Virus structures and conformational rearrangements

Stephen C Harrison

Howard Hughes Medical Institute and Harvard University, Cambridge, USA

Structural virology is a burgeoning subspecialty. Our understanding of the molecular organization of viruses has begun to contribute directly to the analysis of viral attachment and entry, assembly, antigenticity, and even viral pathogenesis, but there are still more puzzles than answers. Recent crystallographic results have helped us to understand the structural changes in viruses that affect their assembly and infectivity.


Introduction

The selection of papers reviewed below focuses on contributions during 1994 that brought us closer to understanding the structural changes that are important for entry of viruses into cells or for interactions that regulate assembly. The emphasis is on molecular details, analyzed crystallographically.

Recognition of an RNA packaging signal: bacteriophage MS2

RNA viruses selectively incorporate their own genomic RNA. The mechanism of this selection often involves recognition of a packaging sequence (or sequences) by the coat protein [1,2]. Crystal structures of icosahedrally symmetric virus particles do not reveal the specific interaction, however, because it occurs at a unique location in the particle, whereas sixty equivalent orientations of the symmetrical particle are present in a crystal lattice. Only if the unique interaction with the packaging sequence were to perturb the symmetry of the particle surface would the crystal contacts ‘feel’ the internal asymmetry and select a single orientation of the particle in the crystal lattice. No case has yet been seen in which such a perturbation occurs.

The capsid (coat) protein of RNA bacteriophages such as MS2 recognizes a sequence of 19 nucleotides containing the start codon of the replicate gene, thereby controlling two processes: genome-specific encapsidation, and translation of replicate protein [3]. One dimer of the coat protein binds to this element, which is known to adopt the stem-loop conformation shown in Fig. 1a [1]. The 19-nucleotide ‘operator’ fragment can by itself stimulate assembly, leading to pseudovirions with 90 copies of the stem-loop element incorporated into an otherwise normal T=3 particle [4].

The structure of MS2 was first determined five years ago [5]. It contains 180 copies of the protein subunit, which forms 90 dimers across twofold and local twofold axes as shown in Fig. 1b. Crystals of recombinant capsids, when soaked in harvest solutions containing the 19-nucleotide operator element, were found to take up the RNA fragment and to incorporate it specifically into the particles. The structure of this complex has been reported recently [6*]. One RNA fragment is bound to each coat protein dimer. Those that are bound to the 60 dimers on local twofold positions (denoted AB' in Fig. 1b) are in a unique orientation, and a model of the complex of the dimer with the 19-mer can be built into an electron density map; those that are bound to the 30 dimers on strict twofold positions (denoted CC' in Fig. 1b) are in two symmetrically equivalent orientations, and a model of part of the fragment can be built and refined with half-occupancy. To the extent that it can be modeled, the CC' complex is the same as the AB' complex, and the interactions leading to specific recognition of the packaging sequence appear to be conserved.

The structure of the protein dimer–RNA operator complex is shown in Fig. 1c. The first three base pairs in the A-type helical stem are not clearly seen and therefore do not appear in the model. The loop assumes a compact conformation, with the cytidine at the -5 position (Cyt-5) and Ade-7 stacked on each other and in turn on Gua-8. The side chain of Tyr85 of monomer A (Tyr A85) stacks on Cyt-5. The essential base of Ade-4 and the required bulged purine base of Ade-10 project outward from the sugar-phosphate backbone to either side of the major groove, and they are almost related to each other by the dyad between the 'top' two base pairs of the stem. The protruding bases insert into similar pockets on the inward-facing surfaces of the A- and B'-subunits. As seen from the view in Fig. 1c, these protein surfaces are

Abbreviations

CCMV—cowpea chlorotic mottle virus; FHV—Flock House virus; TBSV—tomato bushy stunt virus; VP—viral protein.

©Current Biology Ltd ISSN 0959-440X
complementary to the folded surface of the RNA stem-loop, and recognition clearly involves the shape as well as the hydrogen-bonding properties of the RNA.

In the recognition of encapsidation signals by other RNA viral proteins, the folded structure of the RNA signal is likely to be at least as important as it is in MS2. For example, the specific protein–RNA interaction in cowpea chlorotic mottle virus (CCMV) appears to be mediated by a positively charged N-terminal coat-protein segment, which has no ordered structure in the absence of the nucleic acid [7]. That is, like the recognition peptides on the HIV proteins Tat and Rev, the piece of the CCMV coat protein that participates in RNA recognition acquires a defined three-dimensional conformation only when bound to the surface of its RNA target [8–10]. By contrast, in MS2, both surfaces are probably folded before the interaction. The structure of the coat protein in the empty capsid and in the complete virion does not differ from the structure in the complex with bound RNA fragments.

**Proteolytic processing in RNA virus assembly: polio empty capsids**

Polio and other picornaviruses assemble from subunits that are cleaved from a polyprotein precursor [11]. The initial cleavages, carried out by a virally encoded protease, yield species known as viral proteins VP0, VP1, and VP3. Heterotrimeric protomers composed of these three subunits are believed to be initial assembly intermediates, and pentamers of these protomers are thought to follow (Fig. 2). Further cleavage during or after assembly splits VP0 into VP4 and VP2 [12]. This process is probably autocatalytic, and requires the incorporation of RNA. Inhibition of RNA replication leads to the accumulation of empty particles containing VP0, VP1 and VP3, whereas mature virions contain 60 copies each of VP1–4. VP1–3 are similar in size (approximately 30 kDa) and in fold (jelly-rolls β-barrels); they are packed as diagrammed in Fig. 2 [13,14]. VP4 is much smaller (about 8 kDa); it is, in effect, the N-terminal arm of VP0, from which it derives. In mature picornavirus particles, an elaborate internal polypeptide network, formed by N-terminal arms from VP1–3 and by VP4, stabilizes the organization of the virion [13,14]. A recent crystallographic analysis of empty poliovirus capsids, in which the 60 VP0 subunits remain uncleaved and into which no RNA has been packaged, shows that in the absence of the final maturation cleavage, an unexpectedly large part of the internal scaffold is disordered [15]. The scissile bond, between the residues destined to form the C terminus of VP4 and the N terminus of VP2, is displaced by about 20 Å from the positions these residues occupy in the mature virion, and this displacement appears to prevent other parts of the internal network from forming. The empty capsid is probably not itself an intermediate in virus assembly, but rather a moderately stable association of (VP0,VP1,VP3)5 pentamers. The folded β-barrel cores of the subunits and their packing in the icosahedral shell are identical in the empty capsid and in the virus, but part of the network...
of interactions in the mature virus requires cleavage of VP0 in order to form important internal links between pentamers. The immature (uncleaved) pentamer is obviously poised to form these links, but its VP0 component must cleave in order to accomplish it.

![Diagram of subunits in picornaviruses](image)

**Fig. 2.** Packing of subunits in picornaviruses. Shown are the positions of VPs 1–3; VP4 is entirely internal. VP0, VP1, and VP3 are derived by cleavage from a precursor, and autolytic cleavage of VP0 to VP4 and VP2 occurs during assembly. Along with VP4 (which is effectively an extension of the VP2 arm), the N-terminal arms of VPs 1–3 form an elaborate inner scaffold. The recent structure of empty poliovirus particles composed of uncleaved VP0, together with VP1 and VP3, shows that VP0 cleavage leads to a major rearrangement and stabilization of this inner scaffold [15*]. Bold line encloses one protomer; solid pentagons, triangles, and ovals indicate fivefold, threefold, and twofold axes, respectively. (Diagram courtesy of J Johnson, Purdue University.)

How is cleavage of VP0 linked to RNA incorporation? And what is the mechanism of the cleavage itself? The new structure eliminates some old hypotheses and suggests, but does not prove, some new ones [15*]. Disordering of much of the internal scaffold leaves a cavity around the icosahedral threefold axes unoccupied, on the inward-facing surface of the viral shell. This three-lobed hollow contains segments of the VP1 and VP2 arms in the mature virion. A similar cavity is in the same position in comoviruses, such as beanpod mottle, where partly ordered RNA can be seen within it [16]. Basavappa et al. [15*] suggest that this cavity in the picornavirus pro-capsid might likewise be a site for RNA interaction and that assembly could be initiated by association of three pentamers with an appropriate RNA structure. They also suggest how formation of such an initiation complex might lead to an autocatalytic VP0 cleavage. Tests of their proposals, by mutation of key residues, have yet to be performed.

**Participation of RNA in a conformational switch: nodaviruses**

RNA encapsidation is also linked to a maturation cleavage in nodaviruses, T=3 RNA viruses that infect insects, fish, and mammals. The structure of Flock House virus (FHV) shows that RNA binding is, in addition, part of a conformational switch in assembly [17*]. The first plant–virus structures illustrated that one way of achieving a T=3 'quasi-equivalent' design is to construct the shell from 90 chemically identical dimers: 30 disposed across strict twofold axes with little local curvature and 60 across local (quasi) twofold axes, with strong local curvature [18,19]. This packing is represented diagrammatically in Fig. 3a. In the plant viruses, the folding back of an N-terminal arm of the subunit serves as the conformational switch [20]. When ordered, this arm runs along the line between threefold and twofold axes, forcing the dimer into a 'low-curvature' conformation; when disordered, it leaves room for domains of the two subunits to close in on each other and generate a 'high-curvature' position. The assembly pathway must provide for correct setting of this conformational switch as each dimer adds to the growing shell [20]. In FHV, the switching role appears to be shared by an N-terminal arm and a double-helical stem of RNA, which runs beneath it and across the icosahedral twofold axis (Fig. 3b,c). The protein–RNA interactions involve only the sugar-phosphate backbone, and the image of course shows an average of 30 different double-helical segments, presumably of somewhat different lengths [17*]. It is possible that assembly is initiated by specific recognition of one such segment, and that other stem-like segments present in the viral RNA are recruited into symmetry-related positions as the shell forms.

The nodaviruses are in their design something of a hybrid between the plant viruses and the picornaviruses: a post-assembly proteolytic cleavage between a C-terminal, α-helical segment and the main jelly-roll domain of the subunit is required for infectivity [21]. RNA encapsidation stimulates this cleavage, and subunits in empty capsids remain largely unprocessed. The C-terminal segment is indeed a principal RNA contact (Fig. 3b), and it will be interesting to see whether the mechanism of RNA-triggered proteolysis is similar in nodaviruses and picornaviruses.

What is the importance of an RNA-stimulated proteolytic cleavage? In picornaviruses, the empty capsid is significantly less stable than the mature virion, presumably because the internal network of N-terminal arms is incompletely formed. The virion is an extremely robust structure: the poliovirus particle is, for example, resistant to treatment with SDS. Nonetheless, binding of the virus to its receptor triggers an irreversible conformational change. In the case of polio, this change involves expansion, loss of VP4, and externalization of part of the N-terminal arm of VP1 [11,22,23]. In the case of
other picornaviruses, receptor binding can lead directly to disassembly [24]. Thus, the cleavage of VP0 primes the virion for a receptor-catalyzed conformational transition. A detailed picture of this transition remains to be worked out. It has been suggested that the altered state of the polio virion may resemble the expanded state of certain T=3 plant viruses — a state that has been studied in tomato bushy stunt virus (TBSV) [25] and CCMV [26*], as described in the next section. The jelly-roll β-barrel cores of picornaviral VPs 1-3 resemble the shell domain of the TBSV subunit, and the packing of these domains in the viral surface is very similar in the two viruses (compare Fig. 2 to Fig. 3a). Mutations believed to affect the conformational transition in poliovirus map to subunit contacts that correspond to those containing the Ca²⁺-binding sites in TBSV, and the expansion of TBSV may provide a rough model for the change in polio and other picornaviruses that occurs in response to receptor attachment [15*,23].

CCMV and its expanded form

The recently reported CCMV structure differs from that of other T=3 viruses seen so far, such as TBSV and FHV, in the orientation and arrangement of its β-barrel domains and in some other aspects of particle design [26*]. Nonetheless, the underlying features of all these T=3 structures are actually closely related. The β-barrel of the CCMV subunit is very similar to the one in southern bean mosaic virus (SBMV) and TBSV. Indeed, the structure was determined by a modified molecular replacement approach, in which the placement of the barrels was determined by fitting a low-resolution model from cryo-electron microscopy [26*]. The CCMV shell assembles from dimers of its subunit, so that it may (like TBSV) be described as composed of 60 AB dimers and 30 CC dimers. The twofold interactions are stabilized by interchange of C-terminal arms (about 12 residues), clamped in place between the N-terminal arm and one
of the walls of the partner subunit. The orientations of the \( \beta \)-barrels are such that the strands are roughly parallel to the nearest five- or quasi-sixfold axis; moreover, the barrels make strong contacts within the five- and sixfold clusters, giving rise to the typical 'capsomeric' appearance of pentameric and hexameric morphological units. It is important to bear in mind, however, that these units probably do not have independent identity outside the viral shell, because the assembly unit is a dimer. The sixfold clusters are further stabilized by a \( \beta \)-annulus composed of segments from the N-terminal arms of the surrounding subunits; no corresponding structure is seen on the fivefold cluster, and the equivalent residues are disordered.

The CCMV structure reveals likely Ca\(^{2+}\)-binding sites at the interfaces that radiate out from the quasi-threefold axis. Chelation of Ca\(^{2+}\) and elevation of pH above 7 leads to expansion, presumably because of the repulsion of apposed negative charges, leading to dissociation of the Ca\(^{2+}\)-stabilized interfaces. Similarly located divalent cation sites govern TBSV expansion, which was analyzed crystallographically a number of years ago [25]. A combination of crystallography and cryo-electron microscopy has led to a model for the expansion of CCMV [26*], in which contacts within the hexameric and pentameric morphological units are conserved. The dimer interactions are retained, but shifted in their geometry. These interactions are mediated by the C-terminal arms, and flexibility is therefore built into their character. In TBSV, the dimers must likewise adjust: in that case, invariant projecting domain contacts clamp the dimer together, and flexion of the hinge between shell and projecting domains allows for the adjustment [25]. Thus, despite a number of differences in their subunit interactions, CCMV and TBSV expand by remarkably similar mechanisms.

Fig. 4. Sialyl lactose (NeuNAc\(_{2} 3\)Gal\(_{1} 4\)Glc) in the receptor-binding site of polyoma virus VP1 [28**].
Polyoma virus receptor binding

The structure of simian virus (SV)40, reported in 1991, showed an even more dramatic tying-together of subunits by C-terminal arms than the one just described in CCMV [27]. The structure of its close homolog, polyoma virus, has now been determined at 3.65 Å resolution, in a complex with an oligosaccharide that is believed to be a specific component of its receptor [28*]. Polyoma can infect mouse cells that have surface glycoproteins carrying sialylated oligosaccharides with an α2,3 linkage between a terminal sialic acid and a penultimate galactose [29]. Sialyl lactose is therefore a good homolog of the receptor fragment. The complex of sialyl lactose and polyoma virus shows that the receptor binds in a shallow pocket at the outer margin of the VP1 subunit (Fig. 4). Several hydrogen bonds, including an arginine/sugar carboxylate pair, appear to determine sialic-acid recognition. The rest of the fragment is probably accommodated specifically by virtue of its shape, which is complementary to the shallow groove in which the oligosaccharide rests. An α2,6 linkage from the sialic acid, for example, would lead to a shape for the oligosaccharide incompatible with the outline of the groove. Influenza virus also uses sialic acid as its receptor, and the nature of its specific recognition by the hemagglutinin (HA) has recently been analyzed in particularly fine detail [30*]. The binding pocket for sialic acid on the

HA ‘top domain’ is not related to the one on polyoma VP1.

The fusion-active form of influenza hemagglutinin

After binding to its sialic acid bearing receptor, influenza virus enters the endocytic pathway through clathrin-coated vesicles [31]. Acidification of the intracellular compartment in which the virus arrives triggers a conformational change in the HA trimer, resulting in a form of the molecule that can mediate fusion of viral and cellular membranes. Recent publication of the crystal structure of an HA fragment, isolated from low pH treated protein, has allowed the world to see how dramatic this conformational change really is [32*].

The HA is composed of two chains, HA1 and HA2, cleaved from a precursor, HA0 [33]. The N-terminal residues of HA2, just distal to the cleavage site, have a critical role in the fusion mechanism. The structure of the native HA trimer (Fig. 5a) shows that the HA1 chain folds to form a receptor-binding ‘top’ domain, which is borne about 100–135 Å above the viral membrane by a helical stalk formed by HA2. Each HA2 chain in the trimer contributes a long α-helix to a threefold cluster at

![Image](image_url)

**Fig. 5.** Influenza hemagglutinin. (a) The native influenza HA trimer. HA1 is in white, HA2 in gray. (b) The fusion-active conformation of HA2, aligned so that the central, three-helix bundle (residues 76–105) is at the same height as the corresponding region in (a) (see [c,d]). Note the coil→helix transition in residues 55–75, which translates the N terminus of HA2 to the ‘top’ of the molecule. (c) and (d) Corresponding views of the native and low pH treated monomer. Residue numbers refer to the HA2 chain, and can be used as markers to follow the conformational rearrangement in these figures and in (a,b). (Courtesy of F Hughson, P Bullough, J Skehel, DC Wiley.)
the center of the stalk and a short α-helix at the outside. An extended loop joins the shorter helix to the longer one. The fusion peptide at the N terminus of HA2 is tucked into the stalk near its midpoint.

Treatment of the trimer at low pH (to mimic the acidification that occurs intracellularly) triggers a rearrangement, of which a number of features have been suggested by earlier experiments. Cross-linking of the ‘top’ domains prevents fusion activation, showing that these domains must come apart [34]. Moreover, only by trimming away the hydrophobic fusion peptide can the low-pH form be solubilized, showing that this peptide must withdraw from its protected position in the native trimer [35]. A proteolytic fragment of the low pH treated HA, from which not only the fusion peptide but also most of the HA1 chain have been removed, can be crystallized; its structure is shown alongside that of the native HA, in Fig. 5b. The most striking aspect of the conformational change is a loop→helix transition in the segment linking the shorter and longer helices of ‘native’ HA2. This transition (seen more readily in Fig. 5c,d) was predicted by studies on model peptides with sequences that span the loop [36]. An important consequence of the formation of a continuous helix is that the N-terminal part of HA2 (including the fusion peptide, which is lacking in Fig. 5b) is transported to one end of the molecule, probably facing the target membrane. Another, unexpected aspect of the conformational change is the introduction of a sharp bend at the center of the long helix. This break in the continuity of the helix occurs essentially at the point where the fusion peptide was tucked into the stalk. The bend of nearly 180° causes the structural elements at the C terminus of HA2 to lie alongside the ‘upper’ part of the original long helix. Considerable disruption of the organization of the polypeptide chain must occur in the region just proximal to the transmembrane segment (which is, of course, missing from the crystalline fragment). Thus the picture that emerges from a comparison of native and low-pH structures suggests possible events at both the target (host-cell) membrane and the viral membrane.

Perspectives

Virus particles are not passive objects. The transformations they undergo can be analyzed with increasing incisiveness by X-ray crystallography and allied methods. The new structures summarized here reveal some molecular switches governing specific assembly as well as examples of conformational changes important for entry and subsequent disassembly. Any of these transformations are potential targets for intervention if specific inhibitors can be discovered or designed.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:
- of special interest
- of outstanding interest


Describes how the coat protein dimer of MS2 recognizes an RNA operator sequence, which also serves as an assembly signal.


Crystallographic analysis of the poliovirus empty capsid shows that the final organization of the inner scaffold of arms depends on cleavage of VP0.


The Flock House virus structure, as summarized in Fig. 3 of this review,


    CCMV resembles other T=3 structures, but with novel features that account for its ‘capsomeric’ appearance in the electron microscope.


    The first molecular-resolution structure of a complex between an icosahedral virus and a receptor fragment.


    Analysis in particularly fine detail of the recognition of sialic acid and derivatives by the hemagglutinin of influenza.


    The dramatic rearrangement of influenza HA2 triggered by low pH is documented by a striking crystal structure.


SC Harrison, Howard Hughes Medical Institute and Department of Molecular and Cellular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138, USA.