

Structure of the reovirus core at 3.6 Å resolution

Karin M. Reinisch*, Max L. Nibert† & Stephen C. Harrison*‡

* Harvard University and ‡ Howard Hughes Medical Institute, Fairchild Building, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

† Department of Biochemistry and Institute for Molecular Virology, 1525 Linden Drive, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA

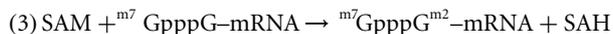
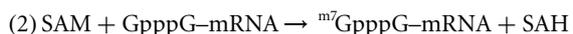
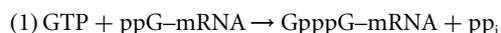
The reovirus core is an assembly with a relative molecular mass of 52 million that synthesizes, modifies and exports viral messenger RNA. Analysis of its structure by X-ray crystallography shows that there are alternative, specific and completely non-equivalent contacts made by several surfaces of two of its proteins; that the RNA capping and export apparatus is a hollow cylinder, which probably sequesters its substrate to ensure completion of the capping reactions; that the genomic double-stranded RNA is coiled into concentric layers within the particle; and that there is a protein shell that appears to be common to all groups of double-stranded RNA viruses.

Reoviruses, like other plant and animal double-stranded (ds) RNA viruses, contain a transcriptionally competent internal capsid particle (ICP), which remains intact after viral penetration¹. The reovirus ICP, known as the core, transcribes plus-strand copies from each of the ten genomic segments packaged within it, adds a methylated guanosine cap to the 5' end of each transcript and exports the mature mRNA into the cytoplasm of the infected cell. The core is thus an elaborately organized molecular machine.

About 700 Å in overall diameter, the reovirus core (Fig. 1) contains five of the eight proteins that make up a complete virion². Three of the five—λ1, λ2 and σ2—are symmetrically arranged in the icosahedral particle, and most of their amino-acid residues (4,500 in the icosahedral asymmetric unit) are clearly resolved in our refined structure. The reovirus core resembles in several respects the core of bluetongue virus (BTV)³, a member of a different genus of dsRNA viruses, but it has a number of distinctive features.

An unusual aspect of protein contacts in the reovirus core is the prevalence of non-equivalent interactions at many interfaces. The contacts within the λ1 shell, and even more strikingly those between this shell and the σ2 subunits that decorate it, show that when the pressure for genetic economy is strong enough, evolution can build at least two totally different specificities into a single protein surface. We draw similar conclusions from the published structure of the core of BTV³.

The sequence of modifications to the 5' end of the mRNA is incorporated into the spatial arrangement of enzymatic domains in the λ2 pentamer, a hollow cylinder that forms projecting turrets around each of the icosahedral fivefold axes. The reovirus mRNA cap consists of a 7-*N*-methyl guanosine linked by three phosphate groups to a 5'-terminal guanosine. The turrets catalyse the addition of this last guanosyl moiety to the mRNA as well as the transfer of a methyl group from *S*-adenosyl-*L*-methionine (SAM) to both the N7 of the added guanosine and to the 2' O of the first template-encoded nucleotide, which is also guanosine in all ten reovirus mRNAs:



where SAH is *S*-adenosyl-*L*-homocysteine⁴. Each of these reactions corresponds to a domain of λ2, which therefore resembles other complex enzymes that encode a biosynthetic pathway by concatenating domains to carry out successive reactions. The active

sites of the domains all face the interior of the pentameric λ2 turret, and the residence time of the mRNA 5' terminus within this hollow is probably sufficient to ensure completion of the cap.

Genomic dsRNA is coiled tightly within the reovirus core, in a manner reminiscent of the packing of dsDNA in phage heads⁵. The smooth inward-facing surface of the λ1 shell organizes the packed RNA into roughly concentric layers. Each segment of RNA is likely to be associated with a transcriptase complex tethered near a fivefold vertex^{6,7}.

The λ1 shell has structural relatives in other dsRNA viruses^{3,8} and it appears to be a common feature of nearly all viruses in this family. The other proteins of the ICP and especially those of the outer shell appear to vary among dsRNA viruses, corresponding to different strategies for infection and spread.

Structure determination

The very large unit cell (space group *F*432, *a* = 1,255 Å), the small crystal dimensions (about 150 μm) and the failure of cryopreservation

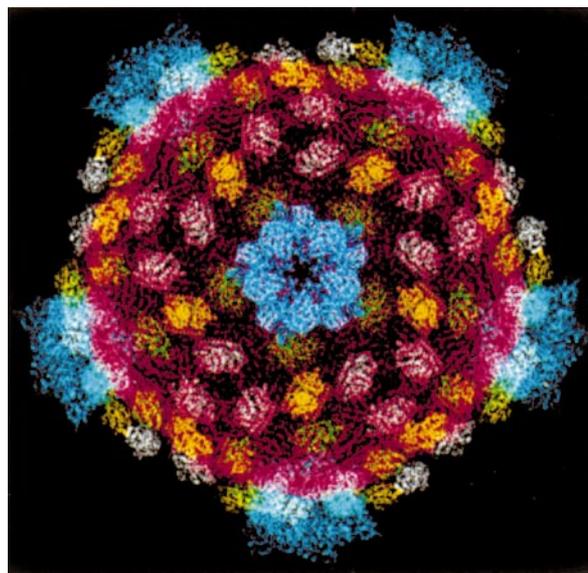


Figure 1 The reovirus core particle, represented by Cα traces of the constituent subunits. λ1 (relative molecular mass (*M*) 142K (ref. 46), 120 copies; shown in red) forms the shell that packages RNA and defines the symmetry and size of the particle. σ2 (*M*, 47K, 150 copies; shown as yellow, green and white nodules) stabilizes the λ1 shell. λ2 (*M*, 144K, 60 copies; shown in blue) forms turret-like structures around the fivefold axes that cap the nascent mRNA and organize its extrusion.

dictated our data-collection strategy. We determined phases at 27 Å resolution using a map reconstructed from electron cryomicroscopy (cryoEM)⁹ and extended using the fivefold non-crystallographic symmetry and solvent flattening.

The λ1 shell

The outer surface of the λ1 shell is relatively smooth, except for low ridges that border the binding sites of λ2 and σ2. Five monomers (set A) radiate from the fivefold axis like the petals of an inverted buttercup, and members of a further set of five (set B) interdigitate with the first. Twelve such decameric units together form a complete protein shell (Fig. 2).

Residues 240–1,275 of λ1 form a large, plate-like structure with no strongly defined domain boundaries (Fig. 3). Comparison of the A and B conformers reveals that two subdomains undergo a simple shift relative to one another: subdomain II (residues 482–922) is an insert into subdomain I (residues 253–470 and 923–1,260; Fig. 3). Each of the subdomains behaves as a rigid body when λ1A is transformed into λ1B, with minor local adjustments in loops at the perimeter of the subunit. Two of the peripheral loops in subdomain II—residues 560–568 and 774–794—block the opening through which plus-strand transcripts, destined to become mRNAs, are likely to pass from the central cavity into the λ2 turret. We believe that these loops will bend out of the way during active RNA synthesis. Direct evidence for their flexibility comes from inspection of λ1B, where residues 563–570 are disordered.

The two conformational states of λ1 allow it to fit into two distinct local environments. λ1A and B make similar contacts only along one side of subdomain I, where λ1A from one decamer faces λ1B from another decamer across a local dyad. This is the only instance of quasi-equivalence in the structure. The interactions of the two λ1 conformers, both with each other and with σ2, are otherwise very different (Fig. 2).

The first 240 residues of λ1 are not part of the tightly folded, plate-like structure just described. Indeed, the amino-terminal residues are largely disordered, except for five copies per decamer (rather than ten) of each of three segments, containing residues 1–12, 40–167 and 181–240, respectively. The three ordered segments are part of a network of arms lying beneath the main part of the shell, and it is evident that the N-terminal residues of λ1 form variably disposed, extended structures that may help to tie the shell together. The way in which the ordered segments connect with each

other is not uniquely determined by well defined electron density, but the positions of the ends of these segments restrict possible connectivities, as summarized in Fig. 2c.

Residues 181–208, in the last of the three ordered arm segments just described, fold into a classical (Cys)₂(His)₂ zinc finger, as predicted from inspection of the sequence. Clear density for residues 181–208 and 215–240 and a lower-resolution bridge between them show that the ordered finger belongs to a λ1B subunit (Fig. 2c). The module is tucked into a pocket formed by three λ1 subunits. The N-terminal end of its α-helix, which contains the residues for specific DNA or RNA recognition in other zinc fingers, is accessible, but the surrounding parts of the λ1 shell would probably interfere with close approach of any genomic dsRNA segment. This finger would, however, be exposed at intermediate stages in assembly, and its variably positioned counterpart in the λ1A subunit is a candidate for RNA interaction, even in the mature core.

The σ2 clamp

The compact, globular monomer σ2 binds at three distinct locations within each icosahedral asymmetric unit (Figs 4 and 5). One of these sites (the position of σ2-i) lies over the middle part of λ1A. The second (σ2-ii) bridges from the middle of λ1B across to the carboxy-terminal portions of λ1B from another decamer. Although they receive the same face of σ2, these two sites are similar only in that each is a very shallow nest on the surface of the λ1 shell. Moreover, the interacting parts of σ2 undergo only modest conformational adjustments to fit tightly into these completely non-equivalent positions. The largest conformational difference is in residues 32–50, near the centre of the interface with the λ1 shell, where helix 39–46 unravels in σ2-ii. The third site for σ2 is directly across the icosahedral dyad over the C-terminal portions of λ1A. It contains a single σ2 (σ2-iii), in one of two equally likely, twofold-related orientations. The larger part of this interaction is identical to that of σ2-ii with λ1B.

The two unrelated σ2 binding sites are both of high affinity, as the σ2 subunit has only limited other contacts within the core but no tendency to dissociate. Its contacts with λ1, like those between neighbours in the λ1 shell, are marked examples of non-equivalent but strong and specific bonding, where a given surface of one subunit interacts with two completely different surfaces of another. When faced with the powerful pressures imposed by the compactness

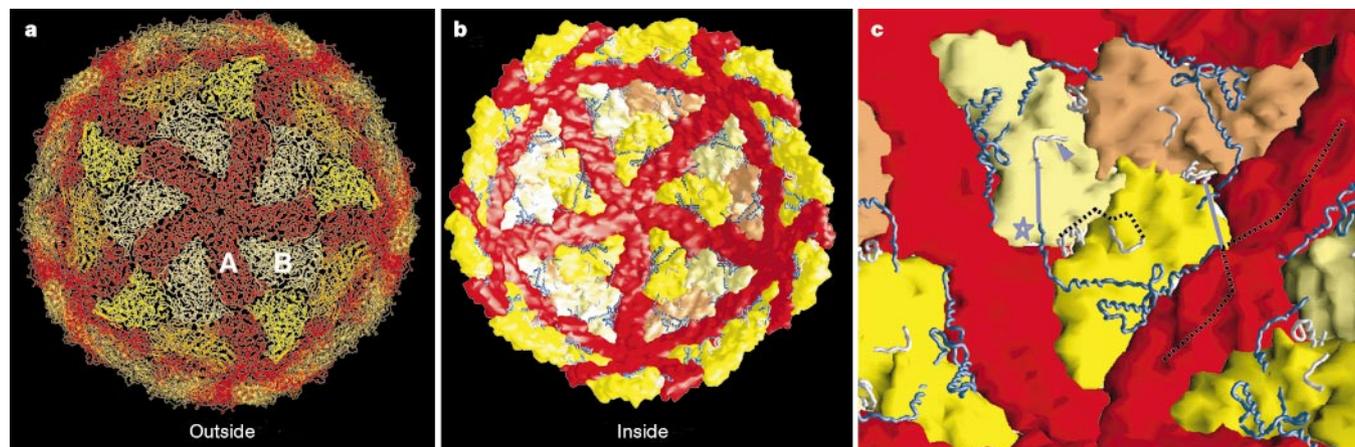


Figure 2 The λ1 shell. **a**, CryoEM trace, seen from the exterior. Five copies of λ1 (set A, in red) radiate from the icosahedral fivefold axis, and members of a second set (set B, in varying shades of yellow) interdigitate with the first to form a decamer. The interactions of the two λ1 conformers with each other are very different in all but one short interface across a local dyad. **b**, Schematic of the reovirus core from the interior. With the exception of three disconnected fragments, the first 240 residues of λ1 are disordered. These fragments include residues 1–12 (tentative assignment), 40–167 and 181–240. Residues 1–12,

40–167 and 181–208, a zinc finger, are shown. Residues 40–167 cannot belong to the λ1B subunit under which they lie, as the distance between 167 and 181 is too great. **c**, A small portion of **b**, magnified. A possible connectivity is shown in purple. In this scheme, residues 1–181 form an arm that ties together the λ1B subunits around a threefold axis. An arrow and a star mark the positions of residues 1 and 40, respectively. Dotted lines indicate other connectivities possible on the basis of distance criteria.

of viral genomes, proteins can evidently evolve to interact stably and specifically in two or more unrelated ways. The non-equivalent interdomain and intersubunit contacts in HIV reverse transcriptase¹⁰ appear to be the outcome of similar evolutionary ingenuity. In most virus structures analysed at high resolution, equivalent or quasi-equivalent interactions predominate. Even in polyoma¹¹ and SV40 (ref. 12), where 60 of the 72 VP1 pentamers are in 6- rather than 5-coordinated positions, use of arms to tie the structure together maximizes the local similarity of interpentamer contacts, despite the variable long-range geometry. As discussed below, interactions in the VP3 shell of BTV³ have the same non-equivalent characteristics as those in the λ 1 shell of reovirus, but there is no σ 2 equivalent. In the orbiviruses and other dsRNA viruses lacking a σ 2-like protein, the proteins that constitute the innermost shell (VP2 and VP3 in rota- and orbivirus, respectively) can self-assemble into icosahedral particles^{3,13,14}. When reovirus λ 1 is expressed in mouse L fibroblasts¹⁵ or in insect cells (M.L.N., unpublished data), however, no icosahedral particles form unless σ 2 is also expressed. The structure confirms the conclusion drawn from these results that σ 2 is a stabilizing clamp.

The λ 2 turrets are mRNA-capping complexes

The λ 2 pentamer is a hollow and slightly flared cylinder (Fig. 6). The cavity in its centre, through which mRNA must pass, is 70 Å in diameter at its widest, but it narrows to 15 Å at the top, where five flaps form a lid. The total volume of the channel is $2 \times 10^5 \text{ \AA}^3$, which can accommodate up to about 300 nucleotides of nascent mRNA. The pentamers interact almost exclusively with λ 1A. Expressed alone λ 2 is monomeric^{16,17}, and it must therefore require a scaffold to assemble into turrets.

The λ 2 monomers have a series of seven domains, strung together with the most N-terminal domain at the base of the turret and the most C-terminal domain at the top (Fig. 6). The long axis of the subunit is at about 45° to the fivefold axis. The first 385 residues form a cup-like structure with its interior open into the hollow of the turret and its side resting on the λ 1 shell. Limited proteolysis of λ 2 has identified this domain as a guanylyltransferase (GTase). Analysis of a series of Lys-to-Ala mutations shows that Lys 190 forms the covalent nucleotidyl-enzyme intermediate and that Lys 171 also participates in GMP transfer¹⁷; both project into the active-site cavity (Fig. 6d). A narrow channel gives access to the active site for nucleotides from outside the turret. The structure of

the reovirus GTase appears to be a new fold; it differs from the known structure of PBCV-1 GTase¹⁸.

Residues 386–433 and 690–802 contribute to a small domain that bridges the GTase and the methylases. It is a four-strand sheet, protected by α -helices facing toward the outside of the turret, and it probably functions as a structural support for the three enzyme domains.

Residues 434–691 form a methyltransferase domain (methylase-1). Helices sandwich a seven-strand β -sheet of mixed polarity in a variant of the SAM-binding domain seen in numerous other methyltransferases (Fig. 6)^{19,20}. The SAM site is conserved among

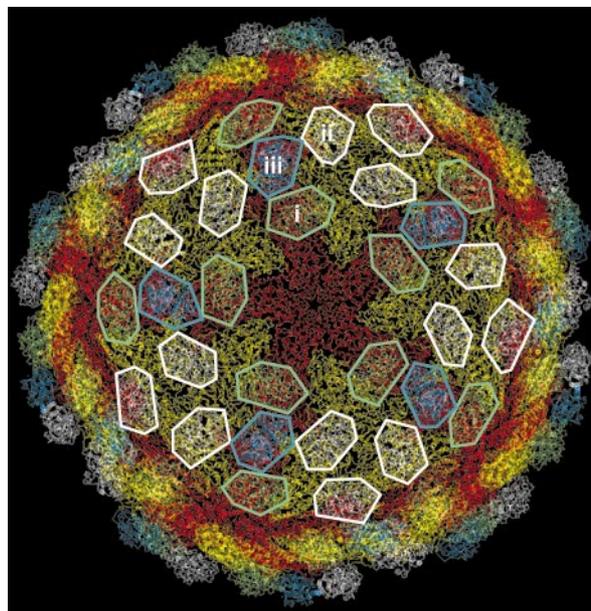


Figure 4 Subunit σ 2 has 150 binding positions on the λ 1 surface. There are 60 copies each of sites i (green outlines) and ii (white outlines) and 30 of site iii (blue outlines). Site iii lies on an icosahedral dyad, and σ 2 in that position adopts two opposite orientations in 50:50 ratio. Sites i and ii involve very different residues of λ 1; each identical half of site iii (on λ 1A) is the same as part of site ii (on λ 1B). Most of the contacts of σ 2 are with the λ 1 shell, but σ 2-iii has sparse contacts with both σ 2-i and ii, and σ 2-i has a modest contact with λ 2.

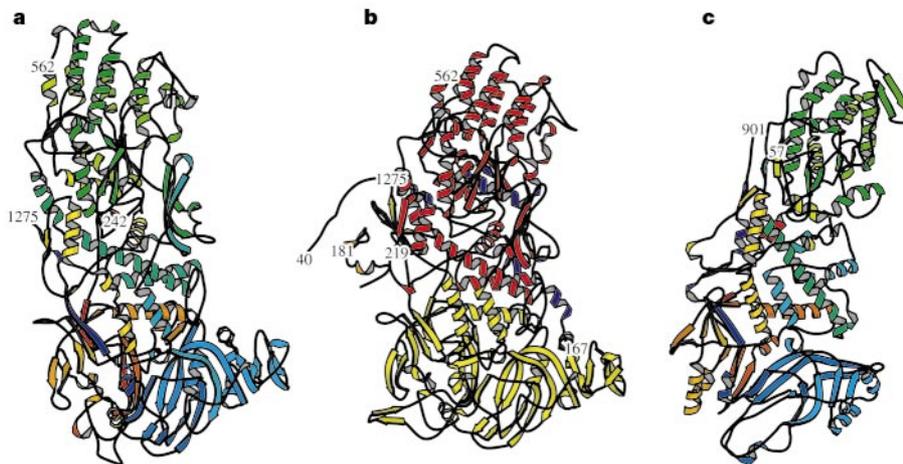


Figure 3 Two conformations of λ 1 and comparison with BTV VP3. **a**, Course of the polypeptide chain of λ 1A, in colours graded from red at the N terminus to blue at the C terminus. λ 1 is mostly α -helical, with significant regions of β -sheet present only in the part furthest from the fivefold axis. **b**, λ 1B; subdomains I and II are yellow and red, respectively. λ 1A and B differ with respect to a conformational shift about a pivot between

the two subdomains. The zinc finger (residues 181–208) is gold, and residues 40–167 from a neighbouring λ 1 subunit are blue. **c**, VP3-A, the λ 1A equivalent from BTV, in colours graded from red at the N terminus to blue at the C terminus³. Although λ 1 and VP3 have no sequence similarity and details of their folds differ, their overall conformations are clearly closely related.

these enzymes, and we can thus predict where SAM will bind to methylase-1 (Fig. 6). Difference maps computed from diffraction data taken after diffusion of SAH into the core crystals confirm that SAH binds in this location, in a channel that connects the interior and the exterior of the turret. A conformational change in residues 519–524 and 579–587, such that the 580s loop blocks off the exterior of the channel, accompanies SAH binding. The mRNA cap must bind close by, in the cavity between methylase-1 and the GTase domain of an adjacent $\lambda 2$ monomer. Sharing of the active site between two monomers explains the absence of methylase activity in monomeric $\lambda 2$ (ref. 16).

Residues 804–1,022 form the second methylase domain (methylase-2, Fig. 6). The order and directionality of the strands and the positions of some α -helices are as in the consensus fold; residues 926–962 and 983–1,017 represent an insertion, forming a three-strand parallel sheet and two helices. The SAM-binding site, identified from the diffusion experiment described above, lies as anticipated from homologous enzymes in another channel to the exterior of the turret. SAH binding does not induce a conforma-

tional change at this site.

The last 250 residues of the $\lambda 2$ polypeptide chain form a three-domain flap at the top of the turret. In virions and intermediate subviral particles (ISVPs), this flap is a cantilever to anchor the cell-attachment protein, $\sigma 1$ (ref. 9). In cores, it may serve as a gate to retard exit of the 5' terminus of the mRNA. The three domains all have structures related to the immunoglobulin fold. The first two resemble the V and C regions of an antibody light chain, although they lack disulphide bonds; the third is a sandwich of two three-strand sheets, folded like a truncated V domain.

Which domain is the 7'-N- and which the 2'-O-methylase? And what determines the order of the modification reactions? The view of two adjacent monomers in Fig. 6f shows that although there is no groove to guide the mRNA terminus from one active site to another, the spatial arrangement of sites in the turret may reflect the temporal sequence of the capping reactions. Guanylyl transfer occurs first, near the base of the turret, where newly synthesized transcript enters the multi-enzyme cavity. The GTase site from one monomer (for example, monomer A in Fig. 6f) is closer to the

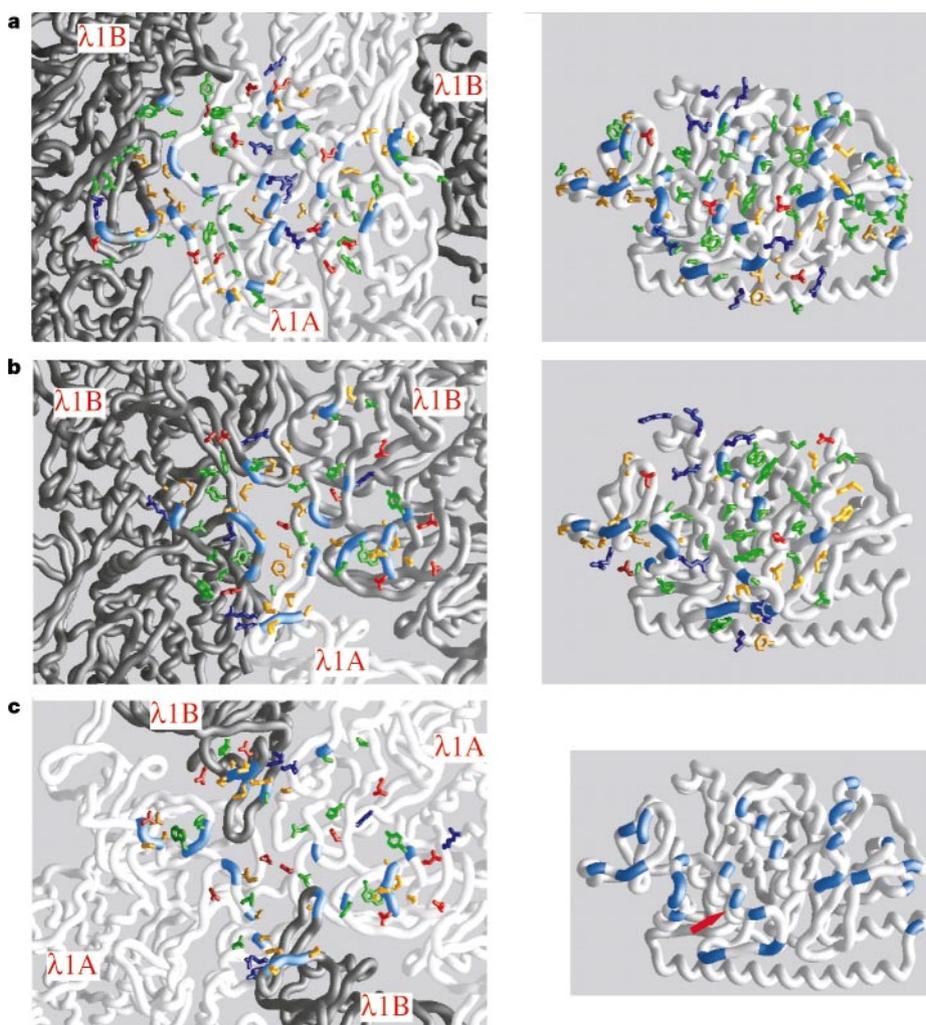


Figure 5 The $\lambda 1$ – $\sigma 2$ interface. **a, b**, Left, the $\lambda 1$ surfaces onto which bind $\sigma 2$ -i and $\sigma 2$ -ii, respectively. Right: the corresponding contacting face of $\sigma 2$ -i and $\sigma 2$ -ii. The $\lambda 1$ surfaces are drawn to receive $\sigma 2$ -i and $\sigma 2$ ii in the same orientation; the images of $\sigma 2$ on the right may be flipped around the axis parallel to the long side of the page to orientate them correctly over the corresponding images of $\lambda 1$ on the left. **c**, Left, site iii on the $\lambda 1$ surface. Right, $\sigma 2$ -i, with a red arrow indicating an α -helix that unravels in $\sigma 2$ -ii. Different $\lambda 1$ subunits are painted in different shades of grey except in **c**, where both $\lambda 1$ As are white and both $\lambda 1$ Bs are dark grey. Main-chain segments within 4 Å of the interacting surface

are coloured light blue. Side chains that fall within 4 Å of the interacting surface are blue (positively charged), red (negatively charged), green (neutral polar) or orange (hydrophobic). $\sigma 2$ -i binds mostly to $\lambda 1$ A, whereas $\sigma 2$ -ii binds mostly to two copies of $\lambda 1$ B and acts as a clamp. The zinc finger (not shown) lies inside the core under $\sigma 2$ -ii but makes no contacts with it. $\sigma 2$ -iii binds at the icosahedral twofold axis; note the similarity of the $\lambda 1$ surface with which it interacts (left) to the white and light grey portions in **b**. $\lambda 1$ residues marked in the last panel on the left are inferred, as $\sigma 2$ -iii has not been built explicitly into the density.

methylase-1 active site in the clockwise neighbour (monomer B in Fig. 6f) than to any other methylase-1 site, so the newly guanylylated 5' end is more likely to pass from the GTase on one subunit to methylase-1 on the neighbour than to any other methylase site. As N7 methylation precedes 2'O methylation⁴, we suggest that methylase-1 is the N7-modifying enzyme and that methylase-2 is the O-methyltransferase. In this tentative assignment, the sequence of domains in the subunit corresponds to the order of reactions in the pathway. N7 methylation would take place in the middle of the turret, and the last reaction, 2'O methylation, would take place at the top, near the exit pore. It is reasonable to suppose that the 2'O methylase, like the related enzyme from vaccinia virus²¹, accepts

termini only when the base of the 5'-guanylyl group has been methylated at N7.

Containment, to allow the 5' end of the transcript to visit various sites, perhaps nearly randomly, but also to ensure that modification will be complete before the cap emerges into the cytoplasm, appears to be a principle embodied in the design of the λ 2 turret. The turret is a container with several, densely arrayed, inward-facing active sites. Similar arrangements have been seen in chaperonins such as GroEL and in intracellular proteases such as ClpAP and the 26S proteasome (see ref. 22 and references therein). In GroEL, release of substrate is timed by ATP hydrolysis; in λ 2, release of the transcript 5' terminus is more likely to be triggered simply by filling of the

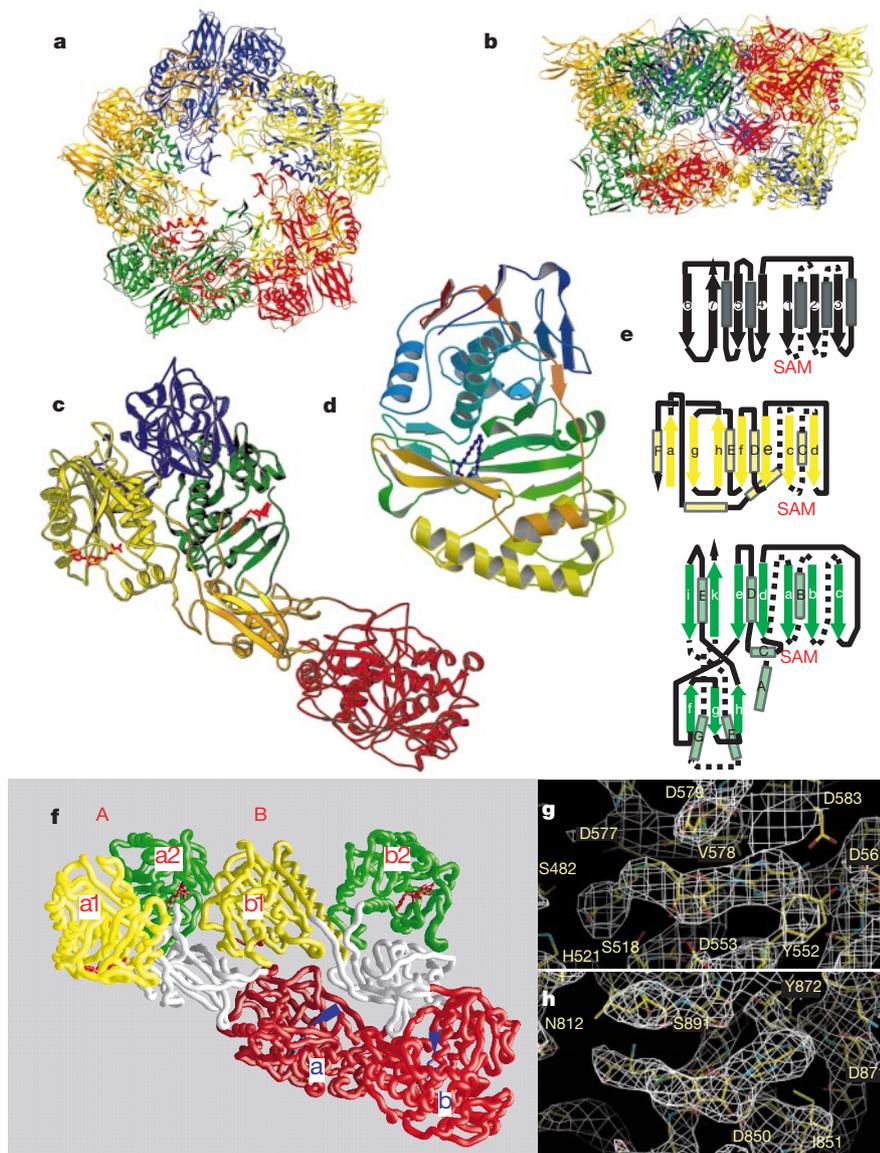


Figure 6 The capping complex. **a**, **b**, λ 2 turret (about 120 Å diameter and 80 Å tall) viewed from the top and side, respectively. The five elongated λ 2 monomers, each shown in a different colour, wrap around the outer surface, with their long axes at about 45° to the radial direction. **c**, The λ 2 monomer, viewed from the inside of the pentamer (the blue monomer in **b**). The GTase domain is red, methylase-1 is yellow, methylase-2 is green and the immunoglobulin-like domains are blue. Red SAH molecules mark the SAM-binding sites. **d**, GTase domain at 90° to **c**, in graded colours with the N terminus red and the C terminus blue. Side chains of K190 and K171 are shown. **e**, Diagrams of SAM-binding domains. The 'universal' SAM-binding domain is shown in black^{19,20}; methylase-1 is yellow and methylase-2 is green. Secondary structural elements are aligned vertically

with their equivalents in the universal fold. The SAM-binding position with respect to the β -sheet is labelled. **f**, Two monomers of λ 2 (labelled A and B) with the immunoglobulin-like domains detached and viewed from the interior of the turret. Monomer A includes GTase 'a' and methylases 'a1' and 'a2'; B includes 'b', 'b1' and 'b2'. Coloured as in **c**. Red SAH molecules indicate SAM-binding locations. Blue arrows indicate the GTase active site. **g**, SAH density for methylase-1 in a 4 Å, 2F_o-F_c map made with data from crystals soaked in 2 mM SAH. SAH binding is accompanied by conformational changes in residues 519–524 and 579–587. Some residues that may interact with SAH are labelled. **h**, SAH density for methylase-2.

cavity. Capping enzyme complexes that have only a single copy of each capping enzyme and that do not sequester the nascent RNA within a λ 2-like container may have some other guiding mechanism to order the progress of the 5' terminus from one active site to the next.

Interior of the core

The inward-facing surface of the λ 1 shell defines a cavity with a mean outer radius of 245 Å. The N-terminal extensions of λ 1 are the most significant irregularities in the otherwise smooth cavity boundary. Reconstructed cryoEM images of empty virions and cores show, however, that there are structures projecting inwards near the fivefold axes—presumably the transcriptase complexes⁶. The presence of a single such complex near each fivefold axis will leave little trace in a high-resolution map.

When packaged within a hollow shell, double-helical nucleic acids form uniform coils that minimize kinks and sharp curves while maximizing the interhelix distance (and hence minimizing electrostatic repulsion)^{5,23–27}. Assuming that the entire interior of the core (after allowing for the volume occupied by non-icosahedrally ordered protein) is uniformly occupied by RNA with local hexagonal packing, the mean spacing between adjacent segments of A-form dsRNA will be about 30 Å. This figure agrees with the spacing predicted from the 26 Å diffraction maximum seen in solution X-ray scattering from intact type 3 reovirus particles²⁴. A similar observation has been reported for BTV cores⁷.

Although we do not see a strong 26 Å diffraction ring, in low-resolution maps we observe three or four shells of density, spaced at 26 Å and layered inwards from the interior surface of λ 1. The shells are weak in the region beneath each fivefold vertex and stronger elsewhere. We interpret these observations, together with those from earlier experiments²⁴, as indicating that dsRNA is coiled within the shell. We expect that an outer layer of RNA will press against λ 1 and that subsequent layers will derive their approximate coherence from the balance of forces described above. Thus, the RNA exhibits significant statistical order, but there is no evidence for unique three-dimensional packing. The presence of a transcriptase complex tethered near each fivefold vertex will tend to exclude RNA from the vicinity of the fivefold axes, and these complexes may even act as spools for coiling individual RNA segments²⁸. Thus, we expect the shells of density to be weak beneath each turret, as observed.

Comparisons with other dsRNA viruses

The transcriptionally competent ICPs from members of the nine known genera of plant and animal dsRNA viruses in the family *Reoviridae* all appear to share a common, λ 1-like framework, which packages the genome and organizes the polymerization and capping enzymes. The organization of this shell may extend to fungal and bacterial viruses with dsRNA genomes^{3,28,29}. The equivalent framework in the orbivirus core, which has been seen at high resolution³, consists of 120 copies of VP3, which is arranged identically to λ 1. Moreover, λ 1 (residues 240–1,275) and VP3 have the same plate-like shape, and although they have no sequence similarity, they have similar overall folds with corresponding parts of their polypeptide chains forming corresponding portions of each molecule. In each protein, an α -helical bundle lies closest to the icosahedral fivefold axis, a β -sheet 'domain' lies most distally from that axis, and a lattice of criss-crossing α -helices lies in the middle. The details of the fold differ, however, and it is not possible meaningfully to superimpose any significant substructure of λ 1 onto VP3. Furthermore, whereas the two conformers of λ 1 differ in the relative orientations of subdomains I and II about one pivot, the two conformers of VP3 differ in the relative orientations of three subdomains about two such pivots, neither of which corresponds to the λ 1 pivot.

Beyond the interior framework, the structures of the dsRNA viruses diverge, reflecting different mechanisms for viral maturation and entry. The mammalian orthoreoviruses have turreted ICPs: the capping complexes (λ 2) are ordered structures displayed on the outside of the λ 1 shell. There are also clamp subunits (σ 2) to reinforce λ 1 interactions. Members of four other dsRNA genera also have turret-like structures that are likely to be homologues of λ 2. Two of these genera, the aquareoviruses³⁰ and the cypoviruses^{8,31}, have been studied by cryoEM. They have nodules at positions corresponding to σ 2-i and σ 2-ii, where the protein at position ii is probably a clamp required for assembly⁸. In contrast, the orbiviruses (and two other genera) have non-turreted ICPs coated with a complete $T = 13$ lattice of a non-enzymatic protein (VP7 and VP6, respectively)^{3,32}, and their capping enzymes lie inside the common shell, perhaps in close association with the polymerases.

Discussion

We draw four conclusions from the above data. First, there is a

Table 1 Crystallographic statistics

Native data					
Resolution range (Å)	R_{sym} (%)	R_{cum} (%)	Completion (%)	Redundancy	I/σ
70–9.82	8.9	8.9	99.4	6.8	6.5
9.82–6.95	13.1	11.1	99.7	6.6	5.3
6.95–5.67	25.0	14.9	99.7	6.3	2.9
5.67–4.91	25.9	17.9	99.7	6.2	2.9
4.91–4.39	25.7	20.0	99.6	6.0	2.9
4.39–4.01	35.7	22.9	99.4	5.6	2.1
4.01–3.71	47.9	25.3	97.4	4.1	1.6
3.71–3.47	53.0	26.0	70.5	2.1	1.4
SAH soaking data					
Resolution range (Å)	R_{sym} (%)	R_{cum} (%)	Completion (%)	Redundancy	I/σ
70–10.44	7.4	7.4	57.7	1.5	8.3
10.44–7.38	11.5	9.7	57.7	1.5	6.4
7.38–6.03	28.0	14.4	56.9	1.5	2.7
6.03–5.22	33.8	18.8	55.3	1.5	2.2
5.22–4.67	31.7	21.6	51.2	1.4	2.3
4.67–4.26	35.7	23.9	45.0	1.4	2.1
4.26–3.95	49.7	26.1	37.7	1.3	1.5
Refinement: R -value for working set					
Resolution range (Å)	20–3.6	(3.64–3.6)			
R (%)	20.81	(30.06)			
No. of reflections	831,984	(23,226)			

$R_{\text{sym}} = \frac{\sum_i \sum_j |I_i(h) - I_j(h)|}{\sum_i \sum_j I_i(h)}$, where $I_i(h)$ is the i th measurement of reflection h and $\langle I(h) \rangle$ is the weighted mean of all measurements of h . $R = \frac{\sum_n |F_{\text{obs}}(h) - k|F_{\text{calc}}(h)|}{\sum_n |F_{\text{obs}}(h)|}$, where k is a scale factor.

striking and unexpected degree of non-equivalence in the inter-subunit contacts that stabilize the shell. Second, the RNA-capping complex contains redundant copies of each active site, facing into the cavity of a hollow cylinder. Temporary retention of the 5' terminus within the cavity probably ensures efficient capping. Third, genomic RNA is tightly packed within the interior of the core, leading to a statistical ordering. Repeated passage of each genome segment past a polymerase tethered near the exit point for the transcript will require a concerted set of activities and motions. Finally, known dsRNA viruses are constructed around a conserved framework, the homologue of $\lambda 1$, with considerable variability in the way other viral proteins, which function in entry and exit, associate with the common shell. □

Methods

We prepared cores of reovirus strain F18 as described^{2,33}. For crystallization, cores (6 mg ml⁻¹ in 0.5 M NaCl, 0–0.3 M NaOAc, 50 mM MgCl₂, 50 mM Hepes, pH 7) were equilibrated against 1 M NaCl, 0–0.6 M NaOAc, 100 mM HEPES, pH 7.0. Crystals, 150–200 μ m at their largest dimensions, grew in 9–12 months. They were transferred to a solution containing 0.8 M NaOAc (but otherwise identical to the well solution) and mounted in capillaries. The space group is *F*432, *a* = 1,255 Å. Data were collected at 0 °C at CHESS F1. One 0.12° wedge of data was collected from each crystal volume, defined by a 0.1 mm collimator, during an exposure of 40 s. We recorded data between 65 and <3.6 Å on two Fuji image plates mounted in a specially constructed double-cassette holder placed 700 mm from the crystal. More than 500 such exposures contributed to the final data set; many others were discarded. Each double image was indexed and integrated with Denzo³⁴; intensities were scaled using Scala and Postref in the CCP4 program suite³⁵. Reflections more than 75% complete were scaled up. Statistics are shown in Table 1. The merged data were sharpened with a *B* factor of –55.

We used a 27 Å cryoEM reconstruction of the core⁹ as a phasing start. To minimize the effect of contrast-transfer-function errors, we converted the cryoEM map into a binary map to represent protein/non-protein regions. We used MAVE (RAVE software package³⁶) to position the model in the *F*432 asymmetric unit. The translation and orientation of each particle was fixed by packing and space group. Each particle was centred at (1/4, 1/4, 1/4) and related positions, with icosahedral two- and threefold axes aligned with crystallographic ones. There are two such orientations, and they correspond to a choice of origin. One pentameric moiety of the icosahedral core lies in each *F*432 asymmetric unit. We adjusted the magnification of the model to maximize correlation and to minimize *R* factor between *F*_{calc} and *F*_{obs} in a resolution range 65–27 Å (*R* = 46.5%, Corr = 59.8%). RAVE was used both to construct an envelope³⁷ and, with CCP4, for all averaging and phase extension. Minor modifications of RAVE and CCP4 were necessary. We carried out phase extension in steps of one reciprocal lattice point. Unmeasured reflections, including all at resolutions lower than 65 Å, were substituted with calculated values, and weights for the terms in *2F*_o–*F*_c maps were assigned in SIGMAA. To extend from 4 to 3.6 Å, some portions of the asymmetric unit ($\sigma 2$ -i and ii, and subdomains I and II of $\lambda 1$) were averaged tenfold. The resulting maps were unambiguous, and models for $\lambda 2$ and two versions of $\lambda 1$ and $\sigma 2$ were built using program O³⁸; a partial model of $\sigma 2$ -ii was docked into $\sigma 2$ -iii density during refinement. We determined Pt sites from 1.5° of an 8 Å data set obtained from crystals soaked in 1 mM of K₂PtCl₄; they correspond to methionine and histidine positions in the structure. Positional refinement, torsion angle dynamics and restrained individual *B*-factor refinement were carried out in CNS³⁹, maintaining strict fivefold restraints. To speed calculations, we carried out refinement in a primitive cell with smaller unit-cell dimensions (*a* = *b* = *c* = 887.42 Å, α = β = γ = 60°) rather than the face-centred cell³⁹ (R. Grosse-Kunstleve, personal communication). Portions of the structure were also refined in small P1 cells using artificial *F*_{obs} calculated from portions of an averaged *2F*_o–*F*_c map in *F*432 (ref. 40). The cross-validation set was not used in calculating these averaged maps. The final *R* factor is 20.8% (*R*_{free} = 20.6%; Table 1). The r.m.s. deviation from ideal bond lengths is 0.009 Å; that from ideal angles is 1.3°. Of all residues, 99.4% fall in either the most favoured or allowed region of the Ramachandran plot, and only 3 out of 4,733 fall in the disallowed region⁴¹.

To determine SAH-binding sites, we soaked crystals in mounting buffer supplemented with 2 mM SAH for several days. A partial data set (60–4 Å) was collected; *2F*_o–*F*_c and *F*_o–*F*_c difference maps were calculated and averaged fivefold using a liberal envelope. Strong density appeared in the SAM-binding locations. Difference density also clearly indicated a backbone movement for residues 579–587 and 519–524 in the methylase-1 domain.

We computed low-resolution maps for visualizing RNA density using additional data between 200 and 50 Å collected on BioCARS beamline 14-ID-B at the Advanced Photon Source. Phase extensions starting at 33 Å and at 27 Å and using a variety of models and final resolution limits yielded similar results: density ripples spaced at 25–26 Å intervals from the inner surface of the core and interrupted under the fivefold axes.

Figures were prepared in Ribbons⁴², Molscrip⁴³, Raster3D⁴⁴, GRASP⁴⁵ and O³⁸.

Received 6 January; accepted 15 March 2000.

- Fields, B. N. in *Fields Virology* (eds Fields, B. N., Knipe, D. M. & Howley, P. M.) 1553–1555 (Lippincott-Raven, Philadelphia, 1996).
- Nibert, M. L., Schiff, L. A. & Fields, B. N. in *Fields Virology* (eds Fields, B. N., Knipe, D. M. & Howley, P. M.) 1557–1596 (Lippincott-Raven, Philadelphia, 1996).

- Grimes, J. M. *et al.* The atomic structure of the bluetongue virus core. *Nature* **395**, 470–478 (1998).
- Furuichi, Y., Muthukrishnan, S., Tomasz, J. & Shatkin, A. Mechanism and formation of reovirus mRNA 5'-terminal blocked and methylated sequence ^{m7}GpppC^{m7}pC. *J. Biol. Chem.* **251**, 5043–5053 (1976).
- Earnshaw, W. C. & Harrison, S. C. DNA arrangement in isometric phage heads. *Nature* **268**, 598–602 (1977).
- Dryden, K. A. *et al.* Internal structures containing transcriptase-related proteins in top component particles of mammalian orthoreovirus. *Virology* **245**, 33–46 (1998).
- Gouet, P. *et al.* The highly ordered double-stranded RNA genome of bluetongue virus revealed by crystallography. *Cell* **97**, 481–490 (1999).
- Hill, C. L. *et al.* The structure of cypovirus and the functional organization of dsRNA viruses. *Nature Struct. Biol.* **6**, 565–568 (1999).
- Dryden, K. A. *et al.* Early steps in reovirus infection are associated with dramatic changes in supramolecular structure and protein conformation: analysis of virions and subviral particles by cryoelectron microscopy and image reconstruction. *J. Cell Biol.* **122**, 1023–1041 (1993).
- Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A. & Steitz, T. A. Crystal structure of 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256**, 1783–1790 (1992).
- Stehle, T., Yan, Y., Benjamin, T. L. & Harrison, S. C. Structure of murine polyomavirus complexed with an oligosaccharide receptor fragment. *Nature* **369**, 160–163 (1994).
- Liddington, R. C. *et al.* Structure of simian virus 40 at 3.8 Å resolution. *Nature* **354**, 278–284 (1991).
- Labbe, M., Charpilienne, A., Crawford, S. E., Estes, M. K. & Cohen, J. Expression of rotavirus VP2 produces empty core-like particles. *J. Virol.* **65**, 2946–2952 (1991).
- Moss, S. R. & Nuttall, P. A. Subcore- and core-like particles of Broadhaven virus (BRDV), a tick-bourne orbivirus, synthesized from baculovirus expressed VP2 and VP7, the major core proteins of BRVD. *Virus Res.* **32**, 401–407 (1994).
- Xu, P., Miller, S. & Joklik, W. K. Generation of reovirus core-like particles in cells infected with hybrid vaccinia viruses that express genome segments L1, L2, L3, and S2. *Virology* **197**, 726–731 (1993).
- Mao, Z. & Joklik, W. K. Isolation and enzymatic characterization of protein $\lambda 2$, the reovirus guanylyltransferase. *Virology* **185**, 377–386 (1991).
- Luongo, C. L., Reinisch, K. M., Harrison, S. C. & Nibert, M. L. Identification of the guanylyltransferase region and active site in reovirus mRNA capping protein 12. *J. Biol. Chem.* **275**, 2804–2810 (2000).
- Håkansson, K., Doherty, A. J., Shuman, S. & Wigley, D. B. X-ray crystallography reveals a large conformational change during guanyl transfer by mRNA capping enzymes. *Cell* **89**, 545–553 (1997).
- Schluckebier, G., O'Gara, M., Saenger, W. & Cheng, X. Universal catalytic domain structure of AdoMet-dependent methyltransferases. *J. Mol. Biol.* **247**, 16–20 (1995).
- Hodel, A. E., Gershon, P. D., Shi, X. & Quijcho, F. A. The 1.85 Å structure of the vaccinia protein VP39: a bifunctional enzyme that participates in the modification of both mRNA ends. *Cell* **85**, 247–256 (1996).
- Hu, G., Gershon, P. D., Hodel, A. E. & Quijcho, F. A. mRNA cap recognition: dominant role of enhanced stacking interactions between methylated bases and protein aromatic side chains. *Proc. Natl Acad. Sci. USA* **96**, 7149–7154 (1999).
- Wickner, S., Maurizi, M. R. & Gottesman, S. Posttranslational quality control: Folding, refolding, and degrading proteins. *Science* **286**, 1888–1893 (1999).
- Earnshaw, W. C., King, J., Harrison, S. C. & Eiserling, F. A. The structural organization of DNA packaged within the heads of T4 wild-type, isometric and giant bacteriophages. *Cell* **14**, 559–568 (1978).
- Harvey, J. D., Bellamy, A. R., Earnshaw, W. C. & Schutt, C. Biophysical studies of reovirus type 3: iv low-angle x-ray diffraction studies. *Virology* **112**, 240–249 (1981).
- Harrison, S. C. Packaging of DNA into bacteriophage heads: a model. *J. Mol. Biol.* **171**, 577–580 (1983).
- Cerritelli, M. E. *et al.* Encapsidated conformations of bacteriophage T7 DNA. *Cell* **91**, 271–280 (1997).
- Booy, F. P. *et al.* Liquid crystalline, phage-like packing of encapsidated DNA in herpes simplex virus. *Cell* **64**, 1007–1015 (1991).
- Butcher, S. J., Dokland, T., Ojala, P. M., Bamford, D. H. & Fuller, S. D. Intermediates in the assembly pathway of the double-stranded RNA virus $\phi 6$. *EMBO J.* **16**, 4477–4487 (1997).
- Cheng, R. H. *et al.* Fungal virus capsids, cytoplasmic compartments for the replication of double-stranded RNA, formed as icosahedral shells of asymmetric Gag dimers. *J. Mol. Biol.* **244**, 255–258 (1994).
- Shaw, A. L., Samal, S. K., Subramanian, K. & Prasad, B. V. V. The structure of aquareovirus shows how the different geometries of the two layers of capsid are reconciled to provide symmetrical interactions and stabilization. *Structure* **4**, 957–967 (1996).
- Zhang, H. *et al.* Visualization of protein–RNA interactions in cytoplasmic polyhedrosis virus. *J. Virol.* **73**, 1624–1629 (1999).
- Prasad, B. V., Wang, G. J., Clerx, J. P. & Chiu, W. Three-dimensional structure of rotavirus. *J. Mol. Biol.* **199**, 269–275 (1988).
- Coombs, K. M., Fields, B. N. & Harrison, S. C. Crystallization of the reovirus type 3 Dearing core. *J. Mol. Biol.* **215**, 1–5 (1990).
- Otwiński, Z. & Minor, W. in *Macromolecular Crystallography A* (eds Carter, C. W. & Sweet, R. M.) 307–326 (Academic, New York, 1997).
- CCP4. The CCP4 suite: programs for X-ray crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Jones, T. A. in *Molecular Replacement* (eds Dodson, E. J., Gover, S. & Wolf, W.) 91–105 (SERC Daresbury Laboratory, Warrington, 1992).
- Kleywegt, G. J. & Jones, T. A. in *From First Map to Final Model* (eds Bailey, S., Hubbard, R. & Waller, D.) 59–66 (SERC Daresbury Laboratory, Warrington, 1994).
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
- Brunger, A. T. *et al.* Crystallography & NMR System: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
- Jacobsen, D. H., Hogle, J. M. & Filman, D. J. A pseudo-cell based approach to efficient crystallographic refinement of viruses. *Acta Crystallogr. D* **52**, 693–711 (1996).
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).

42. Carson, M. in *Macromolecular Enzymology* (eds Carter, C. W. Jr & Sweet, R. M.) 493–505 (Academic, New York, 1996).
43. Esnouf, R. M. An extensively modified version of Molscript that includes greatly enhanced coloring capabilities. *J. Mol. Graphics* **15**, 133–138 (1997).
44. Merrit, E. A. & Murphy, M. E. P. Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr. D* **50**, 869–873 (1994).
45. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11**, 281–296 (1991).
46. Harrison, S. J. *et al.* Mammalian reovirus L3 gene sequences and evidence for a distinct amino-terminal region of the lambda1 protein. *Virology* **258**, 54–64 (1999).

Acknowledgements

We thank T. F. Severson, R. L. Margraf and S. J. Harrison for growing reovirus-infected cells; T. S. Baker and S. B. Walker for providing us with an unpublished cryoEM reconstruction of cores from reovirus reassortant F18; D. Thiel and the staff of CHESS and

McCHESS and R. Pahl and the staff of the BioCars beamline 14-ID-B for assistance; members of the Harrison laboratory, especially J. Cruzan, for scanning image plates; W. Minor for help with data collected with an early version of the double image plate cassette holder; B. Miller for constructing an improved vacuum-based version; S. Ray for advice on computing strategies; P. Adams and R. Grosse-Kunstleve for advice on refinement; and R. Grosse-Kunstleve for implementing refinement in the smaller rhombohedral cell. K.M.R. thanks D. W. Rodgers for help in the initial cryopreservation attempts, for suggesting a vacuum-based design for the double image plate cassette holder and for advice. We acknowledge support from the NIH (to S.C.H. and M.L.N.). S.C.H. is an investigator in the Howard Hughes Medical Institute. M.L.N. receives support from a grant of the Lucille P. Markey Charitable Trust to the Institute for Molecular Virology and, as a Shaw Scientist, from the Milwaukee Foundation. This work was started with the stimulus and collaboration of the late B. N. Fields.

Correspondence and requests for material should be addressed to S.C.H. (e-mail: harrison@crystal.harvard.edu). The atomic coordinates have been deposited in the RCSB Protein Data Bank under accession number 1EJ6.