A virus particle is a structure that has evolved to transfer nucleic acid from one cell to another. It is, in effect, an extracellular organelle. The nucleic acid may be either RNA or DNA, and in both cases particles of varying complexity are found. A fundamental distinction arises between enveloped viruses—those with a lipid-bilayer membrane—and nonenveloped viruses. The distinction corresponds to a difference in the way the virus leaves and enters a cell. The internal structures of an enveloped virus are often similar to those of a nonenveloped virus. This chapter deals with the structures of viruses without lipid-containing envelopes as well as with the internal structures of enveloped viruses; Chapter 4 treats the organization of viral membranes.

KINDS OF STRUCTURES

There are two basic ways in which nucleic acid is packaged: into rod-like or filamentous structures and into roughly spherical ("isometric") ones. In rod-like or filamentous particles, protein subunits bind in a periodic way along the nucleic acid, winding it into a helical path. The best known examples of such structures are plant viruses, e.g., tobacco mosaic virus (TMV). The RNA of myxoviruses and of rhabdoviruses is also incorporated into filamentous structures by association with subunits of the N protein, forming a particle known as the nucleocapsid. The nucleocapsids are then coiled inside the viral membrane. In isometric particles, the nucleic acid is condensed within the virus in a manner geometrically independent of the organization of the shell. For example, in papovaviruses, the double-stranded, circular DNA binds histones in conventional nucleosomal structures. In the small plant viruses, the only cases where we have a relatively close-up view of the way in which nucleic acid (RNA) is packed, the links from protein shell to RNA are sufficiently flexible that few constraints are placed on the precise manner in which the RNA folds. Thus an important chemical difference distinguishes helical and isometric structures—regular, periodic interactions between capsid and nucleic acid in the one case, less restricting links between the shell and the condensed nucleic acid in the other.

An extensive nomenclature has developed to describe virus structures. The following terms, corresponding to common biochemical usage, are used in this chapter.

Subunit or protein subunit: a single folded polypeptide chain (e.g., the 17,000-MW protein of TMV).
Structure unit: a collection of one or more non-identical protein subunits that together form the chemical building block of a larger assembly (e.g., VP1, VP2, VP3, and VP4 in poliovirus or E1, E2, and E3 in Semliki Forest virus).

Morphological unit: the apparent lumps or clusters seen on the surface of a particle by electron microscopy. Because morphological units do not necessarily correspond to chemically unique entities, this is a term of convenience only, for use in describing electron micrographs.

Capsid: the protein shell directly complexed with viral nucleic acid. The word "coat" or "shell" is often just as clear in context.

Nucleocapsid: the complete protein-nucleic acid complex that is the packaged form of the genome in a virus particle (e.g., the core of Sindbis virus or the filamentous complexes of N protein and RNA in VSV, influenza, etc.). This is a useful term principally in cases where the nucleocapsid is a definite substructure of a more complex virus particle.

Viron: the entire virus particle.

The molecular architectures of a number of simple structures are described in this chapter. For a detailed discussion of the chemistry and biosynthesis of individual components, and for descriptions of the more complex structures, the reader is referred to the chapter on the particular viral type in question.

PRINCIPLES OF VIRUS STRUCTURE

The principles of virus structure depend on the biochemical characteristics of efficient and error-free assembly and on the requirements for controlled disassembly when the virus enters a host cell. These are reflected in three properties of the organization of most virus particles. The first is that viral substructures assemble from protein subunits that are usually specified by the viral genome. Genetic economy then dictates that these substructures be made of many identical copies of one or a few kinds of protein (18).
Repeated occurrence of similar protein–protein interfaces requires symmetrical arrangement of the subunit. Otherwise, the same sets of amino acid side chains would have different patterns of noncovalent bonding in different places, leading to ambiguities in assembly. Thus rod-like and filamentous viruses, as well as the filamentous nucleocapsids of many enveloped viruses, have helical symmetry, and simple isometric viruses have icosahedral symmetry. In very large viruses the symmetry is sometimes concealed by amplification of small distortions (on the ~1 Å scale of local motions in folded proteins). Long filamentous viruses may appear flexible, and large spherical viruses may not have a very rigid shape. However, the principle of specific intersubunit bonding appears to be true for the key, assembly-determining interactions in all well-characterized structures.

A second significant property of larger and more complex virus structures is that particles are constructed from distinct subassemblies. The most dramatic illustration is found in the complex bacteriophages, e.g., T4 (14). Heads, tails, and tail fibers assemble independently in pathways that have defined, sequential character (Fig. 1). Within a pathway, a particular intermediate serves to nucleate addition of the next component. For example, tail core subunits in T4 phage do not associate with each other unless assembly is initiated on a baseplate. Likewise, huddling of Sindbis or Semliki Forest virus occurs only around preassembled cores (see Chapter 4).

A third important property of virus particles is that incorporation of viral nucleic acid is specific but independent of most of its base sequence. In the case of helical structures, e.g., TMV, each subunit may interact with a definite number of nucleotides in a sequence-independent manner, but there is likely to be preferential binding to an initiation sequence for nucleating assembly. In the case of isometric particles, the condensation of a nucleic acid molecule must be independent of details of its secondary or tertiary structure, as well as of its sequence (21,28). In single-stranded RNA viruses, for example, it appears that no definite secondary or tertiary fold is needed for the RNA, aside from the restriction that it fit within the shell. What determines specific incorporation in most isometric viruses is uncertain, but a requirement
for one or more defined “packaging sequences” is likely. Clearly, the mechanism depends in part on whether nucleic acid is packed into a preformed shell (as in DNA phages) or whether protein and nucleic acid co-assemble.

SYMMETRIES OF VIRUS PARTICLES

Helical Symmetry

The symmetry of a helical arrangement is conveniently described by the number of units per turn, \( n \) (not necessarily integral) and the axial rise per unit, \( p \). The pitch of the helix, \( P \), is equal to \( n \times p \). Helical structures can have a rotation axis coincident with the helix axis (e.g., the T4 phage tail, with a sixfold axis) or an array of twofold axes perpendicular to the helix axis (as in DNA). The diagram of TMV in Fig. 2 illustrates that a structure with helical symmetry can also be described by reference to a “surface lattice.” The lattice lines correspond to different helical paths, as shown in Fig. 2. The surface lattice description is particularly useful for structures in which each unit interacts with units one turn above and below it in the helix. One can imagine generating the object from a planar net by rolling the array into a tube (Fig. 3). Different structures with closely similar surface lattices can be generated by closing the tube into helices with zero starts, one start, two starts, and so on.

Icosahedral Symmetry

Most closed-shell virus particles have structures based on icosahedral symmetry (Fig. 4). This symmetry is the most efficient of possible arrangements for subunits in a closed shell, in the sense that it uses the smallest unit to build a shell of fixed size. There are exactly 60 identical elements in the surface of any icosahedrally symmetric structure, related to each other by twofold, threefold, and fivefold rotation axes.

The way in which physical units can pack with this symmetry is shown in Fig. 5c. Commas related by twofold axes make contact head-to-head; those related by threefold axes make contact neck-to-neck; those related by fivefold axes make contact tail-to-tail. Closed-shell symmetries of higher order than icosahedral are, in a strict sense, not possible. Most viruses have more nucleic acid than can be packed within a shell defined by 60 subunits of reasonable molecular weight, and in these particles each of the 60 equivalent structural elements is composed of a number of protein chains. In some cases (e.g., poliovirus) these chains are chemically distinct (3). In many structures, however, they are genetically and chemically the same, and the viral shell is composed of some multiple of 60 identical units. This multiplicity raises two important problems for assembly.
all adjacent vertices in this way produces a shell of 60 subunits with completely identical intersubunit contacts (Fig. 5c). These contacts are similar to those in the original extended net, although curvature of the shell implies that units of finite thickness do not have precisely the same bonds in the flat and the icosahedral structures. By cutting at positions of second-nearest-neighbor vertices, a larger shell can be produced, having 180 instead of 60 subunits. All the contacts are similar, but the subunits in fact fall into three classes, denoted A, B, and C in Fig. 5d. Note that there are three kinds of contact that hold this structure together: head-to-head bonds relating a pair of units, back-to-back bonds relating rings of three, and tail-to-tail bonds relating rings of five or six. All subunits have these three kinds of interaction. Thus, although distinct in details of their packing, A, B, and C units in Fig. 5d are in fact very similar in the way they make contact. For example, all subunits form tail-to-tail bonds—those of A types in rings of five, as in the simple 60-subunit assembly, and those of B and C types alternating in rings of six. There are no tail-to-head bonds. Homologous parts of subunits make similar contacts in all cases. This sort of bonding was called “quasi-equivalent” by Caspar and Klug (16) to emphasize that similar though not identical contacts are made by all the subunits. They pointed out that proteins might have a defined, restricted flexibility of noncovalent contact that would make it possible to accommodate genetically and chemically identical molecules with nearly identical packing interactions. This sort of packing actually occurs in TBSV and several other known structures, although in a way slightly different from the original description. These results confirm the basic postulate of closely related, specific interactions. The TBSV structure also shows a mechanism for controlled switching among the different modes of such interactions—i.e., it shows at least one solution to the second problem described at the beginning of this section.

It is sometimes written that structures such as those in Figs. 5c and 5d are made up of pentamers and hexamers. If the tail-to-tail bonds were especially strong, this could indeed be the case. Note that all contacts except tail-to-tail were disrupted the structure in Fig. 4d would dissociate into 12 pentamers and 30 hexamers. Many such structures do have subunits with a prominent bulge, so located that they appear in the electron microscope to be made of rings of five and six. This has been called hexamer-pentamer “clustering.” An example is CCMV (3a). Such clusters are morphological units only. They reflect the shape of outer parts of the subunits and how these outer
FIG. 5. a: A plane hexagonal lattice. The commas represent subunits. The point marked by a star has been transformed into a five-coordinated position in b. b: Curvature may be introduced by changing a six-coordinated position into a five-coordinated one. The segment indicated in a has been excised and the edges brought together. Because such a process involves removing one-sixth of the subunits, a plane hexagonal net is not likely to be an intermediate in viral assembly, and the construction shown is a purely formal one. c: An icosahedrally symmetric structure with 60 subunits. d: An icosahedrally symmetric structure with 180 subunits. All 180 make similar local interactions, but strictly speaking there are three packing modes, indicated by A, B, and C in the figure.

parts touch each other. In fact, most 180-subunit viruses studied to date (e.g., CCMV) dissociate into dimers when gently disrupted, suggesting that the pairwise (head-to-head) contacts are strongest, despite the prominent projections that give rise to the characteristic appearance in negative stain. The description that emphasizes fivefold and sixfold contacts as having special importance is therefore misleading. Likewise, words such as "capsomere," coined to describe the prominent rings of five or six bulges contrasted by negative stain in the electron microscope, are helpful only when they correspond to chemically significant entities.

Icosahedral packings based on more than 180 subunits with closely related interactions can also be generated. If the triangular net in Fig. 5a is cut at third-nearest-neighbor points, a 240-subunit structure results. By extending this construction, it can be seen that, in general, \( T = h^2 + hk + k^2 \) (\( h \) and \( k \) are integers) gives the permitted multiples of 60 units; for \( T = 7 \) (and many higher triangulations), there are two enantiomorph possibilities. The number, \( T \),
giving the multiple of 60 units in the structure, is called the "triangulation number." Structures corresponding to $T = 1$ (STNV; bacteriophage P22; poliovirus; see below), $T = 3$ (many plant RNA viruses), $T = 4$ (Sindbis, Semliki Forest viruses and relatives) (48), and (probably) $T = 7$ (heads of bacteriophage P22) (13) have been identified. Viruses with still larger shells use different kinds of subunits for fivefold vertices and deviate in other ways from the simple geometries derived from a $p6$ net. Such complex structures need an organized assembly pathway, as in T4 tail assembly illustrated in Fig. 1, with one set of protein subunits acting as a framework or adaptor for positioning another. That is, the simple assumptions that go into deriving "permitted" triangulation numbers are not obeyed in the larger icosahedral particles, many of which form around smaller cores. The subtriangulations of an icosahedral surface lattice do, however, indicate the possible ways of obtaining close packing of units in the surface, even if the local symmetry of their interactions is not hexagonal. In other words, the nodes of a hexagonal net are positions of closest packing, even when the objects do not have sixfold symmetry and when the interactions of these objects with their six neighbors are therefore not all the same. Protein oligomers packed in this way make a good protective shell, and this probably explains why adenovirus particles appear to be $T = 25$ structures, or papovaviruses $T = 7$, even when the actual local symmetry does not turn out to correspond to a folded $p6$ net (see below).

EXAMPLES OF VIRUS STRUCTURES

Most of our knowledge of viral architecture comes from electron microscopy of negatively stained virus particles combined with biochemical analysis of purified components. X-ray diffraction has provided very detailed views of several nonenveloped plant viruses. Two of these—tobacco mosaic virus and the tomato bushy stunt virus—are described here first, as integration of information from microscopy and biochemistry of animal viruses presently depends on analogy with such structures. Crystallographic studies of adenovirus hexon (8) and poliovirus (30) promise soon to give comparatively precise knowledge of these animal viruses. The hemagglutinin and neuraminidase of influenza virus are described in the next chapter.

TMV

The rod-like TMV particle (Fig. 2) consists of about 2,130 subunits (158 amino acid residues) arranged in a helix of 16½ units per turn. The axial rise is 23 Å, and the rods are therefore about 3,000 Å long and 180 Å in diameter. RNA winds coaxially with the protein, with three nucleotides bound to each subunit. The virus can be dissociated by several methods into protein and RNA. The various associated forms of the purified protein have been characterized as a function of pH, ionic strength, and other variables (see below); and a particularly important form, the "disk," has been crystallized (4). The disk is a 34-subunit, two-layer structure: subunits in each layer have the same axial orientation, and the interactions between them are closely related to the interactions present in the virus itself (Fig. 6). Schematic views of a subunit in the disk are shown in Fig. 7. A transition from disk to helix, which can indeed occur, might be pictured as follows. Break the two rings at one point; dislocate to form a "lock-washer," tightening up the helix so formed from 17 to 16½ units per turn. The till of the subunits also changes, becoming uniformly $-10^6$. The important thing to note is that lateral interactions do not change significantly whereas axial interactions change completely. Equally important, a loop of polypeptide chain that forms the innermost part of the subunit is disordered in the disk and ordered in the virus. Because this loop acts as a "clamp" for RNA in the virus (Fig. 7), the disorder–order transition is quite significant. A model for the binding of an RNA trinucleotide to TMV has been proposed by Stubbs et al. (47): the polynucleotide backbone threads between subunits at a radius of ~40 Å, with the bases extended to surround the L-R helix of the upper subunit on three sides. Basic residues are in position to neutralize the RNA phosphates.

The celebrated reconstitution of TMV from separated RNA and protein by Fraenkel-Conrat and Williams (25) demonstrated self-assembly under physiological conditions. This experiment was important in showing that the native virus structure is a free-energy minimum for the purified components and that assembly can proceed intracellularly or in vitro without any additional template. Dissociated TMV protein in fact assembles into a variety of structures, shown in Fig. 8 (9). This "phase diagram" illustrates a fundamental polymorphism: helix-related structures at low pH or in the presence of RNA, disk-related structures at neutral pH or above. The regulation of this polymorphism by pH has been ascribed to a set of carboxyl groups (asp or glu side chains) that titrate anomalously in the virus due to enforced proximity in the helical assembly (15). The strain due to the proximity is relieved in the disk or, of course, by proton bonding. Threading RNA through the
FIG. 6. Subunit arrangement in the TMV protein disk (below) and the virus (above). a: The disk has 34 subunits in two rings of 17; the virus has 16½ subunits per helical turn. b: Sections through the axes of helix and disk, showing the tilt of the subunits. The innermost part of the subunit is disordered in the disk (dashed outline), perhaps to facilitate RNA binding. c: Cross section of several subunits in disk and virus, viewed from outside looking toward the axis. Disk sections are representations from an electron-density map; virus sections are schematic. The major part of each subunit consists of four α-helices ([LS, HS, RH, LR], which run in a roughly radial direction. In the transition from disk to helix, subunits slide over each other by about 10 Å. (From ref. 36, with permission.)

FIG. 7. Section through the TMV protein disk, as in Fig. 6b, but with a more detailed representation of the polypeptide chain. (From ref. 36, with permission.)
structure compensates for the strain, even above neutral pH where the carboxyl groups are fully deprotonated. The significance of proton binding in regulatory assembly comes from the role of the disk as an intermediate. In vitro reassembly proceeds much more rapidly if protein is added as disk rather than as "A protein," reflecting a critical function for disk in nucleating assembly as well as a possible role in elongation (10). The pathway as currently understood is shown in Fig. 9 (11,36,38). An initiation loop of RNA, located ~1,000 nucleotides from the 3' end of the RNA (just 5' to the coat cistron), inserts into the central hole of a disk (33,49). The disorder of the inner loop of polypeptide chain and the tilt of the subunit appears to create a pair of loose "jaws," with the RNA binding site at their interior. The nucleotide sequence has some suggestive repeats, in particular a G every three bases. Binding of some initial stretch of trinucleotides might be sufficient to

stretch the axial subunit contacts from disk-like to helix-like. This would dislocate the disk and create a helical growing point. Additions must then occur in two directions; reassembly experiments clearly show a rapid elongation toward the 5' end, drawing RNA up through the hole, and much slower addition toward the 3' end. As shown in Fig. 9, the disk may also play a role in 5' elongation: 3' elongation must proceed by addition of subunits.

This assembly pathway is remarkably more complicated than the simple addition of subunits at a screw dislocation originally suggested by Watson. It appears to have some important selective advantages. The assembled viral subunit must be able to accommodate any RNA triplet in its binding site. A small preference for certain sequences (e.g., G in one of the three positions) can give adequate specificity if the initiating region is relatively long (31)—hence the significance of the disk in nucleation. During elon-
Fig. 9. Model for TMV assembly. Views of initiation (A) and elongation (B) emphasize the importance of the disk as an intermediate subassembly. (From ref. 36, with permission.)

Initiation, assembling virus is a "melting protein" that unwinds helical stems in the RNA. Presentation of a loop of RNA at the 5' growing end can permit cooperative addition of subunits in order to drive unwinding of a stem at the other end of the axial hole (36).

TBSV and Related Structures

Our best view of an icosahedral viral shell comes from single-crystal X-ray diffraction studies of TBSV (28). This is a structure with triangulation number 3: 180 coat protein subunits (MW ~42,000), probably one chain of an 80,000 MW protein, and a molecule of single-stranded RNA (~4,800 nucleotides). The general architecture of TBSV is summarized in Fig. 10.

The coat subunit, containing 386 amino acid residues, folds into distinct modules: a projecting domain (P), a domain forming a tightly bonded shell (S), a connecting arm (a), and an internal domain (R). The three symmetrically distinct environments for this subunit are denoted A, B, and C. The polypeptide accommodates to these three packing positions by flexion at the hinge between S and P and by an ordering or disordering of part of the arm. Units at positions A and B (60 of each) have one hinge configuration, and the entire N-terminal region (arm and R-domain) appears to be spatially disordered. Subunits at position C (60 in all) have another hinge
configuration, and the connecting arm is folded in an ordered way along the bottom of the S domain. The R domain is not fixed with respect to the rest of the subunit, and so it cannot be seen in a high-resolution electron density map, but it may well be folded in a regular way. Thus our best picture of the N-terminal part of the subunit is of a well-structured R domain flexibly tethered to the S domain, which is held firmly in the viral shell. The disordered part of this tether is very long on A and B (the entire connecting arm) and much shorter on C (probably just a few residues). The packing in the interior of the virus is tight enough that arms, R domains, and RNA are probably not actually moving about (41).

The most remarkable feature of the TBSV structure is the way in which the C subunit connecting arms interdigitate to form an internal framework. As shown in Fig. 11, these arms extend along the inner edge of the S domain and loop around icosahedral threefold axes. Three such C subunit arms contact each other in this way, so that all 60 C subunits form a coherent net. The function of this inner framework is to determine the triangulation number precisely, i.e., to ensure that the viral shell closes around on itself correctly during assembly.

The actual interfaces between domains in TBSV have been analyzed in some detail. In general, they show that homologous parts of subunits make similar contacts, as in the description of "quasiequivalence" given above. Exactly which state of a contact occurs in a given position is, however, selected unambiguously by the order or disorder of the subunit arm. Thus the structure takes advantage of the capacity of the subunit to bond in a specific but slightly flexible way (quasiequivalence) without depending on these properties alone to prevent errors.
Packing of RNA in TBSV cannot be visualized directly in electron-density maps derived from X-ray crystallography due to spatial disorder of the nucleic acid and of the protein R domains. The R domains have a large number of positively charged residues, as do the inward-facing S domain surfaces, and together they neutralize about three-fourths of the RNA phosphates. The fundamental conclusion appears to be that few if any constraints other than compactness need to be placed on the configuration of the nucleic acid chain, as R domains are flexibly tethered to the rest of the shell. This is probably a general feature of RNA packing in spherical viruses. The absence of strong constraints or detailed spatial ordering implies that melting out of secondary structure is not required, an important difference from the arrangement in TMV and other helical structures.

The TBSV particle undergoes a reversible expansion when divalent cations are chelated above pH 7.0 (Fig. 12). Such concerted structural changes are properties of many other viruses; a step of this sort occurs during maturation of heads of double-stranded DNA phages.

Two other plant RNA viruses have been studied at molecular resolution by X-ray crystallography: satellite of tobacco necrosis virus (STNV) (40) and southern bean mosaic virus (SBMV) (1). The coat proteins of all three viruses have remarkable similarities. SBMV is a 180-subunit particle, similar to TBSV; its coat protein has R and S domains and a connecting arm, but no P domain. Its arm and S domain are folded and packed in the virion almost exactly like those of TBSV, and there is reasonable sequence homology. STNV has only 60 subunits, each of which has a short, internally projecting N-terminal sequence and a single domain with a fold very similar to that of TBSV-S or of SBMV-S (although no sequence homology). There is no arm, because there is no need for positional switching that governs correct assembly in the T = 3 structures. Moreover, when it packs into the viral shell, the STNV subunit presents somewhat different surfaces to its neighbors than do the S domains of TBSV and SBMV. The evolutionary relationships among such structures are still very unclear, but one direction of speculation is to ask whether these similar RNA packaging proteins might be related to a normal, cellular RNA packaging structure.

Poliovirus and Relatives

These are a group of small, RNA-containing viruses known as picornaviruses. Their RNA is ~6 kb in length, and the overall particle diameter is ~280 Å. The icosahedrally symmetric design of these viruses is actually simpler than that of TBSV, as it contains just 60 protein “structure units.” Each unit has one copy of each of four polypeptide chains (VP1, VP2, VP3, and VP4 where “VP” stands for virion protein). These chains are cleaved from a precursor; some or all of the cleavages occur after assembly into intermediates, so that considering the four as a single unit is not just a formal classification. Indeed a few uncleaved copies of VP0 (VP2 plus VP4) may be
FIG. 12. Divalent-cation-regulated expansion of IBV. There are sites for binding Ca$^{2+}$ at the intersubunit contacts shown in the diagram of the compact structure (left) and at all symmetry-equivalent positions. Two Ca$^{2+}$ ions are found at each of these interfaces, liganded by aspartic acid residues from the adjacent S domains. When Ca$^{2+}$ is removed and the pH raised above 7.0, these interfaces are destabilized by charge repulsion. The subunits therefore move apart, causing the particle to expand (right). Other S domain contacts are conserved in this process, and the interdigitation of C subunit arms (Fig. 11) is not disturbed (45a).

contained in some particles (44,46). Empty particles composed of 60 chains each of VP0, VP1, and VP3 can be isolated. It is believed that these procapsids may be genuine precursors of virions (32) and that the cleavage VP0 — VP2 + VP4 is required for complete RNA incorporation. Such cleavage steps make assembly of polyoma virus an essentially irreversible process. The 5' terminus of polyoma virus RNA has an attached peptide, called VPg (39). This 22-residue protein is derived by cleavage from the primary product of virus-directed translation (13). It appears to have a definite role in initiating replication (43).

Vesicular Stomatitis Virus

Viral membranes are discussed in detail in the next chapter, but a brief description of the overall organization of the vesicular stomatitis virus (VSV) is presented here.

The virus particle contains three proteins in large quantity (G, N, and M) and two minor components (NS and L). The glycoprotein (G, 65 K), the nucleocapsid protein (N, 50 K), and the "matrix" protein (M, 29 K) are in approximate molar ratio 1:2.3 or 1:2:4 (12). They are organized as shown in Fig. 13a. The overall shape and coherence of the particle probably depends on M, which forms tubular structures in conjunction with the nucleocapsid. These tubes appear to be a shallow helix with about 40 subunits per turn and about 35 turns per particle (Fig. 13b), but the precise structure is still uncertain.

Free nucleocapsid appears to be a more tightly coiled structure, with the aspect of a beaded strand (presumably subunits of N bound to RNA) wrapped into a 150 Å diameter helix. It can often be seen as if extruded from a partly disrupted virion (Fig. 13c). When extended on the carbon film of an electron microscope grid, it can appear sinusoidal but with a local radius of curvature similar to the radius of the coil. Some experiments have suggested that the N protein preferentially assembles on the sequence at the 5' end of the genome (5). The glycoprotein (G) communicates with internal structures across the lipid bilayer via a hydrophobic membrane anchor and a small internal domain (see Ch. 4). It is generally supposed that the internal domain makes contact with M, but direct proof is still lacking. The phenomenon of phenotypic mixing and the detection of pseudotypes (see later chapters) indicate that there may be important flexibility in the interaction between internal structures and glycoprotein tails.

Papovaviruses

Papovaviruses are the simplest of the double-stranded DNA viruses, with three protein species (VP1, VP2, and VP3) forming a coat and closed circular DNA complexes with cellular histones H2a, H2b, H3, and H4. Electron microscopy (Fig. 14a) shows the coat to have 72 morphological units ("capsomeres") (37). These units are centered at the vertices of what appears to be a $T = 7$ icosahedral lattice.
The derivation (see above) of icosahedral shells that are consistent with chemically similar protein contacts indicates that such a structure should contain 420 subunits. If hexamer-pentamer clustering is prominent, there will be 12 pentameric clusters and 60 hexameric ones. Low-resolution X-ray data, as well as molecular weight estimates, indicate that polyoma virus (and by inference other members of the group) contains 72 pentamers of VP1, i.e., 360 copies of VP1 rather than 420 (45). This result implies that VP1 subunits in the shell have six chemically very different environments (Fig. 14b) in apparent contradiction to the principle of closely related, specific bonding. The assembly of these viruses involves proteins VP2 and VP3, however, and these can form an internal framework or scaffold. Pentamers of VP1 appear to be stable subassemblies, and the interactions of VP1 subunits within a pentamer are indeed invariant. If the positioning of pentamers in the virion is determined by VP2 or VP3, interactions between pentamers can be relatively varied while still having accurate assembly. Thus papovavirus structure appears to be a simple illustration of subassembly construction, as in adenovirus (see next section). Evidence that the pentamer of VP1 is a stable, independent structure comes from analysis of electron micrographs of tu-
bular structures found in crude lysates of polyomavirus infected cells. A number of these are tubular arrays of pentamers (2,34). Like the analysis of nonviral structures from TMV protein (Fig. 8), this result shows the utility of understanding variant or polymorphic forms of a molecular assembly.

The DNA of papovaviruses is condensed by histones in a typical nucleosomal arrangement (23,27). The number of nucleosomes does not appear to be precisely determined, and the packaging of the "mini-chromosome" is not dependent on a particular three-dimensional configuration of the DNA/histone complex.

Adenoviruses

The adenovirus particle is a much more complex structure than those just described. It illustrates dramatically how large structures are constructed from subassemblies. The adenovirus particle is shown in Fig. 15, indicating substructures into which it can be resolved (22,26). The outer shell is strikingly icosahedral in shape. At first glance it appears to have a triangulation number of 25. The structures at the fivefold positions ("pentons") are different from the rest ("hexons"); however, and the hexons are chemically trimers rather than hexamers. Thus the structure really does not correspond to a simple subtriangulated icosahedral design at all.

The simplest substructures are the hexons and pentons. Hexons are trimers of a 110K polypeptide (conventionally denoted II). They have been crystalized, and high-resolution X-ray diffraction analysis is nearly complete (7). Assembly of the hexon from newly synthesized protein in vivo appears to require
another factor, the "100 K protein," which is also encoded by the virus. This protein does not form part of the final structure (17). In virions, protein VI is associated with each hexon. Pentons are composed of two polypeptides: one of 85 K forms the base, and the other of 62 K forms the projecting fiber seen at all vertices. A surprising recent finding is that the penton base may be trimere (20). This creates a symmetry-matching problem, as the vertices are five-coordinated. Symmetry-matching problems exist in other structures (e.g., a T4 phage tail with sixfold symmetry attached to a fivefold vertex of the head), but in no case have molecular details been visualized.

Dissociation of adenovirus particles by various methods (mild trypsin treatment, pyridine, deoxycholate plus heat, etc.) yields groups of nine hexons, as shown in Fig. 15. They are derived from virions as shown, and they include all but the peripentonal hexons. The groups of nine appear to be held together by protein IX, which copurifies with these structures. Rotational filtering of electron micrographs of groups of nine (19) show their threefold symmetry as well as the threefold symmetry of the hexons themselves. The way in which interaction with protein IX limits the lateral association of hexons in these structures is not yet understood.

The core of adenovirions contains DNA (~30 kb) and two basic proteins (V and VII). The arginine-rich protein VII is present in about 1/100 copies/particle, and it can neutralize about 50% of the DNA phosphates. Isolated cores are compact particles but without any very striking substructure (6).

Groups of nine can be assembled into closed structures that are similar to virus coats lacking pentons and peripentonal hexons (42). It has been suggested that DNA might be inserted into preformed shells during virus assembly in the nucleus, but empty precursors have not been detected. The formation of shells from groups of nine does indicate that the bonding properties of these complexes of proteins II, VI, and IX are sufficient to determine the angular interaction at an icosahedral edge and accurate closure into a correctly formed shell. The groups of nine are thus significant substructures of the particle, reflecting sets of important protein interactions, but they are probably not assembly intermediates.

**VIRAL ENTRY AND DISASSEMBLY**

This chapter has emphasized the significance of structure for requirements for specific and accurate assembly. Entry into cells and correct disassembly.
both properties of the organized particle, are equally essential for function. A number of enveloped viruses are now known to have fusion activity, dependent on a viral glycoprotein, that causes the viral membrane to fuse with a membrane of the cell to be infected. Structural aspects of fusion are discussed in the next chapter. In many cases fusion activity is activated by low pH, and uptake by receptor-mediated endocytosis must first occur (29). This pathway ensures that internal structures, such as nucleocapsids, are discharged into the cytoplasm. It does not do by itself, however, explain why they assemble spontaneously during budding and disassemble after infection. Entry and disassembly of nonenveloped structures, such as picornavirus or adenovirus, is even more puzzling. Adenovirus particles seem to be taken up by endocytosis, first appearing in clathrin-coated vesicles and then in uncoated endosomes (24). Undissociated virions can subsequently be detected in the cytoplasm. A mechanism for disrupting the endosomal membrane and a trigger for subsequent disassembly are properties of the virus structure not yet understood.

REFERENCES


