Part I
The Viral Machine

Chapter 1
Introduction: The Structural Basis of Virus Function

Mauricio G. Mateu

Abstract Viruses may be regarded as dynamic nucleoprotein assemblies capable of assisted multiplication within cells, and of propagation between cells and organisms. Infectious virus particles (virions) assembled in a host cell are dynamic, generally metastable particles: They are robust enough to protect the viral genome outside the cell, but are also poised to undergo structural changes and execute mechanochemical actions required for infection of other cells. This chapter provides an introduction to the structural and physical biology of viruses by including: (i) an elementary overview on virions and the structural basis of virus function; (ii) a concise summary on basic techniques used in structural or physical virology; (iii) brief structure-based general descriptions of the different stages in the virus cycle, especially those in which virions and/or their components are involved. These contents may facilitate a better understanding of the specialized subjects treated in the rest of the book. This chapter is also intended as a "road map" to help interconnect and integrate in a single picture the different topics described in depth in the 21 monographic chapters in this book.

Keywords Virus • Virus cycle • Infection • Capsid • Viral genome • Capsid subunits • Capsid building blocks • Oligomerization • Self-assembly • Assisted assembly • Assembly intermediates • Scaffolding proteins • Conformational stability • Conformational dynamics • Nucleic acid packaging • Capsid-nucleic acid condensation • Virus maturation • Virus stability and dynamics • Virus-antibody recognition • Virus-receptor recognition • Virus entry • Fusion • Uncoating • Antivirals • Vaccines • Biotechnology • Nanotechnology

M.G. Mateu (✉)
Centro de Biología Molecular "Severo Ochoa" (CSIC-UAM)
and Department of Molecular Biology of the Universidad Autónoma
de Madrid, c / Nicolás Cabrera 1, Campus de Cantoblanco, 28049 Madrid, Spain
e-mail: mgarcia@cbm.uam.es

M.G. Mateu (ed.), *Structure and Physics of Viruses: An Integrated Textbook*,
Subcellular Biochemistry 68, DOI 10.1007/978-94-007-6552-8_1,

AAV	Adeno-associated virus
AFM	Atomic force microscopy
CBB	Capsid building block
CP	Capsid protein
cryo-EM	Cryo-electron microscopy
DNA	Deoxyribonucleic acid
ds	Double-stranded
EMDB	Electron Microscopy Data Bank
Fab	Antigen-binding antibody fragment
FMDV	Foot-and-mouth disease virus
HIV-1	Human immunodeficiency virus type 1
HRV	Human rhinovirus
HSV-1	Herpes simplex virus type 1
ICTV	International Committee on Taxonomy of Viruses
MD	Molecular dynamics
mRNA	Messenger RNA
MS	Mass spectrometry
MVM	Minute virus of mice
NMR	Nuclear magnetic resonance
PDB	Protein Data Bank
PV	Poliovirus
RNA	Ribonucleic acid
RT	Reverse transcriptase
SAXS	Small-angle X-ray scattering
ss	Single-stranded
TMV	Tobacco mosaic virus
VLP	Virus-like particle

1.1 The Structure and Physics of Viruses

Viruses are biological entities capable of assisted multiplication within cells and of propagation between cells and organisms. Virus multiplication and propagation is generally a cyclic process: An infectious viral particle (*virion*) introduces its genome into a host cell, new virions are formed in the cell and released, and these in turn may infect other host cells. This cycle of infection is often called the *virus* "*life*" *cycle*. There has been a largely philosophical debate on whether viruses are alive or not. We use the term *virus life cycle* as a synonym of *infectious cycle*; we are not making the statement that viruses are "living" organisms. In this book viruses are contemplated as macromolecular complexes that, through biological evolution, came into existence and were endowed with the capacity to make copies of themselves by using the genetic instructions they enclose and the host cell machinery.

Because of the effects many viral infections cause on living beings, viruses are frequently considered only as pathogens causing disease and human suffering, economic losses and social problems. However, since the times of the "Phage Group" and the advent of molecular virology more than half a century ago, scientists coming from different areas have become increasingly aware that viruses also provide outstanding, relatively simple models to explore biomolecular structure-function relationships using a combination of physical, chemical and biological approaches. The knowledge thus acquired has been decisive not only to combat viral disease, but also in the quest to understand in physico-chemical terms the molecular machinery of life.

Specific reasons for studying the structure, dynamics and physical and (bio) chemical properties of virus particles include the following:

(i) Virus particles constitute excellent models to understand and learn to manipulate molecular self-assembly.
(ii) Virus particles are paradigms to understand structure-function relationships in biomacromolecular assemblies and biological machines.
(iii) A profound knowledge of virus structure, dynamics and properties is essential to understand the life cycles of viruses.
(iv) Virus particles, their components and the processes in which they participate provide novel targets for the design of antiviral agents.
(v) Understanding the structural determinants of virus stability, dynamics and function may facilitate the rational manipulation of virus particles to develop new or improved vaccines, gene therapy vectors, and nanoparticles for drug delivery or other biomedical or bio/nanotechnological uses.

This chapter provides an introduction to structural and physical virology and is intended mainly for M.Sc. students, Ph.D. students and postdoctoral researchers in physics, chemistry, biology or related areas who are interested in viruses, but who may be relatively unfamiliar with the subject. It intends also to provide a "road map" to help the reader integrate in a general picture the topics treated in depth in the other chapters in this book. To achieve these aims, the present chapter includes:

(i) Brief explanations on some elementary concepts and terms in molecular, structural and physical virology; a detailed description of the basic architecture of viruses will follow in the accompanying Chap. 2 in Part I of this book.
(ii) Some broad guidelines on the applicability of most of the different techniques described in Part II, Chaps. 3, 4, 5, 6, 7, 8 and 9, and Chaps. 14, 19, to characterize the structure, dynamics and physical properties of virus particles.
(iii) A brief overview of a generic virus cycle and of the major roles in the cycle of virions and their components.
(iv) Brief accounts of general structural concepts and descriptions regarding each of the different stages of the viral cycle in which virions or their components are involved, and of relevant properties of virus particles. These short, elementary accounts may facilitate those readers with little background in molecular and structural virology a fullest and better integrated understanding of the

contents of Part III, Chaps. 10, 11, 12, 13, 14, 15, 16 and 17; each of these chapters deals in detail with structural aspects of one of those stages of the virus cycle.

(v) A schematic overview of novel, physics-based approaches to study virus structure, dynamics and properties, as a brief introduction to detailed accounts of some physical virology methods (Part II, Chaps. 8, 9), studies (Part III, Chaps. 18, 19) and applications (Part IV, Chap. 22).

(vi) A very brief overview of applied studies in structural virology, to put into a general context the particular applications described in detail in Part IV, Chaps. 20, 21 and 22.

1.1.1 Structural Virology

Our knowledge of the molecular structure and function of viruses has grown spectacularly in the last decades, largely because these entities have uniquely and increasingly attracted the interest of biologists, (bio)chemists and physicists alike. Viruses have, thus, been rediscovered as organized complexes of biomolecules that act as minute machines. These nanomachines are continuously being modified and diversified through mutation and biological adaptation. However, they are invariably determined by the laws of physics and chemistry to blindly perform sophisticated mechanochemical actions, including penetration into host cells and self-assembly from their molecular components after these have been replicated in the cell. The application of physical and physicochemical techniques has led to the determination of the high-resolution molecular structures of many viruses; the interplay of this approach with (bio)chemical and biological approaches have allowed in many cases the elucidation of the structural basis of viral function in unprecedented detail. *Structural Virology* has permeated other virological disciplines to provide a molecular view of viruses and their biology. The detailed structural knowledge on viruses and their components has made, and will surely continue to make, decisive contributions in the fight against viral disease.

1.1.2 Physical Virology

In the last years, the advent of nanoscience and nanotechnology, and the increasing awareness on the outstanding features of virus particles as solid-state nanodevices are leading to a renewed look at viruses from the physicist's point of view. Theoretical physicists have begun to tackle at the most fundamental level different aspects of the architecture, self-assembly and physical properties of virus particles. Also, the development of atomic force microscopy (AFM), optical tweezers and other techniques to study individual molecules have opened up new possibilities for the experimental study of the structure, properties and mechanochemical actions of

viruses and their components. A new term, *Physical Virology*, has recently been coined to encompass theoretical and experimental physics-oriented studies of viruses. This novel approach is slowly beginning to merge with long-standing structural virology approaches based on other physical or physico-chemical techniques such as electron microscopy, X-ray crystallography and many others. As a consequence, viruses are currently being investigated for new developments not only in biomedicine and biotechnology but also in nanotechnology, including nanomaterials and nanoelectronics.

1.2 Virions and Their Structural Components

1.2.1 Molecular Composition of Viruses

From a structural point of view virions may be generally regarded as nucleoprotein assemblies. They all include a nucleic acid genome and many copies of one or more proteins. However, virions present remarkable differences in size, shape, molecular composition, structural organization and complexity (Fig. 1.1). Considering its molecular composition only, viruses are generally classified into two large groups, *non-enveloped viruses* and *enveloped viruses*, depending on the absence or presence of an outer lipid layer.

Non-enveloped Viruses

The simplest non-enveloped virions are composed just of a protein shell, or *capsid* (sometimes called *coat*) made of multiple copies of one or more proteins, that contains the viral nucleic acid (Fig. 1.1, left). In less simple non-enveloped virions, the capsid may contain not only the viral genome, but also other proteins and/or other macromolecules, which may be organized in subassemblies. These additional biomolecules or subassemblies can be enclosed in the capsid shell or externally attached to its surface (Fig. 1.1, center) (see Chaps. 2, 10, 11, 17).

Enveloped Viruses

In enveloped virions, the capsid and/or other internal structures are typically surrounded by a lipid bilayer, or *envelope*, in which some proteins are embedded (Fig. 1.1, right); some enveloped virions have a complex multi-layer structure made of organized lipid, protein and/or nucleoprotein layers (see Chaps. 2, 11).

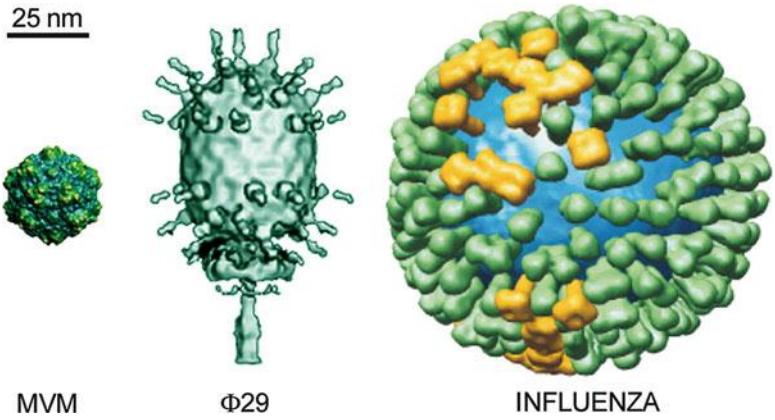

Fig. 1.1 Major types of viruses by molecular composition. *Left* and *center*, non-enveloped viruses; *left*, a very simple virus (the parvovirus MVM); *center*, a complex virus (the tailed bacteriophage ϕ29). *Right*, an enveloped virus (the orthomyxovirus influenza virus). They are reproduced at (approximately) the same scale, indicated by the *horizontal bar* (*top left*). The MVM structural model [Agbandje-McKenna M, Llamas-Saiz AL, Wang F, Tattersall P, Rossmann MG (1998) Structure 6:1369–1381] was obtained from VIPERdb [Carrillo-Tripp M, Sheperd CM, Borelli IA, Sangita V, Lander G, Natarajan P, Johnson JE, Brooks III CL, Reddy VS (2009) Nucleic Acids Res 37:D436–D442]. The models of ϕ29 and influenza virus are respectively reproduced from [Wikoff WR, Johnson JE (1999) Curr Biol 9:R296–R300] and [Harris A, Cardone C, Winkler DC, Heymann JB, Brecher M, White JM, Steven AC (2006) Proc Natl Acad Sci USA 103:19123–19127], with permission. (Figure kindly provided by M.A. Fuertes)

1.2.2 The Virus Capsid

The capsid plays a fundamental role in both the architecture and the biological function of a virus. A virus capsid can be generally described as a hollow symmetric protein oligomer or multimer made from several tens to many hundreds of copies of one or a few different types of folded polypeptides, the *capsid protein* (CP) subunits. Most CPs (or their individual structural domains if formed by more than one domain) can be ascribed to one of a very limited number of protein architectures or *folds* that can assemble into a limited number of quaternary structures (Chap. 2).

In each virus, oligomerization of the CPs during capsid assembly normally leads to a defined type of symmetric quaternary structure. Only very few types of capsid symmetry are frequent. The basic types are *helical* (Fig. 1.2, top) and *icosahedral* (Fig. 1.2, bottom left and right). In some viruses, capsids adopt an elongated (prolate) icosahedral architecture (Fig. 1.1, right). Other capsids, such as those of retroviruses or poxviruses, are made of less simple arrangements of CP subunits (see Chaps. 2, 10, 11). Structural models (density maps) of virions and other virus particles determined by electron microscopy (Chap. 3) are available at the Electron Microscopy Data Bank (EMDB) (http://www.ebi.ac.uk/pdbe/emdb). Atomic coordinates and high-resolution structural models of virions or other virus particles

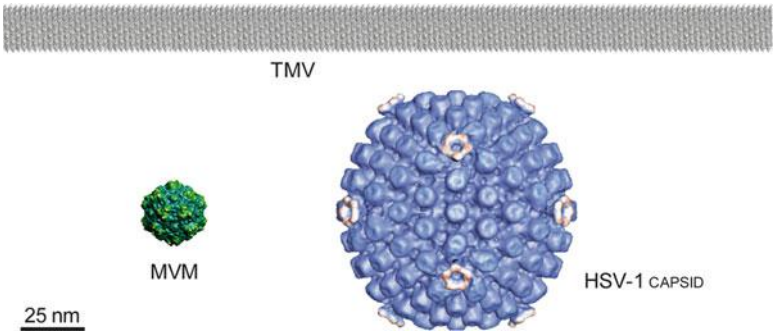

Fig. 1.2 Major types of viral capsid symmetry. *Top image*: helical; the capsid of TMV is shown. *Bottom images*: icosahedral; *left image*, the simple icosahedral capsid of the parvovirus MVM; *right image*, the complex icosahedral capsid of the herpesvirus HSV-1. Images of TMV, MVM and HSV-1 are respectively adapted or reproduced from [Clare DK, Orlova EV (2010) J Struct Biol 171:303–308], [Agbandje-McKenna M, Llamas-Saiz AL, Wang F, Tattersall P, Rossmann MG (1998) Structure 6:1369–1381], and [Grünewald K, Desai P, Winkler DC, Heymann JB, Belnap D, Baumeister W, Steven AC (2003) Science 302:1396–1398], with permission. (Figure kindly provided by M.A. Fuertes)

whose structures have been determined by X-ray crystallography (Chap. 4) are available at the Protein Data Bank (PDB) (http://www.rcsb.org).

Helical Capsids

Helical capsids (Fig. 1.2, top) are extremely simple, theoretically infinite multimers and (in principle) could be made as long as required to encapsidate a nucleic acid genome irrespective of its length; thus, there is no theoretical limitation on the amount of genetic information these capsids could store. However, physical and biological restrictions limit the length of helical capsids, which are much less frequent than icosahedral capsids; regular helical capsids are found in about 10 % of virus families. Chapter 2 provides a description of the architecture of helical capsids; examples of helical viruses can be found in other chapters of this book.

Icosahedral Capsids

In contrast to helical capsids, strictly icosahedral capsids (Fig. 1.2, bottom left) must be made of exactly 60 structurally identical components (CP monomers or oligomers) in order to fulfill intrinsic geometric constraints. This fact could severely limit the size of the nucleic acid genome and, hence, the amount of genetic information that could be enclosed (see Chap. 2). However, evolution has led to different structural solutions to make very large viral capsids with icosahedral symmetry, made of hundreds of chemically identical CP subunits, that can enclose

very large genomes (Fig. 1.2 bottom right). Many large capsids containing many subunits are made of only one type of CP that is capable to adopt different *quasi-equivalent* (similar) conformations (given by the *triangulation number* T); this feature allows the CP molecule to fit into T non-equivalent positions in the icosahedral capsid (see Chap. 2 for a detailed description). This evolutionary strategy minimizes the amount of genetic information required to encode the capsid. Icosahedral capsids are extremely frequent; they occur in about half the virus families. See Chap. 2 for an in-depth description of the architectures of icosahedral viral capsids; Chaps. 2, 10, 11 and others in this book provide many examples of icosahedral capsids and viruses. A database (VIPERdb) contains abundant structural information on viruses with icosahedral capsids whose structure has been determined to high resolution (http://viperdb.scripps.edu).

Virion Architecture

In non-enveloped viruses with a helical or icosahedral capsid, the basic architecture of the capsid determines also the basic architecture of the virion. In enveloped virions, the situation is more complex. Many enveloped virions containing icosahedral capsids and/or other types of compact (nucleo)protein complexes tend to adopt a relatively flexible, frequently spheroidal shape which may differ in size between individual particles.

For example, in herpesviruses (*e.g.*, herpes simplex virus type 1, HSV-1) a large icosahedral capsid (Fig. 1.2, bottom right) containing the viral genome is enclosed in an outer layer of proteins (*tegument*), surrounded in turn by a lipidic envelope; the virion that results is considerably larger than the capsid and spheroidal in shape. In influenza virus (Fig. 1.1, right), several roughly helical nucleocapsid complexes (ribonucleoprotein particles) are directly surrounded by a protein layer (the *matrix*) and the lipid envelope, and this virus is clearly pleomorphic. Many other variations in virion architecture, some of them very complex, do exist. In some cases, two concentric capsids are found; in others, internal lipidic envelopes are present. The structures of some complex viruses are described in detail in Chaps. 2, 11. Other examples of complex virion architectures can be found in several chapters in this book.

1.2.3 Types of Viral Nucleic Acid

The type of genomic nucleic acid used by different viruses is of prime importance to determine the mechanisms of genome replication and expression during the metabolically active phase of the viral cycle. From the perspective of this book the type of nucleic acid, especially whether it is single-stranded (ss) or double-stranded (ds), is extremely relevant for the assembly of many virus capsids (Chaps. 2, 10, 11, 19), the mechanism of nucleic acid packaging (Chap. 12), the organization of the nucleic acid inside the capsid (Chap. 12), virus particle maturation (Chap. 13), and some properties and functions of the virion (Chaps. 2, 10, 11, 12, 13 and 18).

RNA Viruses

Riboviruses (*RNA viruses*) use RNA as genetic material. Some of them use ssRNA (*ssRNA viruses*), others use dsRNA (*dsRNA viruses*). In turn, two types of ssRNA viruses can be distinguished depending on the polarity of their RNA genome strand (*ssRNA*(+) *viruses* and *ssRNA*(−) *viruses*). A genomic single-stranded nucleic acid is considered of positive (+) polarity if its sequence corresponds to that in the viral messenger RNAs (mRNA), and of negative (−) polarity if its sequence corresponds to the complementary of the mRNA sequences. In some viruses, the virion encloses more than one copy of the genome (*e.g.*, retroviruses), or the genome is *segmented*, split into several nucleic acid molecules (*e.g.*, influenza virus). Most of the virus species known are RNA viruses.

DNA Viruses

Deoxyviruses (*DNA viruses*) use DNA as their genetic material. Some of them use ssDNA (*ssDNA viruses*); others use dsDNA (*dsDNA viruses*).

1.2.4 Host Cells and Organisms

Viruses can be grouped also by the kind of cell they can infect (Table 1.1). Many known viruses infect eukaryotic cells (*eukaryotic viruses*). Of them, many infect animals such as vertebrates (including humans) or insects (*animal viruses*); others infect fungi (*fungal viruses* or *mycoviruses*), plants (*plant viruses*) or protista (*protist viruses*). Many other viruses infect prokaryotic cells such as bacteria (*prokaryotic viruses*). Bacterial viruses are usually called *bacteriophages* or, simply, *phages*. Archea are also infected by viruses (*archeal viruses*). Most viruses have evolved to infect one or a few species of organisms (limited *host range*) and one or a few cell types (specific *tropism*). The type of cell a virus can infect is, in part, the consequence of the structural and functional features the virus particle has evolved to enter a particular cell. In general, infection by an animal virus or bacteriophage depends largely on the receptor molecule(s) on the cell surface that the virion can specifically bind (see Sect. 1.4.5 and Chaps. 15, 16, 17). Although some archeal, protist and fungal viruses are providing fascinating cases for structural study (a few of which are mentioned in this book), the vast majority of viruses that have been the subject of structural and molecular virology studies infect animals, plants or bacteria (see Table 1.1). Detailed information on the molecular biology of animal viruses, plant viruses and/or phages can be found also in the books listed at the end of this chapter [1–10].

Table 1.1 Some virus families and species grouped according to type of host and viral genome

Host[a]	Viral genome[b]	Virus order[c]	Virus family[d]	Some virus species[e]
Bacteria	dsDNA		*Corticoviridae*	PM2
		C	*Myoviridae*	P1, P2, P4, SPO1, T4, μ, ϕ92, 8a, 44RR
		C	*Podoviridae*	BPP-1, K1F, P22, Sf6, T7, ε15, ϕ15, ϕ29
		C	*Siphoviridae*	HK97, p2, SF6, SPP1, TP901-1, T5, λ
			Tectiviridae	PRD1, P23-77
	ssDNA		*Microviridae*	ϕX174
			Inoviridae	fd, M13
	dsRNA		*Cystoviridae*	ϕ6, ϕ12
	ssRNA(+)		*Leviviridae*	MS2, Qβ
Archea	dsDNA		*Ampullaviridae*	ABV
			Guttaviridae	SNDV
			Lipothrixviridae	AFV-1
			Rudiviridae	SIRV-2
Protista	dsDNA		*Phycodnaviridae*	PBCV-1
Plant	ssDNA		*Geminiviridae*	BGYMV
	ssRNA(+)		*Bromoviridae*	AMV, BMV, CCMV, CMV
		T	*Flexiviridae* (*Alphaflexiviridae*)	PapMV
			Potyviridae	BYMV
		P	*Secoviridae*	BPMV
			Sobemovirus (*unassigned genus*)	RYMV, SBMV
			Tombusviridae	CPMV, RCNMV, STNV, TBSV, TNV A
		T	*Tymoviridae*	PhMV, TYMV
			Virgaviridae	TMV
	dsDNA-RT		*Caulimoviridae*	CaMV
Fungi	dsRNA		*Partitiviridae*	*Atkinsonella hypoxylon* virus
			Chrysoviridae	*Penicillium chrysogenum* virus
Animal	dsDNA		*Adenoviridae*	(h)Ad
			Asfarviridae	ASFV
			Baculoviridae	AcNPV, cytoplasmic polyhedrosis virus, WNPV
			Iridoviridae	Frog virus 3
		H	*Herpesviridae*	HCMV, HSV-1, VZV
			Mimiviridae	*Acanthamoeba polyphaga* mimivirus
			Papillomaviridae	BPV, HPV 1
			Polyomaviridae	SV40, (human) polyomavirus
			Poxviridae	Vaccinia virus, Variola virus
	ssDNA		*Circoviridae*	PCV2
			Parvoviridae	AAV-2, CPV, FHV, H1-PV, MEV, MVM, PPV

(continued)

Table 1.1 (continued)

Host[a]	Viral genome[b]	Virus order[c]	Virus family[d]	Some virus species[e]
	dsRNA		*Birnaviridae*	IBDV
			Reoviridae	BTV, Rotavirus (A)
	ssRNA(+)		*Astroviridae*	Mamastrovirus 1
			Caliciviridae	NV (Norovirus), RHDV
		N	*Coronaviridae*	SARS virus
			Flaviviridae	Dengue virus, HCV, TBEV, yellow fever virus
			Hepeviridae	HEV
			Nodaviridae	FHV, PaV
		P	*Picornaviridae*	Coxsackievirus, echovirus, FMDV, HAV, HRV, PV
			Tetraviridae (*Alphatetraviridae*)	NωV
			Togaviridae (*genus Alphavirus*)	Chikungunya virus, Semliki Forest virus, Sindbis virus
	ssRNA(−)		*Arenaviridae*	LCMV
			Bunyaviridae	Bunyamwera virus, Hantaan virus, RVFV
		M	*Filoviridae*	Ebola virus, Marburg virus
			Orthomyxoviridae	Influenza virus (A, B, C)
		M	*Paramyxoviridae*	HeV, hMPV, hPIV, measles virus, mumps virus, NDV, NiV, Sendai virus, RSV(Respiratory Syncytial virus)
		M	*Rhabdoviridae*	Rabies virus, VSV
	ssRNA(+)-RT		*Retroviridae*	BLV, EIAV, HIV-1, HIV-2, HTLV-I, MLV, MMTV, MoMuLV, MPMV; RSV(Rous sarcoma virus), SIV
	dsDNA-RT		*Hepadnaviridae*	HBV

[a]Classified according to kingdom. Different virus species of some families may infect organisms from different kingdoms. For example, some species of the *Reoviridae*, *Bunyaviridae*, *Rhabdoviridae* may infect animals and other species may infect plants

[b]*RT* indicates that virus replication involves a reverse transcriptase

[c]The six virus orders are: *C Caudovirales*, *H Herpesvirales*, *M Mononegavirales*, *N Nidovirales*, *P Picornavirales*, *T Tymovirales*. All other virus families listed have not been assigned to any order (see text)

[d]The virus families listed include nearly all of those mentioned in this book and a few others. The current ICTV virus classification includes 94 families, and several independent genera that have not been ascribed to families (see text). The families *Flexiviridae* and *Tetraviridae* have recently been split into several families each; according to the current ICTV classification, the virus species listed for those families belong to the *Alphaflexiviridae* and *Alphatetraviridae* families, respectively

[e]The virus species listed as examples include nearly all of those mentioned in this book and a few additional ones. Virus species abbreviations: *AAV* adeno-associated virus, *ABV Acidianus* bottle-shaped virus, *AcNPV Autographa californica* nuclear polyhedrosis virus, *AFV Acidianus* filamentous virus, *AMV* alfalfa mosaic virus, *ASFV* african swine fever virus, *BGYMV* bean gold yellow

(continued)

Table 1.1 (continued)

mosaic virus, *BLV* bovine leukemia virus, *BMV* brome mosaic virus, *BPMV* bean pod mottle virus, *BPV* bovine papillomavirus, *BTV* bluetongue virus, *BYMV* barley yellow mosaic virus, *CaMV* cauliflower mosaic virus, *CCMV* cowpea chlorotic mosaic virus, *CMV* cucumber mosaic virus, *CPMV* cowpea mosaic virus, *CPV* canine parvovirus, *EIAV* equine infectious anemia virus, *FHV* flock house virus, *FMDV* foot-and-mouth disease virus, *FPV* feline parvovirus, *H1-PV* parvovirus H1, *(h)Ad* (human) adenovirus, *HAV* hepatitis A virus, *HBV* hepatitis B virus, *HCMV* human cytomegalovirus, *HCV* hepatitis C virus, *HeV* Hendra virus, *HEV* hepatitis E virus, *HIV* human immunodeficiency virus, *hMPV* human metapneumovirus, *HPV* human papillomavirus, *hPIV* human parainfluenza virus, *HRV* human rhinovirus, *HSV* human herpesvirus, *HTLV* human T-lymphotropic virus, *IBDV* infectious bursal disease virus, *LCMV* lymphochoriomeningitis virus, *MEV* mink enteritis virus, *MLV* murine leukemia virus, *MMTV* mouse mammary tumor virus, *MPMV* Mason-Pfizer monkey virus, *MVM* minute virus of mice, *NDV* Newcastle disease virus, *NiV* Nipah virus, *NV* Norwalk virus, *NωV Nudaurelia capensis* ω virus, *PapMV* papaya mosaic virus, *PaV* Pariacoto virus, *PBCV Paramecium bursaria Chlorella* virus, *PCV* porcine circovirus, *PhMV Physalis* mottle virus, *PPV* porcine parvovirus, *PV* polio virus, *RCNMV* red clover necrosis mosaic virus, *RHDV* rabbit haemorrhagic disease virus, *RSV* respiratory syncytial virus, *RSV* Rous sarcoma virus, *RVFV* Rift Valley fever virus, *RYMV* rice yellow mottle virus, *SARS* virus, severe acute respiratory syndrome virus, *SBMV* southern bean mosaic virus, *SIRV Sulfolobus islandicus* rod-shaped virus, *SIV* simian immunodeficiency virus, *SNDV Sulfolobus neozelandicus* droplet-shaped virus, *STMV* satellite tobacco mosaic virus, *STNV* satellite tobacco necrosis virus, *SV40* simian virus 40, *TBEV* tick-borne encephalitis virus, *TBSV* tomato bushy stunt virus, *TMV* tobacco mosaic virus, *TNV* tobacco necrosis virus, *TYMV* turnip yellow mosaic virus, *VSV* vesicular stomatitis virus, *VZV* varicella-zoster virus, *WNPV Wiseana* nuclear polyhedrosis virus

Animal Viruses

Animal viruses are extremely diverse in terms of size, shape, presence or absence of envelope, capsid architecture, type of nucleic acid genome, and specific mechanisms employed to complete the different stages of the viral cycle as required for infection, multiplication and propagation. Some animal viruses are among the structurally simplest viruses known. These include the parvoviruses (*e.g.*, the minute virus of mice, MVM (Fig. 1.1, left) or the adeno-associated viruses, AAV; see Chap. 10). Others are structurally much more complex, such as the adenoviruses, herpesviruses (*e.g.*, HSV-1), retroviruses (*e.g.*, the human immunodeficiency virus type 1 (HIV-1) or the giant mimivirus (see Chap. 11)). The study of the structure, function, biology and pathogenicity of many animal viruses has been or is still hampered by a number of difficulties, including their structural and/or functional complexity, technical problems to grow and/or manipulate them and/or safety issues. However, despite all the difficulties, and mainly because of the biomedical or socioeconomic importance of a large number of animal viruses that are pathogenic for humans or livestock, many of these are among the most intensively studied viruses of all. This is reflected also in the many human and animal viruses used as examples and case studies in most chapters of this book. Many particular aspects related to the molecular biology and structure of many different animal viruses can be found also in the books listed at the end of this chapter [1–6, 9–16].

Bacteriophages

Bacteriophages show widely diverse structures and types of nucleic acid genomes; they have helical or icosahedral capsids, and may or may not include a lipid envelope. Phages range from very small and simple non-enveloped icosahedral viruses (*e.g.*, ϕX174) and long but simple helical viruses (filamentous phages), to large and complex viruses (*e.g.*, tailed DNA phages such as ϕ29; Fig. 1.1, center). Since the origins of Molecular Biology over half a century ago, and continuing through several decades, some bacteriophages were found to present important advantages as model systems for molecular and genetic studies compared to most animal and/or plant viruses. Their advantages include the facility to grow phages and their bacterial hosts in large amounts; the relative structural and functional simplicity and ease of handling of bacterial cells compared to eukaryotic cells; and the possibility to readily obtain certain types of mutant viruses to investigate virus structure and function. These and other reasons led to the intensive use of phages as paradigms for many molecular biology studies on nucleic acid replication, gene expression and their regulation. Several bacteriophages have also provided and continue to provide model systems for studying molecular recognition and self-assembly during the morphogenesis of biomolecular complexes. As a consequence, the structure and function of some bacteriophages, and most stages in their life cycles, are known in great detail (see Chaps. 11, 17). Many particular aspects of the molecular biology of different bacteriophages can be found in several books listed at the end of this chapter. For a book on bacteriophage molecular biology see [7].

Plant Viruses

Most of the very abundant plant viruses are non-enveloped ssRNA(+) viruses with a slender helical capsid or a relatively small icosahedral capsid. The structure and/or function of a few plant viruses, such as the tobacco mosaic virus (TMV; Fig. 1.2 top) have been intensively studied for many decades, in many cases because of some advantages of those viruses as model systems; for example, the facility to grow plant viruses in very large quantities by simply infecting host plants. In addition, studies on many plant viruses have been greatly stimulated because of the economically important diseases they cause in crop plants. Recently, some of the advantages of phages and plant viruses referred to above have led to their preferential use as platforms for many bio/nanotechnological developments (see Chap. 22). However, generally speaking plant viruses have been the subject of fewer studies than animal viruses or bacteriophages, and many stages of the life cycles of the former remain less well known than those of the latter. For a book on plant virus molecular biology and structure see [8].

1.2.5 Classification of Viruses

Comparisons of the sequences of viral genes and genomes have led to the establishment of phylogenetic relationships between many viruses. In addition, comparison of the tertiary structure of viral proteins, especially CPs, has allowed the tentative proposal of distant evolutionary relationships among different viruses, or at least between some of their genes (see a brief description in Chap. 7). It must be noted here that genetic recombination and horizontal gene transfer between even very different, unrelated viruses are frequent. Thus, viruses of widely different origins could share some evolutionarily closely related genes, and of course the proteins these genes encode. It is not yet possible to solidly establish a general phylogenetic-based classification of viruses. In 1973 the International Committee on Taxonomy of Viruses (ICTV) was established, and a general database on viruses was created later (http://www.ncbi.nlm.nih.gov/ICTVdb).

Known viruses have been classified into seven major groups based on the type of nucleic acid genome (*Baltimore classification*). These groups are: I: dsDNA viruses; II: ssDNA viruses; III: dsRNA viruses; IV: ssRNA(+) viruses; V: ssRNA (−) viruses; VI: ssRNA(+) virus whose replication involves the action of a reverse transcriptase (RT) that synthesize DNA from a RNA template; VII: dsDNA viruses whose replication involves the action of an RT (see Table 1.1).

In addition to their classification according to type of nucleic acid genome, viruses have been classified by ICTV in a number of taxonomic groups (*taxons*): viral *order*, *family*, *subfamily*, *genus*, and *species*. The most useful taxon in virus classification is the *family* (see Table 1.1). Viruses in a same family probably share a not too distant evolutionary relationship, as established mainly by comparative sequence analysis. There are currently 94 recognized virus families; only 22 of these families have been grouped in 6 orders (*Caudovirales*, *Herpesvirales*, *Mononegavirales*, *Nidovirales*, *Picornavirales*, *Tymovirales*; see Table 1.1); the remaining families have not been assigned to any order yet. Also, several virus genera have not been assigned to any family yet.

Virus family latin names (italicized) include the suffix –*viridae*. Very frequently, the family english name (non-italicized), which include the suffix –virus (plural - viruses) is used instead of the latin name. However, this may occasionally cause confusion on whether one is referring to a viral family or genus, unless this point is specified. Virus species are usually referred to by their english names, and most have been given standard abbreviations. Most bacteriophages are named according to a code of latin letters, greek letters and/or numbers.

In this book, viruses will be generally identified by type of nucleic acid, by family (latin or english name) and/or by species. For example, in different chapters repeated mention is made to the human immunodeficiency virus type 1 (HIV-1). HIV-1 is a virus species of the *Retroviridae* (retrovirus family), which belongs to a group of ssRNA(+) viruses whose replication involves a RT (group VI). A list of

nearly all virus families and species mentioned in this book, along with a few other important families and species, is included in Table 1.1. This table is mainly intended to help the reader navigate among the multitude of virus names that will inevitably appear along the different chapters of this book (or of any other virology book). In addition, an index of names of nearly all virus species and families mentioned in this book can be found at the end of the book.

Viruses in each family do share many features and present many structural and functional similarities, apart from the type of nucleic acid genome. Of the utmost relevance for the subject of this book is that viruses from a same family generally show a clear similarity regarding virion and capsid structures. However, it must emphatically be noted also that even viruses of a same family may sometimes *drastically* differ in some or many structural features, properties, functions and underlying mechanisms, cells and organisms they can infect, pathogenic effects, etc. Substantial phenotypic differences may occur even between some individual viruses from a same clone, and some of these differences can be due to a single amino acid substitution in some viral protein. The converse is also true: some viruses from different families may share many structural and functional similarities, probably as an evolutionary consequence of similar "lifestyles" (either by conservation of basic structural and/or functional features, or by convergent evolution; see, for example, Chaps. 2, 7 regarding structural similarities in CPs and Chaps. 2, 19 for fundamental physics-based similarities in architecture and some properties of icosahedral virus capsids). Figure 1.3 shows very simple general schemes of virion structures for members of some important families of animal viruses, many of which are mentioned in this book.

1.3 Techniques Used to Study the Structure and Physics of Viruses

Today, the structural virologist has at his/her disposal a vast array of sophisticated and powerful techniques to study virus structure, dynamics (including conformational rearrangements, assembly and disassembly) and physicochemical properties. Most of the major structural and many biophysical experimental techniques in current use to study isolated viruses are covered in Part II (Chaps. 3, 4, 5, 6, 7, 8 and 9). Several experimental techniques used to study viral structural components and complete virus particles inside the cell are described in Chap. 14. Several physics-based theoretical approaches to study the structure, dynamics and properties of viruses are covered in Chap. 19.

There are no easy recipes to decide which one(s) of those techniques a researcher should use for solving a specific problem in structural virology. The decision will depend, of course, on many scientific, technical and practical circumstances and considerations. Some options may be easily discarded, but several other options

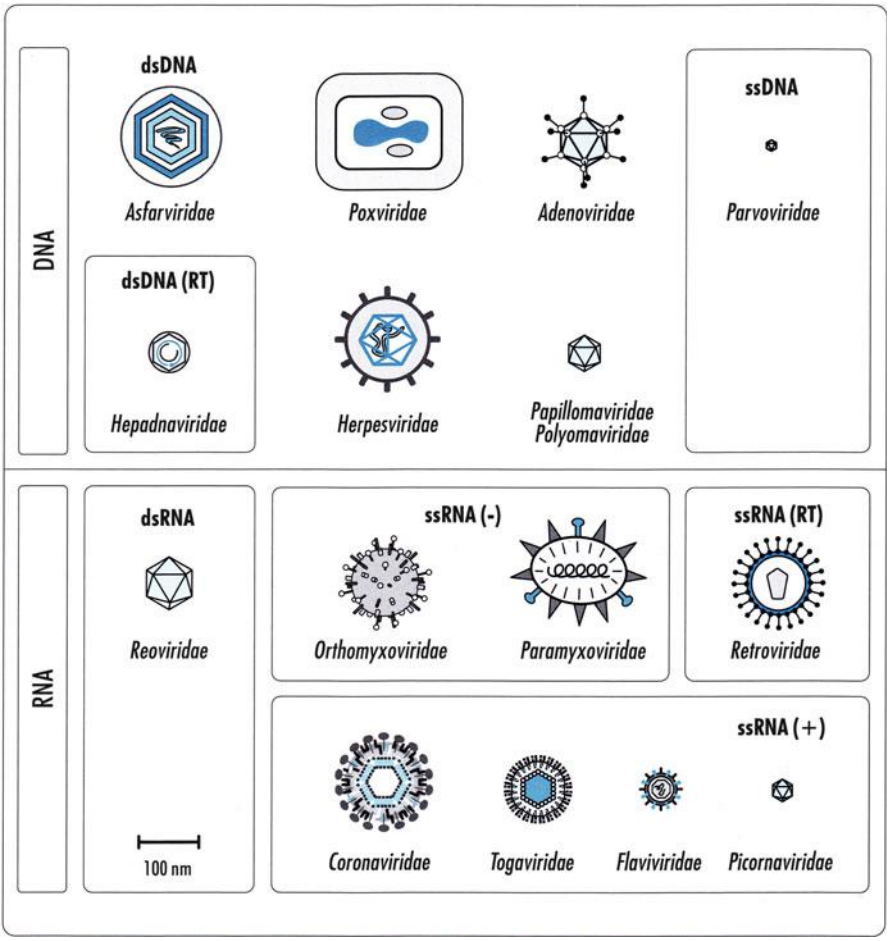

Fig. 1.3 General morphologies of virions belonging to some important animal virus families. The families are grouped according to viral DNA type (Figure reproduced from [Carrasco L, Almendral JM (ed) (2006) Virus Patógenos. Ed. Hélice and Fundación BBVA, Madrid], with permission from Ed. Hélice and Fundación BBVA)

may, in principle, be adequate for any intended goal, model virus and defined situation. Asking oneself a number of questions when considering a structural virology project should greatly help in defining the technical choices. A number of relevant considerations (including a few obvious ones) are listed next:

Some scientific considerations:

(i) What are the researcher's specific scientific goals?
(ii) What information on the chosen virus is already available?

(iii) Is the researcher interested in determining the structure of the complete virion, or only of some structural component(s)? Are the virus particles asymmetrical, or pleomorphic? Are they enveloped?
(iv) Is the researcher interested in studying some dynamic process in which the viral particle could be involved? Dynamic studies may be focused on fluctuations of the viral structure, or externally induced conformational rearrangements, including the structural characterization of intermediates of virus particle assembly or disassembly.
(v) Does the researcher contemplate the study of some aspect of virus structure or dynamics inside a host cell?
(vi) Is the researcher interested in characterizing some physicochemical property of the virus particle? Which one(s)? What for?
(vii) Is the researcher interested in determining the structure or properties of a virus particle or component bound to a ligand? Which one?
(viii) Is the researcher interested in trying to relate the structural information to be acquired on the viral particle or component or virus-ligand complex with some viral function? Which one(s)? Does the researcher aim at identifying the molecular mechanism involved?
(ix) Is the researcher interested in trying to relate the structural information to be acquired with the molecular biology of the virus? With pathogenic effects? With viral evolution? Any other aspect?
(x) Does the researcher intend to develop any particular application based on the structural information acquired? Which application?

Some technical considerations:

(xi) Can the virus in question be produced by infection of cultured cells? or can the viral CPs be produced in some expression system, and will they form viral capsids or virus-like particles (VLPs)? can isolated viral components of interest be obtained in a stable form?
(xii) Can the virions, capsids or VLPs be purified? Which level of purity must be reached? What amounts of sufficiently pure viral particles or components can be reasonably obtained?
(xiii) Do the viral particles or components aggregate at high concentrations? Can they be crystallized?
(xiv) Are the viral particles or components stable enough in the conditions required for structure determination by the technique(s) being considered?
(xv) Should the viral particles or components be chemically or genetically modified? Is there an infectious DNA clone of the virus genome available to allow site-directed genetic modifications?

A few practical considerations:

(xvi) Which structural virology expert(s), if any, could be contacted for possible collaboration? What structural facilities are available to the researcher and/or his collaborators?

(xvii) What about safety issues? Will containment facilities be available, if required? Permissions needed?

(xviii) How much time and what resources will be available for the project? What partial goals should be achieved along the project? In which order should they be tackled? By whom? How long will each goal possibly take? What are the alternatives if something fails or takes (much) more time than expected? These latter considerations are indeed common to all scientific projects, but the very complex nature of many structural techniques applied to study a large molecular assembly, and the time, effort and uncertainties involved, make such considerations particularly relevant in this case (and difficult to decide upon).

Given the very large number of considerations and uncertainties, it would be illusory to try and provide some general guidelines for choosing the "right" technique(s) for analysing virus structure, dynamics or physical properties. However, a brief comparison on the general applicability, strengths and limitations of each of the different structural techniques for different generic studies in structural and physical virology is possible and is included next. Relevant examples on the applications of the different structural and physical techniques to specific problems in virology are provided in most chapters of this book.

1.3.1 Techniques for Studying the Structure of Virus Particles or Their Components

Several powerful techniques are being used to determine the high-resolution structure of virions or their structural components. The detailed structures of many different virus particles have provided both critical answers and relevant new questions in molecular virology, and have led to a spectacular advance in our knowledge of viruses and biological processes, and our options to combat viral disease.

Transmission Electron Microscopy Using Negative Staining

Conventional *transmission electron microscopy* (EM) (Chap. 3) remains an all-important imaging technique, especially for initial structural characterization of isolated virus particles, and for many in-cell studies of virions or their components (Chap. 14). It generally requires very small amounts of virus particles that do not need to be very pure, and is extremely fast. The particle general morphology and architecture can be determined.

However, the sample must be fixed and covered (stained) with a heavy metal layer (opaque to electrons) to increase contrast. This treatment may introduce artifacts, including deformation of flexible particles, alterations of their fine

structure and blurring of their images. The resolution achieved on stained viral particles is, in general, relatively limited. In addition, only the surface topography of the particle can be imaged (see Chap. 3).

Cryo-Electron Microscopy

Cryo-electron microscopy (cryo-EM) (Chap. 3) is a variant of EM, and one of the most important techniques for imaging isolated virus particles and virus-ligand complexes. The sample is vitrified and kept at extremely low temperature; the virus particles remain hydrated, and no staining is generally required. The possibility of artifacts is greatly reduced compared to conventional EM. Even more important, cryo-EM is invariably used together with three-dimensional reconstruction techniques that average the images of thousands of virus particles visualized in different orientations, and result in a volumetric (three-dimensional) structural model of the imaged particles. Because the images of so many individual particles are averaged, different artifacts affecting different individual particles make very little contribution to the averaged model, and the signal-to-noise ratio is greatly increased. Also, as no staining agent needs to be used, the structural model obtained reveals also the anatomy (internal structure) of the viral particle. The typical resolution achieved on virus particles was, until relatively recently, about 1 nm or less, but in the last years cryo-EM models of some virus particles with a resolution around 0.5 nm, or even higher, have been obtained, thus approaching atomic resolution. In these high-resolution models, the tracing of the polypeptide chains in the viral capsids and other fine structural details, including the orientation of some amino acid side chains, can be observed (see Chap. 3). Cryo-EM has been repeatedly used also to obtain the structures of virus particle-ligand complexes which have not been generally crystallized and whose structure, thus, could not be directly determined by X-ray crystallography.

However, in cryo-EM thousands or tens of thousands of good-quality images must be taken, processed and analyzed; all things considered, a cryo-EM project may sometimes take months to be completed. Reaching a very high resolution is not guaranteed, especially with complete virus particles and other large biomolecular complexes. Because the process involves image averaging, the technique is best applied to symmetric viruses, and is not suitable for pleomorphic viruses (see Chaps. 3, 7).

Cryo-Electron Tomography

Cryo-electron tomography (Cryo-ET) (Chap. 3) is a variant of cryo-EM in which an individual microscopic object can be imaged using various angles of view. As a result, a three-dimensional reconstruction of the structure of a single virus particle (even inside a cell, see Chap. 14), and not a reconstruction based on many different particles, can be achieved. Because of this feature, cryo-ET is able to image pleomorphic virions such as influenza virus (Fig. 1.1, right), retroviruses and

many others, for which cryo-EM together with three-dimensional reconstruction cannot be used. A major disadvantage of cryo-ET is the much lower resolution achieved (say, about 3 nm) when compared to cryo-EM (see Chap. 3).

X-ray Crystallography

X-ray crystallography (Chap. 4) is a crucially important technique for the determination of the high-resolution structure of both virus particles and their isolated molecular components. Atomic or near-atomic resolution is generally achieved and there are, in principle, no limitations in the size or shape of the particle whose atomic structure is to be determined. The structures of very large and structurally complex virions and capsids, and of very large and asymmetric biomolecular complexes such as the ribosome have been already determined by X-ray crystallography. Molecular crystals contain a large proportion of water, the virus particles or proteins in them are fully hydrated, and the structures obtained are physiologically relevant.

A major disadvantage is that adequate molecular crystals of the virus particles or their components must be obtained, which is not always easy, or even achieved at all. For example filamentous, enveloped, or pleomorphic viruses will not crystallize; adequate crystals of flexible multimeric complexes between virus particles and large protein ligands such as cellular receptors or antibody fragments have very rarely been obtained. Also, relatively large amounts of very pure virus particles or proteins are most often required. Crystals must be well-ordered, of sufficient size, and stable enough under an intense X-ray beam to allow the collection of adequate data. The question of how to determine some necessary data (the *phase problem*, see Chap. 4) must be considered. In general, determination of the crystal structure of a virus particle, or even of a viral subassembly or individual protein is not an easy task and the project may take many months, or even several years.

Nuclear Magnetic Resonance (NMR) Spectroscopy

Solution NMR spectroscopy (Chap. 5) complements X-ray crystallography for the determination of the atomic or close to atomic resolution structure of molecular components of virus particles. The molecules remain in solution, generally in close to physiological conditions, and the uncertainties of the crystallization process are avoided. The major disadvantage of NMR for structural determination is a rather severe limitation in the size of the biomolecule whose structure can be solved. At present, molecules for structural determination by NMR must be no larger than about 35 kDa. However, by using state-of-the-art instrumentation and sophisticated technical approaches the structures of some larger proteins (50 kDa and more) have been solved. Unfortunately, direct structural determination of complete viral particles, most viral subassemblies and many multidomain or oligomeric proteins are still way beyond the reach of solution NMR spectroscopy. Large amounts of

purified sample and, usually, isotopic labelling of the sample (see Chap. 5) are required too, and the biomolecules must not aggregate during the slow acquisition of data at moderate temperature. *Solid-state NMR spectroscopy* has been applied for structural studies of a few viral particles (see Chap. 5).

Atomic Force Microscopy

In *atomic force microscopy* (AFM), small objects are not probed by electrons or photons as in EM, X-ray crystallography or NMR spectroscopy. Instead, the objects are probed by "touching" them with a very fine stylus (tip) that is used to trace the topography of the object (Chap. 8). AFM has been already applied to image a number of virus particles (see Chap. 8). Its greatest advantage as a static virus imaging technique is that the structure of single particles can be determined in liquid, in close to physiological conditions. Most importantly, AFM can also be used as a *dynamic* imaging technique, where structural changes can be followed in real time (see below). A clear disadvantage is that, like conventional EM but unlike cryo-EM, cryo-ET, X-ray crystallography and NMR spectroscopy, only the surface topography of the object (not its internal structure), can be determined. Because of basic geometric considerations regarding the tip-sample interaction, care has to be exercised in interpreting the images obtained (see Chap. 8). Also, flexible particles could be deformed by adsorption to the solid substrate. The resolution on virus particles is currently limited to about 1 nm at best. In special conditions (ultra-high vaccuum) with special instrumentation, atomic resolution has already been achieved using AFM on small organic molecules. See Chap. 8 for a detailed description of this *single-molecule technique* and its application to imaging viruses.

Combined Structural Approaches

Some limitations of X-ray crystallography, NMR spectroscopy and cryo-EM have been overcome by combining these techniques (see Chap. 7). For example, the atomic structures of a virus particle and its receptor molecule can be independently solved by X-ray crystallography, and the lower resolution structure of the virus particle-receptor complex can be determined by cryo-EM; by combining the structural models obtained, a high-resolution pseudo-atomic model of the complex can be obtained. In another example, if a complete viral capsid cannot be crystallized, the atomic structure of the isolated CP subunit can be solved by NMR spectroscopy or X-ray crystallography, and the lower-resolution structure of the complete capsid can be determined by cryo-EM; again, by combining the structural models obtained, a detailed pseudo-atomic model of the complete capsid can be constructed. The power of combining these and other structural, biophysical and biochemical techniques for structural virology studies is fully illustrated in Chap. 7, and can be also appreciated in examples described in other chapters of this book.

1.3.2 *Biophysical Techniques for Studying the Dynamics and Physical Properties of Virus Particles*

The high-resolution "photographs" obtained by applying the techniques just mentioned do provide invaluable structural information; however, there is a clear need to complement those static images with "movies" on the dynamics of virus particles using the highest possible spatial and temporal resolution. This is a difficult task, and the available techniques that have been applied so far to viral particles have generally provided relatively limited views of their dynamics. However the results already obtained have certainly established that virus particles are highly dynamic complexes, and provided abundant evidence on the critical importance of the local and global dynamics of virus particles and/or some of their components during the infection cycle (see Chaps. 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 and 19 for many specific examples).

Two types of dynamism in virus particles can be distinguished: equilibrium dynamics (*e.g.*, *breathing*), in which the structure of the virus particle in any particular state fluctuates around a minimum energy conformation; and reversible or (frequently) irreversible conformational transitions between two states of the viral particle, induced by external physical and/or chemical external factors *in vivo* and/or *in vitro* (see Sect. 1.4.4).

In the case of conformational rearrangements, the initial and final states and any reaction intermediates that happen to be stable enough may, in principle, be structurally characterized using the techniques already mentioned in Sect. 1.3.1 and described in detail in Chaps. 3, 4, 5, 7. In addition, other biophysical or biochemical techniques can be used to detect changes in secondary, tertiary and/or quaternary structure in the viral capsid. These techniques include *circular dichroism spectroscopy* and *fluorescence spectroscopy*, and these have been used also to probe in solution some particular aspects of the structure and properties of virus particles, including chemical and conformational stability (see Chap. 6). In some cases, unstable intermediates of dynamic processes involving virus particles have been stabilized by kinetic trapping using biochemical or genetic methods, and characterized using some of the structural techniques referred to above. A few transient intermediates of assembly/disassembly have been identified by using *mass spectrometry* (MS) techniques (Chap. 6). Small-angle X-ray scattering (SAXS) (Chap. 7) can be used to determine the shape of virus particles and contributes to structural determination in combination with other techniques; in addition, SAXS has started to be used to study virus dynamics including assembly and disassembly (see Chap. 7). *Differential scanning calorimetry* can also be used to follow conformational transitions in virus particles (see Chap. 7).

Single-molecule physical techniques such as AFM (Chap. 8) and *optical tweezers* (Chap. 9) have recently begun to be applied to study dynamic processes in viral particles. AFM offers the possibility not only to structurally characterize initial and final states and stable intermediates in dynamic processes, but to follow in real-time conformational rearrangements or the disassembly of viral particles, and to probe

some of their physical properties (see Chaps. 8, 18). Optical tweezers has been used to follow and quantify in real-time the dynamics of DNA packaging in bacteriophage capsids (see Chap. 9).

Equilibrium dynamics of a number of viral proteins has been probed by *NMR spectroscopy* (see Chap. 5). Global or local equilibrium dynamics (breathing) of complete virus particles has been probed by different biophysical or biochemical techniques including *NMR spectroscopy* (Chap. 5), *limited proteolysis/MS* (Chap. 6) and *hydrogen exchange/MS* (Chap. 6).

Physics-based models and computational simulations, including *molecular dynamics* (MD) and other approaches are being used to predict some aspects of virus particle dynamics. However, at present even the smallest virus particles are too complex for long enough all-atom MD simulations; thus, highly simplified (coarse-grained) structural models of simple viral capsids are used instead. Recently, a few short all-atom MD simulations of some aspect of the dynamics of very simple viral capsids have been carried out. These and other physics-based theoretical approaches to study virus particle dynamics, including conformational changes, assembly and disassembly are described in Chap. 19.

1.3.3 Techniques for Studying Structure-Function Relationships in Viral Particles

Once the molecular structure and/or conformational dynamics of a virus particle or structural component are known, a vast array of biophysical, biochemical, genetic, cell biology and other techniques can be used to investigate the relationship between viral structure and viral properties or function. An account of those techniques and how they can be used to study virus structure, properties and function are outside the scope of this book. However, some of these techniques and their uses in structural virology are briefly mentioned in Chaps. 7, 14. Examples of the results obtained in structure-function analyses of virus particles or their components can be found also in other chapters of this book (see especially Chaps. 10, 11, 12, 13, 14, 15, 16, 17 and 18).

1.4 The Roles of Virus Particles and Their Components Along the Virus Life Cycle

1.4.1 General Overview of the Virus Cycle

Viruses have evolved different alternatives for their multiplication and propagation. However, these alternatives share a number of general features and involve a number of broadly similar processes, albeit with multiple variations.

Several types of virus infection are known; which one may occur will depend on the virus species, host cell and conditions in which the cells are infected. In the case of animal viruses, many infections lead directly to the production of progeny virions in large numbers and the death of the host cell (*lytic infection*). In contrast, some infections involve the constant production of virus particles in small numbers, without killing the infected cells which are able to divide, and both the infected cells and the virus will persist in time (*persistent infection*). Other infections lead to the indefinite perpetuation in the host cell of the viral genome, which is replicated along with the cellular genome, and propagated by passing to the progeny cells (*latent infections*). In these infections, the viral genome may be propagated by becoming physically integrated in the cellular genome, as in the case of HIV-1, or by being maintained as an independent (*episomal*) nucleic acid molecule, as in the case of HSV-1. However, even latent infections can eventually lead to the production and release of virus particles that will be able to infect other host cells, closing the virus life cycle. Some bacteriophages can also produce both lytic and latent infections, which are viewed as evolved alternative strategies for virus survival; either alternative will be followed depending on the conditions the infected cell and the virus may encounter.

A generic, simplified virus life cycle is illustrated in Fig. 1.4. It must be emphasized again that many variations are possible depending on the virus species, host cell and environmental conditions. However, this simplified virus cycle may be useful in this book for putting into perspective the different processes in which virus particles and their molecular components participate during the infection process. The cyclic existence of a virus physical entity and its descent is traditionally contemplated in textbooks from the perspective of an infected host cell. Thus, a viral cycle is generally described by starting from a pre-existing, already formed virus particle which infects a cell; and ending with the progeny virus particles that are eventually produced in, and released from the infected cell (either directly in a lytic infection, or after a number of cellular generations have gone through a latent infection). From this classic perspective, the major generic steps in the life cycle of a virus (especially adapted for animal viruses, but mostly valid also for bacteriophages) can be listed in the following order, from step (1) to step (12) (see Fig. 1.4, where steps are identified using the same numbering):

Cell recognition by the virion and entry

(1) Recognition of receptor(s) molecule(s) (and eventually co-receptors) on the host cell surface.
(2) Internalization (entry) of the virus particle or of some of their components (invariably including the viral genome) into the host cell.
(3) Trafficking of the viral particle or some viral structural component (invariably including the viral genome) in the host cell.

Viral genome uncoating

(4) Release of the viral nucleic acid in some intracellular location where its replication will take place. Unless only the naked nucleic acid genome enters the cell (as in the case of many bacteriophages), release of the nucleic acid from

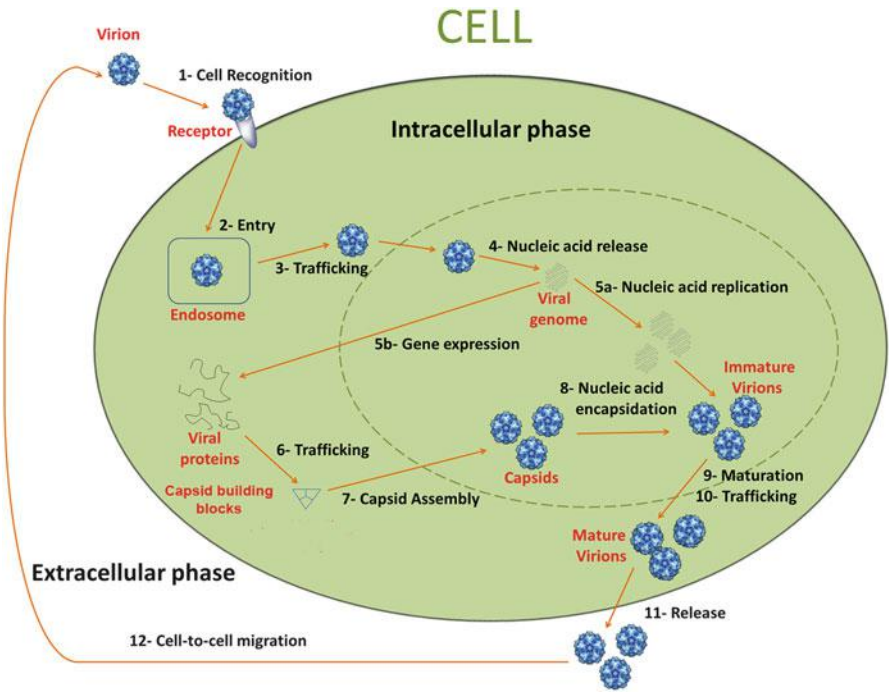

Fig. 1.4 A generic viral cycle. The following stages of the cycle are schematized: *1* Recognition of host cell receptor(s) on the host cell surface. *2* Entry of the virus particle (or of some components) into the host cell. *3* Trafficking of the viral particle (or some components) in the host cell. *4* Release of the nucleic acid. *5a* Replication of the viral genome. *5b* Expression of the viral genome. *6* Trafficking of the synthesized virus components to sites where the morphogenesis of virus particles will take place. *7* Assembly of the virus capsid from the previously synthesized CP subunits. *8* Packaging of the nucleic acid genome inside an assembled capsid (capsid assembly and nucleic acid packaging may occur concomitantly). *9* Maturation of the virus particle. Depending on the animal virus species, stages 4, 5a, 5b (in part), 6, 7, 8 and 9 may occur in either the cytoplasm or the nucleus of the host cell; this is indicated in the figure by encircling those steps by a discontinuous line, which represents the nuclear membrane only in those cases where the processes indicated occur in the cell nucleus. *10* Intracellular trafficking of the virus particles and/or mature virions, which may occur between, during and/or after steps 7, 8, and/or 9. *11* Release of virions from the cell. *12* Navigation of virus particles in the extracellular environment. See text for a more complete description and some important variations (Figure kindly provided by M.A. Fuertes)

the viral particle carrier (viral genome uncoating) must occur there. From this moment until the progeny virus particles are assembled, the virus fully enters the *eclipse phase*. The parental virus loses completely its identity as a molecular complex, and only its genome persist (and, eventually, also some of its viral proteins if required during the intracellular phase of the viral cycle).

Viral gene expression and genome replication

(5) Expression and replication of the viral genome. Unless a latent infection is established, *e.g.*, by integration of the viral genome into the cellular genome, the cellular metabolism is subverted to the virus advantage. Viral genes are transcribed into mRNAs in spatially and temporally controlled processes, and viral proteins required during the different processes of viral metabolism, and structural viral proteins required to form the progeny virus particles are synthesized. In addition, the viral nucleic acid genome is replicated by viral and/or cellular polymerases, following one of a vast array of mechanisms that will depend in part on the type of viral nucleic acid genome and virus species.
(6) Trafficking of the synthesized virus components to intracellular sites where the morphogenesis of virus particles will take place.

Virion morphogenesis

(7) Assembly of the virus capsid from the previously synthesized CP subunits.
(8) Packaging of the nucleic acid genome inside an assembled capsid. For many viruses, capsid assembly and nucleic acid packaging occur concomitantly in a single condensation process.
(9) Maturation of the virus particle to become an infectious virion. For some viruses, maturation occurs in response to nucleic acid packaging, and these processes cannot be clearly separated.

Virion release

(10) Intracellular trafficking of the progeny virus particles and/or mature infectious virions, which may occur between, during and/or after steps 7, 8, and/or 9.
(11) Release of virions from the cell. For some viruses, maturation may occur during, or even after release.

Extracellular existence of the virion

(12) Navigation of virus particles in the extracellular environment, within and between organisms. During this apparently passive phase, the virion is in fact confronted with, and must fight against, multiple external factors that can lead to its disruption or inactivation. These factors include many physical and chemical aggressions in the environment and multiple antiviral biochemical defenses mounted in the host organism: for example, virus neutralization by antibodies and/or other components of the immune system of an animal host.

This book is focused on virus particles, and a major part of it (Part III, Chaps. 10, 11, 12, 13, 14, 15, 16, 17, 18 and 19) deals with each of the major general stages in the virus cycle *in which virus particles or their molecular components participate* (*i.e.*, steps 1–4 and 7–12 above). Thus, we have chosen the "beginning" of the (in fact) endless virus cycle (Fig. 1.4) from the perspective of a virus particle, and not of an infected cell. We start our description after the viral components are synthesized in a host cell and when the viral particle begins its existence by self-assembly from some of those components (Chaps. 10, 11, 12, 13 and 14); and we proceed until the same viral particle ceases to exist by uncoating its nucleic acid

genome in another host cell (Chaps. 15, 16, 17). A logical advantage of this approach is that during morphogenesis of a virion critical features are created that will be needed later by that same virion, for example when it enters the cell or when its genome is uncoated. Examples of such features include metastability of many viruses or their components (Chaps. 10, 11, 12, 13, 15 and 16), or built-in pressure in tailed bacteriophages (Chaps. 9, 12, 17, 19).

1.4.2 Viral Metabolism: Viral Genome Expression and Replication

After a virion infects a host cell, it begins to lose its physical integrity, finally disappearing as such when its nucleic acid is released inside the cell. From this point until the progeny viruses are formed (if they are formed at all), the virus as a discrete physical entity ceases to exist. What we find instead is an infected cell that contains an additional set of (viral) genetic instructions. The viral genome and, in some viruses, a few viral proteins that have entered the host cell along with the genetic material may not only start a number of new metabolic events, but may also alter the cellular metabolic processes and the structural organization of the cell. Eventually, many copies of the viral genome and of the viral proteins required to form a new virus particle are made in the infected cell, and a number of progeny virions are assembled from them.

The diverse viral metabolic processes, such as viral genome replication, recombination, integration, transcription and translation and their regulation, and some of the virus-induced alterations of cellular components and reactions, are generally much better known than most stages in which viral particles participate, such as virus morphogenesis or disassembly and uncoating. Viral metabolism, including some structural and many functional aspects, constitutes the major part of most molecular virology textbooks and is out of the scope of the present textbook, which focuses on the structural biology and physics of virus particles instead. Thus, only a few very general basic facts on viral metabolism will be sketched here. For detailed accounts on viral metabolism the reader is referred at the end of this chapter to a few excellent molecular virology textbooks [1–8], and to the specialized reviews and references contained in them.

Partly because of the different types of nucleic acid molecules that can act as viral genomes, different viruses follow a bewildering diversity of pathways and mechanisms for genome expression and replication. Moreover, genetic information in viral genomes is stored in a remarkably compact ("streamlined") configuration, probably because of size limitations in the viral capsid and other constraints. This fact also contributes to the notorious complexity observed in the regulation of viral gene expression in many viruses.

A central requisite for viral gene expression (as for cellular gene expression) is the synthesis of mRNA molecules (*genome transcription*). The dsDNA genomes of

herpesviruses, adenoviruses, or poliomaviruses (to cite some examples) are replicated by the action of a DNA polymerase in the cell nucleus, and their genes are transcribed into mRNAs by a RNA polymerase much like the genes in cellular chromosomes. In hepadnaviruses, replication of the dsDNA genome involves a RNA intermediate and its reverse-transcription by a RT. The ssDNA genomes of parvoviruses are copied by a DNA polymerase in the nucleus to yield dsDNA molecules, which are transcribed into mRNAs, and also used as intermediates to yield multiple progeny ssDNA genomes that are encapsidated during virus morphogenesis in the nucleus.

ssRNA viruses are extremely abundant and also diverse. In order to be transcribed into mRNAs, viral ssRNA(+) genomes, such as those of picornaviruses or coronaviruses, are replicated in the cytoplasm using a viral RNA replicase to yield complementary ssRNA(−) molecules, which are in turn replicated/transcribed to yield complementary ssRNA(+) molecules that may act as viral genomes and/or mRNAs. Viral ssRNA(−) genomes, such as those of orthomyxoviruses (influenza virus) or paramyxoviruses, are transcribed in the cytoplasm to yield complementary ssRNA(+) molecules that may act as mRNAs and/or are in turn replicated to yield ssRNA(−) molecules that act as viral genomes. In retroviruses, the ssRNA(+) genome is reverse-transcribed by a viral RT into a complementary ssDNA(−) molecule, which is then used as template to form a viral dsDNA molecule which may be integrated in the cellular genome and/or transcribed to yield viral mRNAs. The dsRNA genomes of reoviruses are directly transcribed into mRNAs and replicated in the cytoplasm.

The transcription of viral genes into viral mRNAs inside the infected cell is a carefully regulated process, both in space and time. In many but not all viruses, during a lytic infection some viral genes are preferentially transcribed early in the infection process, and the corresponding mRNAs are translated in the cell ribosomes to yield *early viral proteins*, which will be generally needed to subvert the cell metabolism or control later stages of viral gene expression. At some point, replication of the viral genome starts and genomic nucleic acid molecules accumulate in some location of the infected cell. In later stages, other viral mRNAs are translated to yield *late viral proteins*. These are generally synthesized in much larger amounts than early viral proteins, and usually include, among others, the *structural viral proteins*, those that will form a part of the newly assembled virus particles during virion morphogenesis, the next stage of the viral cycle.

1.4.3 Virion Morphogenesis

Virion morphogenesis is a very complex and tightly regulated process and, except for a few viruses, is still poorly understood in many aspects. Several different stages that sometimes overlap can be distinguished: capsid assembly (Chaps. 10, 11), nucleic acid packaging (Chap. 12) and virus particle maturation (Chap. 13).

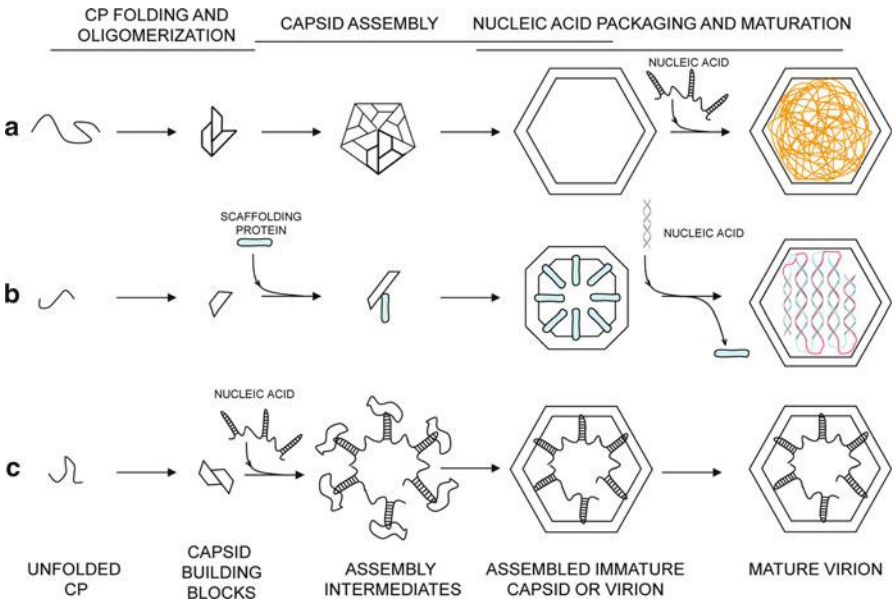

Fig. 1.5 Simplified schemes of three general strategies for assembly of virus capsids. (**a**) Unassisted self-assembly. (**b**) Scaffolding protein-assisted assembly. (**c**) Viral nucleic acid-assisted assembly (see text for details) (Figure kindly provided by M.A. Fuertes and reproduced from [Mateu MG (2013) Arch Biochem Biophys 531, 65–79]. With permission)

Capsid Assembly

Assembly of a virus capsid is a complex oligomerization (some would say multimerization) process that proceeds along a pathway regulated by ordered interactions between the participating macromolecules. The structural complexity of these large oligomers (or multimers) usually requires intricate assembly pathways mediated by allosteric switches, irreversible steps, and the assistance of other molecules including scaffolding proteins or the viral nucleic acid. Three different general strategies for capsid assembly can be recognized (Fig. 1.5): (i) Capsid self-assembly, which strictly requires only CP subunits that self-associate under appropriate conditions (Fig. 1.5a); in some cases, some factors may be required to provide the right conditions for self-assembly. (ii) Scaffolding protein-assisted capsid assembly, which invariably requires not only the CP subunits, but also the assistance of scaffolding proteins to produce a precursor, immature capsid (procapsid) (Fig. 1.5b); (iii) Viral nucleic acid-assisted capsid assembly, which requires the simultaneous interaction of the CP subunits and the viral nucleic acid in a condensation process in which capsid assembly and nucleic acid packaging occur simultaneously (Fig. 1.5c). In this book, and from a morphogenetic perspective only, we have arbitrarily considered as *simple viruses* those that do not require scaffolding proteins for capsid assembly (Chap. 10); and as *complex viruses* those that do require scaffolding proteins and, thus, will follow a more complex capsid assembly pathway (Chap. 11).

Virus capsids are generally assembled from soluble, stable capsid building blocks (CBB) (Fig. 1.5). Depending on the virus, CBBs may be folded CP monomers or, frequently, small oligomers that are formed by association of CPs in a previous step; stable CBBs can be, thus, indistinctly regarded as preformed starting substrates for the capsid assembly reaction itself, or as early intermediates in the complete morphogenetic process from unfolded CPs to native virions (see Chaps. 10, 11). *In vitro* or in the cell, the CBBs may be found in an assembly-incompetent state (depending on virus and conditions), and some environmental change or additional factor may be required to activate them and/or to assist in their association.

Capsid self-assembly. Several important aspects of the thermodynamics and kinetics of capsid self-assembly (Fig. 1.5a) have been experimentally determined using *in vitro* systems; in these systems, simple virus capsids (or virions in a few cases) are assembled from its CBBs under close-to-physiological or, sometimes, non-physiological conditions. Unfortunately, the assembly pathways from CBBs to simple icosahedral capsids have proved to be very difficult to trace experimentally, as populated (stable) intermediates (not counting the initial CBBs) are rarely observed, and transient intermediates can be difficult to detect. Despite these difficulties, scarcely populated, transient assembly intermediates have been identified and characterized during assembly or disassembly of simple capsids, for example using electrospray ionization-MS or ion-mobility-MS. In addition, several stable intermediates have been identified during assembly of a few simple viruses, for example in picornaviruses (see Chap. 10). In contrast, self-assembly and subsequent maturation of more complex icosahedral capsids frequently involve a number of intermediates, including fairly stable intermediates populated at equilibrium and/or transient intermediates, some of which have been identified using different structural or dynamic techniques (see Chaps. 11, 13).

A number of simplified theoretical models and computational simulations on the thermodynamics and/or kinetics of virus capsid self-assembly have been developed (Chap. 19), and contrasted with the results of *in vitro* experiments in which simple virus capsids are self-assembled from their CBBs. In turn, results obtained in these experiments have been used as constraints to develop improved or novel theoretical models. It must be noted that, in general, the theoretical models and simulations are limited to the self-association of CPs or CBBs (although some recent coarse-grained simulations already include the viral nucleic acid). Thus, from a physiological perspective the results thus obtained may be most relevant in those cases in which the capsid self-assembles *in vivo* without the assistance of scaffolding proteins or the viral nucleic acid. However, theoretical and experimental capsid-only approaches may also contribute to (qualitatively) understand the assembly of capsids that, *in vivo*, require scaffolding proteins or the viral nucleic acid. This possibility is supported by the fact that some of these capsids can be correctly assembled *in vitro* without the assistance of any other macromolecules.

A high-order reaction in which 12 pentameric, 20 trimeric or 30 dimeric CBBs collide simultaneously and in the right orientations to form an icosahedral capsid is

unlikely. In fact, theoretical and experimental analyses support the view that simple icosahedral virus capsids can self-assemble from CBBs through a cascade of second-order reactions. Those analyses have also revealed other rather general features: (i) Self-assembly may be thermodynamically considered as two-state: only the dissociated state (soluble CBBs) is populated below a certain CP concentration, and only the fully associated state (complete capsids) is populated above that concentration. (ii) Capsid assembly kinetics can be represented by a sigmoidal curve and includes a lag phase. (iii) The reaction rate is strongly dependent on protein concentration, and this and the previous features suggest that capsid assembly proceeds through nucleation and growth. (iv) At high protein concentrations, free CBBs disappear but only partially assembled capsids are produced. (v) There is hysteresis to dissociation: the capsid disassembles at much lower CP concentrations than those required for assembly. (vi) Off-pathway reactions may occur, leading to aberrant particles, capsids with non-native quaternary structure or polymorphisms. See Chap. 19 for detailed descriptions of theoretical models on capsid self-assembly.

Scaffolding protein-assisted capsid assembly. The proper assembly of many icosahedral virus capsids *in vivo*, and even *in vitro*, requires the assistance of scaffolding proteins (Fig. 1.5b; see Chap. 11). Scaffolding proteins establish specific but transient protein-protein interactions with the CP subunits (CBBs) during the assembly process and are later removed, being absent in the mature virus particle. Thus, scaffolding proteins can be regarded as assembly chaperones. Most, but not all, capsids that require scaffolding proteins for proper assembly have a complex quaternary structure. In some cases the scaffold may self-assemble in the absence of CPs, and is subsequently used as a template for capsid assembly by recruiting CP subunits. In many cases the internal scaffold and the capsid are formed in a co-assembly process where the scaffolding protein subunits promote CP-CP interactions, and the CP subunits promote in turn interactions between the scaffolding protein subunits by eliciting conformational changes in them. Once a precursor, or immature capsid (*procapsid*) is assembled, the scaffolding proteins are removed, either before or during packaging of the viral nucleic acid. In some cases the scaffolding proteins subunits are extruded through openings in the capsid and recycled, thus acting as assembly catalysts. In other cases they are degraded by an encapsidated viral protease.

The use of scaffolding proteins for viral capsid assembly offer additional possibilities for controlling the process. Several functional roles for the scaffolding proteins of different viruses have been identified or suggested: (i) Initiation of capsid assembly by nucleation. (ii) Promotion of capsid assembly by increasing the effective concentration of CPs and/or lowering the energy barrier of a conformational transition that activates the CP or CBB. (iii) Directing assembly; the size of the scaffold would determine capsid size by preventing incorrect CP-CP interactions, and the scaffolding protein-CP interactions would promote the proper conformational switches in the CP subunits being incorporated. (iv) Mediating the

incorporation of other viral structural proteins and subassemblies to the capsid. (v) Preventing the capsid from being filled with some nonspecific proteins which would be difficult to remove. (vi) Stabilization of the assembled procapsid. It must be noted here that other viral proteins or protein assemblies in addition to scaffolding proteins can assist the assembly of some virus capsids, irrespective of its structural simplicity. See Chaps. 10, 11 for detailed descriptions of capsid assembly in different simple or complex viruses.

Nucleic Acid Packaging

Viruses have evolved two general strategies for packaging their nucleic acid genome in the capsid (Fig. 1.5): condensation of the CPs or CBBs and the viral nucleic acid, or insertion of the nucleic acid in a preformed empty capsid (Chap. 12).

Nucleic acid-assisted capsid assembly and genome packaging. In many viruses the viral nucleic acid is recruited to assist the assembly process, to directly yield nucleic acid-containing virions in a combined assembly-packaging process (Fig. 1.5c). In these cases, no empty capsid intermediates are normally formed during virus particle assembly *in vivo*. The experimental observations suggest that, for many ssRNA viruses, the CP subunits (or CBBs) may bind the RNA and influence the acquisition of a defined tertiary structure by the latter, mainly composed of secondary structure elements but not showing a highly compacted, unique fold. In turn, the folded RNA can promote CP oligomerization and influence the geometry of protein-protein interactions between CP subunits to direct the formation of the structurally correct particle. See Chap. 12 for a detailed description of viral capsid and ssRNA co-assembly.

Viral nucleic acid packaging into a preformed capsid. In many other viruses a nucleic-acid-free capsid is assembled first, either without or with the assistance of scaffolding proteins and/or other auxiliary proteins; the viral nucleic acid genome is later packaged into the preformed capsid (Fig. 1.5a, b). This strategy is followed by many dsDNA viruses and dsRNA viruses (see Chap. 12) and also by ssDNA viruses (parvoviruses; see Chap. 10). dsDNA packaging in tailed phages has been extensively studied in exquisite detail. In essence, it occurs as an ATP-driven reaction which involves a molecular motor and a portal protein complex located at one of the vertices of the icosahedron-based capsid, through which the DNA molecule is inserted. See Chap. 12 for detailed descriptions of active dsDNA and dsRNA packaging into preformed viral capsids.

Virus Particle Maturation

Many virus particles are assembled in a non-infectious form. Maturation refers to a generally irreversible reaction, or series of reactions, that convert a non-infectious (immature) virus particle into an infectious (mature) virion (Chap. 13). Maturation reactions may occur before, during and/or after nucleic acid encapsidation. In addition to activation of the virion, a frequent (albeit not universal) effect of

maturation is the physical stabilization of the capsid. Proper assembly of many viral capsids may require weak association energies (see Chap. 19), and the resulting capsids may be thermodynamically unstable. The observed hysteresis to dissociation may be enough, in some cases, to prevent capsid disassembly by extreme dilution during the extracellular phase. However, irreversible post-assembly reactions may have evolved as an additional, or alternative, mechanism to stabilize many virions. Irreversible maturation frequently blocks the capsid in a rather stable (actually metastable, or "spring loaded") state through covalent modification and/or conformational rearrangements, which sometimes involve also the formation of stabilizing covalent bonds and/or binding of additional, "cementing" capsid proteins. Maturation of some other virus particles, for example in retroviruses, may actually destabilize them (see Chap. 18). In considering maturation or any other stage of the viral cycle, it should be taken into account the different selective constraints that may be acting of different viruses depending on their structure and "lifestyle" (*e.g.*, genome protection in enveloped *vs.* non-enveloped viruses, entry of animal viruses *vs.* nucleic acid injection in bacteriophages, etc.).

Probably because of different selective constraints, maturation does take several forms in different virus families and even species. In several families of small, nonenveloped ssRNA viruses, maturation involves a proteolytic cleavage of some capsid subunits. This reaction frequently results in the release of a peptide required for interaction of the virion with cellular membranes during the infection process. In tailed dsDNA phages, maturation overlaps with DNA packaging and is a complex irreversible process that involves capsid conformational changes and achieves different functional goals: removal of the no longer needed scaffolding proteins, capsid expansion, creation of new binding sites for the attachment of additional structural proteins and subassemblies required for virion function, and capsid stabilization. As for simpler icosahedral capsids, the association energy between CPs in tailed phage capsids is low and the procapsid (*prohead*) is rather unstable. The viral dsDNA is packaged into the capsid to a very high (crystalline) density, which results in the capsid becoming pressurized. Capsid pressurization is thought to be required for assisting injection of the viral genome into the infected host bacterium, but will tend to disrupt the unstable procapsid (see Chaps. 9, 17, 18, 19). Thus, different dsDNA phages have evolved different specific strategies for capsid stabilization during and after maturation. See Chap. 13 for a detailed description of these and other maturation strategies and mechanisms in different viruses. Examples of maturation in some other viruses can be found also in a few other chapters in this book (*e.g.*, Chap. 10). Factors contributing to the stability of an already mature virion are briefly described in Sect. 1.4.4.

Host Cell Factors Involved in Virus Morphogenesis

Virion morphogenesis in the infected cell is a very complex process. The cell interior is a molecularly crowded environment, where many competing reactions take place. Macromolecular crowding has been shown to contribute to virus capsid assembly, but it can also promote the formation of aberrant assembly products.

Moreover, replication of the viral genome, capsid assembly, nucleic acid packaging and virus maturation must not only occur, but need to be strictly coordinated in space and time in order to produce an infectious virion. As a result, virion morphogenesis in the host cell generally requires auxiliary viral or cellular factors, including chaperones, and occur in *virus factories*, very complex cellular structures that are formed inside infected cells. Virus factories can be regarded as huge dynamic scaffolds that assist virion replication and assembly; they are currently being studied, for example, by electron microscopy and electron tomography of infected cells. Chapter 14 describes structural, molecular and cell biology techniques that are being used to identify host cell factors that may participate in virus assembly, or for studies on virus morphogenesis in the infected cell; the same chapter describes also some important results obtained on this complex and still poorly known process.

1.4.4 The Virion During the Extracellular Phase

Virus Stability

The many physical and chemical aggressions a virion has to withstand outside cells and organisms have led to multiple structural solutions for virion stabilization; these solutions, however, must not interfere with the capacity of the virion to release its genome into a host cell. Some of the stabilization strategies that different viruses have evolved are summarized below. Studies on the structural determinants of the stability of virus particles are providing further clues on the biological success of viruses; they are also leading to the identification of novel targets for antiviral intervention, and to explore possibilities for the engineering of viral particles of increased stability for bio/nanotechnological applications.

Noncovalent interactions between capsid subunits. Interfaces between CPs or CBBs can differ widely between capsids, or even within a capsid: They can be small or quite large; discontinuous or continuous; quite planar or convoluted; independent of, or overlapping with other CP-CP interfaces; including a hydrophobic core and polar rim, or essentially polar; etc. (see Chap. 10 for a few examples). In general, mutational analyses have identified many specific structural elements and residues in the CPs that are relevant for capsid assembly and/or stability in different viruses. The assembly and/or stability of viral capsids may be critically altered by single amino acid substitutions; in addition to the hydrophobic effect (as estimated by buried surface area), buried electrostatic interactions (hydrogen bonds and charge-charge attractive interactions) may be most important for assembly of a stable capsid in some viruses. Conversely, charge-charge repulsions between CP subunits have been clearly involved in the destabilization of very different virus capsids. Electrostatic repulsions between CP subunits in the capsid can be modulated by protonation/deprotonation both *in vitro* and *in vivo*, and may

contribute to the capsid stability-instability balance and the control of biologically relevant capsid conformational transitions.

Covalent bonds between capsid subunits. Disulfide bonds or other covalent bonds are formed in the capsids of very different viruses; in many cases, they have been shown to transiently or permanently contribute to capsid assembly and/or stability. The reversibility of disulfide bonding (bonds formed in the oxidizing extracellular environment, but likely disrupted in the reducing environment inside the cell) may provide a convenient strategy to meet the conflicting requirements between extracellular stability *versus* intracellular lability for some viruses.

Cementing proteins. The interactions established between scaffolding protein subunits, and between them and the CP subunits, can help to transiently stabilize some immature capsids before they become irreversibly stabilized by maturation. The attachment of bridging or "cementing" proteins during capsid maturation constitutes an evolutionary strategy of some viruses to increase capsid stability by providing additional intersubunit interactions.

Capsid-viral nucleic acid interactions. In many icosahedral ssRNA viruses, the N-terminal and/or C-terminal segments of the CPs are positively charged and normally located in the capsid interior, where they may establish ionic interactions with the negatively charged phosphates of the nucleic acid. In some viruses a fair number of positive residues are located at the capsid inner wall. Other virions contain polyamines. All of these alternatives can neutralize a substantial fraction of the RNA negative charge, and contribute significantly to the nucleic acid-assisted assembly and/or virion stability. In several ssRNA plant and animal viruses and a few ssDNA viruses (parvoviruses), short stretches of the viral nucleic acid molecule conform to the icosahedral symmetry of the capsid and can be directly visualized by X-ray crystallography of the virions. Non-covalent interactions between structured regions at the capsid inner wall and nucleic acid stretches may have a role in RNA-assisted capsid assembly or DNA packaging into a preformed capsid, and may also stabilize the assembled viral particles against a conformational change or dissociation.

Hysteresis to dissociation. For nonenveloped viruses whose capsids are not irreversibly stabilized after assembly through maturation, inbuilt hysteresis to dissociation may be essential to prevent disassembly in the extracellular environment by extreme dilution.

Entropic stabilization. Increased capsid breathing reduces the difference in conformational entropy between the dissociated and associated states. Thus, breathing may help to stabilize the assembled capsid.

Stabilizing ligands. Some ligands (metal ions, antibodies, etc.) may increase the stability of virus particles. For example, small hydrophobic molecules (*pocket factors*) are able to bind hydrophobic pockets in the capsids of some picornaviruses (poliovirus, PV and human rhinovirus, HRV) and can entropically stabilize the virion against conformational changes, inhibiting genome uncoating (see also below).

Covalent attachment of functional groups. Some irreversible enzymatic reactions on viral capsids during maturation stabilize the viral particle, and chemical modifications of specific capsid residues may have different roles in assembly and/or stability.

Mechanical stability of virus particles. Nearly every study on the molecular determinants of viral capsid stability so far has focused on analysis of the resistance of the virion against inactivation of its infectivity (as an indication of structural alterations), or of the viral particle against chemical modification (including proteolytic cleavage), conformational rearrangement or dissociation into subunits by the action of heat or (bio)chemical agents. However, viruses may be also subjected to severe mechanical stress both outside and inside cells, and must be robust enough to withstand the mechanical forces acting on them. Moreover, viruses may have evolved not only to withstand, but also to use those mechanical forces. Very recently, the mechanical stability of virus particles and other aspects of virus mechanics, and their possible biological consequences, have begun to be investigated using AFM (Chap. 18) and physics-based theoretical approaches (Chap. 19).

Virus Dynamics

Virions must achieve enough stability in the extracellular environment without compromising viral genome uncoating in the host cell. This conflict between stability and instability has generally been met by many viruses by evolving a virion that is relatively robust, but conformationally dynamic and metastable.

Breathing. Biochemical and biophysical studies of non-enveloped virions and capsids in solution reveal that that they are not static at equilibrium, but constitute dynamic assemblies which sample a range of similar conformations around a minimum free energy state. This so-called *breathing* tends to transiently increase exposure to solvent of some structural elements in the particle, including functional CP peptide segments, and these events may be biologically required during some steps of the viral cycle. Capsid breathing can be modulated by internal components or capsid-bound external molecules. In the PV and HRV virions a large hydrophobic cavity (pocket) is involved both in breathing, and in facilitating a transition of the particle from a metastable conformation to another conformation during cell receptor recognition and nucleic acid uncoating. Binding of antiviral agents to the hydrophobic pocket (see Chap. 20), or of a virus-neutralising antibody to a region nearby, or mutations in the capsid that may partly fill the pocket were shown to impair viral breathing. The same factors also impair the transition between the metastable and final states of the capsid. The above and other results indicate that global or local conformational dynamics (breathing) may be a general feature in virus particles, and may be required for infection by facilitating structural transitions. Very recent AFM studies have revealed that different regions in viral particles can differ in mechanical stiffness, and indicate that local regions in the particle with high mechanical flexibility may also present high conformational dynamism which may be biologically relevant (see Chap. 18).

Conformational rearrangements in mature virions. Excess stability and rigidity may lead to a biologically inert virion, unable to replicate in a host cell. In addition to conformational dynamism at equilibrium, metastable virions respond to the appropriate stimuli by undergoing controlled conformational transitions required for infectivity. In non-enveloped viruses, the entire viral particle may be conformationally rearranged, although the greatest alterations may sometimes be localized (Chap. 15). In enveloped viruses, a particular structural component (*e.g.*, an envelope glycoprotein on the surface) may change its conformation (Chap. 16). The transitions must be controlled through the action of regulatory factors, to ensure they occur at the appropriate location and time. The same conformational rearrangements caused by physiological factors can be frequently induced *in vitro*, even under non-physiological conditions using different agents (*e.g.*, changes in temperature or pH). These facts have facilitated the *ex vivo* or *in vitro* studies of capsid conformational rearrangements using a vast array of techniques. In some cases, the biological relevance of the transitions observed is not yet clear. In other cases, the identified capsid rearrangements have been shown to be biologically required. These structural changes may be required, for example, for exit of the virion from the cell compartment or the cell where it was formed, entry into another cell, intracellular transport, or uncoating of the viral genome. Specific examples of conformational rearrangements of viral particles or some of their components, and their biological implications, are described in many chapters of this book including Chaps. 10, 11, 12, 13, 15, 16, 17, 18 and 19.

To summarize, the usual end product of virus morphogenesis is a robust enough macromolecular complex able to effectively protect the viral genome during its extracellular existence. However, this complex is also a flexible, dynamic, breathing, metastable particle poised to undergo controlled conformational transitions required to perform biologically critical functions during different stages of the virus life cycle, including entry into cells, intracellular trafficking and viral genome uncoating.

Virion Neutralization by Antibodies and Virus Escape from Neutralization

An extracellular virion must not only be stable enough to withstand physicochemical aggressions without compromising its function; it must also be able to evade the host defenses without compromising virion stability and function. Higher animals, especially mammals, fight viral infection by mounting a complex molecular and cellular immune response (see Chap. 21 for an outline). Antibodies constitute a major part of this immunological mechanism of defense against many viruses. A number of structural studies have focused in the recognition of virions or some of their components by antibodies, the mechanisms by which antibodies neutralize virion infectivity, and the structural strategies viruses follow to escape from antibody recognition. Structural descriptions of complexes between virions or their components and antibodies can be found in Chap. 14 of ref. [15] and Chap. 9 of ref. [17] provides a review on the relationships between the structural basis of virus-antibody recognition and virus escape from neutralization, focusing on foot-and-mouth disease virus (FMDV) as a model system for small, nonenveloped viruses.

From an applied perspective, the information obtained on virus-antibody recognition and virus escape has proven to be critical for the development of improved and novel vaccines including VLPs and chimeric viral particles (Chap. 21)

Virus-antibody recognition. The structures of complexes between individual viral proteins (*e.g.*, envelope proteins of enveloped viruses) and monovalent antigen-binding antibody fragments (Fabs) are generally determined by X-ray crystallography (Chap. 4). However, complexes between non-enveloped virions or capsids and Fabs do not readily crystallize, and combined approaches based in X-ray crystallography of virus particle and Fab separately, and cryo-EM of the virus particle-Fab complex (at lower resolution) are generally used to obtain quasi-atomic models of the ensembles (Chap. 7). In some respects, virus-antibody recognition is structurally similar to virus-receptor recognition and other virus-protein ligand interactions. The reader is referred to Chap. 15 for general structural principles of virus-receptor recognition, some of which can also be applied to virus-antibody interaction (but see below).

In enveloped viruses (such as influenza virus or HIV-1), some antibodies recognise and bind strongly to specific patches (*neutralization epitopes*) on the surface of the envelope proteins. In non-enveloped viruses, some antibodies bind to neutralization epitopes on the capsid surface. In the case of some small non-enveloped viruses (*e.g.*, PV and some types of HRV), a remarkable difference between receptor and antibody recognition is that the receptor-binding site lies in a *canyon*, a depression of the virus surface, while antibodies cannot penetrate very deep into the depression and bind preferentially to capsid protrusions. This is not the case with other similar viruses, including FMDV, whose receptor-binding site is located in a fully exposed, protruding and mobile protein loop (termed the *major antigenic loop*) which also binds many neutralizing antibodies.

On a viral capsid, epitopes may be formed by residues in just one loop of a CP subunit (*continuous epitopes*; *e.g.*, most epitopes in the *major antigenic loop* of some FMDV types), or several segments of a same CP subunit or of several neighboring CP subunits or CBBs (*discontinuous epitopes*; *e.g.*, epitopes identified in other, equally immunodominant antigenic regions in FMDV). If segments from different CBBs are involved, the epitope will be formed only after some intermediate or the complete capsid has been assembled; some will form only after the capsid has been structurally reorganized during virus maturation (Chap. 13). Antibodies against these epitopes can be, thus, used as probes for studying capsid dynamics. It must be noted here that even if the antibody-binding residues are all located in one protein loop, not every residue in a peptide stretch may be directly involved in binding and, strictly speaking, nearly all epitopes are discontinuous. In addition, all viral antigenic epitopes are *conformation-dependent* as, like any other protein ligand, they must adopt a defined conformation either before or, more rarely, on binding the antibody.

About 15–20 residues in the virus capsid or envelope protein may form the *contact epitope*, and interact with a similar number of residues in the antibody *paratope*, contributing to binding strength and specificity through the hydrophobic effect and multiple non-covalent interactions (van der Waals, hydrogen bonds and

charge-charge interactions). However, in many viral epitopes very few residues in the contact epitope, those that form the *energetic epitope*, contribute most of the binding energy. In addition, some residues that may not form a part of the contact epitope may indirectly contribute to binding because, if substituted by mutation, the epitope conformation may be altered, preventing antibody binding. The residues whose mutation affects antibody recognition, located either within the contact epitope or elsewhere in the viral particle, may be considered a part of the *functional epitope*.

Virus neutralization by antibodies. Once bound to a capsid (non-enveloped viruses) or envelope protein (enveloped viruses), antibodies can neutralize virions by one of several different mechanisms. A frequent mechanism is interference with cell receptor or co-receptor recognition. For example, antibodies that bind the major antigenic loop of FMDV overlap extensively with the receptor binding site and sterically interfere with receptor binding (see Chap. 15; [17]). Other mechanisms involve impairment of a conformational rearrangement of a protein or of the virion that would be required at some stage of the viral cycle. For example, some antibodies prevent a conformational transition of the viral fusion protein that is required during entry of enveloped viruses (*e.g.*, HIV-1) into the host cell (see Chap. 16). Some antibodies may exert a neutralizing effect by binding bivalently to the capsid, increasing capsid stability and probably impairing uncoating; they may even act intracellularly after the virus-receptor antibody has been internalized. Other antibodies may exert their neutralizing effect by facilitating virus elimination through *opsonization*. Some mechanisms of antibody-mediated neutralization of viruses have not been elucidated or well characterized yet.

Virus escape from neutralization by antibodies. In general, virus escape occurs through negative selection by circulating antibodies in the organism; it may occur also by genetic drift. RNA virus populations are *quasispecies*, a term that refers to their extreme genetic heterogeneity, being formed by multiple genetic variants in different and sometimes rapidly varying proportions [18]. This generally leads to the frequent emergence of virus variants differing in some phenotypic trait, usually but not necessarily when subjected to the appropriate selection pressure [18]. As a consequence, drug-resistant variants (see Chap. 20) or antibody-resistant variants, especially in the case of RNA viruses (such as FMDV or HIV-1, among many others) will appear most readily; this phenomenon severely hampers viral disease control.

In some RNA viruses, such as influenza A virus, drastic antigenic changes leading to virus escape from antibody recognition may occur abruptly (*antigenic shift*) by genetic reassortment that substitutes at once a glycoprotein type exposed on the viral envelope by an antigenically very different type. This sudden antigenic change poses particularly severe problems for disease control. In addition, in these and many other viruses antigenic variation may occur more gradually (*antigenic drift*), through the accumulation of amino acid substitutions in epitopes located in several antigenic regions of the envelope proteins (in enveloped viruses) or of the viral capsid (in non-enveloped viruses). These amino acid substitutions generally

eliminate some virus-antibody interactions, sterically interfere with the antibody recognition process, and/or distort the epitope conformation.

In PV, some HRV types and other viruses depressions may have been selectively favored as receptor-binding sites (see above) to allow for partial hiding from antibody recognition, thus preserving to some extent receptor binding under antibody pressure (see Chap. 15). However, a clear-cut structural separation between receptor-binding sites (or other functional regions), and antibody-binding sites does not generally occur in small non-enveloped viruses. For example, some antibodies that can neutralize HRV recognise epitopes that include part of the canyon wall and also residues involved in receptor binding. In FMDV the receptor-binding site is exposed and nearly fully overlap with a major antibody-binding site (see above). Thus, additional or alternative mechanisms must operate to preserve cell receptor recognition, and also other virion functions, in the face of a strong antibody selection pressure (see next).

Very thorough genetic mapping of epitopes in small non-enveloped viruses (*e.g.*, FMDV) by selecting many antibody-escape mutants revealed that escape mutations tend to repeatedly occur at very few positions in the contact epitope. These residues may sometimes correspond to those in the energetic epitope, but this is not always the case at all. A paradigmatic case is that of the major antigenic loop of FMDV type C. In this loop, a RGD motif and some neighboring residues are critically involved in binding both the cell receptor and different virus-neutralizing antibodies. FMDV cannot escape from neutralization by any of these antibodies through mutation of nearly any one of the residues critically involved in antibody binding, because these are precisely the ones critically involved also in cell receptor binding. However, the virus manages to readily and repeatedly escape neutralization by adopting either of two non-exclusive alternatives: (i) a drastic one (a lesser kind of antigenic "shift", albeit not due to reassortment, affecting a major antigenic site), by mutation of one of a pair of unique residues that are critical for binding most antibodies that recognise this loop, but less so for binding the receptor; and (ii) a gradual one (a kind of antigenic "drift" in that same antigenic site) by accumulation of mutations of several residues that, individually, are of minor relevance for antibody binding but that, together, severely impair binding of most antibodies that recognise this viral loop. Studies with many FMDV type C field variants isolated during several decades in two continents indicate that antigenic "shifts" in the major antigenic loop involving mutation of those unique residues have occurred repeatedly and independently, but infrequently; this paucity may be related to a significant loss of fitness caused by these "balancing act" mutations. In contrast, antigenic "drifts" in this loop by gradual accumulation of less critical mutations appear to have frequently contributed to FMDV escape in the field.

In small non-enveloped virions, the overlap between antigenic regions and regions involved in virion functions may extend to the entire capsid surface. Competition between antibodies or Fabs for binding to FMDV or other viruses suggest that a large part of the capsid surface is involved in neutralization epitopes. In contrast, viral escape mutations identified in the laboratory or in virus variants isolated in the field collectively tend to cluster in very few discrete regions highly exposed on the capsid surface. Thorough mutational studies with FMDV and other

small viruses indicate that most residues in the capsid surface are involved in virus function: if mutated, they severely impair infection, or at least decrease the biological fitness of the virus. Unlike cellular proteins, the protein capsids of small, nonenveloped viruses may have been functionally *streamlined* through multiple, sometimes conflicting selective pressures imposed on a very simple structure (see Chaps. 10, 11, 12, 13, 14, 15, 16, 17, 18 and 19). As a consequence, only very few residues in the viral capsid can be mutated to evade the antibody response and still allow normal virus infection. Immune evasion of FMDV type C in the field appears to have repeatedly involved mainly combinations of mutations of a surprisingly restricted number of residues in the capsid surface; these capsid residues may be among the few whose mutation has no severe effect on virus function.

1.4.5 Virus-Host Cell Recognition, Entry and Uncoating

After an often long and dangerous voyage in the extracellular environment, in which a few virions may have suceeded in preserving their integrity and functionality while evading the host defenses, one of the surviving virions may eventually encounter another host cell. The virion may then infect the cell by anchoring to appropriate receptor molecule(s), which will trigger the penetration of the entire virion or some of its components (invariably including the viral nucleic acid) into the cell. Unless the nucleic acid enters alone and naked, viral genome replication and expression will require uncoating of the internalized viral nucleic acid. Cell recognition, entry and uncoating frequently overlap and rely on quite different mechanisms depending on the virus species and host type, and are very briefly summarized here separately for animal viruses (see Chaps. 15, 16), bacteriophages (see Chap. 17) and plant viruses.

Cell Recognition, Entry and Uncoating of Animal Viruses

Cell receptor recognition. Like any other cell, animal cells are surrounded by a soft, lipidic plasma membrane. However, unlike bacteria or plant cells, animal cells are not surrounded by a hard, thick cell wall, and this absence facilitates virus penetration. In general, animal viruses start the infection of a host cell by specifically recognizing some *receptor* molecule embedded in the cell membrane (Chap. 15). In the case of non-enveloped viruses, a part of the capsid surface, or a specialized capsid-attached viral protein can bind some cellular molecule (typically a glycoprotein, carbohydrate, or glycolipid) that normally serves some cellular function. In the case of enveloped virions, a viral glycoprotein embedded in the viral lipid envelope can specifically interact with the cell receptor molecule. In some cases,

co-receptor molecules in the cell membrane may also be recruited to complete the recognition process. Virus-receptor binding is an energetically favorable process, and may trigger conformational changes in the bound capsid or viral protein that mediate the entry process.

Entry into the host cell. Attachment of the virus to the receptor molecule starts a complex process in which the entire virus particle, or a part of it (usually the capsid with the nucleic acid inside), enters the cell. Penetration may, in general, occur by direct entry through the cell membrane; or, much more frequently, by *endocytosis* mediated by interaction with the receptor(s), which lead to the engulfment of the virion in an endosome. In this latter case, a subsequent mechanism is required for the release of the viral particle or uncoated genome in the appropriate cellular compartment.

The mechanisms of penetration of non-enveloped animal viruses are in general poorly known, but for some viruses this process is understood in some detail (Chap. 15). For example, cell-receptor recognition by many non-enveloped viruses may activate a process of endocytosis, in which the virion enters the cell largely intact and is transported through one or several different internalization routes, ending with the release of the viral genome in the appropriate cellular compartment. Release of the virus or some component (including the viral genome) in the appropriate compartment may occur after lysis of the endosome and diffusion of the virion or disassembled components (see below); or by permeabilization of the endosome membrane, which may include the transient opening of a channel through which the viral nucleic acid can be translocated.

The mechanism of penetration of enveloped viruses is, in general, better known. Binding of a viral glycoprotein to a cell receptor invariably triggers a conformational change in a viral *fusion protein* that leads to the physical joining (fusion) of the viral envelope and a cellular membrane (Chap. 16). In direct entry (as in HIV-1), after receptor recognition the virus envelope and the plasma membrane are fused. In endosome-mediated entry, after the virus is engulfed in the endosome, acidification leads to the fusion between the virus envelope and the endosome membrane. In both cases, the fusion event allows the diffusion of the viral capsid or nucleocapsid containing the viral genome inside the cell. This is followed by trafficking of the relevant viral component to the appropriate cellular compartment for genome expression and replication.

Genome uncoating. Uncoating is, in general, a very poorly known step in the life cycles of animal viruses. In some non-enveloped virions, such as the picornaviruses PV and HRV, receptor binding and or low pH mediate capsid conformational rearrangements that may open pores or fractures large enough for nucleic acid extrusion. In other non-enveloped virions such as the picornavirus FMDV, acidification at the endosomes dissociates the viral capsid into subunits, and the nucleic acid is released into the cytoplasm. Many details of these processes are still uncertain or unknown (see Chap. 15). In enveloped virions, the viral nucleic acid-containing capsid or nucleocapsid that entered the cell after fusion of the virus and cell membranes is transported to the appropriate cell compartment for viral nucleic acid uncoating. In some cases, the capsid is disassembled and the nucleic

acid just released, either by dilution or by a poorly characterized mechanism that may require the participation of cellular factors (*e.g.*, in HIV-1). In other cases, direct replication and expression of the genome may occur with the viral nucleic acid being still a part of a ribonucleoprotein particle (*e.g.*, in influenza virus) or being still associated with a viral capsid core (*e.g.*, in reoviruses). See Chaps. 15 and 16 for a detailed description of different mechanisms of cell recognition, entry and uncoating in different non-enveloped or enveloped animal viruses.

Cell Recognition and Nucleic Acid Transfer by Bacteriophages

Bacterial cells are surrounded by thick and hard cell walls that bacteriophages cannot penetrate. Like animal viruses, bacteriophages recognize cellular receptors on the cell surface. However, the phage particle does not actually enter the cell. The barrier posed by the cell wall is overcome in different ways. In many dsDNA phages, a tube penetrates the cell wall as if it were the needle of a hypodermic syringe, and the nucleic acid genome is injected inside the cell, driven in part by the internal pressure built-in in the capsid during packaging of the nucleic acid. See Chap. 17 for a detailed description of cell receptor recognition by different bacteriophages and phage nucleic acid delivery into the host bacterium.

Entry of Plant Viruses

Plant cells are surrounded, like bacterial cells, by a rigid cell wall a virus cannot easily penetrate. Nearly all plant viruses are ssRNA(+) viruses with a slender helical capsid or a relatively small icosahedral capsid. Most seem to enter the plant cell in a passive way, through breaches in the cell wall and transient openings in the plasma membrane that may be caused by mechanical injury, for example by the action of herbivorous animals or insects. Once a cell is infected, progeny virions may propagate from cell to cell in the organism through discontinuities in the cell wall crossed by channels that establish interconnections between neighboring cells (*e.g.*, plasmodesmata). A general description of plant virus entry and infection can be found, for example, in the book by Hull [8].

1.5 Experimental and Theoretical Developments in Physical Virology

Physical virology (see Sect. 1.1.2 for a brief introduction to the field) is becoming integrated with classical structural virology and providing novel fundamental insights into the structure, dynamics and properties of virus particles. In this book we have described in detail dominant experimental (Chaps. 8, 9) and theoretical (Chap. 19) methods and several basic studies (Chaps. 18, 19) in Physical Virology, as well as some potential applications (Chap. 22).

Two very important physics-based, single-molecule techniques that have started to be applied to viruses are AFM (Chap. 8) and optical tweezers (Chap. 9). AFM can be used to image viruses and follow structural changes in viral particles in real-time (with the temporal limitations imposed by the time required to acquire each image). In addition, AFM can be used to probe some physical properties of virus particles. In particular, their mechanical characteristics such as particle stiffness, Young's modulus of the capsid material (as a measure of its intrinsic stiffness), tensile strength or resistance to mechanical fatigue are being determined. A dominant method used to study virus mechanics involves the indentation of virions or capsids by "pushing" on defined points on the particle surface with the AFM tip (Chap. 18). AFM is also being used to determine the values of the physical forces involved in molecular recognition between virus particles and different ligands, for example cell receptors or antibodies, even in physiologically relevant conditions, with receptors embedded in the membrane of a cell or liposome. The general approach in these studies involves attaching a virus particle or virus ligand to the AFM tip and the interacting partner to a solid substrate, and "pulling" until the interaction between them is broken. Mechanical aspects of some stages in the viral cycle, such as virus entry into cells, can also be followed by AFM (see Chap. 8). Optical tweezers have been particularly successful in determining physical forces involved in the action of molecular motors during the mechanochemical process of dsDNA encapsidation inside a preformed bacteriophage capsid (see Chap. 9).

Several theoretical physics-based approaches are being applied to understand the physical foundations of virus capsid architecture; material properties of viruses, especially virus mechanics; and different aspects of the infection process in which viral particles are involved, including capsid assembly and maturation and virus entry into cells. In this book we have integrated theoretical studies in physical virology in a single monographic chapter (Chap. 19). In that chapter, a theoretical physics-based perspective is used to contemplate many aspects of virus structure, dynamics and properties that had been described in previous chapters (Chaps. 2, 10, 15 and 18) from an empiric, structural biology-based perspective.

1.6 Applied Structural and Physical Virology

Several decades of intensive research on the structure, dynamics and properties of virus particles, part of which is described in Chaps. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 and 19 of this book, have opened up many possibilities for the exploitation of the knowledge acquired, and of the viral particles themselves, in biomedicine and biotechnology. In addition, the recent advent of nanotechnology and novel studies on the physics of viruses has led to a growing awareness by researchers of physical or chemical disciplines on the multiple possibilities virus particles present also for nanotechnological applications. In addition to their possible direct use, viruses are also providing inspiration for the design of non-viral or

hybrid nanoparticles and nanomaterials. Among others, two extremely important applied areas in human and animal health in which the knowledge acquired on virus structure is having a clear impact are the development of novel antiviral drugs (Chap. 20) and virus vaccines (Chap. 21). Knowledge on the structure and physical properties of virus particles is also being applied for the development of viral-based nanoparticles, nanomaterials and nanodevices for multiple applications (Chap. 22)

1.6.1 Virus Structure-Based Antiviral Research

The knowledge acquired from structural studies on virus morphogenesis (Chaps. 10, 11, 12, 13 and 14) and cell receptor recognition and entry of animal viruses into cells (Chaps. 15 and 16) have provided excellent opportunities to develop new antiviral agents based on the inhibition or misdirection of these processes. Chapter 20 focuses on the general structural basis of antiviral action and describes principles and methods for the structure-based search and design of new antiviral agents. Of particular relevance for the subject of this book, the chapter includes several examples where the target of antiviral research is virus assembly, cell receptor recognition or entry into the host cell.

1.6.2 Applications of Virus Particles

Virus particles have acquired through evolution outstanding properties and functions. To name just a few: capsid self-assembly; capsid metastability and conformational rearrangements; the mechanochemical action of viral molecular motors for nucleic acid packaging; the targeting of specific cells through precise molecular recognition; and chemical and mechanical actions for delivery of the genome into the host cells, including particle penetration and controlled particle disassembly, or nucleic acid injection. These properties have opened up many possibilities to turn viral particles into useful biomaterials and devices in biomedicine, biotechnology and nanotechnology. However, viruses as evolved in nature do not have all of the several properties and functionalities required, or are not optimized, for any of their contemplated applications. Thus, methodologies have to be devised to manipulate virus particles and modify their structure, properties, and/or functions to serve the intended goals. Virus particles can be modified by direct chemical means and/or by protein engineering approaches using genetic techniques (see Chap. 22). Learning to successfully modify virus particles may not only provide useful products, but also lead to a deeper understanding on the structure and function of these and other biomolecular complexes.

A non-comprehensive list of current or under-development applications of the knowledge acquired on virus particles, and of adequately engineered viral particles may include (Fig. 1.6):

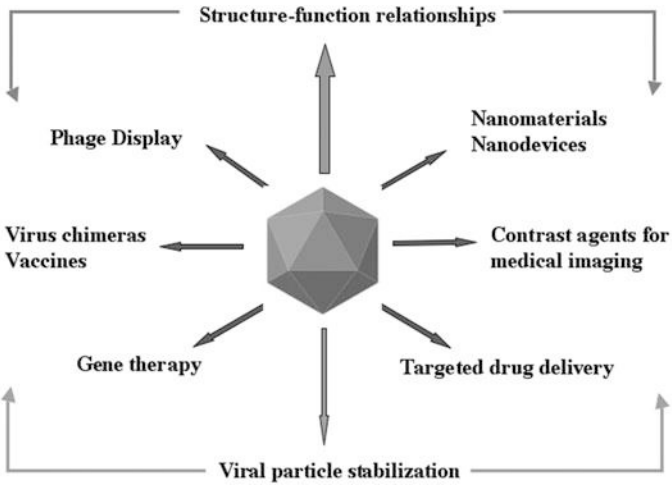

Fig. 1.6 Some applications of virus particles in biomedicine, biotechnology and nanotechnology (see text for an explanation)

(i) Further structure-function studies of viruses oriented to explore, evaluate or implement novel applications.
(ii) *Phage display* techniques for the selection from a combinatorial library of variant proteins or peptides of those with the desired properties (binding to specific ligands, improved binding, chemical or conformational stability, catalysis of specific substrates or improved catalytic efficiency, etc.).
(iii) Novel or improved vaccines, including those based on virus-like particles (VLPs) obtained by production of CPs in expression systems and self-assembly in the cell or *in vitro*, or chimeric VLPs or virions carrying foreign epitopes from a pathogen (Chap. 21).
(iv) Delivery of therapeutic genes into specific cells for gene therapy.
(v) Targeted delivery of drugs using viral particles as nanocarriers or nanocontainers for specific chemotherapies.
(vi) Target-specific contrast agents for molecular imaging and diagnosis.
(vii) Building blocks for the construction of nanomaterials, nanostructures and nanodevices for technological uses, from medicine to electronics. Chapter 22 describes a number of viral-particle-based possibilities being explored for applications in biotechnology and nanotechnology.

1.7 Concluding Remarks

A vast amount of information on the structure, dynamics and properties of virions and their molecular components has accumulated at an ever increasing pace. The interplay between structural, physical and molecular virology has led to a giant leap

in our understanding of virus structure and function; a deeper and better comprehension on the workings of the molecular machinery of life; critical advances in the fight against viral disease; and promising developments in the application of virus particles to medicine and technology. However, there is still much to be learned on how viruses are made and how they work.

I would like to end this chapter with a note intended for the young student: Truly inspiring views of coral reefs are obtained not by snorkeling, but by diving. William Beebe recommended readers of his 1928 book *Beneath Tropic Seas* (Putnam's sons, NY): "Don't die without having borrowed, stolen, purchased or made a helmet of sorts, to glimpse for yourself this new world". This chapter has provided just a quick look from the *surface* at some of what has been learned on virus structure and physics so far. By *diving deep* into the chapters that follow, the reader will be rewarded with amazing close-up views of these fascinating molecular machines.

Acknowledgements I gratefully acknowledge former and current collaborators and members of my group for their invaluable contributions to our studies on virus structure-function relationships, Miguel Angel Fuertes for providing figures and technical assistance, and Pedro J. de Pablo, Julio Gómez-Herrero, David Reguera and Ricardo Arias-Gonzalez for fruitful discussions on the connection between the physics and biology of viruses. Work in the author's laboratory is funded by grants from the Spanish Government (BIO2009-10092 and BIO2012-37649) and Comunidad de Madrid (CM) (S-2009/MAT/1467), and by an institutional grant from Fundación Ramón Areces. The author is an associate member of the Institute for Biocomputation and Physics of Complex Systems, Zaragoza, Spain.

References and Further Reading[1]

Molecular Virology Textbooks[2]

1. Flint SJ, Enquist LW, Racaniello VR, Skalka AM (2009) Principles of virology, 3rd edn. ASM Press, Washington, DC. *Intermediate level, highly recommended book on molecular virology*
2. Dimmock N, Easton A, Leppard K (2007) Introduction to modern virology, 6th edn. Blackwell, Malden. *Introductory text on general principles in molecular virology*
3. Carter JB, Saunders VA (2007) Virology. Principles and applications. Wiley, Chichester. *Introductory text on general principles in molecular virology*

[1] *Because of the elementary, student-oriented nature of this introductory chapter, we have purposefully limited the list of references to some recommended books. For further references to scientific reviews and original articles on structural and physical virology, the reader is referred to the sections on "References and further reading" included in the rest of the chapters in this book.*

[2] These books contain elementary to intermediate information on the molecular biology of viruses, with emphasis on viral genome replication and expression. Reference to some structural aspects directly related with these processes, which are out of the scope of the present book, may be also found in these books.

4. Wagner EK, Hewlett MJ, Bloom DC, Camerini D (2008) Basic virology. Blackwell, Malden. *Introductory text on general principles in molecular virology*
5. Cann AJ (2012) Principles of molecular virology, 5th edn. Academic Press, Waltham. *Introductory text on general principles in molecular virology*
6. Acheson NH (2007) Fundamentals of molecular virology, 2nd edn. Wiley, Hoboken. *Introductory text dealing separately with the molecular biology of many virus families*
7. Kutter E (2005) Bacteriophage: molecular biology and applications. CRC Press, Boca Raton. *A molecular virology book specialized in bacteriophages*
8. Hull R (2009) Comparative plant virology, 2nd edn. Academic Press, Burlington. *A specialized book on the molecular biology of plant viruses, including many structural features*

Advanced Reference Books on Virology

9. Knipe DM, Howley PM (eds) (2007) Fields virology, 5th edn. Lippincott Williams & Wilkins, Philadelphia. *Advanced, extensive coverage of virology; a classic two-volume reference book*
10. Mahy BWJ, van Regenmortel MHV (eds) (2008) Encyclopedia of virology, 3rd edn. Academic Press, Boston. *Advanced, extensive coverage of virology topics based on concise reviews; a five-volume reference book*

Advanced Books on Structural or Physical Virology[3]

11. Casjens S (ed) (1985) Virus structure and assembly. Jones and Bartlett, Boston. *Largely outdated, but very useful for understanding some fundamental principles in structural virology*
12. Chiu W, Burnett RM, Garcea RL (eds) (1997) Structural biology of viruses. Oxford University Press, New York. *Outdated in many respects, but an extremely useful book that describes clearly many principles in structural virology and contains detailed advanced monographs on a variety of structural virology topics*
13. Chiu W, Johnson JE (eds) (2003) Virus structure. Advances in protein chemistry, vol 64. Academic Press, New York. *An excellent collection of advanced reviews on a limited but highly relevant selection of structural virology topics*
14. Stockley PG, Twarock R (eds) (2010) Emerging topics in physical virology. Imperial College Press, London. *An excellent and largely up-to-date collection of advanced reviews on a number of specific topics in theoretical or experimental physical virology. Highly recommended*
15. Agbandje-McKenna M, McKenna R (eds) (2011) Structural virology. RSC Publishing, Cambridge. *An excellent and up-to-date book containing monographs on many structural virology topics. Some chapters in that book complement others in the present book by covering topics such as evolution of viral capsid structures, structural aspects of virus-antibody recognition, or viral vectors for gene therapy. Physical virology-specific techniques (AFM, optical tweezers, physics-based models and simulations) and studies using these approaches are not contemplated. Highly recommended*

[3] These specialized books contain excellent advanced monographs on many structural or physical virology topics.

16. Rossmann MG, Rao VB (eds) (2012) Viral molecular machines. Advances in experimental medicine and biology, vol 726. Springer, New York. *A collection of excellent and up-to-date, quite advanced and detailed reviews on some structural and molecular aspects of viruses considered as molecular machines, with emphasis on nucleic acid packaging. Physical virology-specific topics (see comment on ref. 15) are not contemplated. Highly recommended for advanced, specialised reading*

References on Virus Genetic and Phenotypic Variation and Evolution[4]

17. Sobrino F, Domingo E (eds) (2004) Foot and mouth disease: current perspectives. Horizon Bioscience, Wymondham/Norfolk. *See Chaps. 10 and 9, respectively, for detailed accounts on quasispecies and virus evolution and antibody recognition and viral escape, using a thoroughly studied virus model, FMDV*
18. Domingo E, Biebricher C, Eigen M, Holland J (2001) Quasispecies and RNA virus evolution: principles and consequences. Landes Bioscience, Austin. *An excellent book on RNA virus variation, quasispecies and evolution*

[4] Important to understand the evolutionary concepts and genetic mechanisms behind the origin and continuous modification and adaptation of the viral machines described in the present book.

Chapter 2
The Basic Architecture of Viruses

José R. Castón and José L. Carrascosa

Abstract Viruses are elegant macromolecular assemblies and constitute a paradigm of the economy of genomic resources; they must use simple general principles and a very limited number of viral components to complete their life cycles successfully. Viruses need only one or a few different capsid structural subunits to build an infectious particle, which is made possible because of two reasons: extensive use of symmetry and built-in conformational flexibility. Although viruses from the numerous virus families come in many shapes and sizes, two major types of symmetric assemblies are found: icosahedral and helical particles. The enormous diversity of virus structures might be derived from one or a limited number of basic schemes that has become more complex by consecutive incorporation of structural elements. The intrinsic structural polymorphism of the viral proteins and other observations indicate that capsids are dynamic structures. Study of virus structures is required to understand structure-function relationships in viruses, including those related to morphogenesis and antigenicity. These structural foundations can be extended to other macromolecular complexes that control many fundamental processes in biology.

Keywords Capsid • Nucleocapsid • Helical symmetry • Icosahedral symmetry • Triangulation number • Conformational polymorphism • Quasi-equivalence • Prolate capsid • Metastable capsid • Molecular switch

Abbreviations

3D	Three-dimensional
CP	Capsid protein

J.R. Castón (✉) • J.L. Carrascosa
Department of Macromolecular Structure, Centro Nacional de Biotecnología (CSIC), c/Darwin 3, Campus de Cantoblanco, 28049 Madrid, Spain
e-mail: jrcaston@cnb.csic.es

M.G. Mateu (ed.), *Structure and Physics of Viruses: An Integrated Textbook*, Subcellular Biochemistry 68, DOI 10.1007/978-94-007-6552-8_2,

cryo-EM	Cryo-electron microscopy
cryo-ET	Cryo-electron tomography
ds	Double-stranded
ss	Single-stranded
T	Triangulation number

2.1 Introduction

Viruses are probably the most extreme examples of nature's efficient use of limited coding capacity. Optimization of resources is reflected in their multiplication and dissemination strategies, as well as in their construction. Viruses use simple, general principles to successfully complete their life cycles [1]. They need only one or a few proteins to build a protein shell (the capsid) and package the viral genome inside to yield an infectious particle (the virion). Virions are endowed with a remarkable number of essential functions, including protection of the genome during extracellular transport, entry into the host cell, intracellular trafficking, genome uncoating and, in some instances, viral genome replication. All of these activities are carried out efficiently, even in the face of structural variability of the capsid proteins (CP) due to distinct selective pressures.

Optimization of resources in building a viral particle has been achieved in two ways: extensive use of symmetry and conformational polymorphism of the CPs [2], both of which allow the virus to build a three-dimensional (3D) container of defined size and shape by using many copies of only one or a few proteins. Except in some complex viruses, the two main classes of symmetric capsid structures are icosahedral and helical [3, 4]. In both cases, the need for multiple copies of one or more CP subunit type arranged symmetrically can also be inferred from the nature of the genetic code. A base triplet (codon) has an approximate molecular weight of 1,000 Da and encodes a single amino acid with an average molecular weight of 150 Da. At best, a nucleic acid can encode 15 % of the virus weight as protein; nevertheless, 50–90 % of the virus molecular weight is protein. This high protein: nucleic acid ratio is necessary to allow complete encapsidation of the genome in a large enough, hollow container through CP oligomerization.

The classical vision of viral capsids as inert containers that protect the fragile genome during the extracellular phase of the virion is inappropriate; capsids are dynamic structures whose components have transient conformations related to specific functions in the viral cycle [5, 6]. The capsid is a metastable assembly, and should be considered a compromise: it is robust enough to protect the genome and ensure its propagation during passage from one host cell (or organism) to another, and labile enough to allow genome delivery in the host cell to initiate infection.

CP subunits (a single folded polypeptide chain) may be grouped in protomers (structural units or "capsid building blocks") from which a capsid is built.

Protomers may be identified by biochemical techniques, and may comprise one or more nonidentical protein subunits (*e.g.*, VP1, VP2, VP3 and VP4 in poliovirus, the smallest subset of CP subunits that may reconstitute the complete capsid through symmetry operations). Protomers may cluster into capsomers, morphological units such as hexamers, pentamers, trimers or dimers of protomers. Capsomers correspond to the apparent clusters, knobs or projections seen on the particle surface by electron microscopy. Because capsomers do not necessarily correspond to chemical entities that can be isolated, this term is generally restricted to descriptions of viruses from electron micrographs.

Interactions among protein subunits in the viral capsid are noncovalent and generally weak. There are exceptions, such as the bacteriophage HK97 protein shell, in which pentameric and hexameric capsomers are covalently joined to form rings that are also topologically interlocked, creating a protein chainmail [7]. For the most part, however, as a result of viral particle symmetry, numerous weak interactions (hydrogen bonds, salt bridges, van der Waals forces and hydrophobic interactions) must cooperate to maintain an appropriate 3D structure and to facilitate dynamic structural processes.

How is a closed shell with a defined shape built? This is a complex process that must reconcile two apparently contradictory requisites; energy requirements must be minimal to ensure high assembly efficiency, which decreases the occurrence of aberrant assemblies *in vivo*, and the protein container should be stable and robust. A number of factors increase the effectiveness of assembly. Parts of the tertiary structure of the CP have built-in conformational flexibility, which allows them to maintain the slightly different intersubunit contacts necessary for the assembly of most icosahedral capsids (see below). CPs have also large interaction areas and, even isolated from their natural environment, CP subunits or small CP oligomers (*i.e.*, capsid building blocks) are able to self-assemble spontaneously *in vitro* into complex assemblies similar or related to viral particles (viral-like particles or VLPs; see Chap. 21). This spontaneous assembly indicates that the viral particle is in a free-energy minimum state. To facilitate and direct correct intersubunit interactions, viruses can use host resources such as molecular chaperones and membrane-specific regions, or other viral components such as scaffolding, accessory, cement, and proteolytic proteins.

In this chapter, we describe the principles of the basic structural organization of viruses and discuss numerous examples of capsid architecture. X-ray crystallography and 3D cryo-electron microscopy (cryo-EM) techniques have shown the molecular mechanisms by which identical protein building blocks make use of symmetry and structural polymorphism to assemble a multifunctional container in a biologically feasible period of time.

2.2 How Virus Structures Are Studied

Viruses were first visualized by electron microscopy, which provided the basis for their classification following morphological criteria. Development of electron microscopy techniques has had a central role in the study of viral structure and

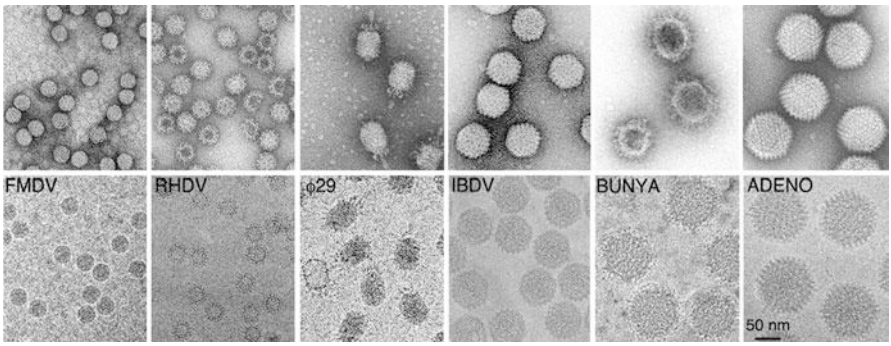

Fig. 2.1 Conventional EM and cryo-EM analysis of several viruses. A collection of representative virus preparations visualized by conventional negative staining (*top*), or directly vitrified without staining and/or fixing agent (*bottom*). FMDV, foot-and-mouth disease virus (sample provided by E. Domingo, CSIC); RHDV, rabbit hemorrhagic disease virus (provided by J. Bárcena, INIA); ϕ29, prolate bacteriophage ϕ29; IBDV, infectious bursal disease virus; BUNYA, the enveloped bunyavirus, Bunyamwera virus (provided by C. Risco, CSIC); ADENO, adenovirus (micrographs by C. San Martín, CSIC)

life cycle (see Chap. 3). The introduction of adequate fixation methods (chemical or physical) and contrast protocols (such as negative staining) for viral samples have allowed observation of isolated viruses and virus-infected cells in thin sections. More recently, cryo-EM has become a major tool; samples are frozen so rapidly that they are immobilized in an amorphous (or vitreous) solid state, without physical damage of the specimen by ice crystal growth [5]. Cryo-EM allows direct visualization of viral particles, as no staining or fixative agents are introduced during sample preparation, and native conformation is preserved, as samples remain hydrated (Fig. 2.1). Combined with image reconstruction techniques, cryo-EM is used widely to establish the 3D structure of viruses. Cryo-EM has contributed notably to demonstration of the metastable nature of viral capsids thoughout their life cycle. Dynamic studies of viruses have shown conformational changes of structural subunits that helped to explain the underlying molecular mechanisms of specific processes. 3D cryo-EM maps are usually calculated at subnanometer resolution, enabling us to infer secondary structure elements in the capsid structural subunits. A few 3D cryo-EM maps have recently been calculated at near-atomic resolution, to trace amino acid backbones (Cα model building) and to define the surface contacts among structural components of the virion [8, 9]. Enveloped virions are flexible (pleomorphic) and less symmetric than naked icosahedral or helical viruses; cryo-electron tomography (cryo-ET) is used to study individual particles of complex enveloped viruses such as orthomyxo- and poxvirus, albeit at much lower resolution than single particle cryo-EM. Cryo-ET is used also to study structural aspects of the viral life cycle directly in infected cells (see Chap. 3).

X-ray diffraction techniques are the other major approach for determining virus structure and interactions at the atomic level. Their major limitation is the need to procure sufficiently large (milligram) quantities of highly pure (and stable) virus

preparations to form suitable 3D crystals for X-ray diffraction. These crystals are relatively easy to obtain from small spherical virus particles, although resolution of complex icosahedral capsids is becoming reasonably common, due to the combination of robotized crystallization procedures and the use of improved X-ray diffraction equipment allowing the structural resolution from very small microcrystals (see Chap. 4). The first high resolution virus structure, that of the ssRNA plant virus tomato bushy stunt virus (TBSV), was determined in 1978 [10]. The first structures of small human viruses, poliovirus [11] and rhinovirus [12], were obtained in the mid-1980s. Atomic resolution maps of viral capsids have been essential in establishing the mechanisms that govern assembly and stability of viral structures and the structural basis of many virus functions.

Despite their minimal sequence similarity, many CP of spherical viruses (in animals and plants) adopt a very limited number of tertiary folds, the most common of which is the antiparallel β-barrel [13]. The β-barrel fold is also known as a jelly-roll or a Swiss-roll β-barrel. This wedge-shaped structure is formed by eight antiparallel β-strands (with two opposing BIDG and CHEF β-sheets) and two α-helices (Fig. 2.2a). CP size varies between 20 and over 70 kDa, with the major differences located in N- and C-terminal ends (usually facing the interior and the outer surface of the capsid, respectively) and in the loop sizes between β-strands (in some cases, more than 100 amino acids that form additional domains). The β-strands usually lie roughly tangential to the capsid surface, defining a protein shell ~30 Å thick, although they can also lie orthogonally relative to the particle surface. The wedge-like shape allows the subunits to pack tightly and, due in part to this built-in flexibility, the jelly-roll motif-based CP can assemble into unrelated capsids of variable sizes (by increasing the T number, see below). The β-barrel is also found in some bacilliform viruses such as alfalfa mosaic virus (AMV).

β-barrel duplications have also been reported. Joint folds were first observed in the adenovirus trimeric capsomer; each monomer consists of two successive jelly-roll motifs, producing a pseudo-hexameric structure [14]. Many other large double-stranded (ds) DNA viruses have similar trimeric capsomers, such as the *Paramecium bursaria chlorella* virus 1 (PBCV1) [15], the dsDNA bacteriophage PRD1 [16], and the sulfolobus-infecting virus, STIV (sulfolobus turreted icosahedral virus) [17]. The large subunit of comoviruses, single-stranded (ss) RNA viruses that infect plants, is also made by fusion of two β-barrel domains [18]. A similar fused fold is found in a transient structural protein in a nonicosahedral virus, vaccinia virus [19].

Crystallographic analysis showed that, in addition to the β-barrel as the most extended fold among CP of spherical viruses [13], a limited number of other folds is used to form viral capsids, including those found in RNA bacteriophage MS2 or in dsDNA icosahedral bacteriophages such as the HK97 phage [20]. The HK97 CP forms an L-shaped, mixed α/β structure with two domains. The peripheral (P) domain consists of a long α-helix and an elongated, three-stranded β-sheet plus extensions (N-arm and E-loop) and the axial (A) domain, composed of two α-helices and an additional β-sheet. The two residues that form covalent crosslinks with neighboring subunits are on opposite sides of the protein (Fig. 2.2b). This structure is optimal for CP self-assembly to form a metastable particle, in which subunit interactions are

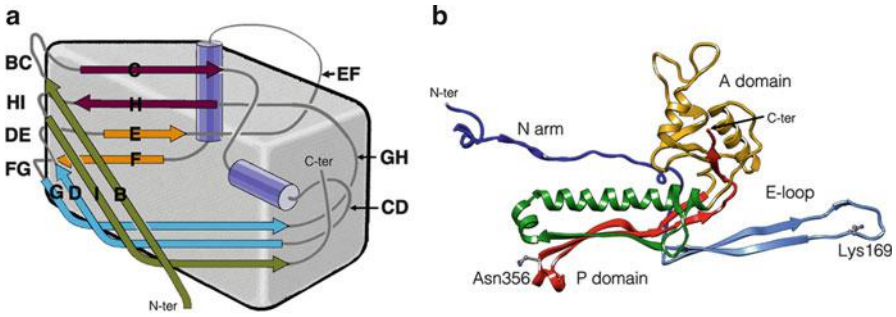

Fig. 2.2 The wedge-shaped β-barrel fold and the HK97 capsid protein fold. (**a**) Scheme showing the connecting loops between secondary structural elements (eight β-strands, arrows labeled B-I, and two α-helices, shown as cylinders). Loops CD, EF and GH generally have large insertions; loops BC, HI, DE and FG are short. This wedge-like domain has large interaction surfaces, appropriate for icosahedral symmetry. (**b**) Structure of the HK97 capsid protein (gp5), color-coded by domains (PDB entry 1OHG). The two domains A and P are indicated, as well as the extended N-arm (*dark blue*) and the E-loop (*cyan*). In the mature capsid, Lys^{169} on the E-loop forms a covalent bond with Asn^{356} on a neighboring subunit

modulated by the CP N-terminal scaffolding domain. A proteolytic process then takes place to remove the scaffolding domain, followed by massive conformational changes and crosslinking, leading to a robust mature capsid [20] (see Chaps. 11 and 13). Some bacteriophages incorporate additional domains in the main shell protein fold, resulting in different assembly mechanisms [21].

X-ray crystallography is used not only to solve atomic structures of entire viral capsids, but also to obtain detailed structural maps of isolated viral proteins such as hemagglutinin and neuraminidase of influenza virus, several viral proteases (from picornaviruses, herpesviruses, hepatitis B virus (HBV), human immunodeficiency virus (HIV), etc.), the HIV reverse transcriptase and several RNA polymerases, among many others. Large viruses and intermediate states of capsid assembly are usually refractory to high-resolution analysis by X-ray crystallography, but are accessible by 3D cryo-EM. Many studies have successfully combined these complementary techniques; pseudoatomic models of the capsids are derived after docking atomic structures of individual capsid components into lower resolution cryo-EM density maps (see Chap. 7). Nuclear magnetic resonance spectroscopy has also contributed valuable structural information by revealing the structure of viral protein domains and the structural dynamics of viral protein domains and capsids (see Chap. 5). Atomic structures of virions, capsids and viral proteins have supplied valuable information for the design of diverse antiviral drugs, including some that alter viral particle stability, interfere with capsid assembly or impair viral enzyme activity (see Chap. 20); they have been also extremely useful for the engineering of virus particles for biomedical, biotechnological or nanotechnological applications (see Chaps. 21 and 22).

2.3 Viral Capsid Symmetry

Molecular crystals are formed from multiple copies of single molecules that establish identical interactions with adjacent molecules, and thus have equivalent environments in the crystal. This is the simplest solution to form a regular or symmetric structure. There are fundamental differences between crystals and viruses, however. Whereas a crystal is a continuous 3D network with no size limitations, viruses build an empty container of defined shape for nucleic acid packaging in a biologically feasible time frame [22]. The capsid shell is therefore seen as a two-dimensional crystal that is limited in size and, in most cases, closed, leaving a cargo space inside. Although they use basic principles of crystal growth, viruses require additional mechanisms to control their dimensions [23].

There are several alternatives for generating 2D crystals within the 17 planar crystallographic groups. Some of these groups are variants of hexagonal lattices. In a hexagonal network, identical subunits can be arranged to maintain them in identical environments. If correct distortions are introduced, two types of closed shells can be made from this planar network, helices and icosahedra. If after bending and binding, the network edges are out of register, a tube is obtained in which layer lines are in register and generate a helix; if the edges are in register, a cylinder with a pile of stacked discs is obtained (Fig. 2.3a). In a helix, structural subunits have identical interactions, except those located at the ends, and the self-assembly of protomers can easily be envisaged. The diameter of the tubular structure depends on the curvature in the subunit interactions, whereas its length is limited by that of the nucleic acid it encloses and by subunit-nucleic acid interactions.

The icosahedron is a closed polyhedron with cubic symmetry, and asymmetric units can therefore be laid on the symmetric polyhedron surface with identical interactions among them. In 1956, Watson and Crick proposed the first principles for construction of spherical viruses [24]. In theory, a limited number of regular solids (termed Platonic polyhedra) allow close packing of repeated (asymmetric) subunits such as proteins. These polyhedra are the tetrahedron (four triangular faces and 12 equivalent positions), the cube (six square faces and 24 equivalent positions), the octahedron (eight triangular faces and 24 equivalent positions), the dodecahedron (12 pentagonal faces and 60 equivalent positions), and the icosahedron (20 triangular faces and 60 equivalent positions). The icosahedron (and the dodecahedron) is the most economic natural structure in terms of genetics, as a larger number of identical subunits encoded by a single gene can be laid in identical positions. In addition, the icosahedron is the cubic symmetrical structure in which the surface/volume ratio is lowest, involving the best component use to generate a hollow polyhedron, that is, a container, of the highest possible capacity per unit of genetic information required to encode it. An icosahedron can be generated from a hexagonal planar lattice by conversion of planar hexagonal nodes (or hexamers) at defined positions in the lattice into convex pentagonal nodes (or pentamers) after removing one subunit/hexagonal node (Fig. 2.3b). Inclusion of pentamers allows

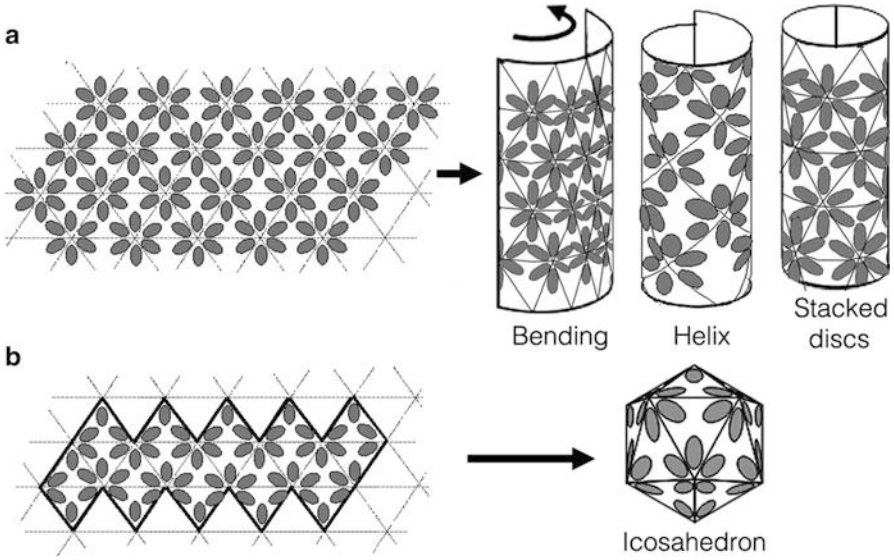

Fig. 2.3 Generation of curved structures such as a helix and an icosahedron from a planar hexagonal lattice. (**a**) Planar hexagonal lattice with subunits arranged in the triangular faces defined by the lattice vectors. (**b**) Icosahedron generated from a defined subset of faces derived from the hexagonal array

curvature in the lattice, which is closed as a polyhedron with 12 pentagonal vertices and 60 identical subunits. The icosahedron has a series of characteristic views depending on its orientation relative to the rotational symmetry of the solid, known as 532 symmetry. An icosahedron has six fivefold axes through the 12 vertices, ten threefold axes through the 20 triangular faces, and 15 twofold axes through the center of each edge (Fig. 2.4). The capsid described (with three identical subunits to form each triangular face) would be the simplest icosahedral capsid formed by 60 equivalent subunits.

2.4 Quasi-Equivalence Theory and Icosahedral Capsid Architecture

When the first icosahedral virus structures were characterized in some detail, it became evident that the limitation to 60 identical CP subunits was not compatible with the high molar mass and size of most viral capsids. The problem to be solved was how an icosahedral capsid could be made using more than 60 identical CP subunits. Based on the conceptual model of the hexagonal planar lattice, some hexamers must be added between pentamers to build larger icosahedra. In 1962, Caspar and Klug developed the quasi-equivalence theory, which constitutes the basis of structural studies of viruses [25].

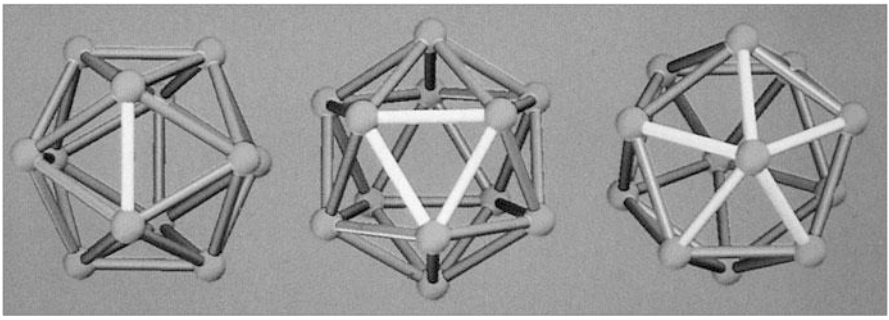

Fig. 2.4 Rotational icosahedral symmetry axes. The icosahedron viewed along two-, three- and fivefold symmetry axes (*left to right*). The twofold axis passes through the middle of each edge, the threefold axis passes through the center of each icosahedral face (an equilateral triangle), and the fivefold axis through the vertices of the icosahedron

The quasi-equivalence theory is based on the assumption that identical CP subunits interact forming nearly similar (quasi-equivalent) bonds for the construction of stable icosahedral capsids with multiples of 60 subunits. There is a degree of flexibility in the intersubunit contacts (conformational polymorphism), which involves minor variations in the 3D structure of the CP subunits. Each conformational state is termed a conformer. For example, the structure of the CP in pentamer subunits is not identical to that of the CP in the hexamer subunits. A number of different-sized icosahedra can be generated, because each has a unique ratio of hexamers to the 12 pentamers in the basic icosahedron. As indicated above, five or six subunits that interact closely are called pentameric or hexameric capsomers, respectively, or simply pentons and hexons. Given that the icosahedron is derived from the planar hexagonal lattice, not every combination of pentamers and hexamers is allowed; conversion of 12 non-adjacent hexamers into pentamers to form a closed shell must take place at regular, correctly separated intervals (Fig. 2.5a). The ratio that systematically defines all possible icosahedra from the hexagonal lattice is termed triangulation number (T), which matches the unit triangles of the elemental planar lattice contained in each triangular face of the icosahedron (facet) (Fig. 2.5a). The smallest viruses follow $T = 1$ geometry. The rules of quasi-equivalence allow only certain values of T; T is given by the formula $T = h^2 + k^2 + hk$, in which h and k are positive integers that define the pentamer positions in the original hexagonal lattice (Fig. 2.5a). Icosahedral viruses larger than $T = 1$ have T values of 3, 4, 7, 9, 12, 13, 16, 19, 21, 25, 27, 28, and so on.

The quasi-equivalence theory, combined with the concept of T, provides a systematic nomenclature with which to describe icosahedral capsid structure. Icosahedra fulfill Euler's polyhedral formula, faces (20) + vertices (12) = edges (30) + 2. Taking T into account, this formula becomes $20\ T + V = 30\ T + 2$. As there must always be 12 pentameric vertices, the total number of hexamers is 10 (T-1), and the total number of subunits is 60 T. T also describes the classes of conformers, or different quasi-equivalent (similar) environments occupied by a subunit (Fig. 2.5b); however, T does not describe the basic morphological unit or

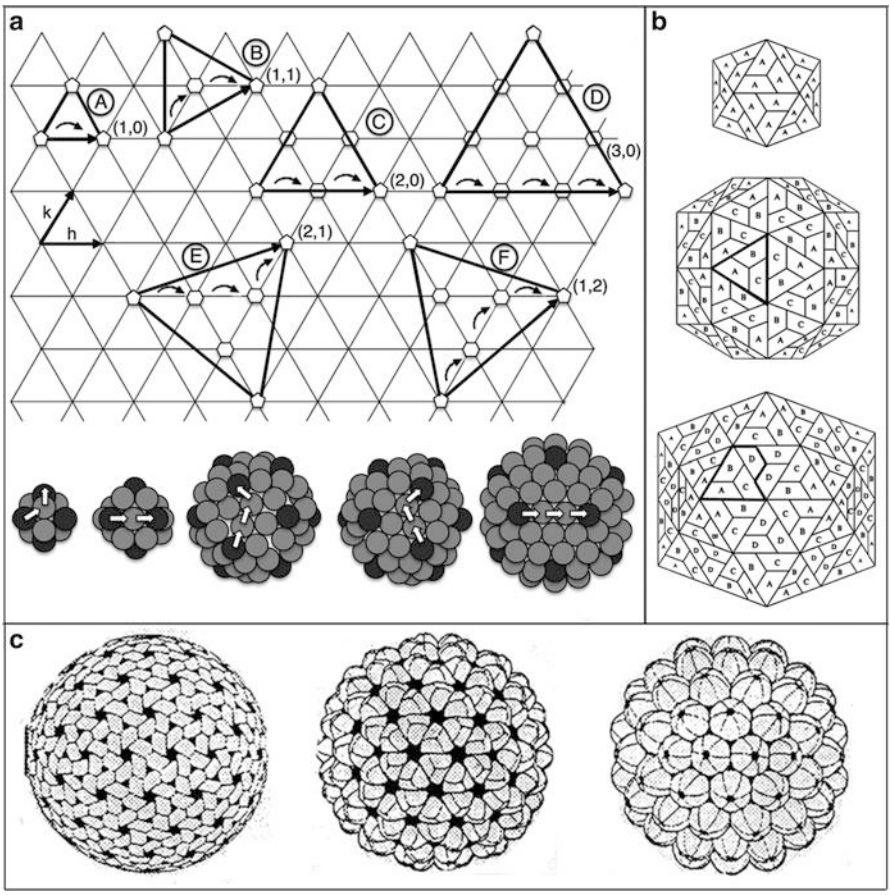

Fig. 2.5 The triangulation number T, and how it is determined in an icosahedral lattice. (**a**) The coordinate system of a planar hexagonal net with axes h and k crossing at a 60 º angle is formed by equilateral triangles; each triangle is formed by three protein subunits (as in Fig. 2.3), and specific hexamer-to-pentamer conversion introduces curvature into the hexamer sheet (see main text). Capsomers are located at the lattice intersection points; hexagons represent hexamers (or planar hexameric positions) and pentagons correspond to (*convex*) pentamers. Several triangular facets of an icosahedron are indicated (*bold outline*), and a facet edge of the icosahedron (*bold vector*) connects the two closest neighboring pentamers. The two unit vectors of the underlying hexagonal net, h and k, define the coordinates of the vector relative to its origin (positions at which hexamers are replaced by pentamers), and determine the parameters of the icosahedral capsid. The number of steps or spacings (*curved arrows*) from one structural unit (a hexamer or pentamer) to the adjacent one is therefore an integer. In example A, a pentamer is separated from the adjacent pentamer by one step along the h axis (*i.e.*, pentamers are in direct contact); hence, $h = 1$, $k = 0$, which corresponds to a $T = 1$ icosahedron (in which case 12 adjacent hexamers are converted to pentamers). Larger icosahedra are built by converting 12 nonadjacent hexamers to pentamers at regular, precisely-spaced intervals. In example B, one pentamer is separated from another by one step along the h axis and one step along the k axis; $h = 1$, $k = 1$, and $T = 3$. Note that T indicates the number of unit triangles contained in the icosahedral facet. In examples C and D, pentamers are separated by two and three steps along the direction of the h axis, and are $T = 4$ and $T = 9$,

repeated oligomer (other than hexamer/pentamer) that is actually used to build the capsid (the capsid building block), which must be determined empirically for each icosahedral virus (Fig. 2.5c).

Viral capsids with T numbers corresponding to $h \neq k$ (and both different from zero) result in skewed lattices with left- ($h > k$) or right-handed ($h < k$) possible configurations (see Fig. 2.5a). For example, for the skewed lattice with the smallest T number the left-handed (T = 7 *laevo* or 7 *l*) and right-handed (T = 7 *dextro* or 7*d*) configurations correspond to (h, k) arrangements (2, 1) and (1, 2), respectively. Besides the possible lattice handedness viral capsids are, as any protein assembly, enantiomeric structures. Therefore, the organization of capsomers (dimers or pentamers) or subunits might show a conspicuous handedness even in nonskewed capsids such as T = 1 or T = 3 (Fig. 2.5a).

Paradoxically, a T = 1 capsid with only one protein subunit is the least frequently observed quaternary structure in native viruses [13]. Only a few small ssRNA viruses with protein subunits of 20–40 kDa, such as satellite tobacco necrosis virus (STNV; 159-residue CP), are known to have this capsid type. Other T = 1 capsids are formed by ssDNA viruses such as parvoviruses and microviruses (*e.g.*, bacteriophage ϕX174) with CP as large as 75 kDa ([26] and references therein).

High-resolution structural studies have provided some clues as to how identical subunits adopt different conformations in a T > 1 icosahedral capsid [8, 9, 13, 27]. The distinct conformational states may be controlled by order/disorder of flexible regions of the protein (*e.g.*, in loops or N- and C- termini), dsRNA or ssRNA segments of the viral nucleic acid, metal ions, protons (pH), or combinations of any of

Fig. 2.5 (continued) respectively. A T = 4 lattice consists of 80 unit triangles, 240 subunits, and 4 quasi-equivalent protein subunits in the icosahedral asymmetric unit. Triangulated icosahedra might have local (or quasi) six-, three- and/or twofold symmetry. When the coordinates h or k are zero, or h = k, the icosahedral face is symmetric relative to the coordinate system (examples A, B, C and D). When h differs from k, the capsid is asymmetric or skewed, as in examples E and F. In a T = 7 icosahedron, the icosahedral face is turned relative to the hexagonal lattice, but the surfaces of the unit triangles gained or lost are compensated and T is a integer. The handedness of a T = 7 capsid is defined by the separation steps between the two closest pentamers, giving the maximum number of steps in the selected direction (this movement is similar to that of a knight in chess). In the T = 7laevo (left-handed) or 7 *l* (as in E), h = 2 and k = 1 (this step is leftward, along the direction of the k axis); similarly, in the T = 7dextro (right-handed) or 7*d* (as in F), k = 2 and h = 1 (a rightward step, along the direction of the h axis). The two T = 7 lattices are enantiomorphs, *i.e.*, mirror images of each other. The three-dimensional models (*bottom*) show how T number is determined; capsomers are shown as balls (*black*, *pentamer*; *grey*, *hexamer*). Note that capsid size is increased at higher T numbers, as capsomers are the same size. (**b**) Icosahedra corresponding to the triangulation numbers T = 1 (*top*), 3 (*middle*) and 4 (*bottom*); CP subunits are chemically identical (shown as trapezia). In a T = 1 capsid, every CP lies in an identical environment (labeled A); in a T = 3 capsid, the icosahedral asymmetric unit contains three CP subunits that occupy slightly different geometrical environments (A, B and C; one such unit is shown in thick lines); in a T = 4 capsid, each asymmetric unit contains four CP subunits (A, B, C and D). (**c**) Three different capsids with the same T number (T = 9) but different capsomers, viewed from a twofold axis; structural subunits are assembled without forming oligomers (*left*), as trimers (*middle*), or as pentamers and hexamers (*right*)

these [28]. The structural elements that participate in different, quasi-equivalent contacts at geometrically different positions in a $T > 1$ capsid are referred to as molecular switches. Removal of the molecular switching mechanism results in assembly of aberrant or non-native structures. In theory, quasi-equivalence involves small differences in subunit interactions and conformations; in practice, for capsids analyzed to date, viruses show broad variation, and the literature describes equivalent, quasi- and non-equivalent capsids. Whereas the structure of cowpea chlorotic mottle virus (CCMV) at atomic resolution showed that building blocks have quasi-equivalent contacts, with allowed variations in bond angles for noncovalent intersubunit interactions that stabilize oligomers, the two classes of dimeric building blocks in TBSV have completely different (strictly non-equivalent) bonding contacts [26].

Quasi-equivalence, which involves distinct spatial conformations, is supported by the intrinsic structural polymorphism of the CP. Structural polymorphism is also fundamental to the maturation process, since it allows in situ transient conformational changes throughout morphogenesis [29] (see Chap. 13). This strategy is optimal, as no additional genetic information is needed to generate these structural variants of the capsid.

Despite the flexibility and multiple possibilities that quasi-equivalence offers, capsid assembly frequently requires additional structural proteins. In many icosahedral viruses made of more than 60 subunits, quasi-equivalence in the strict sense of the term and molecular switches are not used at all. In these viruses, quasi-equivalent positions are occupied by different proteins, as in picornaviruses (see Chap. 10). The picornavirus capsid has 60 copies each of three subunits (VP0 (later processed to VP4 + VP2), VP1, VP3) that are synthesized as a precursor polyprotein (see Chap. 10). There is very little sequence similarity among them, but they all have the same fold, the canonical β-barrel. The additional genetic cost provides advantages difficult (if not impossible) to achieve only by conformational changes in a single protein, such as rapid acquisition of mutations to evade the immune system without altering capsid assembly [11–13]. When capsids are assembled from different gene products, T numbers are no longer valid, and pseudoT (P) numbers are used instead (for example, $P = 3$ in picornaviruses). Another well-known group with $P = 3$ capsids are the comoviruses, whose capsid has two subunits, a small subunit with a β-barrel domain and a large subunit formed by fusion of two β-barrels [18].

Molecular switches may be also insufficient to control assembly of capsids with large T numbers and/or when complex viruses are considered. These capsids may require one or more auxiliary proteins (scaffold, minor capsid or enzymatic proteins) that act as morphogenic factors to trigger structural changes [2, 30] (see Chap. 11).

There are many other variations on the icosahedral theme. For example, geminivirus particles consist of two fused $T = 1$ icosahedra, in which each lacks a pentameric cap [31]. In TBSV capsids, among others, particle curvature results from both pentamers and hexamers [10]. These variations are not restricted to capsids based on the tangential wedge-like β-barrel domain; the CP of HBV forms cores of varying T number and does not share this fold. The proportion of HBV $T = 3$ and $T = 4$ particles is affected by modifying the length of the C-terminus, which is located on the inner surface [32].

There are notable exceptions [30] that do not adhere to quasi-equivalence rules, such as the all-pentamer T = 7d capsid of papovaviruses (papilloma- and polyomaviruses), formed by 72 pentamers (360 subunits and six conformers, which would correspond to a disallowed T = 6) rather than 12 pentamers and 60 hexamers (420 subunits and seven conformers as predicted by quasi-equivalence) [33, 34]. In adenoviruses, which have a T = 25 shell (in fact, a pseudoT = 25, as pentameric subunits differ from hexameric subunits), each hexon is a trimer [14], but each subunit is formed by two similar domains with the β-barrel fold (see Chap. 11). A T = 25 capsid would have 25 conformers, but as the hexameric trimer is the basic unit, only four classes of (trimeric) conformers are actually needed. Another striking exception is the inner core of dsRNA viruses (formally, T = 1) formed by 12 decamers, in which dimers are the asymmetric unit; these capsids would correspond to the forbidden "T = 2" [35–37]. A number of studies have firmly established that this "T = 2" architecture is a common feature for dsRNA viruses ([38] and references therein).

Structural comparison of CP is used to establish evolutionary relatedness between viruses (or at least between their capsid genes) when sequence conservation is limited [13, 39, 40]; such comparisons led to detection of relationships among icosahedral viruses that infect hosts from three different domains of life (bacteria, eukarya and archaea). Icosahedral viruses are currently grouped in four separate lineages, each with common structural elements and assembly principles: these lineages are represented by the PRD1 bacteriophage, HK97 bacteriophage, bluetongue virus (BTV), and picornavirus-like viruses [41] (see also Chap. 7). The PRD1-like lineage includes adenovirus and viruses that infect archaea, such as STIV; the HK97-like lineage has the tailed dsDNA bacteriophages and herpesviruses, and the BTV-like lineage includes the icosahedral cores of dsRNA viruses and phage ϕ6. Although many viruses are not included in these four lineages, the number of folds that satisfy the assembly constraints for a viable viral shell is nonetheless thought to be limited.

2.5 Variations on the Icosahedral Capsid Theme: Multiple Layers and Prolate Icosahedra

Most animal viruses are icosahedral and follow the general principles described above, although many viruses have additional non-symmetric envelopes. The simplest icosahedral viruses include RNA and DNA viruses. Outstanding examples of RNA viruses belong to the *Reo-*, *Birna-*, *Calici-*, *Noda-*, *Tetra-* and *Picornaviridae* families; examples of DNA viruses belong to the *Parvo-*, *Papova-* and *Adenoviridae* families. Some icosahedral viruses have locally broken the symmetry at one of 12 fivefold vertices, due to the presence of the portal and tail complexes in place of a pentamer. To illustrate the variation in size and structural features among different viral capsids, we show 3D reconstructions of several distinct viruses calculated from cryo-EM images, from the 28 nm-diameter capsid of picornaviruses to the ~100 nm rotavirus capsid (Fig. 2.6). 3D maps are

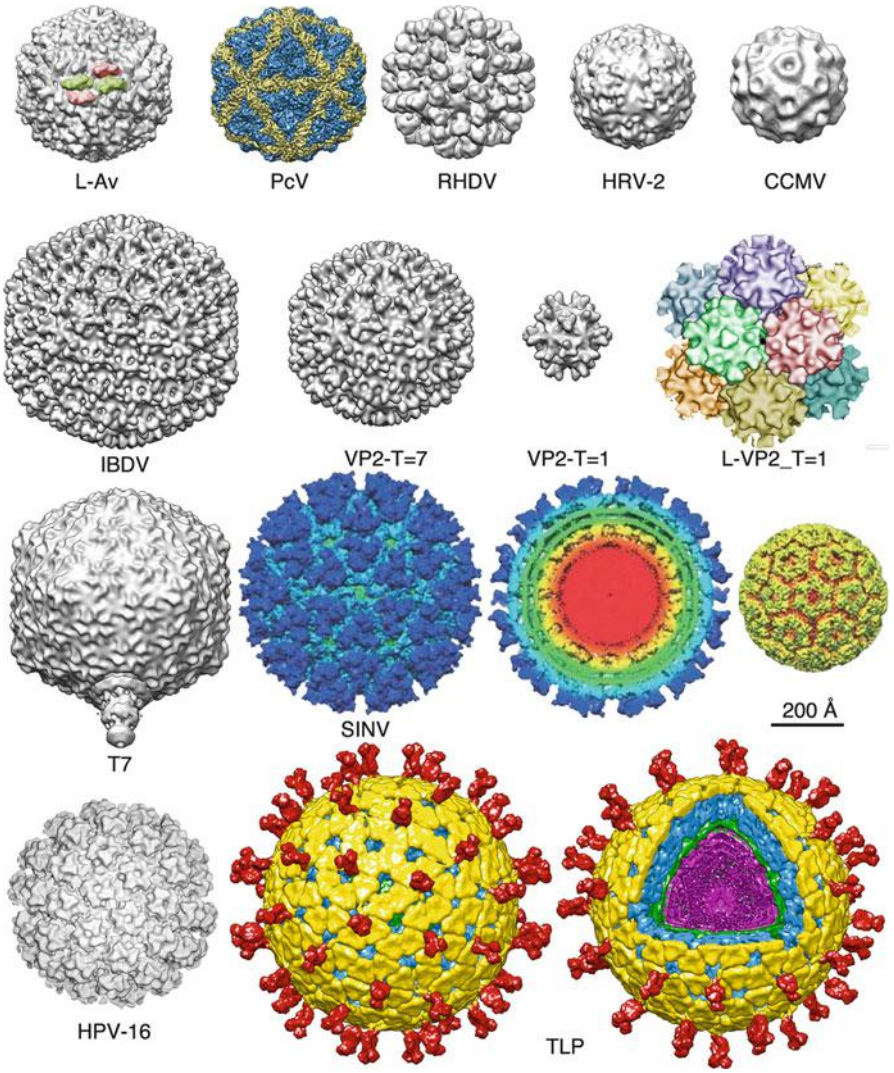

Fig. 2.6 Three-dimensional maps of icosahedral viruses calculated from two-dimensional cryo-electron microscopy images. L-Av, L-A virus of the yeast *Saccharomyces cerevisiae* with a T = 1 (or "T = 2") capsid (430 Å diameter); the basic unit is an asymmetric dimer (*green and red*) that assembles into decamers ([58], PDB entry 1M1C); PcV, Penicillium chrysogenum virus, T = 1, 400 Å; the basic unit is a monomer that has two similar domains (*yellow and blue*) (Adapted from [38]. Reproduced with permission); RHDV, rabbit hemorrhagic disease virus, T = 3, 400 Å; the basic unit is a dimer of the capsid protein (Adapted from [59]. Reproduced with permission); HRV-2 human rhinovirus 2, P = 3, 304 Å; the asymmetric unit is made of three different proteins (unpublished data; D. Blass and J.R. Castón); CCMV cowpea chlorotic mottle virus, T = 3, 300 Å (unpublished data; A. de la Escosura and J.R. Castón). CCMV capsid has 90 CP dimers (*protomers*) that assemble into 12 pentamers and 20 hexamers (*i.e.*, 32 capsomers or morphological units); IBDV infectious bursal disease virus, T = 13 *l* lattice, ~700 Å; the building block is a VP2 trimer (Adapted from [45, 60]. Reproduced with permission); VP2 can assemble into

intentionally presented at low to medium resolution (some are available at much higher resolution) to emphasize the capsomers (di-, tri-, penta- or hexamer) as well as the T numbers. Detection of specific capsomers implies stronger intra- than intercapsomeric interactions; isolated capsomers can be purified for structural and functional studies.

Many dsRNA viruses, including important human, animal and plant pathogens (rota-, reo-, and orbiviruses of the *Reoviridae* family) consist of a multilayered concentric capsid with icosahedral symmetry [42]. This complex organization is based on the role of each protein shell. Outer layers, usually a T = 13 shell, provide protection against chemical, physical and enzymatic damage and participate in receptor interaction and cell entry. The innermost core, a T = 1 shell with a protein dimer as the asymmetric unit, is essential for genome and replicative complex organization, participates in genome transcription and replication, and isolates the virus genome from host defense mechanisms. Fungal dsRNA viruses, which lack an extracellular route, have only a T = 1 capsid.

Capsids with larger T numbers might be considered advantageous, as formation of larger containers would in turn allow the incorporation of larger genomes; the concomitant packaging of specific viral components, such as polymerases and other enzymes, would reduce virus dependence on host cell machinery. Such a container, made at low economic cost, nonetheless presents enormous challenges. For example, in a quasi-equivalent T = 25 capsid that roughly emulates the adenovirus capsid, 1,500 copies of an individual protein must be laid accurately into the shell. In a relatively simpler T = 13 lattice, as in the birnavirus capsid, "only" 780 VP2 proteins must be arranged on the outer capsid surface (see Chap. 13). The VP2 subunit has three domains, an outer domain with an β-barrel fold in which β-strands are oriented radially, a shell domain with a β-barrel in tangential orientation, and an interior helix bundle contributed by the N- and C- termini [43]. The protein distribution problem is reduced in several ways: VP2 subunits form trimers, and five classes of these trimers are needed (rather than 13 conformers) to assemble into pentamers and hexamers. VP2 structural polymorphism is controlled by an inherent molecular switch, a C-terminal segment, that is absent in the mature VP2 form. Proteolytic processing of this switch is a finely tuned process in which viral

Fig. 2.6 (continued) T = 7 (VP2-T = 7; the left-handed form is chosen arbitrarily) and T = 1 (VP2-T = 1) capsids; T = 1 capsids self-assemble into a dodecahedron with 12 T = 1 capsids (L-VP2_T = 1); T7, T7 bacteriophage head, T = 7, ~600 Å [61]; the map shows a tail, located at the same vertex where the connector core is assembled; SINV Sindbis virus, an alphavirus with a double-layered particle, ~700 Å: outer T = 4 capsid (*blue*), a lipid bilayer (*green*), and an inner T = 4 capsid (*yellow*) (Adapted from [62]. With permission); HPV-16 human papillomavirus 16, T = 7d, ~600 Å; this capsid is made of 72 pentameric capsomers ([63]. Courtesy of BL Trus). TLP a rotavirus with a triple-layered particle, ~700 Å (without spikes): VP4 spikes (*red*), VP7 outermost T = 13 *l* layer (*yellow*), VP6 intermediate T = 13 *l* layer (*blue*), VP2 innermost T = 1 layer (*green*), and the internal density of dsRNA and polymerase complexes (*violet*) (unpublished data, courtesy of D. Luque and J.M. Rodríguez)

and cellular proteases participate; these proteolytic events take place in a procapsid-like structure, in which multiple scaffolding protein copies are essential for stabilizing or promoting interactions between precursor subunit trimers within hexamers [44–46].

The largest icosahedral viruses analyzed by cryo-EM include the insect larvae iridovirus CIV (T = 147, 185 nm diameter [47]), the algal dsDNA viruses *Paramecium bursaria chlorella* virus 1 (PBCV-1, T = 169d, 190 nm [48]) and *Phaeocystis pouchetii* virus (PpV01, T = 219d, 220 nm [49]), and the amoebal dsDNA mimivirus (estimated T between 972 and 1,200, 750 nm, [50]); all have a lipid membrane beneath the capsid and use minor CP for stabilization and to determine their assembly. Mimivirus can be "infected" by its own virus, Sputnik (with a T = 27 icosahedral capsid), and its genome size is similar to that of intracellular parasitic bacteria (see Chaps. 11 and 14).

A different strategy to increase the internal volume of the capsid is followed by some icosahedral-like viruses, which elongate the capsid along a fivefold axis with additional subunits in the central band or cylinder. This prolate icosahedral capsid is typical of bacteriophages such as ϕ29 [51] and T4 [52] (with a portal –the DNA packaging machine– and a tail –the DNA injecting machine– in a pentameric vertex), but is also found in several fungus, plant and animal viruses.

Consider an icosahedron made of three parts, the top and bottom caps (with five equilateral triangular facets) and the central band or cylinder (with ten triangles) that connect the caps. These capsids are characterized by two numbers, the T number and the elongation number Q (also referred to as T_{end} and T_{mid}, respectively). As the T number specifies the unit triangles of the triangular facet of an icosahedron (subdivisions into smaller triangles), the Q number also indicates the number of unit triangles in the stretched triangles of the central band, but can be any positive integer (Fig. 2.7). For icosahedral viruses, the total number of subunits is defined by n = 60 T; for prolate capsids, the total number of protein subunits is given by n = 30(T + Q). An isometric, icosahedral particle has Q = T; a prolate particle has Q > T. In the case of phage T4, the capsid can be defined by T = 13 *l* and Q = 20 [52]; in phage ϕ29, T = 3 and Q = 5 [51].

Assembly of prolate capsids based on the elongation of an icosahedron along a threefold or twofold axis of symmetry has also been suggested for some bacilliform viruses such as AMV; geometrical rules have been established for their construction [53].

2.6 Helical Capsids

Helical symmetry facilitates assembly of nucleoprotein complexes without size limitations, as variable-sized genome could, in principle, be coated by identical protein subunits that surround it, forming a cylinder with helical symmetry ([54], and references therein). Although helical symmetry simplifies capsid assembly and offers flexibility for genome packaging, this assembly is relatively infrequent. Less

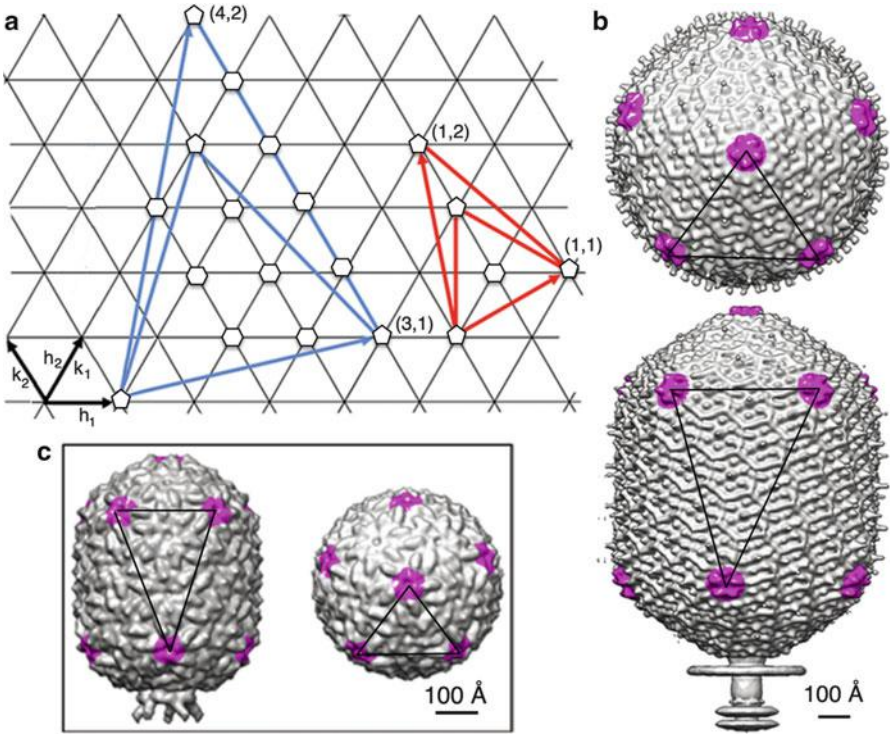

Fig. 2.7 Determination of the triangulation number (T) and the elongation number (Q) in a prolate icosahedral lattice. (**a**) A planar equilateral triangular lattice (as in Fig. 2.5) is used to show the triangulation numbers, T (or T_{end}) and Q (or T_{mid}), of prolate capsids; two unit vector systems, (h_1,k_1) and (h_2,k_2), are used to represent the equilateral triangular facets at the end caps and the elongated triangular facets at the central band (see text). The large equilateral triangle (*blue*) represents an end cap facet of a bacteriophage T4 head, and is defined by the (h_1, k_1) coordinate point [the (h_1, k_1) lattice point corresponds to a (h, k) point in Fig. 2.5]; $h_1 = 3$, $k_1 = 1$, and $T_{end} = h_1^2 + h_1k_1 + k_1^2 = 13$ *l*. The large triangle with unequal sides (*blue*) represents the phage T4 midsection facets and is described by two vectors, one from the origin to the lattice points (h_1, k_1) and the other to (h_2,k_2); $h_2 = 4$, $k_2 = 2$, and the triangulation of the extended facet is $Q = T_{mid} = h_1h_2 + h_1k_2 + k_1k_2 = 20$. For bacteriophage ϕ29, the equilateral facets at the end caps (*red*) are defined by $h_1 = 1$, $k_1 = 1$, and $T_{end} = 3$, and the stretched facets at the midsection triangles by $h_2 = 1$, $k_2 = 2$, and $Q = 5$. (**b**) Surface-shaded representation of the 3D cryo-EM map of the T4 bacteriophage head ([52], cryo-EM map deposited at the EMDataBank; ID EMD_1075), viewed along the fivefold axis of an end cap (*top*) and perpendicular to the fivefold axis (*bottom*). The equilateral and stretched triangles show the icosahedral equilateral and elongated facets defined by pentameric vertices (*pink*). **c** 3D map of ϕ29 bacteriophage prohead ([51], EMD_1117) showing equivalent views as for T4 phage

than 10 % of viral families have helical symmetry and most of them are plant viruses and bacteriophages. The most evident problems associated with this geometry are the disadvantageous surface/volume ratio and the structural instability inherent to long, thin structures. The paradigm of helical viruses is the tobacco mosaic virus (TMV), the first virus to be identified. A helix is defined by its pitch

(P), which includes two parameters: the number of subunits per helix turn (μ) and axial rise per subunit (ρ); thus, P = μ x ρ. For TMV, pitch is 23 Å, μ = 16.3 and ρ = 1.4 Å. TMV particles are rigid, right-handed, rod-like structures, 300 nm long and 18 nm in diameter, with a central 4 nm-diameter hole. Other helical viruses however, such as the Inoviridae (phages M13 and fd) are longer and flexible.

The genome of many animal negative-sense ssRNA viruses is also protected by a helical capsid which, together with the enclosed nucleic acid, forms the nucleocapsid, which is in turn invariably surrounded by a lipid envelope. Unlike many phage and plant viruses, no known animal viruses have a nonenveloped helical nucleocapsid. To provide some examples, rhabdoviruses (rabies and stomatitis vesicular viruses) and paramyxovirus (Sendai, mumps and measles viruses) have left-handed helical nucleocapsids surrounded by a lipid envelope. Orthomyxovirus (influenza virus) have multisegment helical nucleocapsids, referred to as ribonucleoproteins (RNP). As for most enveloped viruses, the influenza virus nucleocapsids are surrounded by an amorphous layer formed by the matrix protein, which also interacts with the external lipid envelope.

2.7 The Viral Nucleic Acid Inside

The mechanisms of nucleic acid packaging and related aspects will be reviewed in depth in Chap. 12. We will focus here only on some general aspects of the organization of the viral nucleic acid in the capsid from the point of view of virus architecture. Whereas the assembly of CP subunits to build a viral capsid follows a few well-defined strategies, and in most cases yields a unique geometric solution as outlined above, the organization of the nucleic acid inside the capsid is more diverse, and is greatly conditioned by capsid geometry. In viruses with helical symmetry, such as the rigid rod-shaped plant viruses (*e.g.*, TMV), ssRNA is co-assembled with the CP, integrated such that it strictly follows the geometrical characteristics imposed by the CP assembly. Other DNA-containing helical viruses (like phage M13) also show a close structural relationship between nucleic acid and protein. In this case, the fit is less tight than in TMV-like viruses, and is based on defined contacts between the DNA and the inner walls of the helical protein shell. A different type of viral helical protein-nucleic acid arrangement is found in flexible nucleoprotein complexes characteristic of Mononegavirales (*Borna-*, *Filo-*, *Paramyxo-* and *Rhabdoviridae* families). These helical assemblies, which are enclosed within lipid-based envelopes, work as an active complex for replication and transcription.

Nucleic acid arrangement in icosahedral capsids shows a wide variety of solutions, from partly disordered to highly compact, quasi-crystalline organization. In many ssRNA viruses, most of the RNA (up to 80 %) is disordered, as the molecules interact with the icosahedral capsid and use only small nucleotide stretches with specific capsid symmetries (as in the case of satellite tobacco mosaic virus or flock house virus. These ordered nucleotide stretches are important, as they

are probably involved in the assembly of the CP subunits by directing and/or stabilizing the assembly intermediates (Chap. 12). At the other extreme, dsDNA bacteriophages show very high nucleic acid density inside the capsid (>500 mg/ml). In these cases, viruses have incorporated a sophisticated machinery to package the DNA molecule as a tight coil that forms concentric shells inside the capsid, with the DNA strands arranged in a parallel pseudo-hexagonal crystalline array (Chap. 12).

2.8 Basic Architecture of Enveloped Viruses

We will refer here briefly to the general structural organization of enveloped viruses in which a lipid membrane constitutes the outermost structural layer. For detailed descriptions of different aspects of the structure and function of enveloped virus, see Chaps. 11, 15, and 16.

Capsid or nucleocapsid assembly is not necessarily the final step in virion morphogenesis, as some viral capsids or nucleocapsids become surrounded by one or several membrane envelopes of cellular origin, in which viral proteins are embedded. The lipid bilayer is acquired during the budding from a cell membrane, which can be the plasma membrane during escape from host cells, or membrane-delimited structures in eukaryotic cells such as the nuclear membrane, endoplasmic reticulum, Golgi apparatus and vacuoles. Enveloped viruses can exit the cell without causing cell death, since the budding process does not alter cell membrane integrity, and can cause either persistent or latent infections. Viral glycoproteins are inserted in the membranes *via* the compartmentalization routes and constant flux of vesicular trafficking. Glycoproteins have an external domain (ectodomain) with sugar chains, which is anchored in the membrane by a transmembrane hydrophobic segment (typically α-helix) that terminates in a small domain (endodomain) facing the cytoplasm. Viral glycoprotein monomers associate to form multimeric spikes (visible by electron microscopy), which are the functional state for at least four activities: identification and binding to cell receptors, membrane fusion, transcription activation, and receptor destruction. These functions can be carried out by one or several glycoproteins. Envelopes studded with glycoproteins usually have minor proteins that make ionic channels. Finally, matrix proteins are internal proteins that mediate association of glycoprotein endodomains with the nucleocapsid (see Chap. 16).

Membrane acquisition differs among viruses. Alpha- and flaviviruses, which have a positive-sense ssRNA genome, have a nucleocapsid that is assembled independently in the host cytoplasm and subsequently interacts with the glycoprotein endodomains. In alphaviruses (of the *Togaviridae* family), lateral interactions among glycoproteins exclude cell membrane proteins, resulting in a three-layered capsid: a glycoprotein outer icosahedral layer, a lipid bilayer, and a protein inner icosahedral layer (both icosahedral shells are in register with a $T = 4$ geometry [55]). Flaviviruses have an outer ordered protein layer above a polygonal lipid bilayer, and an inner disordered nucleocapsid; the outer capsid is formed by 90 closely packed glycoprotein dimers that lie flat on the particle surface, rather than forming protruding

spikes. Dimer arrangement differs from a true T = 3 shell, however, and the capsid should be considered a T = 1 capsid with three monomers per asymmetric unit ([56] and references therein). Whereas budding takes place in the plasma membrane for alphaviruses, flavivirus buds in the endoplasmic reticulum lumen.

Contrary to the well ordered, quasi-crystalline arrangement found in icosahedral capsids, most enveloped viruses lack regular organization in the membrane. Orthomyxoviruses (influenza virus) have an additional protein, the matrix protein, which mediates interactions between helical nucleocapsids (in eight segments) and the viral envelope. Retroviruses have a similar membrane-associated matrix protein. Influenza virus membrane has two glycoproteins, neuraminidase, which cleaves sialic acid groups from receptor glycoproteins, and hemagglutinin, a trimer that contains the binding site for the virus receptor and is responsible for viral and cell membrane fusion. Rhabdo- and paramyxoviruses are similar to orthomyxoviruses, except that the genome is a single helical nucleocapsid. Bunyaviruses have a tripartite genome and are very flexible, since they lack matrix proteins; they have glycoprotein spikes, some ordered in a T = 12 capsid, some locally ordered, and some lacking any order. Herpesvirus virions comprise a T = 16 icosahedral capsid (125 nm diameter) and interact with the envelope (derived from the trans-Golgi network) through a complex asymmetric layer with many viral (and a few cellular) proteins, called the tegument (virion size, 200 nm diameter). The herpesvirus envelope has 12 different glycoproteins and about 800 spikes per virion. Poxviruses (or brick-shaped viruses) such as vaccinia virus encompass multiple structural elements, and are described in detail in Chap. 11.

Retroviruses, such as the lentivirus HIV, are enveloped pleomorphic virions. HIV is about 100 nm in diameter and has two copies of a positive sense ssRNA genome associated with the nucleoprotein basic protein (NC). The dimeric ribonucleoprotein complex is inside a protein shell built by about 1,500 copies of a single protein, CA, which is assembled into a cone-shaped capsid in the mature HIV virion. There is a matrix protein (MA) layer that interacts with the inner leaflet of the viral envelope (MA is a peripheral membrane protein). MA, CA and NC are synthesized as a single precursor, Gag, that is arranged radially in immature, incomplete spherical particles. Immature HIV virions bud at lipid rafts; these membranes have viral spikes formed by the surface (gp120) and transmembrane (gp41) glycoproteins, which are also initially fused. Virions mature after budding by the concomitant activation of the viral protease. CA forms hexamers on a variably curved hexagonal lattice closed by insertion of 12 pentamers, seven at the wide and five at the narrow end of the fullerene cone. CA has two domains, CA^{NTD} and CA^{CTD}; in the mature capsid, CA^{NTD} forms hexameric and pentameric rings, whereas CA^{CTD} dimeric contacts, at base of the protein shell, connect neighboring rings ([57] and references therein) (see also Chap. 5).

Acknowledgements We thank Daniel Luque for preparing several figures used in this chapter and Catherine Mark for editorial help. This work was supported by grants BFU2011-25902 (to JRC) and BFU2011-29038 (to JLC) from the Spanish Ministry of Science and Innovation.

References and Further Reading

1. Harrison SC (2007) Principles of virus structure. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA, Roizman B, Strauss SE (eds) Fields virology, vol 1. Lippincott Williams & Wilkins, Philadelphia, pp 59–98
2. Dokland T (2000) Freedom and restraint: themes in virus capsid assembly. Structure Fold Des 8:R157–R162
3. Flint SJ, Enquist LW, Krug RM, Racaniello VR, Skalka AM (2000) Principles of virology. Molecular biology, pathogenesis, and control. ASM Press, Washington, DC
4. Cann AJ (2012) Principles of molecular virology. Elsevier, Amsterdam
5. Baker TS, Olson NH, Fuller SD (1999) Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. Microbiol Mol Biol Rev 63:862–922
6. Prasad BV, Schmid MF (2012) Principles of virus structural organization. Adv Exp Med Biol 726:17–47
7. Hendrix RW, Johnson JE (2012) Bacteriophage HK97 capsid assembly and maturation. Adv Exp Med Biol 726:351–363
8. Chang J, Liu X, Rochat RH, Baker ML, Chiu W (2012) Reconstructing virus structures from nanometer to near-atomic resolutions with cryo-electron microscopy and tomography. Adv Exp Med Biol 726:49–90
9. Zhou ZH (2011) Atomic resolution cryo electron microscopy of macromolecular complexes. Adv Protein Chem Struct Biol 82:1–35
10. Harrison SC, Olson AJ, Schutt CE, Winkler FK, Bricogne G (1978) Tomato bushy stunt virus at 2.9 A resolution. Nature 276:368–373
11. Hogle JM, Chow M, Filman DJ (1985) Three-dimensional structure of poliovirus at 2.9 A resolution. Science 229:1358–1365
12. Rossmann MG, Arnold E, Erickson JW, Frankenberger EA, Griffith JP, Hecht HJ, Johnson JE, Kamer G, Luo M, Mosser AG et al (1985) Structure of a human common cold virus and functional relationship to other picornaviruses. Nature 317:145–153
13. Rossmann M, Johnson J (1989) Icosahedral RNA virus structure. Annu Rev Biochem 58:533–573
14. Roberts MM, White JL, Grutter MG, Burnett RM (1986) Three-dimensional structure of the adenovirus major coat protein hexon. Science 232:1148–1151
15. Nandhagopal N, Simpson AA, Gurnon JR, Yan X, Baker TS, Graves MV, Van Etten JL, Rossmann MG (2002) The structure and evolution of the major capsid protein of a large, lipid-containing DNA virus. Proc Natl Acad Sci U S A 99:14758–14763
16. Abrescia NG, Cockburn JJ, Grimes JM, Sutton GC, Diprose JM, Butcher SJ, Fuller SD, San Martin C, Burnett RM, Stuart DI, Bamford DH, Bamford JK (2004) Insights into assembly from structural analysis of bacteriophage PRD1. Nature 432:68–74
17. Fu CY, Johnson JE (2012) Structure and cell biology of archaeal virus STIV. Curr Opin Virol 2:122–127
18. Lomonossoff GP, Johnson JE (1991) The synthesis and structure of comovirus capsids. Prog Biophys Mol Biol 55:107–137
19. 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
20. Wikoff WR, Liljas L, Duda RL, Tsuruta H, Hendrix RW, Johnson JE (2000) Topologically linked protein rings in the bacteriophage HK97 capsid. Science 289:2129–2133
21. Veesler D, Johnson JE (2012) Virus maturation. Annu Rev Biophys 41:473–496
22. Zlotnick A (2005) Theoretical aspects of virus capsid assembly. J Mol Recognit 18:479–490
23. Johnson JE (2008) Multi-disciplinary studies of viruses: the role of structure in shaping the questions and answers. J Struct Biol 163:246–253
24. Crick FH, Watson JD (1956) Structure of small viruses. Nature 177:473–475
25. Caspar DLD, Klug A (1962) Physical principles in the construction of regular viruses. Cold Spring Harbor Symp Quant Biol 27:1–24

26. Liljas L (1986) The structure of spherical viruses. Prog Biophys Mol Biol 48:1–36
27. Johnson JE, Speir JA (1997) Quasi-equivalent viruses: a paradigm for protein assemblies. J Mol Biol 269:665–675
28. Johnson JE (1996) Functional implications of protein-protein interactions in icosahedral viruses. Proc Natl Acad Sci USA 93:27–33
29. Chen DH, Baker ML, Hryc CF, DiMaio F, Jakana J, Wu W, Dougherty M, Haase-Pettingell C, Schmid MF, Jiang W, Baker D, King JA, Chiu W (2011) Structural basis for scaffolding-mediated assembly and maturation of a dsDNA virus. Proc Natl Acad Sci USA 108:1355–1360
30. Steven A, Trus B, Booy F, Cheng N, Zlotnick A, Castón J, Conway J (1997) The making and breaking of symmetry in virus capsid assembly: glimpses of capsid biology from cryoelectron microscopy. FASEB J 11:733–742
31. Zhang W, Olson NH, Baker TS, Faulkner L, Agbandje-McKenna M, Boulton MI, Davies JW, McKenna R (2001) Structure of the Maize streak virus geminate particle. Virology 279:471–477
32. Zlotnick A, Cheng N, Conway JF, Booy FP, Steven AC, Stahl SJ, Wingfield PT (1996) Dimorphism of hepatitis B virus capsids is strongly influenced by the C-terminus of the capsid protein. Biochemistry 35:7412–7421
33. Rayment I, Baker TS, Caspar DL, Murakami WT (1982) Polyoma virus capsid structure at 22.5 A resolution. Nature 295:110–115
34. Liddington RC, Yan Y, Moulai J, Sahli R, Benjamin TL, Harrison SC (1991) Structure of simian virus 40 at 3.8-A resolution. Nature 354:278–284
35. Grimes JM, Burroughs JN, Gouet P, Diprose JM, Malby R, Zientara S, Mertens PP, Stuart DI (1998) The atomic structure of the Bluetongue virus core. Nature 395:470–478
36. Reinisch KM, Nibert ML, Harrison SC (2000) Structure of the Reovirus core at 3.6 Å resolution. Nature 404:960–967
37. Naitow H, Tang J, Canady M, Wickner RB, Johnson JE (2002) L-A virus at 3.4 a resolution reveals particle architecture and mRNA decapping mechanism. Nat Struct Biol 9:725–728
38. Luque D, Gonzalez JM, Garriga D, Ghabrial SA, Havens WM, Trus B, Verdaguer N, Carrascosa JL, Caston JR (2010) The T = 1 capsid protein of penicillium chrysogenum virus is formed by a repeated helix-rich core indicative of gene duplication. J Virol 84:7256–7266
39. Bamford DH, Grimes JM, Stuart DI (2005) What does structure tell us about virus evolution? Curr Opin Struct Biol 15:655–663
40. Baker ML, Jiang W, Rixon FJ, Chiu W (2005) Common ancestry of herpesviruses and tailed DNA bacteriophages. J Virol 79:14967–14970
41. Abrescia NG, Bamford DH, Grimes JM, Stuart DI (2012) Structure unifies the viral universe. Annu Rev Biochem 81:795–822
42. Pesavento JB, Crawford SE, Estes MK, Prasad BV (2006) Rotavirus proteins: structure and assembly. Curr Top Microbiol Immunol 309:189–219
43. Coulibaly F, Chevalier C, Gutsche I, Pous J, Navaza J, Bressanelli S, Delmas B, Rey FA (2005) The birnavirus crystal structure reveals structural relationships among icosahedral viruses. Cell 120:761–772
44. Saugar I, Luque D, Ona A, Rodriguez JF, Carrascosa JL, Trus BL, Caston JR (2005) Structural polymorphism of the major capsid protein of a double-stranded RNA virus: an amphipathic alpha helix as a molecular switch. Structure 13:1007–1017
45. Luque D, Saugar I, Rodriguez JF, Verdaguer N, Garriga D, Martin CS, Velazquez-Muriel JA, Trus BL, Carrascosa JL, Caston JR (2007) Infectious bursal disease virus capsid assembly and maturation by structural rearrangements of a transient molecular switch. J Virol 81:6869–6878
46. Irigoyen N, Caston JR, Rodriguez JF (2012) Host proteolytic activity is necessary for infectious bursal disease virus capsid protein assembly. J Biol Chem 287:24473–24482
47. Yan X, Yu Z, Zhang P, Battisti AJ, Holdaway HA, Chipman PR, Bajaj C, Bergoin M, Rossmann MG, Baker TS (2009) The capsid proteins of a large, icosahedral dsDNA virus. J Mol Biol 385:1287–1299
48. 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

49. Yan X, Chipman PR, Castberg T, Bratbak G, Baker TS (2005) The marine algal virus PpV01 has an icosahedral capsid with T = 219 quasisymmetry. J Virol 79:9236–9243
50. 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
51. Tao Y, Olson NH, Xu W, Anderson DL, Rossmann MG, Baker TS (1998) Assembly of a tailed bacterial virus and its genome release studied in three dimensions. Cell 95:431–437
52. 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
53. Luque A, Reguera D (2010) The structure of elongated viral capsids. Biophys J 98:2993–3003
54. Stubbs G, Kendall A (2012) Helical viruses. Adv Exp Med Biol 726:631–658
55. Zhang R, Hryc CF, Cong Y, Liu X, Jakana J, Gorchakov R, Baker ML, Weaver SC, Chiu W (2011) 4.4 A cryo-EM structure of an enveloped alphavirus Venezuelan equine encephalitis virus. EMBO J 30:3854–3863
56. Mukhopadhyay S, Kuhn RJ, Rossmann MG (2005) A structural perspective of the flavivirus life cycle. Nat Rev Microbiol 3:13–22
57. Ganser-Pornillos BK, Yeager M, Pornillos O (2012) Assembly and architecture of HIV. Adv Exp Med Biol 726:441–465
58. Castón JR, Trus BL, Booy FP, Wickner RB, Wall JS, Steven AC (1997) Structure of L-A virus: a specialized compartment for the transcription and replication of double-stranded RNA. J Cell Biol 138:975–985
59. Luque D, Gonzalez JM, Gomez-Blanco J, Marabini R, Chichon J, Mena I, Angulo I, Carrascosa JL, Verdaguer N, Trus BL, Barcena J, Caston JR (2012) Epitope insertion at the N-terminal molecular switch of the rabbit hemorrhagic disease virus T = 3 capsid protein leads to larger T = 4 capsids. J Virol 86:6470–6480
60. Castón JR, Martínez-Torrecuadrada JL, Maraver A, Lombardo E, Rodríguez JF, Casal JI, Carrascosa JL (2001) C terminus of infectious bursal disease virus major capsid protein VP2 is involved in definition of the T number for capsid assembly. J Virol 75:10815–10828
61. 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
62. Jose J, Snyder JE, Kuhn RJ (2009) A structural and functional perspective of alphavirus replication and assembly. Future Microbiol 4:837–856
63. Buck CB, Trus BL (2012) The papillomavirus virion: a machine built to hide molecular Achilles' heels. Adv Exp Med Biol 726:403–422

Further Reading

Chiu W, Burnnett RM, Garcea RL (eds) (1997) Structural virology of viruses. Oxford University Press, New York
Chiu W, Johnson JE (eds) (2003) Virus structure. Advances in protein chemistry, vol 64. Academic Press, Amsterdam
Mahy BWJ, van Regenmortel MHV (eds) (2008) Encyclopedia of virology. Elsevier, Amsterdam
Patton JT (ed) (2008) Segmented double-stranded RNA viruses, structure and molecular biology. Caister Academic Press, Norfolk
Rossmann MG, Rao VB (eds) (2012) Viral molecular machines. Adv Exp Med Biol, vol 726. Springer, New York

Also especially recommended for further reading are references [5, 6, 9, 21, 30, 41] *listed above*.