

Chapter 11

Structure and Assembly of Complex Viruses

Carmen San Martín

Abstract Viral particles consist essentially of a proteinaceous capsid protecting a genome and involved also in many functions during the virus life cycle. In simple viruses, the capsid consists of a number of copies of the same, or a few different proteins organized into a symmetric oligomer. Structurally complex viruses present a larger variety of components in their capsids than simple viruses. They may contain accessory proteins with specific architectural or functional roles; or incorporate non-proteic elements such as lipids. They present a range of geometrical variability, from slight deviations from the icosahedral symmetry to complete asymmetry or even pleomorphism. Putting together the many different elements in the virion requires an extra effort to achieve correct assembly, and thus complex viruses require sophisticated mechanisms to regulate morphogenesis. This chapter provides a general view of the structure and assembly of complex viruses.

Keywords Virus structure • Virus assembly • Symmetry • Capsid • Cementing proteins • Envelope • Symmetry mismatch • Scaffold • Maturation • Virus evolution

Abbreviations

3D	Three-dimensional
ABV	<i>Acidianus</i> bottle-shaped virus
ATV	<i>Acidianus</i> two-tailed virus
cryo-EM	Cryo-electron microscopy
dsDNA	Double-stranded DNA
dsRNA	Double-stranded RNA

C. San Martín (✉)

Department of Macromolecular Structure, Centro Nacional de Biotecnología (CSIC),
c/Darwin 3, Campus de Cantoblanco, 28049 Madrid, Spain
e-mail: carmen@cnb.csic.es

EMDB	Electron Microscopy Data Bank
GON	Group of nine
GOS	Group of six
PBCV-1	<i>Paramecium bursaria Chlorella</i> virus-1
PDB	Protein Data Bank
Sid	Size determination protein
SNDV	<i>Sulfolobus neozelandicus</i> droplet-shaped virus
ssDNA	Single-stranded DNA
SSIP-1	<i>Salisaeta</i> icosahedral phage 1
ssRNA	Single-stranded RNA
STIV	<i>Sulfolobus</i> turreted icosahedral virus

11.1 Introduction

A viral particle consists essentially of a proteinaceous capsid with multiple roles in protection of the viral genome, cell recognition and entry, intracellular trafficking and controlled uncoating. Evolutionary forces have caused viruses to adopt different strategies to achieve these goals. Simple viruses (Chap. 10) generally build their capsids from a number of copies of the same, or a few different proteins, organized into a symmetric oligomer. In the case of complex viruses, capsid assembly requires further elaborations. What are the main characteristics that define a structurally complex virus?

Structural complexity on a virus often, but not necessarily, derives from the need to house a large genome, in which case a larger capsid is required. However, capsid or genome sizes by themselves are not determinants of complexity. For example, flexible filamentous viruses can reach lengths in the order of microns, but most of their capsid mass is built by a single capsid protein arranged in a helical pattern [1]. On the other hand, architecturally complex viruses such as HIV have moderate sized genomes (7–10 kb of single-stranded (ss) RNA) [2]. Structurally complex viruses incorporate a larger variety of components into their capsids than simple viruses. They may contain accessory proteins with specific architectural or functional roles; or incorporate non-proteic elements such as lipids.

The elaborated composition of complex virus particles often involves a rupture of the basic symmetry rules (Chap. 2), from a range of symmetry mismatches in icosahedral shells, to completely asymmetric or pleomorphic capsids. The more subtle departure from symmetry is the case when identical subunits occupy similar but slightly different environments, as in the case of quasi-equivalence in icosahedral shells (see Chap. 2). In other cases, virion components with different symmetry may interact with each other, forming a *symmetry mismatch* at the interface. In the extreme case, identical components may form morphological units with no

symmetry, or even assemble in a completely different manner for each realization of the virion – this property is called *pleomorphism*. Recent advances in cryo-electron microscopy (cryo-EM) (Chap. 3) and X-ray crystallography (Chap. 4) are helping to unveil the organization of complex viruses in great detail, including features that depart from strict icosahedral symmetry. Understanding the architectural details of asymmetric capsids is the most challenging problem, since structural biology techniques heavily rely on the use of symmetries to reach high resolution detail. Cryo-electron tomography (Chap. 3) is helping to advance our understanding of these viruses, although the resolution currently attained is still in the 3–5 nm range.

The presence of many different elements in the virion entails an extra effort to achieve correct assembly. Accordingly, complex virus morphogenesis requires sophisticated mechanisms, tightly regulated in space and time. Here we provide a general view of all these variations in complexity, finishing with a consideration on the evolutionary insights provided by structural studies on complex viruses.

11.2 Molecular Composition of Complex Viruses

11.2.1 *Different Proteins with Specific Roles*

A characteristic feature of complex viruses is the presence of multiple proteins in the virion, playing specific architectural or functional roles during the viral cycle. For example, different proteins may occupy the sixfold and fivefold coordinated positions in the icosahedral net (see Chap. 2). The specific architectural role of proteins at the fivefold vertices is often combined with a specific functional role, as will be described in Sect. 11.3. In icosahedral viruses with triangulation numbers $T > 1$, mobile terminal regions of the capsid proteins may adopt different conformations depending on their position in the capsid. In this way they act as molecular switches, enabling the same protein to occupy the different quasi-equivalent environments (see Chap. 2). In complex viruses, these mobile arms may still exist, but they often appear combined with a variety of minor capsid proteins, required for correct assembly of the virion. These *cementing proteins* can be considered as detached molecular switches, required to modulate the variety of interactions needed for assembly and stability of a complex capsid. One case where the intricate capsid organization includes: (i) biochemically different hexameric and pentameric capsomers; (ii) a network of mobile arms; and (iii) cementing polypeptides, has recently been described in great detail: adenovirus [3, 4].

Adenoviruses infect vertebrates. They enclose their dsDNA genome within a *pseudo* $T = 25$ icosahedral shell with a diameter of 950 Å between vertices. Trimers of the major capsid protein, hexon, constitute all the sixfold capsomers, while the vertices are occupied by pentamers of a protein named penton base. Although there is no sequence similarity between hexon and penton base, both have

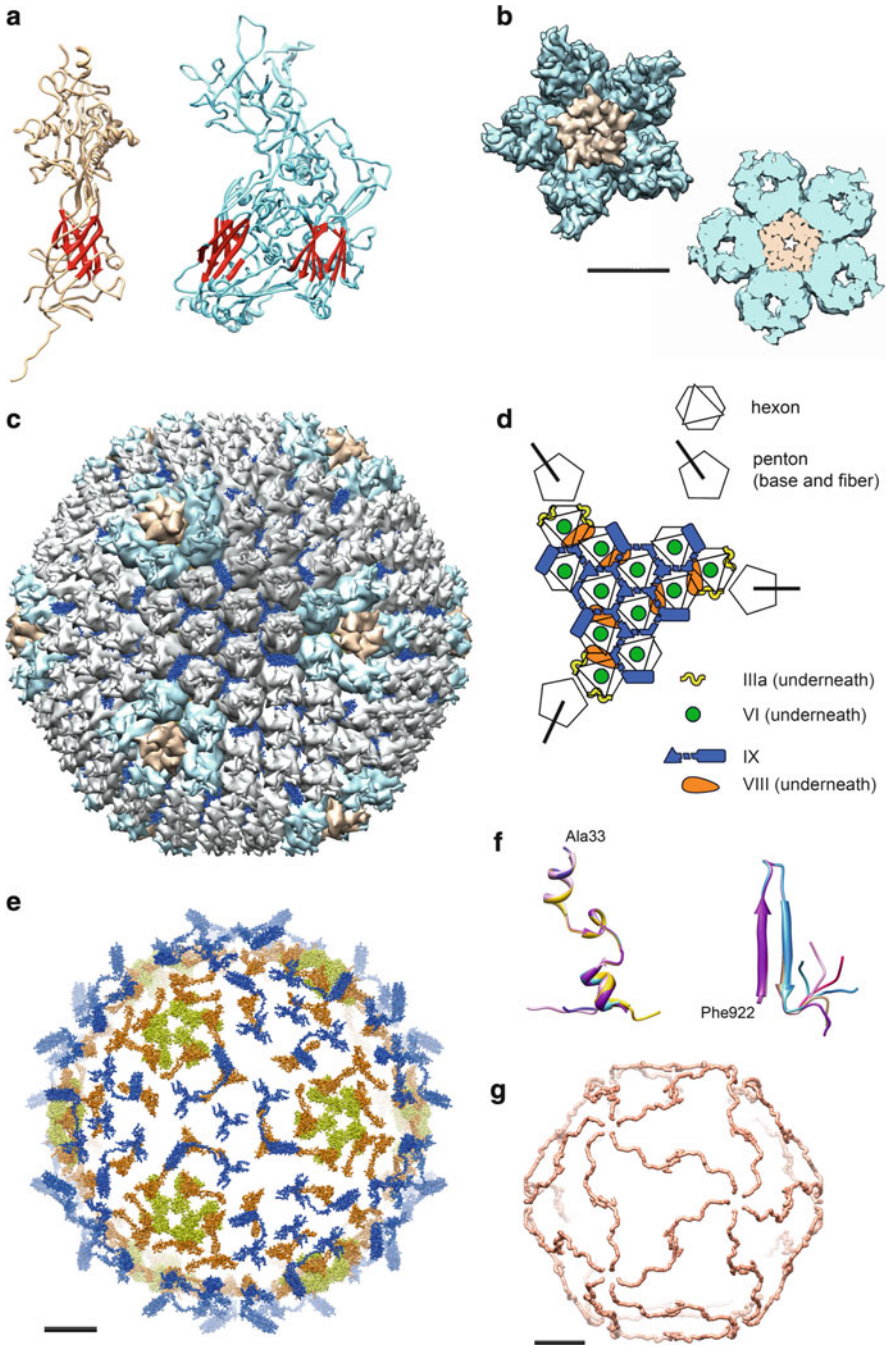


Fig. 11.1 Molecular composition of complex viruses: capsid proteins, cementing proteins, and mobile arms. (a) Structure of adenovirus penton base (*beige*) and hexon (*cyan*) monomers, with

a very similar fold based on the eight-stranded β -barrel motif, or “jelly roll”, which allows appropriate interdigitation of the different capsomers to form the closed shell [5, 6]. Penton base folds as a single jelly roll. The hexon monomer folds as a double β -barrel jelly-roll perpendicular to the capsid surface (Fig. 11.1a). Thanks to this fold, trimers have a pseudo-hexagonal shape and can occupy sixfold symmetry positions in the icosahedron. Each capsid facet is formed by 12 trimers of hexon. The general icosahedral architecture of adenovirus can be described as two different systems of tiles. Nine hexon trimers form the central plate of each facet, known as GON (Group of Nine). The five peripentonal hexon trimers, together with the penton base, form the second tile system, known as GOS (Group of Six) (Fig. 11.1b). Numerous double-stranded (ds) DNA viruses share the basic jelly-roll fold in their capsid proteins and assemble capsids organized as combinations of tiles based on pentagonal and pseudo-hexagonal building blocks. In larger viruses, the tiles composed by the pentameric vertex and surrounding capsomers are known as *pentasymmetrons*, while the triangular tiles centered at the threefold icosahedral axis are called *trisymentrons*. This efficient architectural solution facilitates construction of very large icosahedral capsids; triangulation numbers as large as $T = 169$ have been described, and larger ones are likely to exist [8, 9].

In adenovirus, apart from hexon and penton base, there are at least four other proteins making up the icosahedral shell. Minor capsid proteins IIIa, VI, VIII and IX are required for correct capsid assembly and occupy specific positions in the capsid (Figs. 11.1c, d, e), forming specialized networks that stabilize the two systems of tiles [3]. Polypeptide IX is the only cementing protein located on the outer part of the adenovirus capsid. It has an extended structure and forms a sort of hairnet keeping together the hexon trimers in each GON, and binding GONs to

Fig. 11.1 (continued) the β -barrel jelly roll motifs highlighted in *red*. The molecules are oriented so that the external capsid surface is up. Notice the extended N-terminal arm in penton base directed towards the interior of the virion (Unless otherwise indicated, all ribbon and surface structure representations in this chapter, as well as fullerene models, were prepared using USCF Chimera software [7]). **(b)** One penton base pentamer and the five peripentonal hexon trimers form the GOS, the adenovirus pentasymmetron. *Left*, view of the GOS from outside the capsid. *Right*, a slab showing the tight interdigitation at the base of the capsomers to close the shell. Color scheme as in **(a)**. **(c)** Adenovirus capsid, seen along a threefold icosahedral axis. The GOS is colored as in **(b)**, and the rest of the hexon trimers are in *gray*. These form the central plate of the icosahedral facet, termed GON, or trisymentron in the more general terminology. In *blue*, the cementing protein polypeptide IX. **(d)** Schematics showing the location of cementing proteins in an adenovirus capsid facet. Reproduced from [4]. **(e)** Cementing proteins in the adenovirus capsid. The view is as in **(c)**, with the pentons and hexons removed to reveal both the external and internal networks of accessory proteins. These are colored as in **(d)**. Notice that polypeptide VI has not been traced in the high resolution structure. **(f)** Mobile arms in adenovirus hexon. Different conformations of the hexon N-terminal (*left*) and C-terminal (*right*) regions in the capsid are shown. Panels **(a)**, **(b)**, **(c)**, **(e)** and **(f)** made from atomic coordinates deposited at the Protein Data Bank (PDB; Protein Data Bank is at <http://www.pdb.org>) with entry ID 3IYN. **(g)** Minor capsid proteins as size determinants. The bacteriophage PRD1 “tape measure” protein P30 forms a cage beneath the capsid surface (PDB ID 1W8X). Scale bars correspond to 100 Å

GONs across the icosahedral edges. The N-terminal domains of three IX monomers join *via* hydrophobic interactions at the icosahedral and local threefold axes in the GONs forming triskelion structures. A long, unstructured domain of each monomer runs in a different direction towards the facet edges, where the C-terminal α -helix joins with the C-terminal helices of another three copies of IX, different from those forming the N-terminal triskelion, to create a leucine 4-helix bundle. On the interior of the shell, each GON is further stabilized by copies of polypeptide VIII located around the icosahedral threefold symmetry axis. Also on the inner capsid surface, polypeptide IIIa mediates the interaction between penton base and the peripentonal hexons, to keep each GOS together. Finally, IIIa and VIII cooperate to bind each GOS to its five surrounding GONs. The remaining minor capsid protein, polypeptide VI, has not been unequivocally traced so far, but has been assigned to density within an internal cavity present in each hexon trimer.

Mobile regions of hexon and penton base also play a role in the extensive interaction networks in the adenovirus capsid. Due to their flexibility, these regions could not be traced in the crystal structures of the isolated proteins, but they adopt ordered conformations when they are within the capsid context. The N- and C-termini of the hexon monomer, located at the innermost part of the capsomer, adopt a total of 5 (N-) and 6 (C-) different conformations to establish interactions between neighbouring hexons, or between hexons and minor capsid proteins (Fig. 11.1f). Similarly, for each penton base monomer an N-terminal arm extends away from the β -barrels that form the main body of the protein towards the viral core, interacting with two IIIa monomers along the way, and therefore contributing to anchor the penton within the GOS. Interestingly, some of the interactions between cementing proteins and hexons, and among cementing proteins, occur by β -sheet augmentation. That is, the interaction is mediated by a β -strand from one of the proteins binding to the edge of a β -sheet in the other. This observation tells about the intricate organization of the capsid and makes us wonder about the difficulty of assembling all elements together. The fact that no high resolution structure is available for any of the minor capsid proteins in isolation suggests that they may require the virion context to fold properly.

Although in general it is understood that cementing proteins are required for correct viral assembly, it is difficult to pinpoint their exact role in morphogenesis. Some of them are dispensable for assembly, but required to reach structural stability; this is the case of adenovirus polypeptide IX [10]. Others are thought to play the role of “molecular rulers”, determining capsid size. This role was proposed when the structure of bacteriophage PRD1 was solved by protein crystallography (see also Chap. 17). It was then found that minor capsid protein P30, required for assembly, runs beneath the icosahedral edges, from the vertex to the twofold symmetry axis. Thanks to its extended conformation, two copies of the 83-residue polypeptide can cover the 300 Å length of the capsid edge, and act as a tape measure during morphogenesis [11] (Fig. 11.1g). A further complication for determining the role of minor virion components in assembly comes from the fact that, in keeping with the genetic economy principle, they often play other roles different from the purely architectural one during the viral cycle. A remarkable example of this phenomenon is illustrated by adenovirus polypeptide VI [4]. This protein is

involved in disrupting the endosomal membrane, so that the virus can escape into the cytosol after internalization. It also has a role in facilitating virion traffic to the nucleus along the microtubular network; acts as an activator of the adenoviral gene expression; and promotes transport of newly synthesized hexon to the nucleus. Finally, a C-terminal peptide of polypeptide VI activates the viral protease for maturation.

Other additional proteins may be incorporated to the viral particles and play fundamental roles for viability. Elucidating the organization of these additional components within the virion is not straightforward, since they usually do not follow defined symmetry rules, and their disposition may even change between particles of the same virus. Notable examples are viral proteases, such as the maturation protease VP24 in herpesvirus [12], or the adenoviral protease AVP [4]; and molecular motors involved in nucleic acid translocation, such as dsDNA packaging ATPases in bacteriophage and herpesviruses [13, 14], or the dsRNA packaging ATPase in cystoviruses (bacteriophage Φ 6) [15]. Viruses with RNA genomes must carry their own replication and transcription enzymes, to supply RNA metabolism functions absent in the cell [16]. Viruses that carry out their replication in the cytosol (*e.g.*, vaccinia) must also supply DNA and RNA processing enzymes whose cellular counterparts are only present in the nucleus [17]. Some dsDNA viruses encapsidate basic proteins that help screen the nucleic acid charge repulsion, to facilitate compaction of the genome within the reduced capsid space. These basic proteins can be of cellular origin, such as in Simian Virus 40, which uses histones to pack its minichromosome [18]. Baculovirus [19], adenovirus [20], mimivirus [21] and poxviruses [17] encode their own DNA compacting proteins. The genomes of negative strand ssRNA viruses usually appear in the form of ribnucleoproteic structures [22]. More information on the packaging motors and on the organization of nucleic acids within viral capsids can be found in Chap. 12.

11.2.2 Membranes

Apart from the genome and structural proteins, a large number of viruses incorporate lipidic layers into their architecture. Lipid bilayers (membranes) are widely extended in biological entities such as cells and organelles, and are ideally suited for enclosing a defined volume and separating it from neighboring compartments or the surrounding environment. This is the same function they play in viral capsids. For viruses, membranes are particularly advantageous, since they can readily be taken from the cell, are highly scalable in size, and do not consume coding space in the genome. More detail on how viruses sequester cell membranes for their own use can be found in Chap. 14.

Viruses with external membranes are called *enveloped viruses*. Envelopes form a protective layer, blocking entry of aggressive chemicals or enzymes into the viral particle. They are composed not only by lipids but also by protein and sugars

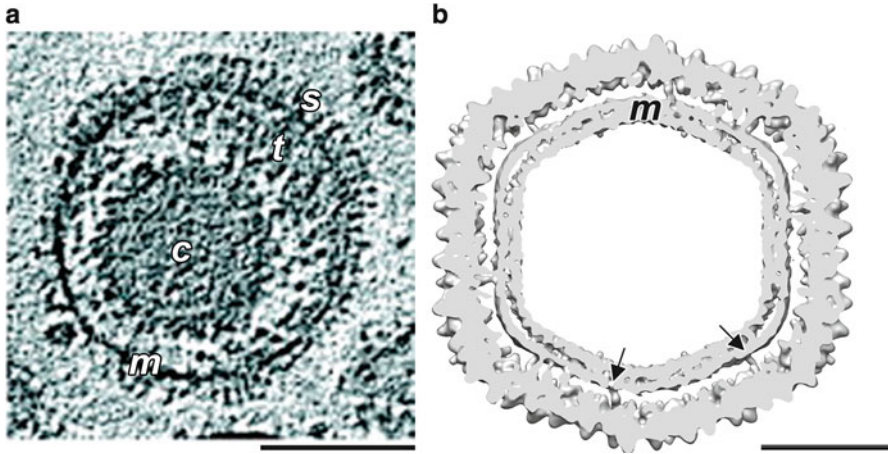


Fig. 11.2 Viruses with lipid bilayers. (a) Slice of a cryo-electron tomography reconstruction of herpes simplex virus-1 showing the icosahedral capsid (*c*), tegument (*t*), membrane (*m*) and glycoprotein spikes (*s*). Scale bar, 1,000 Å (Modified from [25]). (b) Bacteriophage PRD1 has an internal membrane (*m*) that adopts an icosahedral profile due to interactions with capsid proteins (*arrows*). Shown is a slab of the cryo-EM map deposited at the Electron Microscopy Data Bank (EMDB; Electron Microscopy Data Bank is at <http://www.ebi.ac.uk/pdbe/emdb/>) with ID EMD_1013, viewed along a twofold icosahedral axis. The scale bar corresponds to 100 Å

(glycoproteins). An implication of this arrangement is that in enveloped viruses infectivity is linked to membrane integrity, because the viral attachment proteins which recognize the host cell receptors and trigger internalization are in the viral envelope. Such is the case of influenza virus hemagglutinins [23], or Env proteins in retroviruses [2]. In the simpler enveloped viruses, such as Alphaviruses, the glycoproteins completely cover the lipid surface and follow the same icosahedral organization as the inner shell ($T = 4$) [24]. In other cases, such as retroviruses or herpesviruses (Fig. 11.2a), the structure directly in contact with the genome (nucleocapsid or capsid) is not in contact with the membrane, and the organization of glycoprotein spikes is irregular and does not reflect the organization of the virion inner contents [25, 26].

The membrane can also be located beneath the icosahedral shell, such as in Tectiviruses (PRD1) and structurally related viruses [11, 24]. In these cases, the membrane itself adopts an icosahedral layout forced by interactions with the capsid proteins (Fig. 11.2b). In PRD1, some of the mobile arms in its major capsid protein are embedded in the outer leaflet of the membrane, contributing to enforce the icosahedral shape [11]. Icosahedral viruses with internal membranes have a large complement of virion-encoded membrane proteins: approximately half of the 18 proteins present in the PRD1 virion are membrane proteins [27]. The membrane in PRD1 can undergo a large conformational change and protrude forming a tube from one of the vertices. It has been proposed that this tube has a function in injecting the

viral genome into the host cell, similar to that played by tail structures in other bacteriophages [28] (see Chap. 17).

Some of the most architecturally complex viruses are enveloped. For example, herpesviruses have a $T = 16$ icosahedral capsid (1,500 Å in diameter), formed by one major capsid protein and several accessory proteins. This capsid is surrounded by a thick tegument layer, containing at least 13 different viral proteins and also some cellular components. Capsid and tegument are enveloped by a membrane with more than 12 different types of viral glycoproteins [12, 25]. Large dsDNA viruses infecting aquatic eukaryotic microorganisms, including the giant *Acanthamoeba polyphaga* Mimivirus (diameter ~750 nm), have internal membranes like bacteriophage PRD1 [21, 29]. Asfarviruses (African swine fever virus, diameter ~200 nm) have both an internal membrane surrounded by an icosahedral shell and a loose external envelope [30]. Other examples of complex, lipid-containing viruses will be described in Sect. 11.4.

11.3 Departures from Symmetry in Quasi-Icosahedral Capsids

11.3.1 Layers with Different T Numbers

In Sect. 11.2.2, it was pointed out that viruses could be organized in multiple layers, intercalating protein (ordered or not) and lipids. In other cases such as adenovirus, multiple cementing proteins combine to form a single icosahedral capsid, while additional components (*e.g.*, dsDNA condensing proteins) do not show any symmetrical organization. In yet another instance, concentric icosahedral protein shells are formed. Remarkably, these shells may have different triangulation numbers, including some not predicted by the theory of quasi-equivalence. This type of organization is most prominently present in the dsRNA Reoviruses.

Rotaviruses and mammalian orthoreoviruses are the best characterized members of the *Reoviridae* family [16]. The mature rotavirus virion has a diameter of approximately 100 nm, and is organized in three concentric layers composed by four different proteins (Fig. 11.3). The innermost layer, or core shell, is formed by 120 molecules of protein VP2. Because of the number of protein monomers in this shell, it is described as having a $T = 2$ triangulation number, a conformation not allowed in the Caspar and Klug formalism [31] (Chap. 2). In fact, the VP2 monomers adopt two different conformations to yield a $T = 1$ icosahedron of asymmetric dimers. VP2 works as a platform for assembly of the next layer, formed by 260 trimers of protein VP6 in a $T = 13$ lattice. VP6 must be able to adopt not only the different conformations required for the $T = 13$ quasi-equivalence requirements, but also those extra ones required to compensate for the symmetry mismatch with the VP2 layer. The third, and outermost layer, is also a $T = 13$ icosahedron composed by 260 trimers of glycoprotein VP7, plus 60 spikes formed by trimers of VP4 that protrude from pentagonal positions. The VP4 spikes must

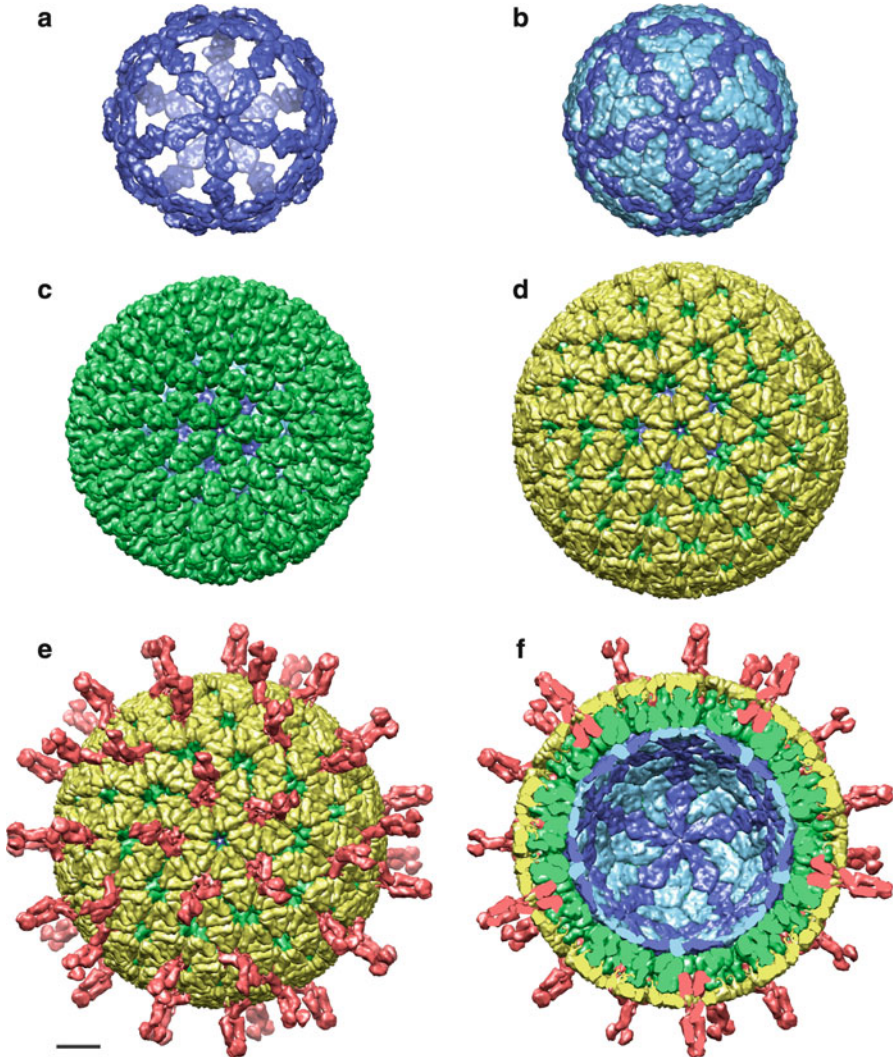


Fig. 11.3 Multiple layered viruses: rotavirus capsid architecture. Panels (a) to (e) show the consecutive building layers of the mature rotavirus virion. (a) Only one type of VP2 monomer in the core shell is represented. (b) The VP2 core shell, a $T = 1$ particle formed by asymmetric dimers. (c) The intermediate $T = 13$ VP6 layer. (d) VP7 glycoprotein layer. (e) The complete virion with VP4 spikes. (f) Cut out representation of the virion depicting the layered organization. The view is along a fivefold icosahedral symmetry axis. The scale bar represents 100 Å. Figure prepared using atomic coordinates from PDB IDs 3N09 and 3IYU

be proteolytically cleaved for the virus to be infectious. Remarkably, after cleavage the fragments remain non-covalently associated on the virion surface, but undergo an intriguing conformational modification, changing from a trimeric arrangement to

a mixture of trimeric, dimeric and monomeric associations. The VP1 and VP3 transcription enzymes are also part of the virion, and are located beneath the core shell surface.

The orthoreovirus capsid is also triple layered, and shares architectural similarities with rotavirus, namely the mismatch between “ $T = 2$ ” and $T = 13$ symmetries. However, the composition is more complex, with six different proteins instead of four. The major differences appear in the outermost layers, probably reflecting differences in the viral entry mechanism. Instead of having 60 short spikes distributed in the icosahedral facet, orthoreovirus displays large turreted structures combined with a long, flexible fiber in each of the 12 vertices.

Why do viruses have different layers? As in the case of membranes, protein layers help to separate different compartments, and most likely also different functions along the infectious cycle. dsRNA viruses need to keep their genome confined within the core shell at all times during infection, to protect it from aggressions by cellular nucleases, and to prevent antiviral reactions triggered by accumulation of dsRNA. In rotavirus, the double layered particle formed by VP2 and VP6 is the transcriptionally competent form of the virus. The external layers carry the viral components in charge of initial interaction with the host: recognition, attachment and entry. These are shed once entry into a new host cell has been accomplished [32]. In reovirus however, the turrets are not lost upon entry like the other external layers, but form part of the double layered, transcriptionally active form of the virus [33].

11.3.2 Symmetry Mismatches

In the previous section, we have seen that in reoviruses there is a symmetry mismatch between two concentric shells with different triangulation numbers. Nevertheless, the two layers still follow icosahedral symmetry, and therefore it has been possible to study their organization at a very detailed level. Symmetry mismatches (two elements with different symmetries in direct interaction) are frequent in icosahedral viral structures, particularly at the vertices, where proteins involved in genome translocation or host attachment reside. Solving the organization of mismatched features represents a remarkable challenge for structural biology techniques, due to the predominance of icosahedral symmetry in the complete virion that obscures them. In the *Cystoviridae* representative bacteriophage $\Phi 6$, a hexameric ssRNA packaging ATPase occupies multiple fivefold vertices of the empty procapsid [15]. In dsDNA bacteriophages, 12-fold portal structures occupy one of the vertices (Sect. 11.3.4 and Chaps. 12 and 17). Host recognition elements often take the shape of elongated fibers protruding from the fivefold capsomers. The oligomerization state and number of fibers per vertex varies, and is usually at odds with the pentameric architecture of the capsomers.

In human adenoviruses, a trimer of the fiber forms a non-covalent complex with a pentamer of penton base (Fig. 11.4a). The C-terminal domain of fiber is the main

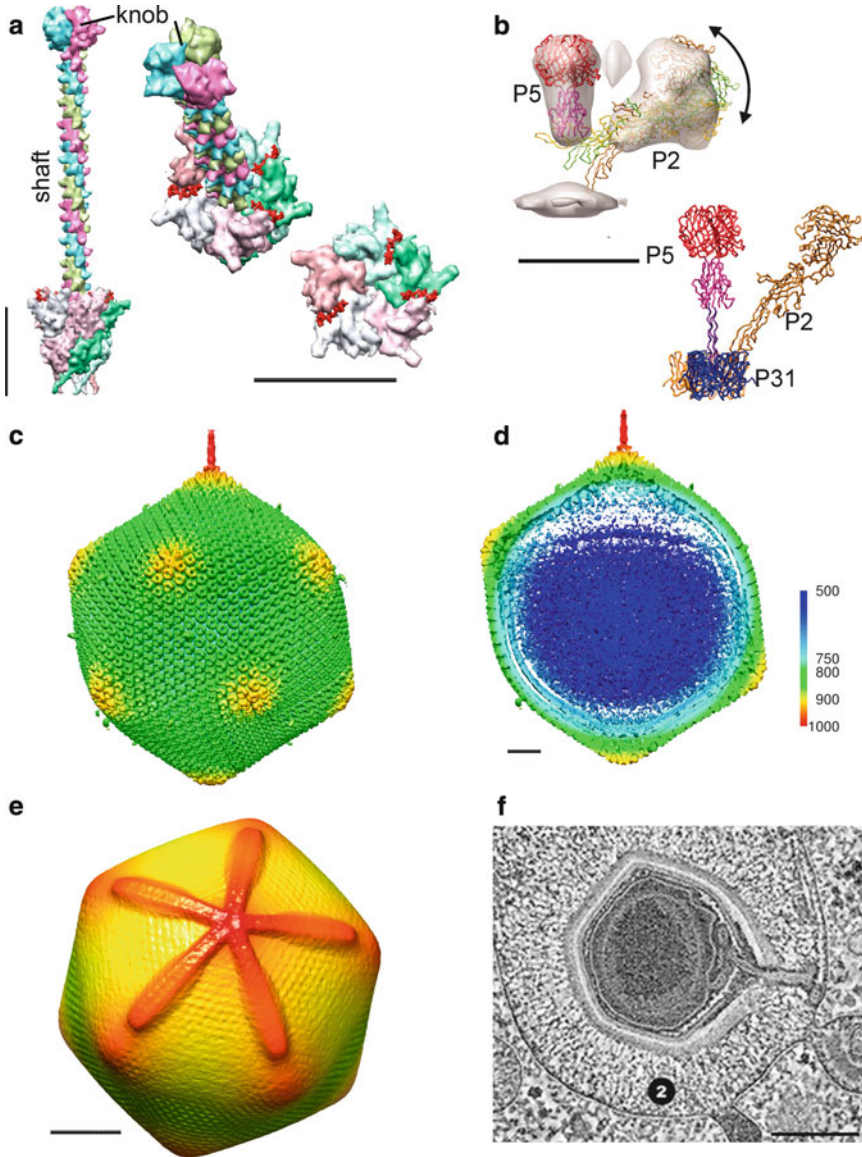


Fig. 11.4 Symmetry mismatches and singular vertices. (a) Adenovirus vertex complex, composed by pentameric penton base (compare with Fig. 1a and b) and the trimeric fiber (knob and distal shaft from PDB ID 1QIU, rest of the shaft modeled). In red, three N-terminal fiber peptides bound to penton base (PDB ID 3IZO). *Left: side view; center, oblique view; right, top view* of the penton base with the fiber shaft and knob removed to highlight the position of the N-terminal fiber peptides. (b) A model for the double fiber complex in bacteriophage PRD1 vertex. *Top:* high resolution structures of trimeric P5 and monomeric P2 fitted to a cryo-EM reconstruction of the vertex region (*side view*). The *arrow* indicates variability in the position of P2. *Bottom:* a model of the complex interactions between the three vertex components. P31 forms the fivefold capsomers

player in trimerization and forms the distal knob, responsible for binding to the adenovirus primary receptor. The knob is followed by a shaft of variable length and flexibility, depending on the serotype. The knob domain has an eight-stranded β -sandwich fold similar to that of the major capsid protein, and the shaft forms a triple β -spiral [36]. Finally, the N-terminal region is responsible for binding to penton base. At the proximal part the shaft becomes frayed and three flexible N-terminal peptides spread out to attach to the penton base pentamer. The crystal structure of penton base in complex with an N-terminal fiber peptide [5] showed the peptide bound to a groove on the outer surface of the pentamer formed by the interface between two penton base monomers, reaching radially from the penton center to the outer rim. Five peptides were observed with equivalent density, implying that all fiber binding sites are equivalent. Therefore, in the virus the three N-terminal tails may adopt two different arrangements: either they occupy three consecutive grooves, or two of them are in neighboring binding sites and the third one is flanked by two empty grooves. In the recently solved cryo-EM atomic structure of the complete virion, density for the proximal part of the fiber shaft was observed protruding from the center of the penton base pentamer [3, 37]. Density for the fiber shaft is blurred by the enforced fivefold symmetry, but it could be observed that its base interacts with a hydrophobic ring at the rim of a narrow channel in the center of the penton. This hydrophobic interaction may allow relative rotation of fiber on penton to accommodate the symmetry mismatch, while the N-terminal tails secure the binding to the penton grooves. Fiber binding to the adenovirus receptor in the cell results in its release from the capsid, which in turn induces a cooperative conformational change in the penton base pentamer. This change is thought to play a role in preparing pentons for release at a later stage of adenovirus entry and uncoating [4].

In orthoreovirus, the receptor binding protein $\sigma 1$ also forms a flexible trimeric spike attached to the fivefold turret [16, 38]. Intriguingly, some viruses can hold two fibers attached to the same vertex: this is the case of fowl adenovirus type-1, where two fibers of different length can be observed bound to the same penton base [39]. Bacteriophage PRD1 also has two different spikes (proteins P2 and P5) attached to a single pentamer of the vertex protein P31 (Fig. 11.4b). Moreover, in PRD1 each spike has a different oligomerization state. P5 is a trimer resembling the structure of the adenovirus fiber; while the other, P2, is a monomer with a pseudo- β propeller

Fig. 11.4 (continued) (pentons) in PRD1. This study showed that the two fibers interact with each other at the icosahedral capsid level, and that P2 can move relatively to P5 (Reproduced from [34]. With permission). (c) The PBCV-1 capsid (EMDB ID EMD_1597) showing the special vertex with a spike, and (d), a central slab where the asymmetry in internal contents can be appreciated. Color key indicates color changes with map radius. (e) 3D map of the Mimivirus capsid (EMDB ID EMD_5039) showing the starfish feature (*red*). (f) Slice of a tomographic reconstruction showing a Mimivirus particle within a phagosome. The viral membrane is extruded through the open stargate (Modified from [35]). Scale bars represent 100 Å in panels (a) and (b); 200 Å in (c) and (d); and 1,000 Å in (e) and (f)

head. The role of each spike in PRD1 host recognition and attachment is not fully clarified [27, 34] (see also Chap. 17).

The biological significance of these symmetry mismatches has long intrigued virologists. For the dodecameric nucleic acid packaging motors, the mismatch may allow conformational changes required for the translocation function (see Chap. 12). In the case of host recognition fibers, it is possible that the mismatch facilitates flexibility to scan for and attach to the viral receptor, as well as fiber removal upon binding, a step required to initiate the cascade of signals in both cell and virion for appropriate entry and/or genome delivery.

11.3.3 *Special Vertices*

Special (or singular) vertices have been found in many icosahedral dsDNA viruses. Singular vertices play key roles in genome packaging and ejection; they may also represent initial or final points in the assembly pathway of the shell. They represent a rupture of icosahedral symmetry (one vertex different from the other 11), and often include a symmetry mismatch (protein with non-fivefold symmetry occupying a fivefold coordinated position in the icosahedral net). Dodecameric proteins involved in genome packaging are found in a single vertex in tailed bacteriophages (see Sect. 11.3.4, and Chap. 12), as well as in herpesviruses [13, 14]. The best characterized case of special vertex is the portal in tailed bacteriophages, which connects the icosahedral head with the conspicuous tail that is characteristic of this viral family.

Giant dsDNA viruses infecting eukaryotic microorganisms also have singular vertices. *Paramecium bursaria Chlorella* virus-1 (PBCV-1) has a 190 nm diameter icosahedral capsid surrounding a lipid bilayer and dsDNA genome. A 250 Å long spike protrudes from one of the capsid vertices [8] (Figs. 11.4c, d). The peripentonal capsomers around the singular vertex seem to be structurally different from the rest. A ring-shaped density is observed near the singular vertex inside the capsid, which may correspond to a portal structure involved in genome packaging; however, there is no indication of symmetry mismatch between the fivefold vertex and this ring [29]. The spike is too thin to be used as a DNA ejection tube; besides PBCV-1 is thought to deliver its genome into the host by fusion of the internal membrane with the host one. It has been proposed that the function of the PBCV-1 spike is to puncture the cell wall to initiate the fusion process. The capsid side holding the spike is disassembled upon attachment to the host [29].

The giant Mimivirus has a 500 nm large icosahedral capsid structurally related to those of adenovirus, bacteriophage PRD1, and PBCV-1, covered by 125 nm long fibers. Early images of Mimivirus showed a starfish-shape feature with five arms reaching from one of the vertices to the five neighbouring ones (Fig. 11.4e). The arms of the starfish are inserted between adjacent facets, opening a gap between them. The starfish is an independent macromolecular assembly that remains together when detached from the virion [21], and is the only part of the capsid

not covered by fibers. When Mimivirus enters the cell by phagocytosis, a remarkable structural change occurs, whereby the five icosahedral facets in contact with the starfish feature open, leading to the structure called “stargate” [35]. The internal viral membrane is extruded through the stargate, to fuse with the phagosome membrane and release the viral DNA into the cytosol (Fig. 11.4f). Tailed phages use their special vertex both for genome packaging and delivery (Sect. 11.3.4; Chaps. 12, 17). In Mimivirus however, the stargate vertex is used for genome delivery, but not for packaging, which occurs instead *via* an aperture located in the icosahedral facet.

Remarkably, the asymmetry originated by the singular vertex in both PBCV-1 and Mimivirus reflects in an asymmetry of the internal virion contents [8, 21]. The viral genome and surrounding membrane do not occupy the full internal volume of the capsid. Rather, a gap exists between the DNA core and the side of the capsid containing the special vertex. This gap forms a pocket where viral enzymes required for cell membrane penetration may be contained. It may also contain structural elements required to precisely determine the asymmetric location and shape of the genome within the virion. However, these elements have not been identified or imaged so far.

Other viruses, such as PRD1 or adenovirus, have been reported to have singular vertices, based on genetic, biochemical and immunolabeling assays [40, 41]. However, for these viruses no structural information on the singular vertex is available yet, possibly due to the lack of large conspicuous features (such as tails) that would help calculation of three-dimensional (3D) maps without imposing full icosahedral symmetry.

11.3.4 *The Extreme Case: Heads, Tails and Baseplates (Tailed Phages)*

Tailed bacteriophages (order *Caudovirales*) are among the best described and more complex of the non-enveloped viruses. Their virions are composed by several functionally specialized morphological units, arranged according to different symmetries and connected *via* multiple symmetry mismatches (see also Chap. 17).

Bacteriophage capsids (heads) contain the dsDNA viral genome. They are icosahedrally ordered, but details vary among the different viruses. For example, HK97 or T7 phages have a strictly icosahedral head with a $T = 7$ net and a single major capsid protein occupying both the sixfold and fivefold coordinated positions. Others, like T4, have elongated, prolate icosahedral heads; two different proteins form the hexameric and pentameric capsomers, and several minor proteins are located at specific positions on the head surface [42] (Fig. 11.5a). One of the 12 vertices in the head is different from the other 11. Instead of a regular penton, it contains the portal structure or connector. This specialized structure is critical during assembly, since it contains the machinery needed to select the viral genome

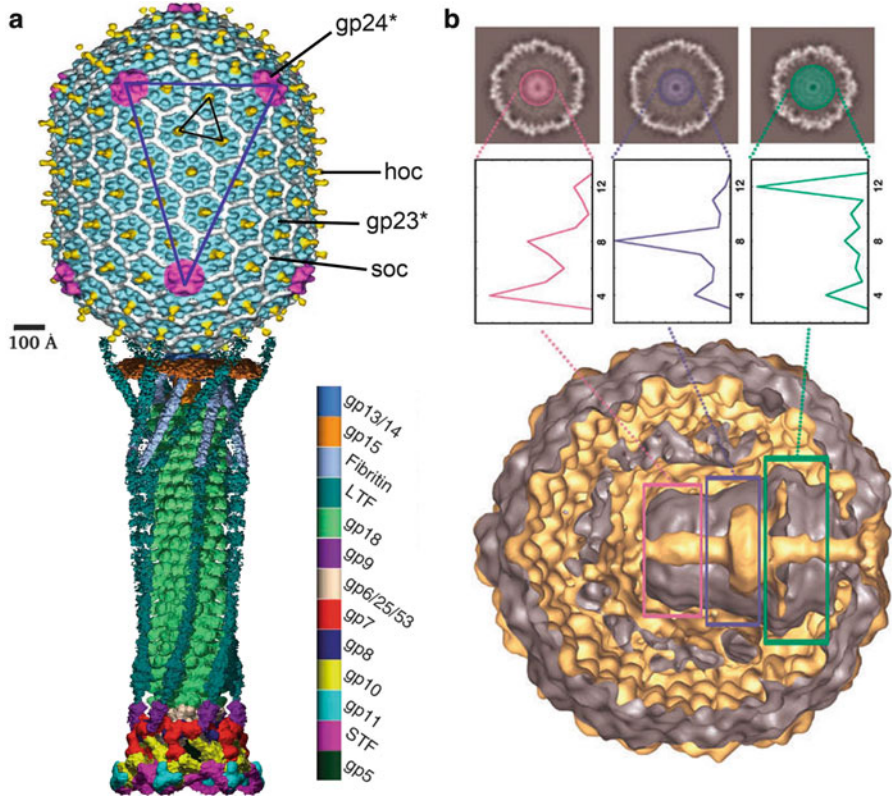


Fig. 11.5 Structure of tailed bacteriophages. (a) Bacteriophage T4 virion. The structures of the head and tail/baseplate have been solved separately, and merged to compose a representation of the complete virion. The tail is shown in its extended conformation (Modified from [43, 44]. With permission). (b) 3D map of the bacteriophage T7 procapsid, showing the internal core complex. The plots show the dominant symmetry for each core region (Reproduced from [45]. With permission). Diameter of the T7 procapsid = 510 Å

and pump it into the head (see Chap. 12). In addition, it links the head to the feature that differentiates these viral families: the tail.

Bacteriophages in the *Caudovirales* order are further classified in three groups, depending on the tail morphology: long, contractile (*Myoviridae*); long, non-contractile (*Siphoviridae*); and short, non-contractile (*Podoviridae*). In *Myoviridae* (representative: T4), the tail is composed by two layers of protein, one of them (the outer one) contractile [46] (Fig. 11.5a). The tail connects the head to a distal structure called the baseplate, formed by at least 16 different proteins in T4. Fibers with different lengths protrude from the baseplate; fibers may also be present at the portal region and the icosahedral head. In siphoviruses (representative: phage λ), the tail lacks the outer contractile sheath. Research on the structure of non-contractile tails has unveiled a crucial structural element: the tape measure protein,

whose length determines that of the tail by limiting the stacking of tail protein rings to a defined number. The baseplate composition is variable among the different viruses in this family, with some having only a simpler element called the tail tip complex. Fibers project both laterally from the periphery or longitudinally from the very tip of the tail tip complex [47]. Podoviruses (representatives: T7, P22) have short tails, with a fiber complement that may include long, thin fibers (T7) or thick spikes (P22) [48]. Fibers, tails and baseplates or tail tips form the complex machinery required to initiate infection by recognizing and attaching to the host, and delivering the viral genome through the many layers protecting the bacterial cell (see Chap. 17).

Tailed bacteriophages deviate from the icosahedral symmetry due to their conspicuous genome delivery apparatus; additionally, they are a compendium of symmetry mismatches. Icosahedral (prolate or not) heads have a singular vertex where a fivefold symmetric capsomer is replaced by a 12-fold ring of the portal protein [13]. The portal complex is connected to the tail, which in general follows sixfold symmetry along the tube and baseplate. In the case of myoviruses however, a further mismatch may exist, since the contractile sheath presents helical symmetry, and it is not yet clear if the inner tube follows the arrangement of the sheath or the sixfold symmetry observed in non-contractile tails [46]. Additionally, some podoviruses such as T7, incorporate an internal proteic structure referred to as the core. This structure grows from the portal vertex towards the capsid center, and is thought to serve as a spindle for wrapping the DNA. In T7, the core presents eightfold and fourfold symmetries [45] (Fig. 11.5b). Finally, another symmetry mismatch may appear when the packaging motor binds to the portal vertex in the prohead during encapsidation. Reported oligomeric states for components of packaging motors include pentamers (T4 gp17, Φ 29 pRNA), octamers (SF6 small terminase) and tetramers (λ terminase). However, for some of these motors it is not clear if the oligomerization states found in recombinant proteins are the same than in the prohead (immature capsid) context [13].

11.4 Asymmetric Virus Particles

11.4.1 Brick-Shaped Viruses

Poxviruses are large, enveloped dsDNA viruses apparently lacking any kind of high order symmetry in their capsids. The representative of the poxvirus family is vaccinia virus (Fig. 11.6a). The complex composition, large size, and asymmetric organization; plus the sensitivity of these viruses to the different preservation methodologies for EM, have caused the architecture of vaccinia to be a subject of debate for a long time. To complicate things even more, vaccinia virions exist in three different infectious forms carrying a different number of envelopes: mature virions, wrapped virions, and extracellular virions. The mature virion of vaccinia is

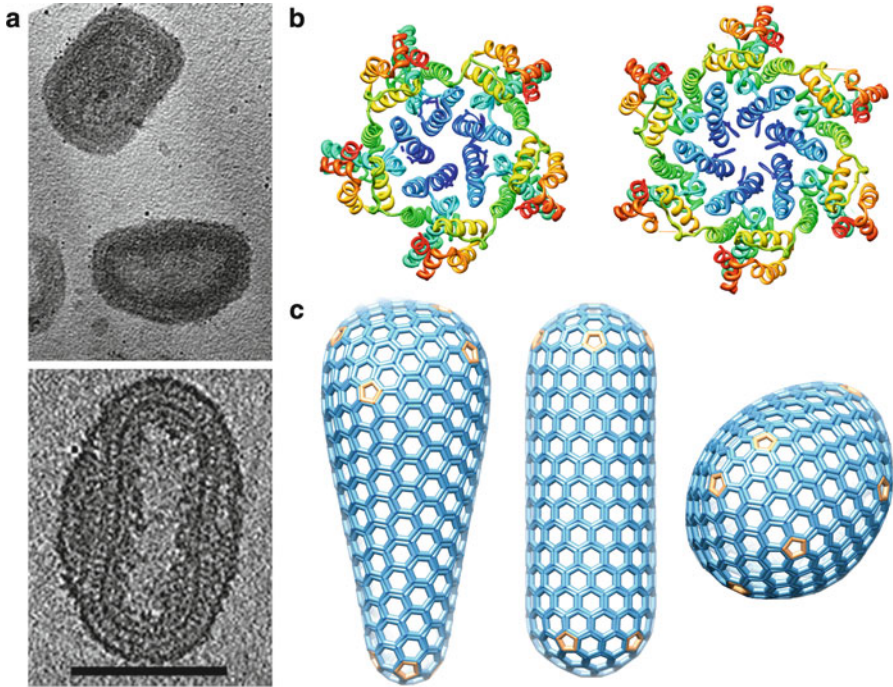


Fig. 11.6 Asymmetric and pleomorphic viruses. **(a)** Cryo-electron tomography reconstruction of mature vaccinia virions. *Top*: central section showing two virions in different orientations. *Bottom*: Central section of a virion where the dumbbell shape of the core can be appreciated. The bar represents 2,000 Å (Reproduced from [49]. With permission). **(b)** Structure of HIV CA protein assembled as a pentamer (*left*; PDB ID 3P05) or a hexamer (*right*; PDB ID 3H4E). **(c)** Example of fullerene-like objects generated from 12 pentamers and a variable number of hexamers

a brick-shaped enveloped particle with approximate dimensions $350 \times 250 \times 150$ nm (varying depending on the imaging technique used), composed by at least 75 different proteins [17]. Mature virions are intracellular forms of the viral particle. Wrapped virions are also found in the cells, and consist of mature virions surrounded by two additional membranes derived from the Golgi cisternae. Wrapped virions leave the cell by fusing with the cell membrane, leaving one of their envelopes behind, to produce the extracellular virion.

The 200 kbp genome of vaccinia is contained in a core with an elongated dumbbell shape, surrounded by a protein capsule (core wall). The core also contains a variety of viral enzymes involved in RNA metabolism, required for the virus to replicate in the cytosol. The outer part of the core wall has striated appearance (palisade layer) while the inner part is smooth. It is not known if these different appearances are due to the existence of two chemically different layers or if there is only one asymmetrically organized layer. The extremes of the dumbbell rest adjacent to the envelope, while the central part is surrounded by electron-dense material (lateral bodies) of unknown function. *In vitro* disruption studies suggest

that the dsDNA in the core is in complex with condensing proteins [50]. However, the condensing proteins have not been identified yet.

11.4.2 Pleomorphic Viruses

Pleomorphic viruses not only do not follow high symmetry rules when forming the infectious particle, but may even adopt a wide range of sizes, shapes and composition from particle to particle, making each virion unique. Because of their intrinsic variability, the structural organization of pleomorphic viruses cannot be deduced from structural biology techniques based on averaging data from many identical virions, such as X-ray crystallography or cryo-EM analyses. The advent of electron tomography to visualize single virus particles (see Chap. 3) has started to reveal the architectural details of this kind of macromolecular machines, which includes many important pathogens for humans. Examples of pleomorphic viruses include retroviruses (HIV); orthomyxoviruses (influenza); coronaviruses (SARS-coronavirus); and paramyxoviruses (measles) [2, 51–53]. In addition, atomic force microscopy can be used for surface visualization of any kind of single virus particles (see Chap. 8) and holds great potential for imaging pleomorphic viruses in liquid, in close to physiological conditions. Possible deformations by adhesion to a solid base should be, however, considered in this case.

Pleomorphism is most pronounced among enveloped viruses, since the lipid envelope readily adapts different shapes and sizes. But also proteins with a tendency to form symmetric aggregations can give rise to pleomorphic capsids. The capsid protein of retroviruses (CA) can assemble into either hexamers or pentamers, in much the same way as capsid proteins of icosahedral viruses (Fig. 11.6b). Recombinant CA forms only hexamers in certain conditions, giving rise to tubular oligomers or flat, ordered sheets; while when pentameric oligomerization is enforced, $T = 1$ icosahedral particles are formed [54, 55]. However, when CA hexamers and pentamers associate to form the closed mature capsid that contains the nucleocapsid complex including the ssRNA genome, they do it in such manner that the pentamers are not distributed regularly within the hexamer lattice. Even if a fixed number of 12 pentamers is incorporated into each capsid, the asymmetry of their distribution results in asymmetrical structures that can adopt shapes ranging from roughly spherical to roughly conical, and can be modeled using the geometrical principles governing fullerene cones (Fig. 11.6c).

11.4.3 A Glimpse of the Weird Shapes of Archaeal Viruses

In the last years numerous new microorganisms living in extreme environments have been described, and with them their corresponding infecting viruses [56]. The most abundant repertoire of archaeal viruses reported so far is that of dsDNA

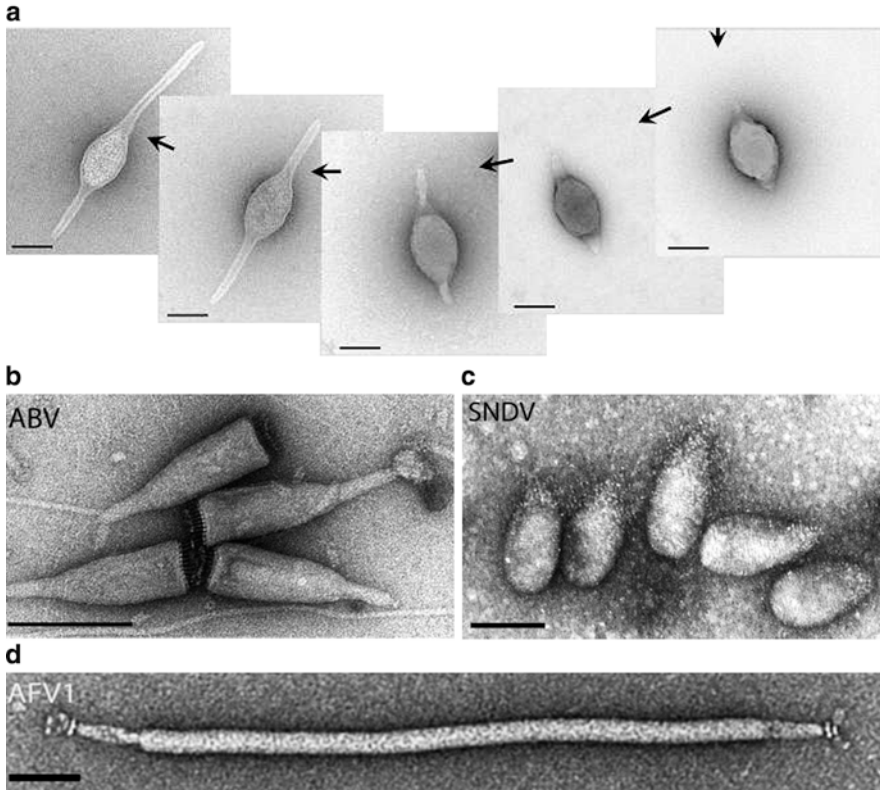


Fig. 11.7 Examples of archaeoviruses with unusual shapes. (a) From right to left: extracellular elongation of the two tails in the fusiform ATV virion. (b) Bottle-shaped and (c) droplet-shaped viruses. (d) A Lipothrixvirus representative (AFV-1) with claw-shaped structures at both ends (Panel (a) reproduced from [60]; (b), (c) and (d) reproduced from [56]. With permission). Scale bars represent 100 nm

viruses. Only a few ssDNA archaeal viruses have been described, enclosing their genome in pleomorphic enveloped particles [57]. It is still not clear if RNA archaeal viruses exist [58]. Some dsDNA viruses infecting archaea follow the general architectural types previously known for bacterial and eukaryotic viruses. For example, *Sulfolobus* turreted icosahedral virus (STIV), *Haloharcula hispanica* SH1, or *Salisaeta* icosahedral phage 1 (SSIP-1) [59] are icosahedral, tailless viruses with an internal membrane, structurally similar to bacteriophage PRD1 [11] or the algae virus PBCV-1 [8]. Some archaeal viruses with tailed bacteriophage morphologies have also been reported [56].

Intriguingly, other viral families discovered in archaea are unique to this branch of life, with morphologies never observed before (Fig. 11.7). According to their overall organization, they are classified into fusiform, droplet or bottle-shaped, and linear viruses. Within each class, viruses with different genome types, sizes, and

bearing no sequence similarity can be found. Fusiform viruses are very abundant in habitats dominated by archaeal microorganisms. They have spindle-shaped virions with tails of variable length protruding from the spindle poles. One fusiform virus, *Acidianus* two-tailed virus (ATV) not only has an unusual shape, but is able to assemble new structural features after leaving the host cell. When propagated at temperatures slightly suboptimal for its host (75 °C), isolated virions look like ~0.2 µm long lemons. However, when temperature is raised in the absence of the host cell, these viruses grow two filamentous tails of variable lengths, one from each pole. This is the only known example of a virus with extracellular assembly, but it is likely that others exist [60]. The tails end in an anchor-like structure, thought to be involved in attachment to the host.

Acidianus bottle-shaped virus (ABV) and *Sulfolobus neozelandicus* droplet-shaped virus (SNDV) are the only known members of the two viral families termed *Ampullaviridae* and *Guttaviridae* [56]. The enveloped ABV virion contains a conical core formed by a supercoiled nucleoprotein filament. A brush of short filaments protrudes from the bottom of the bottle, but host attachment seems to occur at the opposite side of the virion. Little is known about the architecture of the SNDV virion, except for its droplet shape and the presence of a tuft of long fibers at its narrower pole. Finally, linear archaeal viruses can form stiff rods (*Rudiviridae*) or flexible filaments (*Lipothrixviridae*). Rudiviruses are relatively simple in composition, with no envelope and only a few proteins arranged in particles of variable length, usually related to that of the genome. Lipothrixviruses are enveloped, and the ends of their filamentous capsids are capped with structures of varied shapes (spider legs, pincers, bottle brushes), probably involved in attachment to the host.

11.5 Sophisticated Regulation of Assembly and Maturation

In simple viruses, assembly can occur either in a single step where the newly replicated nucleic acid associates with capsid protein subunits during co-assembly, or in a two-step process where an empty capsid is assembled first and the viral nucleic acid is packaged afterwards (see Chaps. 10, 12). For complex viruses, putting together the many different pieces in their proper places at the appropriate time requires elaborated regulation of the morphogenesis process. In the following sections we discuss some of the strategies used by viruses to achieve assembly of complex capsids. The additional steps required by enveloped viruses to coordinate assembly of proteic elements with recruitment of membranes from the cell will be described in Chap. 14.

11.5.1 *Separate Assembly Lines*

In viruses with complex chemical composition, the different morphological components are often built separately, forming subassemblies that will be later

put together along carefully regulated pathways. Some of these subassembly reactions may also require chaperones, either of cellular or viral origin, as is also the case for some simpler viruses (Chap. 10). For example, in adenovirus, capsid protein oligomers are formed in the cytosol before being transported to the nucleus, where viral assembly takes place. This preassembly step includes hexon trimerization, which requires a viral chaperone (L4 100K) [61]; and piecing together the vertex complex, composed by a pentamer of penton base bound to a trimer of the fiber protein [4]. In adenovirus, however, the precise temporal order of incorporation of major and minor capsid proteins is not yet understood. Similarly, in herpesvirus hexamers and pentamers of the major capsid protein VP25 are formed previous to particle assembly.

The best described examples of subassembly formation and integration into a virion come from the order *Caudovirales*, more specifically from the *Myoviridae* prototype bacteriophage T4 [42]. In these long tailed bacteriophages, the head, fibers, and tail form separately (Fig. 11.8). Bacteriophage T4 head assembly starts from an initiation complex containing the portal protein gp20 bound to a cellular membrane. This complex recruits the components of a scaffolding core composed by eight different types of viral proteins, among them the main scaffolding protein gp22 and the viral protease (gp21). This core will later be coated by the hexameric and pentameric capsomers (gp23 and gp24) to form the procapsid, or prohead. Once coating is complete, gp21 is activated and cleaves all components of the procapsid except the portal protein. The prohead is detached from the membrane and most of the cleaved peptides exit the particle, opening the way for entry of the viral genome. Genome packaging leads to the large structural rearrangements involved in capsid maturation (Sect. 11.5.3 and Chap. 13), to produce the head in its final form. Finally, the head completion (or neck) proteins gp13, gp14, gp2 and gp4 attach to the portal to form the interface between the head and the tail.

The tail is in turn formed from several preassembled pieces. For contractile tails, the baseplate is assembled first, and used as a seeding point for assembly of the inner tube and contractile sheath. In bacteriophage T4, tail assembly involves 19 different proteins and seven viral chaperones. To form the T4 baseplate, proteins gp6, gp7, gp8, gp10, gp11, gp25 and gp53 assemble in the form of hetero-oligomeric wedges. Six wedges bind around a central hub containing gp5 and gp27. Proteins gp9 and gp12 (the short tail fiber) are then inserted at the gaps between wedges, and the interface between wedges and hub is sealed by proteins gp48 and gp54. This seal is the starting point from which the gp19 inner tail tube will grow. The length of the tube is controlled by a tape measure protein gp29, which extends from the hub to the tube end where the tail capping protein gp3 will bind. The tail sheath gp18 assembles around the inner tube, and finally the tail terminator protein gp15 binds to gp3 and the last row of gp18 subunits, making the tail ready to bind to the neck proteins in the head.

The final stage of tail assembly is incorporation of the long tail fibers to the base plate. The fibers also assemble independently, even starting from separate

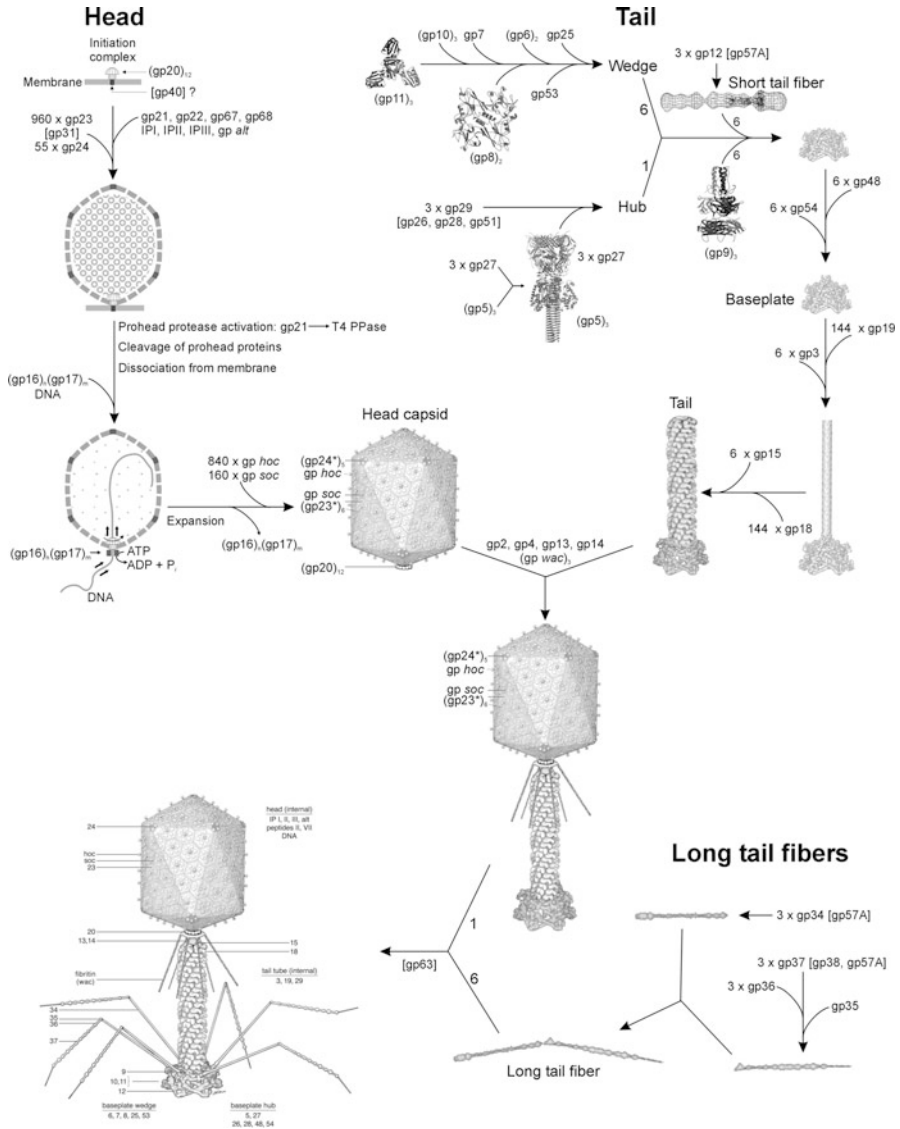


Fig. 11.8 Separate assembly lines. Schematic depiction of the complex bacteriophage T4 morphogenesis pathway (Reproduced from [42]. With permission)

subassemblies. The long tail fiber of bacteriophage T4 is kinked; the proximal part of the fiber before the kink is formed by a single protein (gp34), while the distal part contains three different proteins: gp35, gp36, and gp37. The proximal and distal parts assemble separately, and then join before attaching to the baseplate. All long

tail fiber proteins form trimers, except monomeric gp35 (notice the symmetry mismatch) that sits at the interface between the distal and proximal half fibers. Interestingly, the long fibers cannot join the tail until it is bound to the DNA-filled head. A similar assembly pathway has been described for non-contractile long tails. In podoviruses however, the short tail is not assembled as a separate entity, but it grows outward from the portal vertex on the virion capsid. The careful temporal regulation of these assembly processes is proved by the fact that, when any of the structural proteins is absent, viral morphogenesis is interrupted and the assembly intermediates previous to the disrupted step accumulate in the cell.

11.5.2 Scaffolding

Scaffolding elements are crucial for accurate assembly of large viral capsids. They are present in assembly intermediates (*e.g.* procapsids), but absent in the final, infective product. Their role is to facilitate interactions between capsid elements at the early stages of assembly, by promoting nucleation – that is, putting together the viral proteins that may be highly diluted within the crowded cellular context. Scaffolds are also thought to stabilize weak interactions at the initial stages of assembly [62], while simultaneously allowing flexibility for mistake corrections. This last function is most important in large capsids, where the number of interactions to be checked for errors is correspondingly large. Finally, scaffolds have a role in determining the size and shape of viral capsids.

The most studied scaffold proteins are those present in tailed bacteriophages. For example, bacteriophage P22 (*Podoviridae*) scaffold is a 33 kDa protein. In the early stages of P22 morphogenesis, a procapsid is formed by 415 copies of the capsid protein, with approximately 300 molecules of scaffold inside. Unlike the capsid protein, the scaffold does not follow icosahedral symmetry; therefore, little is known about its organization in the assembly intermediate [45]. Scaffold proteins have been quite refractory to structural studies. Nuclear magnetic resonance and crystallographic studies on the scaffold of P22 and Φ 29 indicate that they have a helical fold. Biophysical analyses indicate that many of them share an extended, rod-like shape and a tendency to dimerize in solution [63]. However, an equilibrium between different oligomeric forms seems to be required to achieve correct capsid assembly. Kinetic studies have revealed that in phage P22, scaffold is predominantly a dimer during assembly, but the presence of free monomers is absolutely required to complete the head. Kinetically trapped intermediates are observed when monomers are depleted by decreasing the ionic strength, while restoring it eliminates the trap and allows elongation to proceed. Phage scaffolds are usually ejected from the procapsid immediately before genome packaging. In P22 and Φ 29, the intact protein exits the shell, and can be recycled in a new round of assembly. In other cases, the scaffold is removed *via* cleavage by a viral-encoded protease.

In spite of their apparently simple organization, some small bacteriophages, such as the Microvirus representative Φ X174 ($T = 1$), encode both internal and external scaffold proteins [63]. The Φ X174 internal scaffold protein (protein B) helps in the early stages of assembly by preventing aggregation of the capsid protein F into aberrant oligomers, and ensuring the recruitment of the vertex spike protein G. The C-terminal region of protein B (24 aminoacids) interacts with the capsid and can be observed in the crystal structure of the procapsid, while the rest is disordered and appears to be largely tolerant to mutations. On the contrary, the external scaffold protein D is highly ordered and sensitive to mutations. Protein D is absolutely required for elongation (to assemble capsid pentamers into a spherical particle), while protein B helps to make assembly efficient but is not strictly required: in the absence of B, viral particles can be formed, but the process requires overexpression of protein D and takes as much as ten times longer than in the presence of both scaffolds. It is believed that scaffold redundancy confers an evolutionary advantage to Φ X174 by facilitating extremely rapid replication cycles.

In the absence of scaffold, many phage capsid proteins have been observed to self-assemble into aberrant oligomers (tubes, elongated shells, $T = 1$ icosahedrons). It follows from these observations that scaffolding proteins are involved in determining the correct curvature in the interactions between capsomers, so that they can form a closed shell of the appropriate size to hold the viral genome. A remarkable proof of this size determination role comes from the P2/P4 phage system [64]. Phage P2 is a member of the *Myoviridae* family, with an assembly pathway similar to P22 that includes a $T = 7$ procapsid formed with the help of an internal scaffolding protein, gpO. Remarkably, P2 gpO combines the function of scaffold and protease, able to cleave itself and other shell components at later maturation stages. In the presence of P4, however, the whole assembly pathway is altered. P4 is a replicon that can exist either as a prophage or a free plasmid, and does not code for any major structural proteins. However, P4 is able to sequester players of the P2 assembly line to form small $T = 4$ capsids where only its smaller genome can fit, excluding that of P2 (Fig. 11.9). This parasitic process is achieved by synthesis of a P4-encoded scaffold protein, Sid (SIze Determination protein). Sid forms an external scaffold on P2 assembly intermediates, forcing a narrower curvature and therefore a smaller icosahedral net. Unlike the internal scaffolds, P4 Sid forms a dodecahedral ordered cage that can be observed in cryo-electron microscopy reconstructions.

Although one can generally speak about scaffolding proteins, and indeed many viruses have such proteins, scaffolding functions can also be performed by flexible regions of the capsid proteins, which establish interactions during assembly that are later removed *via* conformational changes or cleavage by viral proteases. For example, bacteriophage HK97 (*Siphoviridae*) does not encode a scaffold protein. Instead, a 103 residue stretch at the N-terminus of the capsid protein, known as the delta-domain, performs the scaffold function [65]. The delta domain is located towards the interior of the capsid and mediates interactions between capsomers during assembly. Once the procapsid is completed, and before the DNA is packaged, the delta domain is cleaved out by the viral protease, allowing the transition to

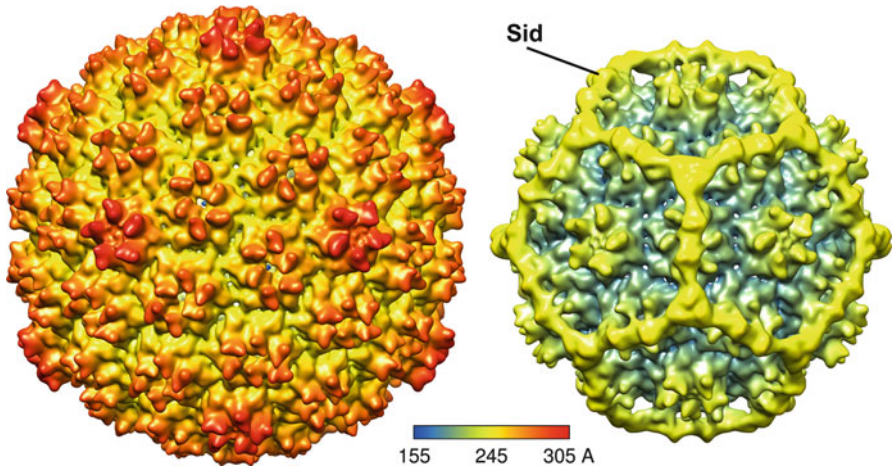


Fig. 11.9 Capsid size determination by scaffolding proteins: the P2/P4 phage system. *Left*: bacteriophage P2 T = 7 procapsid (EMDB ID EMD_5406). *Right*: the parasitic P4 T = 4 procapsid (EMDB ID EMD_5405). Note the Sid external scaffold restricting the shell size. The view is along a twofold icosahedral axis. Original maps have been filtered to a lower resolution for clarity. Color key indicates color changes with map radius

the next assembly intermediate [62]. Adenovirus may use a dual system, including both a separate scaffolding protein [66] and flexible regions of minor capsid proteins removed by the viral protease during maturation [67].

11.5.3 Maturation

In complex viruses, piecing together a number of proteins into a capsule to host the genome is far from producing the final, infectious form. Instead, the newly formed particles (procapsids) need to undergo a series of morphological and/or stability changes to acquire their full infectious potential. This process is known as maturation. There is a double goal for maturation on the viral cycle: first, to produce virions stable enough to protect the genome from aggressive conditions in the extracellular milieu; and second, to prepare the viral particle for correct delivery of the genome into the new host cell.

In dsDNA bacteriophages, maturation encompasses large structural changes and protein rearrangements in the capsid, concomitant with scaffold removal and genome packaging. The capsid changes from a weak, labile object to a highly stable shell, ready to withstand the high internal pressure imposed by the tightly packed DNA inside [62] (see Chaps. 9, 12, 18, and 19). In other viruses, such as polio [68] or adenovirus [67], maturation does not end with an extremely stable shell, but with a metastable one. This difference with respect to the bacteriophage case is likely related to the mode of infection of these eukaryotic viruses. Instead of

ejecting the genome across the plasma membrane leaving the capsid behind, polio and adenovirus are internalized in the cell, and must be disassembled within in a concerted fashion to ensure exposure of the genome at the appropriate place and time for successful replication. Maturation prepares these viruses to start the programmed uncoating sequence upon reception of the appropriate signal, for example attachment to the receptor, or pH changes along the endocytosis pathway. Interestingly, in adenovirus maturation is related to genome packaging in quite a unique way. The adenoviral protease, which is the main maturation agent, is packaged together with the viral genome thanks to its dsDNA binding ability, and uses the dsDNA itself as a cofactor to increase its catalytic activity several orders of magnitude [4].

Maturation processes are not restricted to icosahedral capsids: pleomorphic enveloped viruses such as retroviruses also undergo extensive structural rearrangements to become fully infectious [2]. A more extensive discussion on maturation for several different viruses can be found in Chap. 13.

11.6 Perspectives and Conclusions

In the past decade, structural studies on complex viruses have greatly benefited from technical improvements in structural biology techniques such as those described in Chaps. 3, 4, 5, 6, 7, 8, and 9. As more details are known, it is becoming clear that many complex viruses that infect hosts far apart in evolution share common structural solutions.

For example, adenovirus, which infects vertebrates, has a striking structural similarity to PRD1, a bacteriophage with an internal membrane. The parallels between adenovirus and PRD1 extend from their DNA replication mechanisms, to their capsid architecture and the folding of their major capsid proteins [3, 11]. In the last years, more members of the PRD1-adenovirus family have been described or predicted, and the lineage now extends from viruses infecting bacteria or archaea, to the large nucleo-cytoplasmic DNA viruses such as Asfarvirus, Iridovirus and the giant Mimivirus [9]. All these viruses are built from the same kind of double 8-stranded β -barrel, pseudo-hexagonal capsomers arranged in different tiling systems, with triangulation numbers ranging between $T = 21$ and $T = 169$, and reaching up to $972 < T < 1,200$ for the giant Mimivirus [21]. Intriguingly, even a scaffold protein of the non-icosahedral vaccinia virus folds as a double barrel pseudo-hexamer, indicating a possible common ancestor with icosahedral dsDNA viruses [69].

Adenovirus and PRD1 are not the only cases indicating an evolutionary relationship between animal and bacterial viruses. Herpesviruses, which infect all sorts of animal organisms, share many structural characteristics with tailed bacteriophage [14]. They follow a similar assembly pathway, starting from an empty procapsid formed with the help of scaffold, and maturing to a more angular shell *via* large structural rearrangements upon DNA packaging. Like tailed bacteriophage, one of

the vertices is different from the rest and contains a 12-fold symmetric portal structure involved in genome encapsidation. Herpesvirus capsid protein folds with a topology very similar to that of the HK97 phage family; and like tailed phages, the packed genome forms concentric shells when observed in icosahedrally averaged EM reconstructions. Finally, structural parallels also exist between Cystoviruses (dsRNA bacteriophage, representative $\Phi 6$) and Reoviruses [70].

The fact that many complex viruses with different hosts share a common structural solution has evolutionary implications. First, since the design has been conserved throughout time, even after all traces of sequence similarity have disappeared, it must be a highly efficient arrangement. Second, it suggests that the architecture was established in the early stages of evolution, before the branches of the evolutionary tree diverged into the three kingdoms known today (archaea, bacteria and eukarya). One could wonder, then, if all complex viruses existing today fall within a limited number of structural solutions selected by their success early in evolution [71]. However, discovery of the unique architectures of archaeal viruses indicates that other structural solutions exist. Advances in high throughput virus isolation and structural characterization techniques will contribute to clarify this question.

In summary, complex viruses incorporate a wide range of molecules into their capsids, including specialized host interaction, genome packaging and cementing proteins; and in some cases membranes, either internal or external. Accommodation of the different components often requires deviations from high order symmetry, from mismatches to pleomorphism; and involves complex regulation of the assembly dynamics. Key elements in this regulation are: separate assembly lines, scaffold elements, and maturation processes. Host and virus evolution probably act hand in hand to optimize viral particle structure and morphogenesis.

Acknowledgements Work in the San Martín lab is funded by grants BFU2010-16382 and FIS2010-10552-E/FIS2011-16090-E from the Ministerio de Ciencia e Innovación of Spain. José R. Castón (CNB-CSIC) is acknowledged for careful reading of the manuscript.

References and Further Reading

1. Stubbs G, Kendall A (2012) Helical viruses. *Adv Exp Med Biol* 726:631–658
2. Goff SP (2007) Retroviridae: the retroviruses and their replication. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA (eds) *Fields virology*, vol II. Lippincott Williams & Wilkins, Philadelphia, pp 1999–2069
3. Liu H, Jin L, Koh SB, Atanasov I, Schein S, Wu L, Zhou ZH (2010) Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks. *Science* 329:1038–1043
4. San Martín C (2012) Latest insights on adenovirus structure and assembly. *Viruses* 4:847–877
5. Zubieta C, Schoehn G, Chroboczek J, Cusack S (2005) The structure of the human adenovirus 2 penton. *Mol Cell* 17:121–135
6. Rux JJ, Kuser PR, Burnett RM (2003) Structural and phylogenetic analysis of adenovirus hexons by use of high-resolution x-ray crystallographic, molecular modeling, and sequence-based methods. *J Virol* 77:9553–9566

7. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612
8. Cherrier MV, Kostyuchenko VA, Xiao C, Bowman VD, Battisti AJ, Yan X, Chipman PR, Baker TS, Van Etten JL, Rossmann MG (2009) An icosahedral algal virus has a complex unique vertex decorated by a spike. *Proc Natl Acad Sci U S A* 106:11085–11089
9. Krupovic M, Bamford DH (2008) Virus evolution: how far does the double beta-barrel viral lineage extend? *Nat Rev Microbiol* 6:941–948
10. Colby WW, Shenk T (1981) Adenovirus type 5 virions can be assembled *in vivo* in the absence of detectable polypeptide IX. *J Virol* 39:977–980
11. Abrescia NG, Cockburn JJ, Grimes JM, Sutton GC, Diprose JM, Butcher SJ, Fuller SD, San Martín C, Burnett RM, Stuart DI, Bamford DH, Bamford JK (2004) Insights into assembly from structural analysis of bacteriophage PRD1. *Nature* 432:68–74
12. Roizman B, Knipe DM, Whitley RJ (2007) Herpes simplex viruses. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA (eds) *Fields virology*, vol II. Lippincott Williams & Wilkins, Philadelphia, pp 2501–2601
13. Feiss M, Rao VB (2012) The bacteriophage DNA packaging machine. *Adv Exp Med Biol* 726:489–509
14. Cardone G, Heymann JB, Cheng N, Trus BL, Steven AC (2012) Procapsid assembly, maturation, nuclear exit: dynamic steps in the production of infectious herpesvirions. *Adv Exp Med Biol* 726:423–439
15. Huiskonen JT, Jaalinoja HT, Briggs JA, Fuller SD, Butcher SJ (2007) Structure of a hexameric RNA packaging motor in a viral polymerase complex. *J Struct Biol* 158:156–164
16. Trask SD, McDonald SM, Patton JT (2012) Structural insights into the coupling of virion assembly and rotavirus replication. *Nat Rev Microbiol* 10:165–177
17. Condit RC, Moussatche N, Traktman P (2006) In a nutshell: structure and assembly of the vaccinia virion. *Adv Virus Res* 66:31–124
18. Germond JE, Hirt B, Oudet P, Gross-Bellark M, Chambon P (1975) Folding of the DNA double helix in chromatin-like structures from simian virus 40. *Proc Natl Acad Sci U S A* 72:1843–1847
19. Tweeten KA, Bulla LA, Consigli RA (1980) Characterization of an extremely basic protein derived from granulosis virus nucleocapsids. *J Virol* 33:866–876
20. Giberson AN, Davidson AR, Parks RJ (2012) Chromatin structure of adenovirus DNA throughout infection. *Nucleic Acids Res* 40:2369–2376
21. Xiao C, Kuznetsov YG, Sun S, Hafenstein SL, Kostyuchenko VA, Chipman PR, Suzan-Monti M, Raoult D, McPherson A, Rossmann MG (2009) Structural studies of the giant mimivirus. *PLoS Biol* 7:e92
22. Ruigrok RW, Crepin T, Kolakofsky D (2011) Nucleoproteins and nucleocapsids of negative-strand RNA viruses. *Curr Opin Microbiol* 14:504–510
23. Palese P, Shaw ML (2007) Orthomyxoviridae: the viruses and their replication. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA (eds) *Fields virology*, vol II. Lippincott Williams & Wilkins, Philadelphia, pp 1647–1689
24. Huiskonen JT, Butcher SJ (2007) Membrane-containing viruses with icosahedrally symmetric capsids. *Curr Opin Struct Biol* 17:229–236
25. Grunewald K, Desai P, Winkler DC, Heymann JB, Belnap DM, Baumeister W, Steven AC (2003) Three-dimensional structure of herpes simplex virus from cryo-electron tomography. *Science* 302:1396–1398
26. Forster F, Medalia O, Zauberman N, Baumeister W, Fass D (2005) Retrovirus envelope protein complex structure in situ studied by cryo-electron tomography. *Proc Natl Acad Sci U S A* 102:4729–4734
27. Butcher SJ, Manole V, Karhu NJ (2012) Lipid-containing viruses: bacteriophage PRD1 assembly. *Adv Exp Med Biol* 726:365–377

28. Grahm AM, Daugelavicius R, Bamford DH (2002) Sequential model of phage PRD1 DNA delivery: active involvement of the viral membrane. *Mol Microbiol* 46:1199–1209
29. Zhang X, Xiang Y, Dunigan DD, Klose T, Chipman PR, Van Etten JL, Rossmann MG (2011) Three-dimensional structure and function of the paramecium bursaria chlorella virus capsid. *Proc Natl Acad Sci U S A* 108:14837–14842
30. Tulman ER, Delhon GA, Ku BK, Rock DL (2009) African swine fever virus. *Curr Top Microbiol Immunol* 328:43–87
31. Caspar DLD, Klug A (1962) Physical principles in the construction of regular viruses. *Cold Spring Harb Symp Quant Biol* 27:1–24
32. Baker M, Prasad BV (2010) Rotavirus cell entry. *Curr Top Microbiol Immunol* 343:121–148
33. Danthi P, Guglielmi KM, Kirchner E, Mainou B, Stehle T, Dermody TS (2010) From touchdown to transcription: the reovirus cell entry pathway. *Curr Top Microbiol Immunol* 343:91–119
34. Huiskonen J, Manole V, Butcher S (2007) Tale of two spikes in bacteriophage PRD1. *Proc Natl Acad Sci U S A* 104:6666–6671
35. Zauberman N, Mutsafi Y, Halevy D, Shimoni E, Klein E, Xiao C, Sun S, Minsky A (2008) Distinct DNA exit and packaging portals in the virus *Acanthamoeba polyphaga* mimivirus. *PLoS Biol* 13:e114
36. van Raaij MJ, Mitraiki A, Lavigne G, Cusack S (1999) A triple β -spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein. *Nature* 401:935–938
37. Liu H, Wu L, Zhou ZH (2011) Model of the trimeric fiber and its interactions with the pentameric penton base of human adenovirus by cryo-electron microscopy. *J Mol Biol* 406:764–774
38. Chappell JD, Prota AE, Dermody TS, Stehle T (2002) Crystal structure of reovirus attachment protein $\sigma 1$ reveals evolutionary relationship to adenovirus fiber. *EMBO J* 21:1–11
39. Hess M, Cuzange A, Ruigrok RWH, Chroboczek J, Jacrot B (1995) The avian adenovirus penton: two fibres and one base. *J Mol Biol* 252:379–385
40. Christensen JB, Byrd SA, Walker AK, Strahler JR, Andrews PC, Imperiale MJ (2008) Presence of the adenovirus IVa2 protein at a single vertex of the mature virion. *J Virol* 82:9086–9093
41. Gowen B, Bamford JK, Bamford DH, Fuller SD (2003) The tailless icosahedral membrane virus PRD1 localizes the proteins involved in genome packaging and injection at a unique vertex. *J Virol* 77:7863–7871
42. Leiman PG, Kanamaru S, Mesyanzhinov VV, Arisaka F, Rossmann MG (2003) Structure and morphogenesis of bacteriophage T4. *Cell Mol Life Sci* 60:2356–2370
43. Kostyuchenko VA, Chipman PR, Leiman PG, Arisaka F, Mesyanzhinov VV, Rossmann MG (2005) The tail structure of bacteriophage T4 and its mechanism of contraction. *Nat Struct Mol Biol* 12:810–813
44. Fokine A, Chipman PR, Leiman PG, Mesyanzhinov VV, Rao VB, Rossmann MG (2004) Molecular architecture of the prolate head of bacteriophage T4. *Proc Natl Acad Sci U S A* 101:6003–6008
45. Agirrezabala X, Martin-Benito J, Caston JR, Miranda R, Valpuesta JM, Carrascosa JL (2005) Maturation of phage T7 involves structural modification of both shell and inner core components. *EMBO J* 24:3820–3829
46. Leiman PG, Shneider MM (2012) Contractile tail machines of bacteriophages. *Adv Exp Med Biol* 726:93–114
47. Davidson AR, Cardarelli L, Pell LG, Radford DR, Maxwell KL (2012) Long noncontractile tail machines of bacteriophages. *Adv Exp Med Biol* 726:115–142
48. Casjens SR, Molineux IJ (2012) Short noncontractile tail machines: adsorption and DNA delivery by podoviruses. *Adv Exp Med Biol* 726:143–179
49. Cyrklaff M, Risco C, Fernández JJ, Jiménez MV, Estéban M, Baumeister W, Carrascosa JL (2005) Cryo-electron tomography of vaccinia virus. *Proc Natl Acad Sci U S A* 102:2772–2777

50. Kuznetsov Y, Gershon PD, McPherson A (2008) Atomic force microscopy investigation of vaccinia virus structure. *J Virol* 82:7551–7566
51. Harris A, Cardone G, Winkler DC, Heymann JB, Brecher M, White JM, Steven AC (2006) Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proc Natl Acad Sci U S A* 103:19123–19127
52. Barcena M, Oostergetel GT, Bartelink W, Faas FG, Verkleij A, Rottier PJ, Koster AJ, Bosch BJ (2009) Cryo-electron tomography of mouse hepatitis virus: insights into the structure of the coronavirus. *Proc Natl Acad Sci U S A* 106:582–587
53. Liljeroos L, Huiskonen JT, Ora A, Susi P, Butcher SJ (2011) Electron cryotomography of measles virus reveals how matrix protein coats the ribonucleocapsid within intact virions. *Proc Natl Acad Sci U S A* 108:18085–18090
54. Pornillos O, Ganser-Pornillos BK, Yeager M (2011) Atomic-level modelling of the HIV capsid. *Nature* 469:424–427
55. Cardone G, Purdy JG, Cheng N, Craven RC, Steven AC (2009) Visualization of a missing link in retrovirus capsid assembly. *Nature* 457:694–698
56. Pina M, Bize A, Forterre P, Prangishvili D (2011) The archeoviruses. *FEMS Microbiol Rev* 35:1035–1054
57. Pietila MK, Atanasova NS, Manole V, Liljeroos L, Butcher SJ, Oksanen HM, Bamford DH (2012) Virion architecture unifies globally distributed pleolipoviruses infecting halophilic archaea. *J Virol* 86:5067–5079
58. Bolduc B, Shaughnessy DP, Wolf YI, Koonin EV, Roberto FF, Young M (2012) Identification of novel positive-strand RNA viruses by metagenomic analysis of archaea-dominated yellowstone hot springs. *J Virol* 86:5562–5573
59. Aalto AP, Bitto D, Ravanti JJ, Bamford DH, Huiskonen JT, Oksanen HM (2012) Snapshot of virus evolution in hypersaline environments from the characterization of a membrane-containing salisaeta icosahedral phage 1. *Proc Natl Acad Sci U S A* 109:7079–7084
60. Haring M, Vestergaard G, Rachel R, Chen L, Garrett RA, Prangishvili D (2005) Virology: independent virus development outside a host. *Nature* 436:1101–1102
61. Hong SS, Szolajiska E, Schoehn G, Franqueville L, Myhre S, Lindholm L, Ruigrok RW, Boulanger P, Chroboczek J (2005) The 100K-chaperone protein from adenovirus serotype 2 (subgroup C) assists in trimerization and nuclear localization of hexons from subgroups C and B adenoviruses. *J Mol Biol* 352:125–138
62. Johnson JE (2010) Virus particle maturation: insights into elegantly programmed nanomachines. *Curr Opin Struct Biol* 20:210–216
63. Prevelige PE, Fane BA (2012) Building the machines: scaffolding protein functions during bacteriophage morphogenesis. *Adv Exp Med Biol* 726:325–350
64. Marvik OJ, Sharma P, Dokland T, Lindqvist BH (1994) Bacteriophage P2 and P4 assembly: alternative scaffolding proteins regulate capsid size. *Virology* 200:702–714
65. Huang RK, Khayat R, Lee KK, Gertsman I, Duda RL, Hendrix RW, Johnson JE (2011) The prohead-I structure of bacteriophage HK97: implications for scaffold-mediated control of particle assembly and maturation. *J Mol Biol* 408:541–554
66. Hasson TB, Ornelles DA, Shenk T (1992) Adenovirus L1 52- and 55-kDa proteins are present within assembling virions and colocalize with nuclear structures distinct from replication centers. *J Virol* 66:6133–6142
67. Pérez-Berná AJ, Ortega-Esteban A, Menéndez-Conejero R, Winkler DC, Menéndez M, Steven AC, Flint SJ, de Pablo PJ, San Martín C (2012) The role of capsid maturation on adenovirus priming for sequential uncoating. *J Biol Chem* 287:31582–31595
68. Hogle JM (2002) Poliovirus cell entry: common structural themes in viral cell entry pathways. *Annu Rev Microbiol* 56:677–702
69. Bahar MW, Graham SC, Stuart DI, Grimes JM (2011) Insights into the evolution of a complex virus from the crystal structure of vaccinia virus D13. *Structure* 19:1011–1020
70. Bamford DH, Burnett RM, Stuart DI (2002) Evolution of viral structure. *Theor Popul Biol* 61:461–470
71. Abrescia NG, Bamford DH, Grimes JM, Stuart DI (2012) Structure unifies the viral universe. *Annu Rev Biochem* 81:795–822

Further Reading

- Agbandje-McKenna M, McKenna R (2011) Structural virology. RSC Publishing, Cambridge
- Flint SJ, Enquist LW, Racaniello VR, Skalka AM (2009) Principles of virology. ASM Press, Washington, DC
- Rixon FJ, Chiu W (2003) Studying large viruses. Adv Protein Chem 64:379–408
- Rossmann MG, Rao VB (2012) Viral molecular machines. Adv Exp Med Biol, vol 726, Springer, New York

Also especially recommended for further reading are references [4, 9, 24, 32, 42, 54, 56, 62] listed above.