have to be administered internally. For phages to be used for delivery of drugs, they need to be more precise in their action. Consequently, we need to modify them so that the infectivity will be efficient by replacing the genome with DNA encoding specific enzymes and the adsorption apparatus made more effective. To develop these medical approaches, we need to know the phage organisation and the interactions between protein components at the atomic level. To achieve this hybrid, methods should be used including structural biology, biochemistry and microbiology [21].

Acknowledgements

The authors are grateful to Dr. D. Houldershaw and Mr. Y. Goudetsidis for their computer support. The authors apologise for the incomplete coverage of known phage structures and have drawn on a limited subset, owing to space constraints.

Conflict of interest

The authors declare no conflict of interest.

IntechOpen

IntechOpen

Author details

Helen E. White and Elena V. Orlova*
Institute of Structural and Molecular Biology, Birkbeck College, London,
United Kingdom

*Address all correspondence to: e.orlova@mail.cryst.bbk.ac.uk

IntechOpen

© 2019 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Shors T. *Understanding Viruses*. Sudbury, USA: Jones and Bartlett Publishers; 2008. ISBN: 0-7637-2932-9
- [2] Brussow H, Hendrix RW. Phage genomics: Small is beautiful. *Cell*. 2002;**108**:13-16
- [3] Suttle CA. Marine viruses—Major players in the global ecosystem. *Nature Reviews. Microbiology*. 2007;**5**:801-812
- [4] Breitbart M. Marine viruses: Truth or dare. *Annual Review of Marine Science*. 2012;**4**:425-448
- [5] Coelho J, Woodford N, Turton J, Livermore DM. Multiresistant *Acinetobacter* in the UK: How big a threat? *Journal of Hospital Infection*. 2004;**58**:167-169
- [6] Hanlon GW. Bacteriophages: An appraisal of their role in the treatment of bacterial infections. *International Journal of Antimicrobial Agents*. 2007;**30**:118-128
- [7] Burrowes B, Harper DR, Anderson J, McConville M, Enright MC. Bacteriophage therapy: Potential uses in the control of antibiotic-resistant pathogens. *Expert Review of Anti-Infective Therapy*. 2011;**9**:775-785
- [8] Coates AR, Halls G, Hu Y. Novel classes of antibiotics or more of the same? *British Journal of Pharmacology*. 2011;**163**(1):184-194. DOI: 10.1111/j.1476-5381.2011.01250.x
- [9] Brussow H, Kutter E. Phage ecology. In: Kutter E, Sulakvelidze A, editors. *Bacteriophages: Biology and applications*. Boca Raton, FL: CRC Press; 2005. pp. 129-163
- [10] Malik DJ, Sokolov IJ, Vinner GK, Mancuso F, Cinquerrui S, Vladislavljivic GT, et al. Formulation, stabilisation and encapsulation of bacteriophage for phage therapy. *Advances in Colloid and Interface Science*. 2017;**249**:100-133
- [11] Weinbauer MG. Ecology of prokaryotic viruses. *FEMS Microbiology Reviews*. 2004;**28**:127-181
- [12] Criscuolo E, Spadini S, Lamanna J, Ferro M, Burioni R. Bacteriophages and their immunological applications against infectious threats. *Journal of Immunology Research*. 2017;**2017**:3780697. Published online: Apr 6, 2017. DOI: 10.1155/2017/3780697
- [13] Pietilä MK, Demina TA, Atanasova NS, Oksanen HM, Bamford DH. Archaeal viruses and bacteriophages: Comparisons and contrasts. *Trends in Microbiology*. 2014;**22**:6334-6344
- [14] Orlova EV. Bacteriophages and Their Structural Organisation. In: Kurtböke İ, editor. *Bacteriophages*. IntechOpen; 2012. DOI: 10.5772/1065. ISBN: 978-953-51-0272-4
- [15] Ackermann HW. Classification of bacteriophages. In: Calendar R, editor. *The Bacteriophages*. New York, USA: Oxford University Press; 2006. pp. 8-16. ISBN: 0-19-514850-9
- [16] Drenth J. *Principles of Protein X-Ray Crystallography*. New York: Springer-Verlag; 2007
- [17] Javed A, Christodoulou J, Cabrita LD, Orlova EV. The ribosome and its role in protein folding: Looking through a magnifying glass. *Acta Crystallographica. Section D, Structural Biology*. 2017;**73**(Pt 6):509-521. DOI: 10.1107/S2059798317007446
- [18] Grimes JM, Burroughs JN, Gouet P, Diprose JM, Malby R, Zióntara S, et al. The atomic structure of the bluetongue virus core. *Nature*. 1998;**395**:470-478

- [19] Wikoff WR, Liljas L, Duda RL, Tsuruta H, Hendrix RW, Johnson JE. Topologically linked protein rings in the bacteriophage HK97 capsid. *Science*. 2000;**289**:2129-2133
- [20] Cuniasse P, Tavares P, Orlova EV, Zinn-Justin S. Structures of biomolecular complexes by combination of NMR and cryoEM methods. *Current Opinion in Structural Biology*. 2017;**43**:104-113. DOI: 10.1016/j.sbi.2016.12.008
- [21] Lengyel J, Hnath E, Storms M, Wohlfarth T. Towards an integrative structural biology approach: combining Cryo-TEM, X-ray crystallography, and NMR. *Journal of Structural and Functional Genomics*. 2014;**15**:117-124
- [22] Ram S, Ward ES, Ober RJ. Beyond Rayleigh's criterion: A resolution measure with application to single-molecule microscopy. *Proceedings of the National Academy of Sciences of the United States of America*. 2006;**103**(12):4457-4462
- [23] Orlova EV, Saibil HR. Structural analysis of macromolecular assemblies by electron microscopy. *Chemical Reviews*. 2011;**111**(12):7710-7748. DOI: 10.1021/cr100353t
- [24] Harris JR. *Negative Staining and Cryoelectron Microscopy, The Thin Film Techniques*. Oxford, UK: BIOS Scientific Publishers; 1997. ISBN: 1859961207
- [25] Adrian M, Dubochet J, Lepault J, McDowell AW. Cryo-electron microscopy of viruses. *Nature*. 1984;**308**(5954):32-36
- [26] Ruskin RS, Yu Z, Grigorieff N. Quantitative characterization of electron detectors for transmission electron microscopy. *Journal of Structural Biology*. 2013;**184**(3):385-393. DOI: 10.1016/j.jsb.2013.10.016. Published online: Nov 1, 2013, 10.1016/j.jsb.2013.10.016
- [27] Bartesaghi A, Aguerrebere C, Falconieri V, Banerjee S, Earl LA, Zhu X, et al. Atomic resolution Cryo-EM structure of β -Galactosidase. *Structure*. 2018;**26**(6):848-856.e3. DOI: 10.1016/j.str.2018.04.004
- [28] Wan W, Briggs JA. Cryo-electron tomography and subtomogram averaging. *Methods in Enzymology*. 2016;**579**:329-367. DOI: 10.1016/bs.mie.2016.04.014
- [29] Briggs JAG. Structural biology in situ—The potential of subtomogram averaging. *Current Opinion in Structural Biology*. 2013;**23**:261-267
- [30] Tavares P. The bacteriophage head-to-tail interface. *Sub-Cellular Biochemistry*. 2018;**88**:305-328. DOI: 10.1007/978-981-10-8456-0_14
- [31] Orlova EV, Dube P, Beckmann E, Zemlin F, Lurz R, Trautner TA, et al. Structure of the 13-fold symmetric portal protein of bacteriophage SPP1. *Nature Structural Biology*. 1999;**6**:842-846
- [32] Cingolani G, Moore SD, Prevelige PE Jr, Johnson JE. Preliminary crystallographic analysis of the bacteriophage P22 portal protein. *Journal of Structural Biology*. 2002;**139**:46-54
- [33] Cerritelli ME, Trus BL, Smith CS, Cheng N, Conway JF, Steven AC. A second symmetry mismatch at the portal vertex of bacteriophage T7: 8-fold symmetry in the procapsid core. *Journal of Molecular Biology*. 2003;**327**:1-6
- [34] Trus BL, Cheng N, Newcomb WW, Homa FL, Brown JC, Steven AC. Structure and polymorphism of the UL6 portal protein of herpes simplex virus type 1. *Journal of Virology*. 2004;**78**:12668-12671

- [35] Lorenzen K, Olia AS, Uetrecht C, Cingolani G, Heck AJ. Determination of stoichiometry and conformational changes in the first step of the P22 tail assembly. *Journal of Molecular Biology*. 2008;**379**:385-396
- [36] Hendrix RW, Smith MCM, Burns RN, Ford ME, Hatfull GF. Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. *Proceedings of the National Academy of Sciences of the United States of America*. 1999;**96**:2192-2197
- [37] Bamford DH, Grimes JM, Stuart DI. What does structure tell us about virus evolution? *Current Opinion in Structural Biology*. 2005;**15**:655-663
- [38] Wikoff WR, Conway JF, Tang J, Lee KK, Gan L, Cheng N, et al. Time-resolved molecular dynamics of bacteriophage HK97 capsid maturation interpreted by electron cryo-microscopy and X-ray crystallography. *Journal of Structural Biology*. 2006;**153**(3):300-306
- [39] Huet A, Conway JF, Letellier L, Boulanger P. In vitro assembly of the T = 13 procapsid of bacteriophage T5 with its scaffolding domain. *Journal of Virology*. 2010;**84**:9350-9358. DOI: 10.1128/JVI.00942-10
- [40] Veesler D, Quispe J, Grigorieff N, Potter CS, Carragher B, Johnson JE. Maturation in action: CryoEM study of a viral capsid caught during expansion. *Structure*. 2012;**20**:1384-1390
- [41] Baker TS, Olson NH, Fuller SD. Adding the third dimension to virus life cycles: Three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. *Microbiology and Molecular Biology Reviews*. 1999;**63**(4):862-922. (Review)
- [42] Helgstrand C, Wikoff WR, Duda RL, Hendrix RW, Johnson JE, Liljas L. The refined structure of a protein catenane: the HK97 bacteriophage capsid at 3.44 Å resolution. *Journal of Molecular Biology*. 2003;**334**:885-899
- [43] Atanasova NS, Bamford DH, Oksanen HM. Haloarchaeal virus morphotypes. *Biochimie*. 2015;**118**:333-343
- [44] Baker ML, Jiang W, Rixon FJ, Chiu W. Common ancestry of herpesviruses and tailed DNA bacteriophages. *Journal of Virology*. 2005;**79**:14967-14970
- [45] Huet A, Makhov AM, Huffman JB, Vos M, Homa FL, Conway JF. Extensive subunit contacts underpin herpesvirus capsid stability and interior-to-exterior allostery. *Nature Structural & Molecular Biology*. 2016;**23**:531-539
- [46] Yu X, Jih J, Jiang J, Hong Zhou Z. Atomic structure of the human cytomegalovirus capsid with its securing tegument layer of pp150. *Science*. 2017;**356**:1350
- [47] Lander GC, Evilevitch A, Jeembaeva M, Potter CS, Carragher B, Johnson JE. Bacteriophage lambda stabilization by auxiliary protein gpD: Timing, location, and mechanism of attachment determined by cryoEM. *Structure*. 2008;**16**(9):1399-1406. DOI: 10.1016/j.str.2008.05.016
- [48] Yang F, Forrer P, Dauter Z, Conway JF, Cheng N, Cerritelli ME, et al. Novel fold and capsid-binding properties of the lambda-phage display platform protein gpD. *Nature Structural Biology*. 2000;**7**(3):230-237
- [49] Jiang W, Baker ML, Jakana J, Weigele PR, King J, Chiu W. Backbone structure of the infectious epsilon15 virus capsid revealed by electron cryomicroscopy. *Nature*. 2008;**451**(7182):1130-1134
- [50] Baker ML, Hryc CF, Zhang Q, Wu W, Jakana J, Haase-Pettingell C, et al. Validated near-atomic resolution structure of bacteriophage epsilon15 derived from cryo-EM and modeling.

Proceedings of the National Academy of Sciences of the United States of America. 2013;**110**:12301-12306

[51] Conway JF, Wikoff WR, Cheng N, Duda RL, Hendrix RW, Johnson JE, et al. Virus maturation involving large subunit rotations and local refolding. *Science*. 2001;**292**:744-748

[52] Vernhes E, Renouard M, Gilquin B, Cuniasse P, Durand D, England P, et al. High affinity anchoring of the decoration protein pb10 onto the bacteriophage T5 capsid. *Scientific Reports*. 2017;**7**:41662. DOI: 10.1038/srep41662

[53] White HE, Sherman MB, Brasilès S, Jacquet E, Seavers P, Tavares P, et al. Capsid structure and its stability at the late stages of bacteriophage SPP1 assembly. *Journal of Virology*. 2012;**12**:6768-6777. DOI: 10.1128/JVI.00412-12

[54] Bebeacua C, Lai L, Vegge CS, Brøndsted L, van Heel M, Velesler D, et al. Visualizing a complete Siphoviridae member by single-particle electron microscopy: The structure of lactococcal phage TP901-1. *Journal of Virology*. 2013;**87**(2):1061-1068. DOI: 10.1128/JVI.02836-12

[55] Wang Z, Hardies SC, Fokine A, Klose T, Jiang W, Cho BC, et al. Structure of the marine siphovirus TW1: Evolution of capsid-stabilizing proteins and tail spikes. *Structure*. 2018;**26**:238-248

[56] Tang J, Olson N, Jardine PJ, Grimes S, Anderson DL, Baker TS. DNA poised for release in bacteriophage phi29. *Structure*. 2008;**16**:935-943

[57] Guo F, Liu Z, Fang PA, Zhang Q, Wright ET, Wu W, et al. Capsid expansion mechanism of bacteriophage T7 revealed by multistate atomic models derived from cryo-EM reconstructions. *Proceedings of the National Academy of Sciences of the United States of America*. 2014;**111**(43):E4606-E4614. DOI: 10.1073/pnas.1407020111

[58] Parent KN, Khayat R, Tu LH, Suhanovsky MM, Cortines JR, Teschke CM, et al. P22 coat protein structures reveal a novel mechanism for capsid maturation: stability without auxiliary proteins or chemical crosslinks. *Structure*. 2010;**18**(3):390-401. DOI: 10.1016/j.str.2009.12.014

[59] Chen DH, Baker ML, Hryc CF, DiMaio F, Jakana J, Wu W, et al. Structural basis for scaffolding-mediated assembly and maturation of a dsDNA virus. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;**108**(4):1355-1360

[60] Hryc CF, Chen DH, Afonine PV, Jakana J, Wang Z, Haase-Pettingell C, et al. Accurate model annotation of a near-atomic resolution cryo-EM map. *Proceedings of the National Academy of Sciences of the United States of America*. 2017;**114**:3103-3108

[61] Fokine A, Leiman PG, Shneider MM, Ahvazi B, Boeshans KM, Steven AC, et al. Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry. *Proceedings of the National Academy of Sciences of the United States of America*. 2005;**102**(20):7163-7168

[62] Chen Z, Sun L, Zhang Z, Fokine A, Padilla-Sanchez V, Hanein D, et al. Cryo-EM structure of the bacteriophage T4 isometric head at 3.3-angstrom resolution and its relevance to the assembly of icosahedral viruses. *Proceedings of the National Academy of Sciences of the United States of America*. 2017;**114**:E8184-E8193

[63] Dai X, Zhou ZH. Structure of the herpes simplex virus 1 capsid with associated tegument protein complexes. *Science*. 2018;**360**(6384). pii: eaa07298. DOI: 10.1126/science.aa07298

[64] Holden HM, Ito M, Hartshorne DJ, Rayment I. X-ray structure determination

of telokin, the C-terminal domain of myosin light chain kinase, at 2.8 Å resolution. *Journal of Molecular Biology*. 1992;227:840-851

[65] Xiang Y, Morais MC, Battisti AJ, Grimes S, Jardine PJ, Anderson DL, et al. Structural changes of bacteriophage ϕ 29 upon DNA packaging and release. *The EMBO Journal*. 2006;25:5229-5239

[66] Morais MC, Choi KH, Koti JS, Chipman PR, Anderson DL, Rossmann MG. Conservation of the capsid structure in tailed dsDNA bacteriophages: The pseudoatomic structure of ϕ 29. *Molecular Cell*. 2005;18:149-159

[67] Marchler-Bauer A, Anderson JB, Cherukuri PF, DeWeese-Scott C, Geer LY, Gwadz M, et al. CDD: A conserved domain database for protein classification. *Nucleic Acids Research*. 2005;33:D192-D196

[68] Iwasaki K, Trus BL, Wingfield PT, Cheng N, Campusano G, Rao VB, et al. Molecular architecture of bacteriophage T4 capsid: Vertex structure and bimodal binding of the stabilizing accessory protein, Soc. *Journal of Virology*. 2000;271:321-333

[69] Sun L, Zhang X, Gao S, Rao PA, Padilla-Sanchez V, Chen Z, et al. Cryo-EM structure of the bacteriophage T4 portal protein assembly at near-atomic resolution. *Nature Communications*. 2015;6:7548

[70] Fokine A, Chipman PR, Leiman PG, Mesyanzhinov VV, Rao VB, Rossmann MG. Molecular architecture of the prolate head of bacteriophage T4. *Proceedings of the National Academy of Sciences of the United States of America*. 2004;101:6003-6008

[71] Fokine A, Islam MZ, Zhang Z, Bowman VD, Rao VB, Rossmann MG. Structure of the three N-terminal immunoglobulin domains of the highly

immunogenic outer capsid protein from a T4-like bacteriophage. *Journal of Virology*. 2011;85:8141-8148

[72] Qin L, Fokine A, O'Donnell E, Rao VB, Rossmann MG. Structure of the small outer capsid protein, Soc: A clamp for stabilizing capsids of T4-like phages. *Journal of Molecular Biology*. 2010;395(4):728-741. DOI: 10.1016/j.jmb.2009.10.007

[73] Heymann JB, Cheng N, Newcomb WW, Trus BL, Brown JC, Steven AC. Dynamics of herpes simplex virus capsid maturation visualized by time-lapse cryo-electron microscopy. *Nature Structural Biology*. May 2003;10(5):334-341

[74] Cardarelli L, Lam R, Tuite A, Baker LA, Sadowski PD, Radford DR, et al. The crystal structure of bacteriophage HK97 gp6: defining a large family of head-tail connector proteins. *Journal of Molecular Biology*. 2010;395:754-768

[75] Simpson AA, Tao Y, Leiman PG, Badasso MO, He Y, Jardine PJ, et al. Structure of the bacteriophage ϕ 29 DNA packaging motor. *Nature*. 2000;408:745-750

[76] Guasch A, Pous J, Ibarra B, Gomis-Ruth FX, Valpuesta JM, Sousa N, et al. Detailed architecture of a DNA translocating machine: The high-resolution structure of the bacteriophage ϕ 29 connector particle. *Journal of Molecular Biology*. 2002;315:663-676

[77] Lebedev AA, Krause MH, Isidro AL, Vagin A, Orlova EV, Turner J, et al. Structural framework for DNA translocation via the viral portal protein. *The EMBO Journal*. 2007;26:1984-1994

[78] Olia AS, Prevelige PE Jr, Johnson JE, Cingolani G. Three-dimensional structure of a viral genome-delivery portal vertex. *Nature Structural & Molecular Biology*. 2011;18:597-603

- [79] Agirrezabala X, Martín-Benito J, Valle M, González JM, Valencia A, Valpuesta JM, et al. Structure of the connector of bacteriophage T7 at 8 Å resolution: Structural homologies of a basic component of a DNA translocating machinery. *Journal of Molecular Biology*. 2005;**347**:895-902
- [80] Sun S, Kondabagil K, Draper B, Alam TI, Bowman VD, Zhang Z, et al. The structure of the phage T4 DNA packaging motor suggests a mechanism dependent on electrostatic forces. *Cell*. 2008;**135**:1251-1262
- [81] Oliveira L, Tavares P, Alonso JC. Headful DNA packaging: Bacteriophage SPP1 as a model system. *Virus Research*. 2013;**173**(2):247-259. DOI: 10.1016/j.virusres.2013.01.021
- [82] Chaban Y, Lurz R, Brasiles S, Cornilleau C, Karreman M, Zinn-Justin S, et al. Structural rearrangements in the phage head-to-tail interface during assembly and infection. *Proceedings of the National Academy of Sciences of the United States of America*. 2015;**112**:7009-7014
- [83] Orlova EV, Gowen B, Dröge A, Stiege A, Weise F, Lurz R, et al. Structure of a viral DNA gatekeeper at 10 Å resolution by cryoelectron microscopy. *The EMBO Journal*. 2003;**22**(6):1255-1262
- [84] Lhuillier S, Gallopin M, Gilquin B, Brasilès S, Lancelot N, Letellier G, et al. Structure of bacteriophage SPP1 head-to-tail connection reveals mechanism for viral DNA gating. *Proceedings of the National Academy of Sciences of the United States of America*. 2009;**106**(21):8507-8512
- [85] Plisson C, White HE, Auzat I, Zafarani A, São-Jose C, Lhuillier S, et al. Structure of bacteriophage SPP1 tail reveals trigger for DNA ejection. *The EMBO Journal*. 2007;**26**:3720-3728
- [86] Auzat I, Dröge A, Weise F, Lurz R, Tavares P. Origin and function of the two major tail proteins of bacteriophage SPP1. *Molecular Microbiology*. 2008 Nov;**70**(3):557-569. DOI: 10.1111/j.1365-2958.2008.06435.x
- [87] Cuervo A, Pulido-Cid M, Chagoyen M, Arranz R, González-García VA, Garcia-Doval C, et al. Structural characterization of the bacteriophage T7 tail machinery. *The Journal of Biological Chemistry*. 2013;**288**:26290-26299
- [88] Pintilie G, Chen DH, Haase-Pettingell CA, King JA, Chiu W. Resolution and probabilistic models of components in CryoEM maps of mature P22 bacteriophage. *Biophysical Journal*. 2016;**110**(4):827-839. DOI: 10.1016/j.bpj.2015.11.3522
- [89] Jiang W, Chang J, Jakana J, Weigele P, King J, Chiu W. Structure of epsilon15 bacteriophage reveals genome organization and DNA packaging/injection apparatus. *Nature*. 2006;**439**:612-616
- [90] McElwee M, Vijayakrishnan S, Rixon F, Bhella D. Structure of the herpes simplex virus portal-vertex. *PLoS Biology*. 2018;**16**(6):e2006191. DOI: 10.1371/journal.pbio.2006191
- [91] Lander GC, Khayat R, Li R, Prevelige PE, Potter CS, Carragher B, et al. The P22 tail machine at subnanometer resolution. *Structure*. 2009;**17**(6):789-799
- [92] Tang J, Lander GC, Olia A, Li R, Casjens S, Prevelige P Jr, et al. Peering down the barrel of a bacteriophage portal: The genome packaging and release valve in p22. *Structure*. 2011;**19**:496-502
- [93] Leiman PG, Chipman PR, Kostyuchenko VA, Mesyanzhinov VV, Rossmann MG. Three-dimensional rearrangement of proteins in the tail

of bacteriophage T4 on infection of its host. *Cell*. 2004;**118**:419-429

[94] Isidro A, Santos MA, Henriques AO, Tavares P. The high-resolution functional map of bacteriophage SPP1 portal protein. *Molecular Microbiology*. 2004;**51**:949-962

[95] Newcomb WW, Juhas RM, Thomsen DR, Homa FL, Burch AD, Weller SK, et al. The UL6 gene product forms the portal for entry of DNA into the herpes simplex virus capsid. *Journal of Virology*. 2001;**75**:10923-10932

[96] Cardone G, Winkler DC, Trus BL, Cheng N, Heuser JE, Newcomb WW, et al. Visualization of the herpes simplex virus portal in situ by cryo-electron tomography. *Virology*. 2007;**361**:426-434

[97] Arnaud CA, Effantin G, Vives C, Engilberge S, Bacia M, Boulanger P, et al. Bacteriophage T5 tail tube structure suggests a trigger mechanism for Siphoviridae DNA ejection. *Nature Communications*. 2017;**8**:1953-1953

[98] Pell LG, Gasmi-Seabrook GM, Morais M, Neudecker P, Kanelis V, Bona D, et al. The solution structure of the C-terminal Ig-like domain of the bacteriophage λ tail tube protein. *Journal of Molecular Biology*. 2010;**403**:468-479

[99] Effantin G, Boulanger PE, Neumann E, Letellier L, Conway JF. Bacteriophage T5 structure reveals similarities with HK97 and T4 suggesting evolutionary relationships. *Journal of Molecular Biology*. 2006;**361**:993-1002

[100] Papadopoulos S, Smith PR. The structure of the tail of the bacteriophage phi CbK. *Journal of Ultrastructure Research*. 1982;**80**:62-70

[101] Katsura I. Tail Assembly and Injection. *Lambda II*. Plainview, NY: Cold Spring Harbor; 1983. pp. 331-346

[102] Pell LG, Kanelis V, Donaldson LW, Howell PL, Davidson AR. The phage lambda major tail protein structure reveals a common evolution for long-tailed phages and the type VI bacterial secretion system. *Proceedings of the National Academy of Sciences of the United States of America*. 2009;**106**(11):4160-4165. DOI: 10.1073/pnas.0900044106

[103] Pell LG, Liu A, Edmonds L, Donaldson LW, Howell PL, Davidson AR. The X-ray crystal structure of the phage lambda tail terminator protein reveals the biologically relevant hexameric ring structure and demonstrates a conserved mechanism of tail termination among diverse long-tailed phages. *Journal of Molecular Biology*. 2009;**389**(5):938-951. DOI: 10.1016/j.jmb.2009.04.072. [Epub 2009 May 6]

[104] Chagot B, Auzat I, Gallopin M, Petitpas I, Gilquin B, Tavares P, et al. Solution structure of gp17 from the Siphoviridae bacteriophage SPP1: Insights into its role in virion assembly. *Proteins*. 2012;**80**(1):319-326. DOI: 10.1002/prot.23191

[105] Xiang Y, Leiman PG, Li L, Grimes S, Anderson DL, Rossmann MG. Crystallographic insights into the autocatalytic assembly mechanism of a bacteriophage tailspike. *Molecular Cell*. 2009;**34**:375-386. DOI: 10.1016/j.molcel.2009.04.009

[106] Xu J, Gui M, Wang D, Xiang Y. The bacteriophage ϕ 29 tail possesses a pore-forming loop for cell membrane penetration. *Nature*. 2016;**534**:544-547

[107] Garcia-Doval C, Castón JR, Luque D, Granell M, Otero JM, Llamas-Saiz AL, et al. Conformational changes leading to T7 DNA delivery upon interaction with the bacterial receptor. *Viruses*. 2015;**7**(12):6424-6440. DOI: 10.3390/v7122946

- [108] Steinbacher S, Baxa U, Miller S, Weintraub A, Seckler R, Huber R. Crystal structure of phage P22 tailspike protein complexed with *Salmonella* sp. O-antigen receptors. *Proceedings of the National Academy of Sciences of the United States of America*. 1996;**93**(20):10584-10588
- [109] Steinbacher S, Miller S, Baxa U, Budisa N, Weintraub A, Seckler R, et al. Phage P22 tailspike protein: crystal structure of the head-binding domain at 2.3 Å, fully refined structure of the endorhamnosidase at 1.56 Å resolution, and the molecular basis of O-antigen recognition and cleavage. *Journal of Molecular Biology*. 1997;**267**:865-880
- [110] Chang JT, Schmid MF, Haase-Pettingell C, Weigele PR, King JA, Chiu W. Visualizing the structural changes of bacteriophage Epsilon15 and its *Salmonella* host during infection. *Journal of Molecular Biology*. 2010;**402**(4):731-740. DOI: 10.1016/j.jmb.2010.07.058
- [111] Taylor NM, Prokhorov NS, Guerrero-Ferreira RC, Shneider MM, Browning C, Goldie KN, et al. Structure of the T4 baseplate and its function in triggering sheath contraction. *Nature*. 2016;**533**:346-352
- [112] Zheng W, Wang F, Taylor NMI, Guerrero-Ferreira RC, Leiman PG, Egelman EE. Refined Cryo-EM structure of the T4 tail tube: Exploring the lowest dose limit. *Structure*. 2017;**25**:1436-1441
- [113] Fokine A, Zhang Z, Kanamaru S, Bowman VD, Aksyuk AA, Arisaka F, et al. The molecular architecture of the bacteriophage T4 neck. *Journal of Molecular Biology*. 2013;**425**:1731-1744
- [114] Aksyuk AA, Leiman PG, Shneider MM, Mesyanzhinov VV, Rossmann MG. The structure of gene product 6 of bacteriophage T4, the hinge-pin of the baseplate. *Structure*. 2009;**17**:800-808
- [115] Aksyuk AA, Kurochkina LP, Fokine A, Forouhar F, Mesyanzhinov VV, Tong L, et al. Structural conservation of the Myoviridae phage tail sheath protein fold. *Structure*. 2011;**19**:1885-1894
- [116] Veessler D, Cambillau C. A common evolutionary origin for tailed bacteriophage functional modules and bacterial machineries. *Microbiology and Molecular Biology Reviews*. 2011;**75**:423-433
- [117] Davidson AR, Cardarelli L, Pell LG, Radford DR, Maxwell KL. Long noncontractile tail machines of bacteriophages. *Advances in Experimental Medicine and Biology*. 2012;**726**:115-142. DOI: 10.1007/978-1-4614-0980-9_6
- [118] Veessler D, Robin G, Lichière J, Auzat I, Tavares P, Bron P, et al. Crystal structure of bacteriophage SPP1 distal tail protein (gp19.1): A baseplate hub paradigm in Gram-positive infecting phages. *The Journal of Biological Chemistry*. 2010;**285**:36666-36673
- [119] Xiang Y, Morais MC, Cohen DN, Bowman VD, Anderson DL, Rossmann MG. Crystal and cryoEM structural studies of a cell wall degrading enzyme in the bacteriophage phi29 tail. *Proceedings of the National Academy of Sciences of the United States of America*. 2008;**105**(28): 9552-9557. DOI: 10.1073/pnas.0803787105
- [120] Tang L, Marion WR, Cingolani G, Prevelige PE, Johnson JE. Three-dimensional structure of the bacteriophage P22 tail machine. *The EMBO Journal*. 2005;**24**:2087-2095
- [121] Chang J, Weigele P, King J, Chiu W, Jiang W. Cryo-EM asymmetric reconstruction of bacteriophage P22 reveals organization of its DNA packaging and infecting machinery. *Structure*. 2006;**14**(6):1073-1082

- [122] Tang L, Gilcrease EB, Casjens SR, Johnson JE. Highly discriminatory binding of capsid-cementing proteins in bacteriophage λ . *Structure*. 2006;**14**(5):837-845
- [123] Olia AS, Casjens S, Cingolani G. Structure of phage P22 cell envelope-penetrating needle. *Nature Structural & Molecular Biology*. 2007;**14**(12):1221-1226
- [124] Steinbacher S, Seckler R, Miller S, Steipe B, Huber R, Reinemer P. Crystal structure of P22 tailspike protein: Interdigitated subunits in a thermostable trimer. *Science*. 1994;**265**:383-386
- [125] Sciara G, Bebeacua C, Bron P, Tremblay D, Ortiz-Lombardia M, Lichère J, et al. Structure of lactococcal phage ρ 2 baseplate and its mechanism of activation. *Proceedings of the National Academy of Sciences of the United States of America*. 2010;**107**:6852-6857
- [126] Veesler D, Spinelli S, Mahony J, Lichère J, Blangy S, Bricogne G, et al. Structure of the phage TP901-1 1.8MDabaseplate suggests an alternative host adhesion mechanism. *Proceedings of the National Academy of Sciences of the United States of America*. 2012;**109**:8954-8958. DOI: 10.1073/pnas.1200966109
- [127] Flayhan A, Vellieux FM, Lurz R, Maury O, Contreras-Martel C, Girard E, et al. Crystal structure of pb9, the distal tail protein of bacteriophage T5: A conserved structural motif among all siphophages. *Journal of Virology*. 2014;**88**:820-828
- [128] Zivanovic Y, Confalonieri F, Ponchon L, Lurz R, Chami M, Flayhan A, et al. Insights into bacteriophage T5 structure from analysis of its morphogenesis genes and protein components. *Journal of Virology*. 2014;**88**:1162-1174
- [129] Mitraki A, Papanikolopoulou K, van Raaij MJ. Natural triple β -stranded fibrous folds. *Advances in Protein Chemistry*. 2006;**73**:97-124
- [130] Parent KN, Gilcrease EB, Casjens SR, Baker TS. Structural evolution of the P22-like phages: comparison of Sf6 and P22 procapsid and virion architectures. *Virology*. 2012;**427**:177-188
- [131] Barbirz S, Muller JJ, Uetrecht C, Clark AJ, Heinemann U, Seckler R. Crystal structure of Escherichia coli phage HK620 tailspike: Podoviral tailspike endoglycosidase modules are evolutionarily related. *Molecular Microbiology*. 2008;**69**:303-316
- [132] Rajagopala SV, Casjens S, Uetz P. The protein interaction map of bacteriophage lambda. *BMC Microbiology*. 2011;**11**:213
- [133] Morais MC, Tao Y, Olson NH, Grimes S, Jardine PJ, Anderson DL, et al. Cryoelectron-microscopy image reconstruction of symmetry mismatches in bacteriophage ϕ 29. *Journal of Structural Biology*. 2001;**135**:38-46
- [134] Farley MM, Tu J, Kearns DB, Molineux IJ, Liu J. Ultrastructural analysis of bacteriophage Φ 29 during infection of Bacillus subtilis. *Journal of Structural Biology*. 2017;**197**(2):163-171. DOI: 10.1016/j.jsb.2016.07.019
- [135] Tao Y, Olson NH, Xu W, Anderson DL, Rossmann MG, Baker TS. Assembly of a tailed bacterial virus and its genome release studied in three dimensions. *Cell*. 1998;**95**:431-437
- [136] Leiman PG, Kanamaru S, Mesyanzhinov VV, Arisaka F, Rossmann MG. Structure and morphogenesis of bacteriophage T4. *Cellular and Molecular Life Sciences*. 2003;**60**(11):2356-2370
- [137] Leiman PG, Arisaka F, van Raaij MJ, Kostyuchenko VA, Aksyuk AA, Kanamaru S, et al. Morphogenesis

of the T4 tail and tail fibers.

Virology Journal. 2010;7:355. DOI: 10.1186/1743-422X-7-355

[138] Arisaka F, Yap ML, Kanamaru S, Rossmann MG. Molecular assembly and structure of the bacteriophage T4 tail. *Biophysical Reviews*. 2016;8(4):385-396. DOI: 10.1007/s12551-016-0230-x

[139] Crawford JT, Goldberg EB. The function of tail fibers in triggering baseplate expansion of bacteriophage T4. *Journal of Molecular Biology*. 1980;139(4):679-690

[140] Kostyuchenko VA, Leiman PG, Chipman PR, Kanamaru S, van Raaij MJ, Arisaka F, et al. Three-dimensional structure of bacteriophage T4 baseplate. *Nature Structural Biology*. 2003;10(9):688-693

[141] Yap ML, Rossmann MG. Structure and function of bacteriophage T4. *Future Microbiology*. 2014;9(12):1319-1327. DOI: 10.2217/fmb.14.91

[142] Yap ML, Klose T, Arisaka F, Speirc JA, Veessler D, Fokine A, et al. The role of bacteriophage T4 baseplate in regulating assembly and infection. *Proceedings of the National Academy of Sciences of the United States of America*. 2016;113(10):2654-2659

[143] Kanamaru S, Leiman PG, Kostyuchenko VA, Chipman PR, Mesyanzhinov VV, Arisaka F, et al. Structure of the cell-puncturing device of bacteriophage T4. *Nature*. 2002;415:553-557. DOI: 10.1038/415553a

[144] Shneider MM, Buth SA, Ho BT, Basler M, Mekalanos JJ, Leiman PG. PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature*. 2013;500:350-353. DOI: 10.1038/nature1245

[145] Tavares P, Zinn-Justin S, Orlova EV. Genome gating in tailed

bacteriophage capsids. *Advances in Experimental Medicine and Biology*. 2012;726:585-600. DOI: 10.1007/978-1-4614-0980-9_25. [Review. PMID: 22297531]

[146] Ando H, Lemire S, Pires DP, Lu TK. Engineering modular viral scaffolds for targeted bacterial population editing. *Cell Systems*. 2015;1(3):187-196

[147] Abedon ST. Ecology of anti-biofilm agents II: Bacteriophage exploitation and biocontrol of biofilm bacteria. *Pharmaceuticals (Basel)*. 2015;8(3):559-589. DOI: 10.3390/ph8030559