Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Chapter 3

Virion Structure and Composition

The virus particle or virion represents a virus in its extracellular phase, in contrast to the different intracellular structures involved in virus replication. To ensure survival of a virus, the virion must fulfill two roles: (1) protecting the genome from environmental damage, for example, from heat, desiccation, chemicals; and (2) facilitating the passage of the virus to the next host, that is, from the point of release from the original host, passage through the environment to the point of encountering a new host, followed by entry into the cells of the new host. There are many different ways that different viruses achieve these two roles, and viral genomes and virion structures show enormous variety in both size and composition—yet there are many features or principles of assembly that are shared by most viruses. Notably, many key structures within the virion are assemblies and subassemblies of a large number (usually hundreds) of identical protein subunits that lock together sterically to form a stable shell (capsid or envelope); the employment of large numbers of one or a few different primary units (structural units, capsomeres) allows the genetic coding of relatively large macromolecules by a very small number of different viral genes.

PHYSICAL METHODS FOR STUDYING VIRUS STRUCTURE

Electron Microscopy

The development of electron microscopy was pivotal in the establishment of virology as a scientific discipline. Viruses are smaller than the limit of resolution of the light microscope, which is about 0.3 μm, or 300 nm. Poxviruses were for many years considered amongst the largest of viruses, being just about visible in the light microscope using dark-field optics or certain staining techniques. In recent years, much larger viruses have been discovered but so far none of these have been shown to be human pathogens. Nevertheless, these newly discovered viruses are driving a re-examination of the limits of the concept of virus. For example, the virions of megavirus and mimivirus, both infectious for Acanthamoeba, are about 0.5 μm in diameter, have a DNA genome up to 1.26 Mb in size and code up to 1120 proteins. The virions of pandoraviruses are 1–1.2 μm in size, have a linear DNA genome up to 2.8 Mb, and code up to 2500 proteins. An even larger virus, Pithovirus (recently isolated from Siberian permafrost), is 1.5 μm in size, the size of a small bacterium, with >2500 putative protein-coding sequences, of which only 6% have recognizable relationships with genes from known viruses, microorganisms, or eukaryotes. At the same time, viruses that are smaller than those in previously known taxa have been found, so virologists must now be prepared to work with viruses as large as bacteria and as small as large protein molecules.

The first electron microscopy of viruses by Bodo von Borries, Ernst Ruska, and Helmut Ruska in 1938 employed a simple preparative method that did not show much more than the outline of virions. Later, metal shadow-casting of purified virus preparations improved the visualization of virions, but still not enough. Beginning in the 1950s, ultra-thin sectioning of virus-infected cells became widespread, providing more virion detail and also the beginning of the science of ultrastructural cytopathology—virion morphogenesis, intracellular localization of virus structures and cellular organelles and coincident damage to host cell structures. In 1959, visualization of viral ultrastructure was taken to a still higher level of resolution when Sydney Brenner and Robert Horne developed negative staining. In this method, a solution of potassium phosphotungstate (or other electron-dense salts) is added to a virus suspension on a coated specimen grid; the metal ions surround and fill the interstices in the surface of virions giving a negative image in the electron beam, thus revealing structural details not previously seen. A remarkable diversity of virus structures can be seen in negatively stained preparations. Electron micrographs of virions and infected cells from different families of viruses are shown in the various chapters of this book (see Figs. 1.3 and 1.4 for comparison of ultra-thin sectioning versus negative staining).

In the past few years, the above methods have been complemented by several new microscopy technologies, particularly scanning electron microscopy and cryo-electron microscopy, the latter using computer-based image construction of images of snap frozen, unstained virion
preparations. This technique has the advantage of showing viruses in a hydrated state rather than the desiccated conditions of negative staining in electron microscopy. The resolution of these techniques is rapidly approaching that obtained by X-ray diffraction of crystallized virions and viral substructures (Fig. 3.1).

**X-Ray Crystallography of Viruses**

X-ray crystallography of viruses provides another technique for visualizing virion structural organization showing structural details to near atomic resolution, the location of antigenic sites on the surface of virions, and aspects of virion attachment and penetration into cells. For example, applying this technique to several picornaviruses revealed that the polypeptides of each of the three larger structural proteins are packaged to form wedge-shaped eight-stranded antiparallel β-barrel subassemblies (Fig. 3.2). The overall contour of picornavirus virions reflects the packing of these subassemblies. Relatively unstructured amino acid chains form loops that project from the main wedge-shaped domains. Some loops form flexible arms that interlock with the arms of adjacent wedge-shaped subassemblies, thereby providing physical stability to the virion. Other loops, those involved in virion attachment to the host cell, harbor the antigenic sites (epitopes) that are the targets of the host’s neutralizing antibody response against the virus.

Larger viruses are much more complex in structure, and to study the structure of these viruses it is usually necessary to separate well-defined substructures and examine crystals of these structures by X-ray diffraction. Rotaviruses are an excellent example of this approach, being composed of a core and two capsid layers, each component exhibiting unique structural details, fitting together in a precise fashion to form the complete virion.

One of the pioneering studies of viral structure was the determination by X-ray crystallography of the structure of the hemagglutinin molecule of influenza viruses and the placement and variation of neutralizing epitopes on this molecule. Today, determination of new variations in the amino acid sequence and hence the microstructure of the influenza hemagglutinin is used in the development of updated vaccines. Many individual viral proteins have been analyzed to the level of 2–3 Å resolution, revealing

---

**TABLE 3.1 Relative Sizes of Common Objects in the Biological World**

| Size   | Equivalent Practical Units | Observations                                                                 |
|--------|----------------------------|------------------------------------------------------------------------------|
| 1 m    | 3 ft 3 in.                 | Humans, adult males, are about 1.8 m tall                                   |
| 10 cm  | 4 in.                      | Human adult hand is about 10 cm wide                                        |
| 1 cm   | 1 cm                       | *Aedes aegypti*, adult, mosquito is about 1 cm long                          |
| 1 mm   | 1 mm                       | *Ixodes scapularis* tick, nymphal stage, is about 1 mm long                  |
| 0.1 mm | 100 μm                     | Smallest things visible to the naked eye                                    |
| 0.01 mm| 10 μm                      | Lymphocytes are about 20 μm in diameter                                      |
|        |                            | *Bacillus anthracis*, among the largest of pathogenic bacteria, is 1 μm wide and 5 to 10 μm long |
| 0.001 mm| 1 μm                      | Smallest things visible in light microscope are about 0.3 μm in size         |
|        |                            | Poxviruses, the largest of the viruses of vertebrates, are 300 nm (or 0.3 μm) in their longest dimension |
| 0.1 μm | 100 nm                     | Influenza viruses and retroviruses, typical medium-sized viruses, are about 100 nm in diameter |
|        |                            | Flaviviruses, such as yellow fever virus, typical smaller-sized viruses, are about 50 nm in diameter |
| 0.01 μm| 10 nm                      | Picornviruses, such as polioviruses, typical small viruses, are about 30 nm in diameter |
|        |                            | Circoviruses, the smallest of the viruses of vertebrates, are 17 to 22 nm in diameter |
| 0.001 μm| 1 nm, 10 Å                | Smallest things visible in transmission electron microscope; DNA double helix diameter is 2 nm |
| 0.1 nm | 1 Å                        | Diameter of atoms is about 2 to 3 Å                                         |
Virion Structure and Composition

Chapter 3

Potential targets for new antiviral compounds. Notable is the development of antiviral drugs for the treatment of influenza through analysis of the detailed structure of the viral neuraminidase (Fig. 3.3).

Chemical Composition of Virions

Viruses are distinguished from other macromolecular forms by a possessing rather simple, repetitive chemical composition. The virion, that is the complete infectious virus particle, includes a genome comprising one or a few molecules of either DNA or RNA, surrounded by a morphologically defined protein coat, the capsid; the capsid and the enclosed nucleic acid together constitute the nucleocapsid. A small number of additional proteins may be present within the virion as enzymes. The nucleocapsid of some viruses is surrounded by a lipoprotein bilayer envelope into which are inserted viral proteins (peplomers): these may or may not be glycosylated. Sometimes a matrix protein is also associated with the inner aspect of the viral envelope. The simplest virus (e.g., tobacco necrosis virus satellite, a defective virus that needs a helper virus to

Figure 3.1 Picornavirus structural studies. (A–D) Computer-based virion reconstructions from cryo-electron microscopy and X-ray diffraction images. (A) Poliovirus 1, protein chain model. From Jason Roberts and colleagues, with permission. (B) Poliovirus 1, string model. (C) Coxsackie B3 virus, space-filling model. (D) Human rhinovirus B14, Quemol rendering model. (B–D) from Jean-Yves Sgro, with permission. (E) Model of poliovirus icosahedral capsid showing location of the three proteins making up the capsid surface, VP1, VP2 and VP3 (VP4 is buried on the inner face of the capsid). (F) Poliovirus, negative contrast electron microscopy from Joseph Esposito (deceased), showing little or no virion surface detail—such was the resolution available before X-ray diffraction and cryo-electron microscopy technologies, both of which require massive computer compilation to reconstruct image data.

Figure 3.2 Structure of a typical picornavirus. (A) X-ray crystallographic structure of a native virus particle. A single substructural unit (protomer) is shown in a string-ribbon representation. The viral proteins are colored blue (VP1), green (VP2), magenta (VP3), and orange (VP4), and the approximate positions of the viral 5- and 3-fold axes are indicated by the solid black pentagon and triangle, respectively. (B) The intact, T = 3 capsid structure, again in a string-ribbon representation. The unit shown in panel (A) is highlighted in yellow, and the layout of the virion icosahedron is included as a red line overlay. Adapted from Bakker, S.E., Groppelli, E., Pearson, A.R., Stockley, P.G., Rowlands, D.J., Ranson, N.A., 2014. Limits of structural plasticity in a picornavirus capsid revealed by a massively expanded equine rhinitis A virus particle. J. Virol. 88: 6093–6099.
provide some of its functions) directs the synthesis of only one protein; many important viruses direct the synthesis of five to ten proteins; large viruses, such as the poxviruses and herpesviruses, direct the synthesis of up to 200 proteins and the recently discovered megaviruses up to 2500 proteins: still this is very few relative to the number of proteins involved in the life processes of bacteria (>5000 proteins) and eukaryotic cells (between 250,000 and 1,000,000 proteins). There are many variations on these constructions and diverse additional components are found among larger and more complex viruses (Fig. 3.4).

Viral Nucleic Acids

Viral genes are encoded in either DNA or RNA molecules; both DNA and RNA genomes can be either double-stranded or single-stranded, as well as monopartite (all viral genes contained in a single molecule of nucleic acid) or multipartite (viral genes distributed in multiple molecules or segments of nucleic acid). For example, among the RNA viruses, only viruses of the families Reoviridae, Birnaviridae, and Picobirnaviridae have a double-stranded RNA genome and these genomes are segmented (Reoviridae: 10, 11, or 12 segments, depending on the genus; Birnaviridae and Picobirnaviridae: two segments). All viral genomes are haploid, that is, they contain only one copy of each gene, except for retrovirus genomes, which are diploid. When carefully extracted from the virion, the nucleic acid of viruses of certain families of both DNA and RNA viruses is directly infectious; that is, when transfected into a cell there is sufficient genetic information to initiate a complete cycle of viral replication and produce a normal yield of progeny virions.

The sequence in which the various virus families are described in Part II of this book reflects the essential characters and diversity of viral genomes. The remarkable variety of

**FIGURE 3.3** Crystal structure of the hemagglutinin (HA) protein of the influenza 1918 virus, and comparison with other human, avian, and swine HAs. (A) Overview of the HA0 trimer, represented as a ribbon diagram. Each of the three monomers is colored differently. Carbohydrates are colored orange and labeled with the asparagine to which each is attached. The basic patch is indicated in the light blue ellipse and consists of HA1 residues. The locations of the three receptor binding sites, and cleavage sites, are shown for only one of the three monomers. (B) Structural comparison of different HA0 monomers, showing influenza 1918 HA0 (red), human H3 (green), avian H5 (orange), and swine H9 (blue). Reproduced from MacLachlan, N.J., Dubovi, E.J., 2011. Veterinary Virology, fourth ed. Academic Press, London (Fig. 1.4), with permission.
Virion Structure and Composition  Chapter | 3

Viruses is reflected in diverse ways in which the information encoded in the viral genome is transcribed into mRNA, then translated into proteins, and the ways in which the viral nucleic acid is replicated (see Chapter 4: Virus Replication).

Viral DNA Genomes

The genomes of all DNA viruses of vertebrates are monopartite, consisting of a single molecule that is, double-stranded except in the case of the paroviruses, anelloviruses, and circoviruses. DNA genomes may be linear or circular, depending on the virus family. The DNA of papillomaviruses, polyomaviruses, hepadnaviruses, anelloviruses, and circoviruses is circular. Additionally, the circular DNA of the papillomaviruses and polyomaviruses is supercoiled. The DNA genome of hepadnaviruses is partially double-stranded, although the single-stranded gap is closed during replication to form a covalently closed circular DNA (cccDNA; see Chapter 22: Hepadnaviruses and Hepatitis Delta).

Most linear viral DNAs have characteristics that can facilitate a circular configuration, a requirement for replication by a rolling circle mechanism (see Chapter 4: Virus Replication). The two strands of poxvirus DNA are covalently cross-linked at each terminus (forming hairpin ends), so that upon denaturation, the molecule becomes a large single-stranded circle. The linear double-stranded DNA of several DNA viruses and the linear single-stranded RNA of retroviruses contain repeat sequences at the ends of the molecule that permit circularization. Adenovirus DNA contains inverted terminal repeats; these are also a feature of the single-stranded DNA of paroviruses. Another type of terminal structure occurs in adenoviruses, hepadnaviruses, and paroviruses (and some single-stranded RNA viruses such as the picornaviruses and caliciviruses); all of these viruses contain a protein covalently linked to the 5′-terminus (the 5′ cap) with an essential function in priming replication of the genome.

The size of viral DNA genomes ranges from 1.7 kilobases (kb) for some circoviruses to over 200 kilobase pairs (kbp) for the double-stranded DNA of herpesviruses and poxviruses—and up to 2.8 mbp for the double-stranded DNA of the giant pandoraviruses and pithovirus of amoeba. As 1 kb, or for double-stranded DNA 1 kbp, contains enough genetic information to code for about one average-sized protein, it might be surmised that viral DNAs contain anywhere between two and 200 genes, coding for some two to 200 proteins in poxviruses and more than 2500 proteins in the giant viruses of amoeba. However, the relationship between any particular nucleotide sequence and its protein product(s) is not straightforward. On the one hand, the DNA of most of the larger viruses, similar to that of mammalian cells, contains what appears to be redundant information in the form of repeat sequences, thus the coding capacity of large viral genomes may be overestimated. On the other hand, coding capacity might be underestimated: first, a given DNA or mRNA sequence may be read in up to three alternate reading frames, producing up to three proteins with different amino acid sequences; second, both strands of a double-stranded viral DNA molecule may be transcribed, each transcript yielding a different protein; third, genes may overlap, yielding various transcripts and protein products; and finally, a single primary RNA transcript may be spliced or cleaved in several different ways to yield a number of distinct mRNAs, each of which may be translated into a different protein, or a single polyprotein translation product may be subsequently cleaved by proteolysis to yield multiple discrete proteins.

Viral DNAs contain several kinds of non-coding sequences essential for genome expression and replication, some of which have been conserved throughout evolutionary
time; these include DNA replication initiation sites, RNA polymerase recognition sites, sites for the initiation and termination of translation, RNA splice sites, promoters, enhancers, etc.

**Viral RNA Genomes**

With the exception of the reoviruses, birnaviruses, and picobirnaviruses the genomes of all vertebrate RNA viruses are single-stranded. The RNA may be monopartite or multipartite: for example, the retroviruses, paramyxoviruses, rhabdoviruses, filoviruses, coronaviruses, arteriviruses, picornaviruses, togaviruses, and flaviviruses, all have monopartite genomes. In contrast, the orthomyxoviruses, bunyaviruses, and arenaviruses have multipartite genomes. The genomes of the arenaviruses consist of two segments; the bunyaviruses three; the orthomyxoviruses six, seven, or eight (depending on the genus), the birnaviruses and picobirnaviruses two; and the reoviruses 10 to 12, again depending on the genus. Each RNA molecule in these viruses is unique, often encoding a single protein. There is no vertebrate virus RNA genome that is a covalently linked circle, with the exception of the very small circular single-stranded RNA of delta hepatitis virus which has a structure resembling that of plant viroids. However, the single-stranded RNAs of the bunyaviruses and arenaviruses appear circular owing to hydrogen bonding between complementary ends. The genomes of single-stranded RNA viruses have considerable secondary structure, regions of base-pairing causing the formation of loops, hairpins, etc., that serve as signals controlling nucleic acid replication, transcription (especially initiation and termination), translation, and/or packaging into the capsid.

Single-stranded RNA genomes can be defined according to coding sense (also called polarity). If the genomic RNA is of the same sense as mRNA, that is, it can direct the synthesis of protein, it is said to be of positive-sense (also called plus sense). This is the case with the picornaviruses, caliciviruses, togaviruses, flaviviruses, coronavirus, and retroviruses. If, on the other hand, the genomic nucleotide sequence is complementary to that of mRNA, it is said to be negative-sense. Such is the case with the paramyxoviruses, rhabdoviruses, filoviruses, bornaviruses, orthomyxoviruses, arenaviruses, and bunyaviruses. All RNA virus genomes carry a gene for an RNA-dependent RNA polymerase (transcriptase), with the exception of retroviruses and hepadnaviruses which encode an RNA-dependent DNA polymerase. This enzyme directs the synthesis of positive-sense RNA in the infected cell using the viral genome as template. Mammalian cells do not contain an RNA-dependent RNA polymerase, and therefore this virally encoded non-structural protein has to be synthesized early in infection before rounds of replication can be completed.

Both RNA segments of the member viruses of the family Arenaviridae and one of the three RNA segments of members of one genus of the family Bunyaviridae are ambisense, that is, part positive-sense, part negative-sense (Chapter 29: Bunyaviruses and Chapter 30: Arenaviruses). If the viral RNA is of a positive-sense, it is usually polyadenylated at the 3′-end (in picornaviruses, caliciviruses, togaviruses, coronavirus, and arteriviruses, but not in flaviviruses) and capped at the 5′-end (togaviruses, flaviviruses, coronaviruses, and arteriviruses).

The size of single-stranded viral genomes varies from 1.7 to 32 kb and the double-stranded RNA viruses from 18 to 27 kbp—a much smaller range than found among the double-stranded DNA viruses. Accordingly, these viruses encode fewer proteins than many DNA viruses, generally less than a dozen. Most of the segments of the genomes of orthomyxoviruses and reoviruses are individual genes, each coding for one unique protein.

**Anomalous Inclusion of Nucleic Acids Within Virions**

Viral preparations often contain some particles with an atypical content of nucleic acid. Several copies of the complete viral genome may be enclosed within a single virion, or virions may be formed that contain no nucleic acid (empty particles) or that have an incomplete genome (defective particles and in some cases defective interfering particles by virtue of modulating replication). Moreover, host cell DNA may sometimes be incorporated into virions (for example, in papillomaviruses and polyomaviruses), while ribosomal RNA is found in orthomyxovirus and arenavirus virions. However, there is no evidence that ribosomal RNA within virus particles has any function.

**Viral Proteins**

Some virus-coded proteins are structural, that is, they are used to construct the capsid, envelope, and other components of the virion. Other proteins are non-structural; these are not present in the virion but are involved in various viral replication processes including virion assembly. Many non-structural proteins are enzymes, which may be involved in nucleic acid replication, transcription, and translation, as well as the shutdown of host cell functions, the inhibition of innate immunity, and the subversion of cellular machinery to viral synthetic activities. There are many kinds of viral enzymes described in this book; among these are various types of (1) replicases (also called polymerases; e.g., DNA-dependent DNA replicase or polymerase) and other enzymes involved in viral nucleic acid replication, (2) transcriptases that transcribe mRNA from viral DNA or RNA genomes, and (3) various proteases, helicases, and ligases. Reverse transcriptase, an enzyme that transcribes DNA from an RNA template, is found uniquely in retroviruses and hepadnaviruses. Other
enzymes found only in retroviruses are involved in the integration of the DNA product of reverse transcription into cellular chromosomal DNA (integrase). Poxviruses, which replicate in the cytoplasm and therefore have less access to cellular machinery, duplicate a number of host enzymes that function in processing RNA transcripts and replicating the DNA genome.

**Capsid Structure**

The capsid is built up of identical, non-covalently linked structural subunits, which may be discernible by electron microscopy as protrusions or depressions on the surface of virions (Fig. 3.4). These may correspond to individual viral proteins or aggregates of proteins. Only the simplest virions are assembled from the primary products of protein synthesis; that is, individual polypeptides; in most cases virion capsids are constructed from distinct assembly units (“capsomeres”) of several components, themselves often derived by modification or cleavage of precursors. Each subassembly may contain more than one polypeptide. One crucial step in virion assembly is the incorporation of the viral nucleic acid into the nascent virion—several different mechanisms driving this process have been recognized including the presence of packaging signals within the nucleic acid sequence of the genome.

**Viral Envelope Lipids**

In most cases the integrity of the envelope is necessary for viral infectivity. The envelope is acquired when the nucleocapsid is extruded through one of the cellular membranes—this process is known as budding. This process is not limited to the outer cell membrane: many viruses mature on internal, cytoplasmic membranes or even bud through the host nuclear membrane. The lipids of the viral envelope are derived directly from the cellular membrane, but the major proteins associated with the envelope are virus-coded.

Most lipids found in enveloped viruses are constructed into a typical lipid bilayer, in which the virus-coded glycoprotein peplomers and in some cases other viral proteins are embedded. As a consequence, the composition of the lipids of particular viruses differs according to the composition of the membrane lipids of the host cells in which replication takes place. The composition of the membrane lipids of viruses also varies with the particular membrane system employed for virion budding. For example, the lipids of paramyxoviruses, which bud from the plasma membrane of host cells, differ from those of bunyaviruses and coronaviruses that bud through the membranes of intracytoplasmic organelles. Lipids constitute about 20 to 35% of the dry weight of most enveloped viruses; some 50 to 60% of viral envelope lipid is phospholipid and most of the remainder consists of cholesterol.

The envelopes of some viruses also contain a variable content of host proteins, for example, MHC molecules, which may be retained after the virus buds through the cellular membrane. On occasion, the presence of host antigenic material on the viral surface can cause unexpected effects on the pathogenesis of disease and in diagnostic tests.

**Viral Envelope Proteins**

Most external viral envelope proteins are glycoproteins, occurring as membrane-anchored peplomers (peplos, Gk = a loose outer garment) or spikes, often assembled as dimers or trimers. These can be seen in electron micrographs extending outward from the envelope of enveloped viruses such as orthomyxoviruses, paromyxoviruses, rhabdoviruses, filoviruses, coronaviruses, bunyaviruses, arenaviruses, and retroviruses. However, the virions of some of the more complex viruses also contain glycosylated internal or outer capsid proteins. Recognition sites for cellular receptors are often located at the furthest domain from the viral envelope (distal end) whereas proximal domains interact with the lipid bilayer of the envelope (Fig. 3.5). Oligosaccharide side-chains (glycans) are attached by N-glycosidic, or more rarely O-glycosidic, linkages. Since these are synthesized by cellular glycosyl transferases, the sugar composition of these glycans is analogous to that of host cell membrane glycoproteins.

In contrast, matrix proteins are non-glycosylated and are found as a layer on the inside of the envelope of orthomyxoviruses, paramyxoviruses, rhabdoviruses, filoviruses, and retroviruses, but not coronaviruses, bunyaviruses, and arenaviruses. The presence of a matrix protein provides structural rigidity to a virion; for example, the helical nucleocapsid of rhabdoviruses is closely apposed to a rather rigid layer of matrix protein, and this in turn is tightly bound to the viral envelope by hydrogen bonding to the internal domains of the surface glycoprotein peplomers.

**VIRION SYMMETRY**

Individual protein structures are asymmetric. As discussed above, virions are assembled from multiple copies of one or a few kinds of protein subunits—the repeated occurrence of similar protein–protein interfaces leads to assembly of the subunits into symmetrical capsids. Folded polypeptide chains, specified by the viral genome, comprise protein subunits; assemblages of these protein subunits comprise structural units and in turn sets of these structural units comprise assembly units (capsomeres) that are the major intermediates in the formation of viral capsids. This efficiency of design depends upon principles of self-assembly, wherein structural units and capsomeres are brought into clusters through random thermal movement and are held in place through weak electrostatic bonds.
Viruses come in a variety of shapes and sizes, depending on the shape, size, and number of protein subunits and the nature of the interfaces between these subunits; however, only two kinds of symmetry have been recognized, **icosahedral** and **helical**. Asymmetrical virus structure, such as that of poxviruses, is said to be **complex**.

### Icosahedral Symmetry

In the mid-1950s Francis Crick and James Watson argued, using theoretical principles, that small spherical viruses were constructed out of identical subunits in identical environments as in true crystals. However, the maximum number of subunits that can be arranged in a sphere is 60—when more than 60 identical subunits are assembled an icosahedron is formed. The icosahedron is the optimum solution to the problem of constructing, from repeating subunits, a strong structure enclosing a maximum volume with the minimum amount of energy. The icosahedron is one of the five classical **Platonic or cubic solids**; the icosahedron, that is, the form with **icosahedral symmetry**, has 12 vertices (corners), 30 edges, and 20 faces, each face describing an equilateral triangle. Icosahedra have axes of two-, three-, and fivefold rotational symmetry that pass through the edges, faces, and vertices of the icosahedron, respectively (Fig. 3.6). Evidence from electron microscopy and biophysical
techniques confirmed that most spherical viruses had far more than 60 protein subunits. Drawing inspiration from Buckminster Fuller’s geodesic domes, Donald Caspar and Aaron Klug discovered a solution to the problem—they proposed that spherical viruses were structured like miniature geodesic domes, that is, the identical subunits are asymmetrical and are bonded together not in equivalent, but in quasiequivalent ways wherein there can be slight deformation of optimum bonding distances and angles wherein far more subunits can be accommodated to make larger structures. They also developed the idea of self-assembly after considering the viral assembly process as a type of crystallization process.

Only certain arrangements of structural units can form the faces, edges, and vertices of viral icosahedra. The structural units on the faces and edges of adenovirus virions, for example, bond to six neighboring capsomers and are referred to as hexons or hexamers; those at the vertices bond to five neighbors and are called pentons or pentamers (Fig. 3.6). In the virions of some viruses both hexons and pentons are composed of the same polypeptide(s), whereas the hexons or pentons of other viruses are formed from different polypeptides. The arrangements of structural units on the surface of virions of two icosahedral model viruses are shown in Fig. 3.6. Because of variations in the arrangement of the structural units on different viruses, some viruses appear rather hexagonal in outline and some appear nearly spherical. Even within rather smooth overall surface configurations, at higher resolution functional protrusions, bulges, and projections (often housing cellular attachment ligands and neutralizing epitopes) can be seen. There are also depressions, clefts, and canyons that may also house attachment ligands but usually not neutralizing epitopes.

**Helical Symmetry**

The nucleocapsid of many RNA viruses self-assembles very differently, forming a cylindrical structure in which the protein structural units are spatially arranged as a helix, hence the term helical symmetry. The occurrence of identical protein–protein interfaces on the structural units promotes the symmetrical assembly of the helix. In helically symmetrical nucleocapsids, the RNA genome forms a spiral within the nucleocapsid (Fig. 3.7). Many plant viruses with helical nucleocapsids are rod-shaped, flexible or rigid, and non-enveloped. The helical structure of tobacco mosaic virus was among the first viral structures determined by negative staining electron microscopy—its detailed structure was resolved by X-ray crystallography. However, in all viruses of vertebrates with helical symmetry, the nucleocapsid is wound into a secondary coil and enclosed within a lipoprotein envelope, for example rhabdoviruses (see Fig. 27.1); and paramyxoviruses (see Fig. 26.2).

**The Function of Viral Capsids and Envelopes**

Viral capsids and envelopes are not just inert coverings—these must be sufficiently stable in the environment to protect the contained nucleic acid genome, and at the same time play multiple roles in the interaction between the virion and host cell. Different kinds of envelope-associated proteins are associated with at least four crucial activities:
binding to receptors, membrane fusion, uncoating, and receptor modification. For example, fusion proteins are involved in both viral entry and viral release, in many cases promoting the fusion of viral envelope with cellular membranes at virus entry and promoting virus “pinching off” at virus exit by budding. Moreover, before entry into the cell, viruses may be converted to a primed state to facilitate uptake and infection of target cells. This primed state usually involves conformational rearrangements of the virion surface proteins, making these structures responsive to various triggers. For example, upon entry into the host the hemagglutinin (HA0) of influenza viruses is cleaved at a specific site by the extracellular enzyme tryptase Clara, generating a primed modified structure composed of two unique subunits (HA1 and HA2). After entry of the virion into the host cell via receptor-mediated endocytosis, the primed hemagglutinin molecule then is activated when exposed to the low pH within the endosome. Activated hemagglutinin mediates endosome membrane disruption or fusion, thereby allowing the release of the viral RNA genome into the cytoplasm of the host cell.

This process of virion attachment and entry into the host cell is one of the most important stages of the virus–host relationship. In this context, the terms receptor and ligand have often been used in imprecise ways. The term receptor is properly used to designate specific molecule(s) or structure(s) on the surface of host cells that are involved in virus attachment. The term ligand is used for the molecule(s) on the surface of the virus that bind to the receptor. For example, the hemagglutinin of influenza virus is the ligand that binds to the receptor on the host cell surface, in this case a glycoconjugate terminating in N-acetylgalactosamine.

**Stability of Viral Infectivity**

In general, viruses are more sensitive than bacteria or fungi to inactivation by physical and chemical agents, but there are important exceptions. Knowledge of specific viral sensitivity to environmental conditions and particular physical and chemical agents is therefore important for preserving the infectivity of viruses as reference reagents and in clinical diagnostic specimens. Knowledge of stability is also vital for deliberate inactivation, for example in sterilization, disinfection, and the production of inactivated vaccines.

When a virus-containing sample is being inactivated, it is important to appreciate that individual virus particles within a population successively will lose infectivity, at an overall rate determined by the physical conditions and the properties of the particular virus. Thus, infectivity of the sample will be lost progressively, at a specific rate, for example, 1 log₁₀ loss of titer over a specific period of time. This means that the time for complete inactivation of a sample is critically dependent on the starting titer of virus as well as physical conditions, for example, temperature. It also explains why scientists often give a qualified and even evasive answer to the common question, “How long does HIV survive outside the body?”

**Temperature**

The principal environmental condition that may adversely affect the infectivity of viruses is temperature. In most cases, viral envelope proteins are denatured within a few minutes at temperatures of 55 to 60°C, with the result that the virion is no longer capable of normal cellular attachment, penetration, and/or uncoating. Many viral capsid proteins are only slightly
more heat resistant. At ambient temperature, the rate of decay of infectivity is slower but significant, especially in the summer or in the tropics. In order to preserve infectivity, viral preparations must therefore be stored at low temperature; 4°C (on wet ice or in a refrigerator) is usually satisfactory for a day or so, but long-term preservation requires much lower temperatures. Two convenient temperatures are −70°C, the temperature of solid CO₂ (dry ice) and of many laboratory mechanical freezers, and −196°C, the temperature of liquid nitrogen.

As a rule of thumb, the half-life of most viruses can be measured in seconds at 60°C, minutes at 37°C, hours at 20°C, days at 4°C, and years at −70°C or lower. In general, the enveloped viruses are more heat-labile than the non-enveloped viruses.

Enveloped virions, for example, respiratory syncytial virus, are susceptible to repeated cycles of freezing and thawing, probably as a result of disruption of virions by ice crystals. This poses problems for the collection and transportation of clinical specimens. The most practical way of avoiding such difficulties is to deliver specimens to the laboratory as rapidly as practicable, packed without freezing on wet ice or packed with cold (not frozen) gel packs.

In the laboratory, it is often necessary to preserve virus stocks for years. This is achieved in one of two ways: (1) rapid-freezing of small aliquots of virus suspended in medium containing protective protein and/or dimethyl sulfoxide, followed by storage at −70°C or −196°C, (2) freeze-drying (lyophilization), that is, dehydration of a frozen viral suspension under vacuum, followed by storage of the resulting powder at 4°C or −20°C. Freeze-drying significantly prolongs viability even at ambient temperatures, and is commonly used in the end-stage manufacture of attenuated virus vaccines.

In contrast to the general principles of the temperature lability of viruses, prions (which are not viruses but commonly fall within the general work domain of virologists), are amazingly stable under virtually all environmental conditions, surviving boiling, freezing, many physical and chemical insults, and even large doses of γ-irradiation (see Chapter 38: Prion Diseases).

**Ionic Environment and pH**

On the whole, viruses are best preserved in an isotonic environment at physiological pH, but some tolerate a wide ionic and pH range. For example, most enveloped viruses are inactivated at pH 5–6, rotaviruses and many picornaviruses survive the acidic pH of the stomach without loss of infectivity. A 1 M solution of magnesium cations has been used to stabilize enteroviruses, for example, in stocks of poliovirus vaccine.

**Lipid Solvents and Detergents**

Lipid solvents such as ether or chloroform or detergents such as sodium deoxycholate readily destroy the infectivity of enveloped viruses—these agents must be avoided in laboratory procedures concerned with maintaining the viability of viruses. On the other hand, mild detergents are commonly used by virologists to solubilize viral envelopes and liberate proteins for use as vaccines or serological reagents.

**FURTHER READING**

Harrison, S.C., 2013. Principles of Virus Structure. In: Knipe, D.M., Howley, P.M. (Eds.) Fields Virology (sixth ed.), Lippincott Williams and Wilkins, Philadelphia, PA.